BIOMETRICS 60, 216–224
March 2004

Exact Analysis of Dose Response for Multiple Correlated Binary Outcomes

Karen E. Han,[1,*] Paul J. Catalano,[1,2] Pralay Senchaudhuri,[3] and Cyrus Mehta[1,3]

[1]Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115, U.S.A.
[2]Department of Biostatistical Science, Dana Farber Cancer Institute, Boston, Massachusetts 02115, U.S.A.
[3]Cytel Software Corporation, Cambridge, Massachusetts 02139, U.S.A.
*email: khan@jimmy.harvard.edu

SUMMARY. The neurotoxicity of a substance is often tested using animal bioassays. In the functional observational battery, animals are exposed to a test agent and multiple outcomes are recorded to assess toxicity, using approximately 40 animals measured on up to 30 different items. This design gives rise to a challenging statistical problem: a large number of outcomes for a small sample of subjects. We propose an exact test for multiple binary outcomes, under the assumption that the correlation among these items is equal. This test is based upon an exponential model described by Molenberghs and Ryan (1999, *Environmetrics* **10**, 279–300) and extends the methods developed by Corcoran et al. (2001, *Biometrics* **57**, 941–948) who developed an exact test for exchangeably correlated binary data for groups (clusters) of correlated observations. We present a method that computes an exact p-value testing for a joint dose–response relationship. An estimate of the parameter for dose response is also determined along with its 95% confidence bound. The method is illustrated using data from a neurotoxicity bioassay for the chemical perchlorethylene.

KEY WORDS: Binary data; Exact test; Multiple outcomes; Network algorithm.

1. Introduction

Analyzing multiple outcomes is an issue that arises quite frequently in biomedical research. When conducting animal bioassays to assess the toxicity of a substance, subjects are often evaluated on many different outcomes. The results must then be used to establish an overall assessment of risk and, if any, its relationship with dose. Such is the case when conducting experiments to evaluate whether or not a substance has an adverse effect on neurological function. Exposure to neurotoxins in environmental and occupational settings is a serious public health concern. Examples include the chemical *n*-hexane, which may produce severe reactions including irritation of airway membranes, dizziness, headache, and nausea. Chronic exposure produces weakness, numbness, and even paralysis in outer extremities. Chronic lead exposure is also known to affect central nervous system activity through interference of neurotransmitter function. The United States Environmental Protection Agency (U.S. EPA) has issued guidelines for neurotoxicity risk assessment that provide a methodology for the acquisition of information for assessing the risks to humans from exposure to chemicals that may produce neurotoxicity (Dorman, Breneman, and Struve, 1999). Controlled investigations of the generalized neural effects of substances usually take the form of an animal bioassay. The functional observational battery (FOB) is a neurotoxicity screening assay composed of approximately 30 descriptive endpoints designed to evaluate the neurophysiologic effects of a test chemical (Moser, 1989; Baird et al., 1997). This battery of tests has been used to compute NOAELS (no adverse effect levels), dose–response models, and to perform quantitative risk assessment.

Neurotoxicity assays require animals to be tested on each element in the FOB, as the large number of endpoints may be necessary to account for the numerous possibilities of neurological effects measured by behavioral responses. Obviously, some of these outcomes may be more closely related to each other. It is likely that some endpoints describe similar effects, thus leading to redundancy. Moser (1992) has introduced a method for grouping the individual endpoints into six neurophysiologically based functional domains, each of which contains a variety of discrete and continuous endpoints, indicating target areas of neurological dysfunction. These domains are, in theory, mutually exclusive in the sense that each measures a separate and distinct component of neurotoxicity, and are described in more detail in Appendix A. If the endpoints are grouped properly the result may be a lower false negative rate and thus increased efficiency. However if the groupings are inappropriate, the averaging of responses may decrease the ability to identify true effects, and dependence across domains may increase.

1.1 *Background: Univariate, Latent Variable, and Random Effects Models*

Early analysis of the FOB data includes the work of Creason (1989) who proposed a univariate test on the raw data from the FOB, which in its original form is a mixture of

continuous and discrete outcomes. This method gives a test statistic for each item in the FOB, resulting in overall type I error higher than the type I error for each test, warranting a correction for multiple observations. Such a correction method (i.e., Bonferroni) could reduce the critical value to an overly conservative significance level given the number of outcomes and small sample sizes (Lefkopoulou and Ryan, 1993). Blackwell and Catalano (1999) developed a random effects latent variable model for ordinal data and applied their model to FOB neurotoxicity studies under the domain setting. The six classification domains can be viewed as distinct, independent latent variables, explained by the relationships between the manifest variables, or the observed endpoints of the FOB.

Determining the extent of correlation between endpoints, within a domain, and across domains is important for the development of any multivariate model for the FOB. The large number of endpoints, repeated measures, and multiple data types evaluated in the FOB complicate statistical analysis with complex correlation patterns. One way to achieve this could be by modeling the latent variable as a function of dose. Alternatively, a random effects model could also be used to analyze the effect of dose on response. Unfortunately, many of these models are computationally complex and rely heavily on asymptotic results for inference. One way to address this issue is to reduce the dimensionality of the parameters of these models. Not all parameters are useful in terms of the ultimate goal of making inferences on the dose–response relationship between an exposure and its potential neurotoxic effects. For example, we may not have as much interest in the value of the intercepts or correlation parameters as in the coefficient for dose. Conditional inference methods have been used to eliminate unnecessary parameters from models, and to develop methods for exact tests more appropriate for small sample sizes. In order to do this, we use the sufficiency principle and condition on the sufficient statistics of the nuisance parameters to eliminate them from the likelihood. Unfortunately, due to the complexity of the latent variable and random effects models and the fact that they do not have exponential family form, determining the sufficient statistics for the nuisance parameters in these models is not feasible. Instead, we extend an exact conditional method developed previously to test trend for correlated binary data (Corcoran et al., 2001). In the following sections we describe the model and proceed to develop an exact analysis for dose response. The algorithm to obtain an exact p-value, an estimate of the dose–response parameter, and its 95% confidence interval is detailed in Appendix B and examples using data obtained from FOB studies are presented.

2. An Exponential Family Model for Binary Data

The Molenberghs and Ryan (1999; MR) model developed for clustered multivariate binary data provides a useful framework for exact inference. The model is straightforward and less complex than two-stage latent variable or random effects models, and because it is a member of the exponential family its sufficient statistics are easily obtained. We can therefore condition out unnecessary nuisance parameters and focus on the testing and estimation of the one of interest, the dose–response parameter.

Based on the biology of the problem, previous analysis of the FOB's six domains focused on modeling them as independent latent variables. Similarly, we divide the data based on these groupings, and analyze outcomes from each individual domain. Due to computational requirements of our proposed method, the analysis is performed one domain at a time. Correlation between endpoints exists both within and across domains and we make the assumption that within a domain, outcomes are equally correlated. Relying on the validity of the domain classification, it is reasonable to believe that the correlation between any two outcomes in the same domain is in fact similar. Following a modification of Moser's (1992) classification system, we treat each of the FOB responses as binary outcomes, corresponding to response/nonresponse indicators of toxicity.

2.1 *The Model*

Suppose we have $i = 1, \ldots, N$ animals, each measured on $j = 1, \ldots, M$ outcomes. Let y_{ij} be the binary outcome for the ith animal on outcome j. Each animal has an associated nominal covariate, d_i, which in our case is the dose level for animal i. Hypothetical results of such an experiment are given in Table 1.

The MR model specifies the density of $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{iM})$ as

$$P(\mathbf{Y}_i) = \exp\left\{\sum_{j=1}^{M}(\tau_j + \beta d_i)y_{ij} + \sum_{j<j'}\omega y_{ij}y_{ij'} - A_i(\tau_j, \beta, \omega)\right\},$$

where ω represents the association between any two outcomes and $A_i(\tau_j, \beta, \omega)$ is the normalizing constant. The coefficient for the binary outcomes is a linear function of dose, allowing for a separate intercept for each of the M outcomes and β is the parameter describing the effect of dose response.

Let $z_i = \sum_j y_{ij}$, the number of positive responses for animal i. After some algebra and rearranging of terms, the density can be expressed as

$$P(\mathbf{Y}_i) = \exp\left\{\sum_j \tau_j y_{ij} + \beta\sum_j d_i y_{ij} + \omega z_i(z_i - 1) - A_i(\tau_j, \beta, \omega)\right\}.$$

Our interest is in the dose–response parameter β. The τ_j's and ω are nuisance parameters which we choose to condition out of the full distribution.

Table 1

Contrived data for a hypothetical experiment with four animals measured on two binary outcomes, and associated dose d_i

Animal (i)	Dose (d_i)	Outcome (y_{ij})		z_i	$z_i(z_i-1)$
1	0	1	0	1	0
2	50	0	0	0	0
3	100	1	1	2	2
4	500	1	0	1	0
		$s_1 = 3$	$s_2 = 1$		$r = 2$

3. The Exact Test

To formulate the exact test based on the algorithms developed by Mehta, Patel, and Senchaudhuri (1992; MPS), we first eliminate all unnecessary parameters from the likelihood. The unconditional likelihood is expressed as

$$P(\mathbf{Y}) = \exp\left\{\sum_{j=1}^{M}\tau_j s_j + \beta\sum_{i=1}^{N} d_i z_i + \omega\sum_{i=1}^{N} z_i(z_i-1) - A(\tau,\omega,\beta)\right\}, \quad (1)$$

where $s_j = \sum_i y_{ij}$ and $z_i = \sum_j y_{ij}$. Because this density is of exponential family form, each τ_j has sufficient statistic $s_j = \sum_i y_{ij}, j = 1,\ldots,M$, and sufficient statistics for β and ω are $t = \sum_i d_i z_i$ and $r = \sum_i z_i(z_i-1)$, respectively.

We can eliminate each of the M nuisance intercept parameters (τ_j) and ω to obtain the exact distribution of $\mathbf{Y}$ under the null hypothesis, by conditioning on $\mathbf{s} = (s_1, s_2, \ldots, s_M)$ and r. We must first define the reference sets that correspond to all possible values of the data under these constraints, which can be represented as

$$\Gamma_{\mathbf{s}} = \left\{\mathbf{y}^* : \sum_{i=1}^{N} y_{ij}^* = s_j, j = 1,\ldots,M\right\}$$

and

$$\Gamma_{\mathbf{s},r} = \left\{\mathbf{y}^* \in \Gamma_{\mathbf{s}} : \sum_{i=1}^{N} z_i^*\left(z_i^*-1\right) = r\right\},$$

where $\mathbf{y}^*$ is any generic table of the form shown above, and $z_i^* = \sum_{j=1}^{M} y_{ij}^*$.

The conditional likelihood given $\mathbf{s}$ and r is then

$$P(\mathbf{Y}\,|\,\mathbf{s},r) = \frac{\exp\left\{\sum_{j=1}^{M}\tau_j s_j + \beta t + \omega r + A(\tau,\omega,\beta)\right\}}{\sum_{\mathbf{y}^*\in\Gamma_{\mathbf{s},r}}\exp\left\{\sum_{j=1}^{M}\tau_j s_j + \beta t^* + \omega r + A(\tau,\omega,\beta)\right\}}. \quad (2)$$

Cancellation eliminates all nuisance parameters, reducing (2) to

$$P(\mathbf{Y}\,|\,\mathbf{s},r) = \frac{\exp\{\beta t\}}{\sum_{\mathbf{y}^*\in\Gamma_{\mathbf{s},r}}\exp\{\beta t^*\}}. \quad (3)$$

In order to obtain the exact p-value for β we must find the distribution of t, the sufficient statistic for β, conditional on the observed values of $\mathbf{s}$ and r. Let $C(t\,|\,\mathbf{s},r)$ be the count of the binary sequences $\mathbf{y}$ such that $\mathbf{y}\in\Gamma_{\mathbf{s},r}$ and $\sum_i d_i z_i = t$. The conditional distribution of t is then

$$P(t\,|\,\mathbf{s},r) = \sum_{\substack{\mathbf{y}\in\Gamma_{\mathbf{s},r}:\\ \sum_{i=1}^{N} d_i z_i = t}} P(\mathbf{Y}\,|\,\mathbf{s},r) = \frac{C(t\,|\,\mathbf{s},r)\exp(\beta t)}{\sum_{u\in\Gamma_{\mathbf{s},r}} C(u\,|\,\mathbf{s},r)\exp(\beta u)}. \quad (4)$$

Under the null hypothesis, $H_0 : \beta = 0$, this likelihood is free of all unknown parameters. Estimation of β will be discussed in Section 5. An exact test is formulated by assigning a test statistic to each ordered table in $\Gamma_{\mathbf{s},r}$. That is, conditional on the observed values of $\mathbf{s}$ and r, we find the distribution of the statistic

$$T(\mathbf{Y}) = \sum_{i}^{N} d_i z_i, \quad \text{where } z_i = \sum_{j}^{M} y_{ij}.$$

Denoting a realization of T as t, and an observed table as t_{obs}, let Γ_t be the set of all possible realizations such that $t_{\text{obs}} \le t$. Under the null, the exact probability is thus

$$P(T > t_{\text{obs}}\,|\,H_0,\mathbf{s},r) = \frac{\sum_{t^*\in\Gamma_t} C(t^*\,|\,\mathbf{s},r)}{\sum_{u\in\Gamma_{\mathbf{s},r}} C(u\,|\,\mathbf{s},r)}. \quad (5)$$

Defining t_α as the smallest value such that $P(T > t_\alpha\,|\,H_0, \mathbf{s}, r) \le \alpha$, a one-sided α-level test rejects H_0 when $t_{\text{obs}} > t_\alpha$.

The complexity of computing the p-value lies in the calculation of the denominator of (5), as the numerator is a subset of the denominator. Explicitly enumerating $\Gamma_{\mathbf{s},r}$ is impractical, so we shall directly extend the network algorithm developed by MPS and Corcoran et al. (2001) in order to identify $\Gamma_{\mathbf{s},r}$ implicitly. These authors have proposed an exact trend test for exchangeably correlated (clustered) binary data, based on the same exponential model.

4. Algorithm

4.1 *Network Representation*

We take the same approach as Corcoran et al. (2001) by dividing the task of building the network into two parts. First, we construct the representation of $\Gamma_{\mathbf{s}}$, and then modify this network to meet the additional constraints of $\Gamma_{\mathbf{s},r}$, a subset of the original network. For the initial network, any given node has successor nodes enumerated using the closed-form solution to a linear programming problem identified by MPS. Once the initial network is complete, a second network representing $\Gamma_{\mathbf{s},r}$ is constructed by eliminating the elements of the first network that do not satisfy the additional constraint of $\sum_i z_i(z_i-1) = r$. This algorithm is not closed form, however a method for recursively calculating the largest and smallest possible values of the secondary sufficient statistic r_k, over all partial paths beginning at the initial stage and ending at stage k was developed by Corcoran et al. (2001). Our method is based on this algorithm with some modification. Appendix B includes the details of these algorithms.

4.2 *Monte Carlo*

As the sample size and number of outcomes increase, the computational complexity and demand on memory and processing time also increases. Even on an exceptionally fast computer, data on 40 animals with four to five outcomes may take more than several hours to run. One way to ease some of this burden is to incorporate Monte Carlo estimation. Monte Carlo sampling can be used on the completed network that

represents the set $\Gamma_{s,r}$, resulting in a reduction of time necessary to obtain results. Tables are sampled in proportion to the density of (3) to obtain an unbiased estimate of the exact tail p-value. A sample of K tables is taken, and for the lth table k_l we let $Z_l = I\{T(k_l) \geq t_{\text{obs}}\}$. An unbiased Monte Carlo estimate of the exact p-value is thus $p_{mc} = \sum_{l=1}^{K} Z_l/K$.

5. Exact Parameter Estimation and Confidence Intervals

The algorithms described above provide a way to generate the true permutation distribution of t conditional on s and r in (4), thus allowing us to obtain an exact test for the null hypothesis of no dose effect. In addition, we can use this permutation distribution to compute an exact estimate of the parameter β and its corresponding $(1 - \alpha)$ level confidence interval.

5.1 *Conditional Maximum Likelihood Estimate*

Our approach is one of exact conditional inference, and any nuisance parameters are eliminated from the likelihood by conditioning on their sufficient statistics. Therefore a point estimate $\hat{\beta}$ is a conditional maximum likelihood estimate (CMLE) for β, and has an interpretation that is conditional on the observed values of the sufficient statistics for the nuisance parameters. It can be viewed as an estimate of a common slope on dose across the multiple outcomes and is obtained by finding the value which maximizes (4).

Let t_{obs} be the observed value of the sufficient statistic for β. The exact conditional probability that $t = t_{\text{obs}}$ given s and r is

$$f_\beta(t_{\text{obs}} \mid \mathbf{s}, r) = \frac{C(t_{\text{obs}} \mid \mathbf{s}, r)\exp(\beta t_{\text{obs}})}{\sum_{u=t_{\min}}^{t_{\max}} C(u \mid \mathbf{s}, r)\exp(\beta u)}, \tag{6}$$

where $t_{\min}$ and $t_{\max}$ are the minimum and maximum values that t can assume conditional on s and r. This probability only depends on the parameter β, as all possible values of the function C can be obtained as the network representing $\Gamma_{s,r}$ is processed to compute the exact p-value. Once the final network has been processed, the resulting set of records contains the values of u and $C(u \mid \mathbf{s}, r)$ for all $u \in [t_{\min}, t_{\max}]$. This reduces the problem of finding $\hat{\beta}$ to one of maximizing the conditional likelihood function (6) by choice of β. We use an empirical, iterative approach to determine the extremum, similar to a bisection search, known as the Golden Section search (Press et al., 1988, Chapter 10).

It is well known that when data are sparse or the sample size is small one cannot rely on methods which are based on asymptotic normality, such as in unconditional multivariate logistic regression (Hirji, Mehta, and Patel, 1987). Although our approach is computationally intensive, it does provide an unbiased estimate for the dose–response parameter in the presence of a relatively small number of subjects.

5.2 *Confidence Intervals*

Mehta and Patel (Cytel, 1996) caution that exact point estimates may be of limited use for predictive purposes, and recommend to rely more on confidence intervals when making inferences. In addition to providing the CMLE for β, we also compute its corresponding $(1 - \alpha)$ level confidence bound. Define the right tail of the distribution of t conditional on s and r at the observed value $t = t_{\text{obs}}$ as

$$G_\beta(t_{\text{obs}}) = \sum_{u=t_{\text{obs}}}^{t_{\max}} f_\beta(u \mid \mathbf{s}, r) = \sum_{u=t_{\text{obs}}}^{t_{\max}} \frac{C(u \mid \mathbf{s}, r)\exp(\beta u)}{\sum_{u^*=t_{\min}}^{t_{\max}} C(u^* \mid \mathbf{s}, r)\exp(\beta u^*)}. \tag{7}$$

The left tail can be defined similarly. A one-sided confidence bound is generated based on whether the left tail or the right tail probability is smaller, and the entire α error is used in computing the bound. In our case, we are interested in the hypothesis $H_0 : \beta > 0$, and therefore we determine a lower confidence bound β^- such that $G_{\beta^-}(t_{\text{obs}}) = \alpha$. The algorithm to compute the confidence bound is similar to that of determining the CMLE for β.

6. Example

To illustrate the use of the exact method for multiple binary outcomes, we apply the model to data obtained from a neurotoxicity bioassay. In FOB studies it is assumed that an adverse effect of a substance is manifested as a distinct change in an animal's physiological or behavioral response. We subscribe to the domain approach (Moser, 1992), which separates the outcomes into six mutually exclusive groups, each focusing on a distinct area of neurological functioning (Appendix A).

The data are from a study of the acute neurotoxic effects of perchlorethylene, a manufactured chemical which is most commonly known for its use in the dry cleaning industry. Exposure to PERC is via inhalation of contaminated air or consumption of contaminated food or water (U.S. EPA, 1994). Data are provided by Dr. Virginia Moser at the U.S. EPA. The data contain outcomes from an FOB collected at several time points, and our analysis uses those measured at the time peak effect, or four hours post exposure. Each of the 40 animals received one of five dose levels, 0, 0.15, 0.5, 1.5, or 5 g/kg of perchlorethylene, resulting in eight animals per dose group. The original data are ordinal on a scale from 1 to 4, where a score of 1 indicates a typical response for a control animal, and a score of 4 is given when an animal displays a most severe reaction. For the present analysis, the data have been recoded so that the corresponding binary outcome is 1 if an animal exhibits any adverse reaction to the exposure, and is 0 otherwise.

Table 2 provides a summary of the observed data, giving the number of animals in each dose group that showed a response for each outcome, over all functional domains. There are eight animals per dose group unless otherwise indicated. Table 3 displays the results of our method. The first column of results gives the exact p-value for testing the null hypothesis of no dose effect for each domain. In addition $\hat{\beta}$, the CMLE for β, and its corresponding 95% lower confidence bound is displayed. Unconditional MLEs and p-values obtained using generalized estimating equations (GEEs) similar to the method used in Hudson et al. (2001) are also included for comparison. On average, the exact method produces more conservative results than the asymptotic based method (i.e., slightly larger

Table 2
Counts of events. Values are out of eight animals per dose group unless otherwise indicated.

Domain	Outcome	Dose (g/kg) 0	0.15	0.5	1.5	5
Autonomic	Lacrimation	0	2	0	3	7
	Salivation	0	0	0	0	3
	Pupil response	1	1	0	3	8
	Defecation	1	1	0	1	0
	Urination	4	2	4	2	3
Sensorimotor	Approach response	0	1	2	4	8
	Click response	1	1	1	1	7
	Tail pinch response	2	1	3	3	8
	Touch response	0	3	0	2	7
CNS excitability	Handling	2	5	1	4	7
	Clonic convulsions	4	3	3	3	3
	Arousal	4	3	4	5	8
CNS activity	Home cage posture	0	0	0	1	6
	Rearing	3	1	1	4	3
Neuromuscular	Gait	0	1	0	5	8
	Foot splay	2	2	1	2	8
	Forelimb grip[a]	3	4	3	6	–
	Hindlimb grip[a]	3	4	4	8	6 of 6
	Righting reflex	0	1	1	3	8
Physiological	Piloerection	0	0	0	0	1
	Weight	2	3	4	4	3
	Temperature	2	6	4	4	8

[a] All eight outcomes were missing for forelimb grip and two outcomes were missing for hindlimb grip at dose 5 g/kg. The analysis for the neuromuscular domain is based on complete data, i.e., 32 animals at four doses.

p-values), which is not surprising (Corcoran et al., 2001). Note that for two of the domains, MLEs could not be estimated due to the small sample size of the FOB data. The next three columns of results display the exact p-value for each individual outcome within its domain. This corresponds to results using exact logistic regression, which can also be obtained by using the software LogXact. The individual estimates for β, the dose–response parameter for each univariate case, are displayed in the final column of results. Note that the estimates of β are negative for urination, defecation, and clonic convulsions and have nonsignificant p-values. Had these been significant, it would suggest that a slight trend in the opposite direction exists, implying that as dose increases, the probability of an adverse reaction would decrease. Tests that had no variability (removal, tonic convulsions, palpebral closure) do not contribute to the domain level p-value and are not displayed.

Each domain except the marginally significant physiological domain indicates that there is a significant dose–response relationship. These results support the notion that there is indeed some underlying mechanism of toxicity that is related to dose. Upon further investigation of the individual outcomes, we see that although two yield nonsignificant p-values, all have an expected positive dose–response pattern. These results indicate that as the dose increases, there is a higher probability of having a negative reaction to perchlorethylene. However, for some individual outcomes a dose–response trend is not clear. For example urination, clonic convulsions, and weight all have p-values around 0.5. The observed p-value of 0.77 for defecation may suggest a nonsignificant trend in the opposite direction for that outcome, although it is not strong enough to prevent an overall marginally significant trend for that domain. The results for individual outcomes should be taken into consideration when drawing conclusions for the domain level p-value. It may be unwise to rely on an estimate of $\hat{\beta}$ at the domain level when there are discrepancies among the estimates of $\hat{\beta}$ for individual outcomes. Although it appears that the domain level $\hat{\beta}$ is roughly the average of the individual $\hat{\beta}$, this may not always be the case. For example, the estimated effect for the sensorimotor domain is much smaller than the univariate estimates. This could be due to the fact that the underlying model implicitly accounts for correlation among the outcomes, which factors into the domain level estimate of $\hat{\beta}$. The FOB was originally designed as a screening battery, but due to the expense of administration it is important to make use of the data for identification and evaluation of neurotoxins. In that regard, chemicals showing marginally significant p-values based on our method may warrant further investigation.

7. Discussion

Our model has a significant advantage over previous work done to analyze FOB data as it is the first to propose a method appropriate for multiple outcomes given the small number of animals used in these assays. Given that asymptotic

Table 3
Results for the chemical perchlorethylene

Domain	Multivariate exact p-value (uncond. p-value)	Exact $\hat{\beta}$ {95% LCB[a]} (uncond. $\hat{\beta}$)	Individual outcome	Univariate exact p-value	$\hat{\beta}$
Autonomic	.052[b]	0.165	Lacrimation	.00002	0.86
		{−.007}	Pupil	<.00001	1.4
			Urination	.581	−0.03
			Defecation	.773	−0.46
	(<.0001)	(.5)	Salivation	.006	0.69
Sensorimotor	.0002	.322	Approach	<.00001	1.69
		{.122}	Click	.0001	0.78
			Tail pinch	.0001	0.98
	(<.0001)	(.9)	Touch	.0001	0.77
CNS excitability	.053	.157	Handling	.006	0.51
		{−.005}	Clonic	.601	−0.04
	(NA)	(DNC)[c]	Arousal	.002	0.81
CNS activity	.003	.450	Posture	.00001	1.04
		{.145}	Rearing	.204	0.14
	(NA)	(DNC)[c]			
Neuromuscular	.002[b]	.797	Gait	<.00001	2.63
		{.187}	Foot splay	.0001	0.9
			Forelimb[d]	.077	0.93
			Hindlimb[d]	.0002	2.45
	(.0001)	(1.4)	Righting	<.00001	1.6
Physiological	.067	.139	Piloerection	.20	0.31
		{−.014}	Weight	.454	0.02
	(.007)	(.3)	Temperature	.004	0.67

[a] Lower confidence bound.
[b] Estimate obtained using Monte Carlo sampling on 100,000 tables.
[c] Sample size too small to obtain an unconditional estimate.
[d] All eight outcomes were missing for forelimb grip at dose 5 g/kg, two outcomes missing for hindlimb grip at dose 5 g/kg. The domain level p-value is based on the complete data. When analyzing only the three outcomes for which complete data at all dose levels were available (i.e., gait, foot splay, and righting), the domain level p-value was <0.001. The estimates of $\hat{\beta}$ for individual outcomes are exact and obtained using LogXact.

methods may fail to produce unbiased estimates for small samples, the exact procedure presented here is a valuable tool for testing dose–response in this setting. Estimates for the domain level dose–response parameter have the potential to be extremely useful in risk assessment. For example, it could be a basis for a benchmark dose, which has been widely used as a statistical lower confidence bound on a dose that produces a predetermined increase in an adverse response rate (Crump, 1984).

Other methods used to analyze neurotoxicity data have their advantages in theory but break down in practice. Assuming that the functional domain structure is plausible, it is important that any statistical analysis testing the effect of dose consider all outcomes believed to effect the domain, which a univariate approach does not provide. Blackwell and Catalano developed a sophisticated two-stage latent variable model which allows for all outcomes to be analyzed simultaneously and adjusts for different correlation among the outcomes within a subject. However, estimation of parameters are obtained by using a version of the EM algorithm which at best produces unreliable results when sample sizes are small.

Acknowledgements

This work was supported by grants T32 ES07142-18 and ES06900 from the National Institute of Environmental Health Sciences. We are grateful to Dr Chris Corcoran for his valuable advice and technical expertise. We thank Dr Virginia Moser for providing the data used in the example.

Résumé

On utilise souvent des tests biologiques animaux pour évaluer la neurotoxicité d'une substance. Dans la batterie d'observations fonctionnelles, on expose une quarantaine d'animaux à la substance étudiée et on apprécie la toxicité à partir de réponses multiples : jusqu'à 30 items différents par animal. Nous proposons un test exact pour des réponses multiples binaires à corrélation constante. Ce test repose sur un modèle log-linéaire décrit par Molenberghs et Ryan (1999). Il étend les méthodes de Corcoran et al. (2001) qui ont développé un test exact pour des observations binaires à structure de corrélation échangeable. Nous calculons une valeur exacte de p pour le test d'une relation dose-réponse conjointe. Nous estimons également la pente de la relation dose-réponse et son intervalle de confiance à 95%. Nous

illustrons la méthode sur les données d'un test biologique de la neurotoxicité du perchloréthylène.

References

Baird, S. J., Catalano, P. J., Ryan, L. M., and Evans, J. S. (1997). Evaluation of effect profiles: Functional observational battery outcomes. *Fundamental and Applied Toxicology* **40**, 37–51.

Blackwell, B. and Catalano, P. J. (1999). *A random effects latent variable model for ordinal data*. Technical report. Dana-Farber Cancer Institute, Boston, Massachusetts.

Corcoran, C., Ryan, L., Mehta, C. R., Senchaudhuri, P., Patel, N. P., and Molenberghs, G. (2001). An exact trend test for correlated binary data. *Biometrics* **57**, 941–948.

Creason, J. P. (1989). Data evaluation and statistical analysis of functional observational battery data using a linear models approach. *Journal of the American College of Toxicology* **8**, 157–169.

Crump, K. S. (1984). A new method for determining allowable daily intakes. *Fundamental and Applied Toxicology* **4**, 854–871.

Cytel Software Corporation (1996). *LogXact-4 for Windows User Manual*.

Dorman, D. C., Brenneman, K. A., and Struve, M. F. (1999). Neurotoxicology research at CIIT. *CIIT Activities: Chemical Industry Institute of Toxicology* **19**, 1–6.

Hirji, K. F., Mehta, C. R., and Patel, N. R. (1987). Computing distributions for exact logistic regression. *Journal of the American Statistical Association* **82**, 1110–1117.

Hudson, J. I., Laird, N. M., Betensky, R. A., Keck, P. E., and Pope, H. G. (2001). Multivariate logistic regression for familial aggregation of two disorders. I and II. *American Journal of Epidemiology* **153**, 500–514.

Lefkopoulou, M. and Ryan, L. M. (1993). Global tests for multiple binary outcomes. *Biometrics* **49**, 975–988.

Mehta, C. R., Patel, N. R., and Senchaudhuri, P. (1992). Exact stratified linear rank tests for ordered categorical and binary data. *Journal of Computational and Graphical Statistics* **1**, 21–40.

Molenberghs, G. and Ryan, L. M. (1999). An exponential family model for clustered multivariate binary data. *Environmetrics* **10**, 279–300.

Moser, V. C. (1989). Screening approaches to neurotoxicity: A functional observational battery. *Journal of the American College of Toxicology* **8**, 85–93.

Moser, V. C. (1992). Applications of a neurobehavioral screening battery. *Journal of the American College of Toxicology* **10**, 661–669.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1988). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, U.K.: Cambridge University Press.

United States Environmental Protection Agency, Office of Pollution and Toxics (1994). Chemicals in the environment: Perchlorethylene. *OPPT Chemical Fact Sheet* EPA 749-F-94-020.

Received October 2002. *Revised June* 2003. *Accepted August* 2003.

Appendix A

Functional Domain Classification

Autonomic	Neuromuscular	Sensorimotor
Lacrimation (R)	Gait score (R)	Tail pinch response (R)
Salivation (R)	Mobility score (R)	Click response (R)
Pupil response (B)	Landing foot splay (C)	Touch response (R)
Defecation (C)	Forelimb grip strength (C)	Approach response (R)
Urination (C)	Hindlimb grip strength (C)	
	Righting reflex (R)	

Physiological	Central nervous system (CNS) excitability	CNS activity
Body weight (C)	Ease of removal (R)	Home cage posture (D)
Body temperature (C)	Handling reactivity (R)	Palpebral closure (D)
Piloerection (B)	Clonic movements (D)	Rearing (C)
	Tonic movements (D)	Motor activity (C)
	Arousal (R)	
	Vocalizations (B)	

Types of data: B = binary, C = continuous, R = rank, D = descriptive. *Source*: Moser, Applications of a Neurobehavioral Screening Battery, 1992.

Appendix B

Algorithm Details for the Network Representations

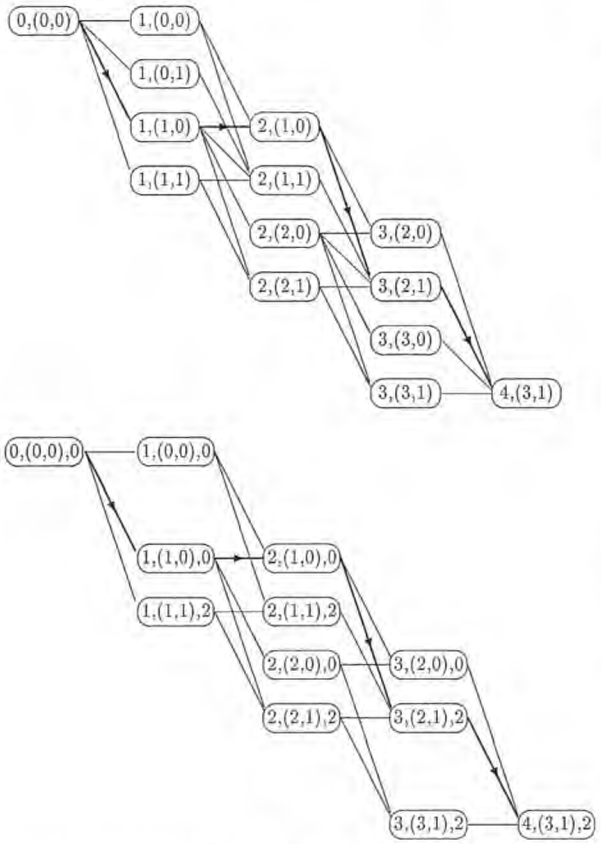
Each of the networks Γ_s and $\Gamma_{s,r}$ consist of a series of nodes and arcs, divided into $N+1$ stages. The nodes in Γ_s are of the form $\{k, (s_{k1}, s_{k2}, \ldots, s_{kM})\}$, where k represents the kth animal, and $s_{kj} = \sum_{l=1}^{k} y_{lj}$ is one possible value of the partial sum of responses from the first k animals for outcome j. The initial node in this network is $\{0, (0, 0, \ldots, 0)\}$ and the terminal node is $\{N, (s_1, s_2, \ldots, s_M)\}$. In the network representing $\Gamma_{s,r}$, each node will be of the form $\{k, (s_{k1}, s_{k2}, \ldots, s_{kM}), r_k\}$, where r_k represents one possible value of the partial sum $\sum_{l=1}^{k} z_l(z_l - 1)$. All paths correspond to exactly one table within its respective reference set, Γ_s or $\Gamma_{s,r}$.

Defining the set of successor nodes to a node $(k-1, (s_{(k-1)1}, \ldots, s_{(k-1)M}))$ as $\mathcal{R}(k-1, (s_{(k-1)1}, \ldots, s_{(k-1)M}))$, constraints are imposed which satisfy

$$\mathcal{R}\{k-1, (s_{(k-1)1}, \ldots, s_{(k-1)M})\} = \{k, (u_1, u_2, \ldots, u_M)\} : \max\left\{s_{(k-1)j}, s_j - \sum_{l=k+1}^{N} 1\right\} \le u_j \le \min\{s_j, s_{(k-1)j} + 1\}\}$$

for $j = 1, \ldots, M$.

Each node in Γ_s has 2^M possible arcs and each path satisfies $\sum_{l=1}^{N} s_{lj} - s_{(l-1)j} = s_j$, for $j = 1, \ldots, M$. For every path in the



Appendix Figure B1. Network representations of Γ_s and $\Gamma_{s,r}$ for the hypothetical data in Table 1.

second network, we are further ensured of

$$\sum_{l=1}^{N}\left(\sum_{j=1}^{M} s_{lj} - s_{(l-1)j}\right)\left(\left(\sum_{j=1}^{M} s_{lj} - s_{l-1,j}\right) - 1\right) = r.$$

For each $(k+1, (u_1, u_2, \ldots, u_M)) \in \mathcal{R}(k, (s_{k1}, \ldots, s_{kM}))$ the length of the arc connecting nodes is defined as

$$l_k = \left\{\sum_{j=1}^{M}\left(s_{(k+1)j} - s_{kj}\right)\right\}\left\{\sum_{j=1}^{M}\left(s_{(k+1)j} - s_{kj}\right) - 1\right\},$$

which is a possible contribution to z_k, the sum of outcomes for animal k. Let $SP_2(k, (s_{k1}, \ldots, s_{kM}))$ and $LP_2(k, (s_{k1}, \ldots, s_{kM}))$ represent the shortest and longest path lengths, $\sum_{i=0}^{k} l_i$, over all partial paths that originate at node $(0, (0, \ldots, 0))$ and terminate at node $(k, (s_{k1}, \ldots, s_{kM}))$. Note that

$$SP_2(k, (s_{k1}, \ldots, s_{kM})) = \min_{\{(k-1,\mathbf{u}) \in \mathcal{P}(k, \mathbf{s}_k)\}} \{SP_2(k-1, (u_1, \ldots, u_M)) + l_{k-1}\}$$

and

$$LP_2(k, (s_{k1}, \ldots, s_{kM})) = \max_{\{(k-1,\mathbf{u}) \in \mathcal{P}(k, \mathbf{s}_k)\}} \{LP_2(k-1, (u_1, \ldots, u_M)) + l_{k-1}\},$$

where $\mathcal{P}(k, \mathbf{s}_k)$ denotes the set of predecessor nodes to node $(k, \mathbf{s}_k)$.

The original network is now modified to meet the constraints imposed by $\Gamma_{s,r}$. The terminal node $(N, (s_1, \ldots, s_M), r)$ is created and the network is processed in reverse. At the kth stage:

1. Choose a node $(k, \mathbf{s}_k = (s_{k1}, \ldots, s_{kM}), r_k)$
2. For each node $(k-1, \mathbf{u}) \in \mathcal{P}(k, \mathbf{s}_k)$:
   (a) If $l_{k-1} + SP_2(k-1, \mathbf{u}) > r_k$ or $l_{k-1} + LP_2(k-1, \mathbf{u}) < r_k$, then the node cannot be a member of the network for $\Gamma_{s,r}$, and is dropped along with its associated arcs.
   (b) If the previous shortest path and longest path tests are passed, then the node $(k-1, \mathbf{u}, r_{k-1})$ is stored in the network, where $r_{k-1} = r_k - l_{k-1}$.

Iterations of the above steps continue until all nodes at the kth stage are exhausted. The exact distribution of t, the sufficient statistic for β for an observed table, conditional on $\mathbf{s}$ and r can now be calculated. For each node $(k, (u_1, \ldots, u_M), r_k)$ that is a feasible predecessor of $(k+1, (s_{(k+1)1}, \ldots, s_{(k+1)M}), r_{k+1})$, the rank length is defined as

$$L_k = d_{k+1}\left\{\sum_{j=1}^{M}\left(s_{(k+1)j} - u_j\right)\right\},$$

where d_{k+1} is the value of the dose for animal $k+1$. The value of the test statistic for a particular path is the sum of rank lengths, $\sum_{k=0}^{N-1} L_k$. $SP(k, (s_{k1}, \ldots, s_{kM}), r_k)$ is defined as the shortest path length in terms of rank, $\sum_{i=k}^{N-1} L_k$, over all paths that begin at $(k, (s_{k1}, \ldots, s_{kM}), r_k)$ and end at $(N, (s_1, \ldots, s_M), r)$, and LP is defined similarly.

The probability length for each path determines how likely a path is to be observed (MPS). Under H_0, the probability of having observed the corresponding table for a particular path is its probability length divided by the sum of probability lengths for all paths in the network. Each path is equally likely under the null, therefore all arcs in the network representing $\Gamma_{s,r}$ have a probability value of 1. The total probability, $TP(k, (s_{k1}, \ldots, s_{kM}), r_k)$, is the sum of the probability lengths for all partial paths that exist between that node and the terminal node. $TP(0, (0, \ldots, 0), 0)$ is thus the sum of all possible paths in the network and the denominator of $P(t \mid \mathbf{s}, r)$ in (4).

Note that during the processing of the initial network, updates to the values of SP, LP, and TP for each node were made. If a node $(k-1, (u_1, \ldots, u_M), r_{k-1})$ was already determined to be a predecessor of another node at the kth stage, then

$$SP(k-1, \mathbf{u}, r_{k-1}) = \min\{SP(k-1, \mathbf{u}, r_{k-1}), SP(k, (s_{k1}, \ldots, s_{kM}) + L_{k-1}\}$$

$$LP(k-1, \mathbf{u}, r_{k-1}) = \max\{LP(k-1, \mathbf{u}, r_{k-1}), LP(k, (s_{k1}, \ldots, s_{kM}) + L_{k-1}\}$$

$$TP(k-1, \mathbf{u}, r_{k-1}) = TP(k-1, \mathbf{u}, r_{k-1}) + TP(k, (s_{k1}, \ldots, s_{kM}), r_k),$$

If a node had not already been previously stored, then

$$SP(k-1,\mathbf{u},r_{k-1}) = SP(k,(s_{k1},\ldots,s_{kM}) + L_{k-1}$$

$$LP(k-1,\mathbf{u},r_{k-1}) = LP(k,(s_{k1},\ldots,s_{kM}) + L_{k-1}$$

$$TP(k-1,\mathbf{u},r_{k-1}) = TP(k,(s_{k1},\ldots,s_{kM}),r_k)$$

To calculate the exact p-value, the final network is navigated breadth first, stage by stage, eliminating nodes through which no path contributes to the critical region. The algorithm to process the network a final time to calculate the exact p-value was also developed by MPS and Corcoran et al. (2001), which we follow with slight modification. Representations of the full and reduced networks for the hypothetical data in Table 1 appear in Appendix Figure B1.

(44)

CASWELL FILE

004855

RELEASABLE

Glyphosate / Tox

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

DEC 12 1985

MEMORANDUM:

OFFICE OF
PESTICIDES AND TOXIC SUBSTANCES

SUBJECT: EPA Reg. #: 524-308; Roundup; Glyphosate; Pathology Report on Additional Kidney Sections
Caswell No. 661A
Accession No. 259621

TO: Robert Taylor
Product Manager (25)
Registration Division (TS-767)

THRU: Robert P. Zendzian, Ph.D. 12/12/85
Acting Head, Review Section IV
Toxicology Branch
Hazard Evaluation Division (TS-769)

FROM: William Dykstra, Ph.D. William Dykstra 12/12/85
Toxicology Branch
Hazard Evaluation Division (TS-769)

12/12/85

Requested Action:

Review pathology report on additional kidney sections.

Background:

Glyphosate was considered oncogenic in male mice causing renal tubule adenomas, a rare tumor, in a dose-related manner. The incidence of this tumor was 0, 0, 1, and 3 in the control, low-, mid-, and high-dose groups, respectively.

Additional evaluation of all original renal sections identified a small renal tubular adenoma in one control male (animal No. 1028) which was not diagnosed as such in the original pathology report.

Subsequently, Toxicology Branch recommended that additional renal sections be cut and evaluated from all control and glyphosate treated male mice.

This review contains the evaluation of the submitted results of the additional sectioning and pathological data.

004855

Conclusion:

The results of the additional pathological evaluation on re-cut kidney sections in male mice demonstrated no additional tumors were present. The significance of this finding will be determined later by the Ad Hoc committee.

Review:

1. The pathology report of additional kidney sections submitted by the registrant (Monsanto) showed that the renal tubule adenoma incidence in male mice was as follows:

Dose (ppm)	0	1000	5000	50,000
Animal number			3023	4029,4032,4041
Renal tubule adenoma	0	0	1	3
No. examined	49	49	50	50

The additional tumor in the control group which had been diagnosed from the re-evaluation of the original slides was not present in the re-cut kidney sections.

Toxicology Branch's pathologist (report attached) stated that the control tumor "does not represent a pathophyioloically significant change".

Statistical analysis of the tumor results showed no significant ($P<0.05$) difference in the incidence of renal tubule adenoma between control and treated groups.

However, the test for linear trend in proportions resulted in a $p=0.016$ which is statistically significant.

According to the registrant's pathology report, non-neoplastic kidney lesions did not reveal evidence of an ongoing chemically induced neprotoxicity.

Based on the original report and the new report, Toxicology Branch concludes that chronic interstitial nephritis occurred in compound-related manner in males at the high-dose as is shown below:

Males (Chronic Interstitial Nephritis)				
Dose (ppm)	0	1000	5000	30,000
Incidence				
Original report	5/49	2/49	7/50	12/50
New report	5/49	1/49	7/50	16/50

7.0 Summary of qualifications:

Curriculum Vitae

WARREN G. FOSTER, Ph.D., FCAHS

July 31, 2017

PERSONAL:

Business Address:
McMaster University
Department of Obstetrics & Gynecology
1280 Main Street West, HSC-3N52D
Hamilton, Ontario, Canada
L8S 4K1

Telephone: [redacted]
Facsimile:

E-mail: [redacted]

Citizenship: [redacted]

EDUCATION:

1991	Ph.D.	McMaster University, Hamilton, ON	Health Sciences
1986	M.Sc.	University of Guelph, Guelph, ON	Biomedical Sciences
1979	B.Sc.(Hon.)	University of Guelph, Guelph, ON	Human Biology

MEMBERSHIP IN PROFESSIONAL SOCIETIES:

1. Society of Toxicology, 1999 – present.
    - Councilor, Lake Ontario Regional Chapter, 2016 – present.
    - Past President, Lake Ontario Regional Chapter, 2014 – 2015.
    - President, Lake Ontario Regional Chapter, 2012 – 2014.
    - Secretary Treasurer, Reproductive Developmental Toxicology Specialty Section, 2010 – 2012.
    - Member, Scientific Program Committee, 2010 – 2012.

2. Society for the Study of Reproduction, 1994 – present.
    - Member, Membership Committee, 2012 – 2015.
    - Member, Scientific Program Committee, 2010 – 2012.
    - Chair, Committee on Reproduction and the Environment (CoRE), 2010 – 2012.
    - Member, Committee on Reproduction and the Environment (CoRE), 2009 – 2012.
    - Member, Animal Care Committee, 1995.

3. The Canadian Fertility and Andrology Society, 1991 – present.
    - Member, Scientific Program Committee, 2007 – 2011.
    - Past President 2008 – 2009.
    - Chair, Scientific Program Committee, 2007 – 2008.
    - President 2007 – 2008.
    - Vice-President, 2006 – 2007.
    - Industrial Liaison Committee, 2005 – 2006.

4. World Endometriosis Society, 2010 – present.
    - Member, Scientific Organizing Committee, World Congress of Endometriosis, Vancouver BC, 2012 – present.

5. European Society for Human Reproduction and Embryology, 2007 – present.

6. American Society for Reproductive Medicine, 2004 – present.

7. Association of Professors of Obstetrics & Gynaecology, 2003 – present.

8. Society of Obstetricians & Gynaecologists of Canada, 2001 – present.

9. Society of Toxicology of Canada, 1992 – present.
   - Vice-President (President Elect), 1998 – 1999.
   - Chairperson, Scientific Program Committee, 1997.
   - Member, Scientific Program Committee, 1996.

10. American Association for the Advancement of Science, 1988 – present.

11. International Federation of Placenta Associations, 2007.
    - Abstract Reviewer, 2007.

HONOURS AND AWARDS:

1. Elected to Fellowship in the Canadian Academy of Health Sciences, 2016, with citation for "Demonstrated leadership, creativity, distinctive competencies and commitment to advance academic health sciences." *The Canadian Academy of Health Sciences recognizes the full breadth of academic health science including all of the medical and allied health sciences and ranging from fundamental science to social science and population health. Members elected to the Academy will be well recognized by their peers nationally and internationally for their contributions to the promotion of health science and will have demonstrated leadership, creativity, distinctive competencies and commitment to advance academic health sciences. Such individuals are elected to the organization after a nominating and rigorous peer review procedure, which seeks to recognize those who are marked by a record of substantial accomplishment. Election to the Academy is considered one of the highest honours for members of the Canadian health sciences community and carries with it a covenant to serve the Academy and the future well-being of the health sciences irrespective of the member's specific discipline.*

2. Award of Excellence in Reproductive Medicine, 2016. *An honour reserved for members of the Canadian Fertility and Andrology Society for an outstanding contribution to, and leadership in, the field of Reproductive Medicine and Science. The recipient must demonstrate tremendous dedication to advancing the field in Reproductive Medicine and Science in research, clinical activity, teaching and consulting activities. Recognized as an innovator and scholar by peers—as shown by invitations to participate in professional societies and associations, providing service to the field in international, national or provincial public advisory groups, scholarly activities, such as symposia or workshops; applied experimental and clinical results to the benefit of society; shared research and clinical experience willingly with other members of the*

medical and scientific community through various activities, such as offering expert advice, demonstrating new techniques, communication with lay support groups and related public relations and, conveying original research and clinical results lucidly and in a time manner to peers via journals and scientific meetings.

3. Postdoctoral Fellow Supervision Award, Faculty of Health Sciences, McMaster University, 2016.

4. Graduate Student Supervision Award, Faculty of Health Sciences, McMaster University, 2016.

5. Postdoctoral Fellow Supervision Award, Faculty of Health Sciences, McMaster University, 2015.

6. Senior author of the best CFAS paper (oral presentation) in basic science category. Wessels JM, Leyland NA, Agarwal SK, **Foster WG**. Estrogen regulation of brain-derived neurotrophic factor in the uterus and the link to endometriosis. 60th Annual Meeting of the Canadian Fertility & Andrology Society, Quebec City, QC, 2014.

7. Graduate Student Supervision Award, Faculty of Health Sciences, McMaster University, 2014.

8. Mid-career Award, CIHR/Ontario Women's Health Council. In recognition of excellence and contributions to women's health and reproductive toxicology. Five-year salary award, ($75,000/year) 2006 – 2010.

9. Career Award, Ontario Women's Health Council. In recognition of expertise and scientific contributions to the field of reproductive toxicology. One-year award, ($100,000/year) 2005.

10. Co-author of the best CFAS paper (oral presentation) in basic science category and Alpha Award. Neal MS, Petrik J, **Foster WG**, Holloway AC. *In utero* and lactational exposure to nicotine: ovarian effects. Conjoint American Society for Reproductive Medicine and the 51st Annual Meeting of the Canadian Fertility & Andrology Society, Montreal, QC, 2005.

11. Co-author of the best CFAS paper (oral presentation) in basic science category. Van Vugt DA, Krzemien A, Roy BN, **Foster W**, Lundhal S, Marcus S, and Reid RL. Photodynamic ablation in non-human primates. 42nd Annual Meeting of the Canadian Fertility & Andrology Society, Lake Louise, AB, 1996.

12. Senior author of the best CFAS paper (oral presentation) in basic science category. **Foster WG**, Rice DC, McMahon A. Suppression of luteal function in the chronically lead exposed cynomolgus monkey (*Macaca fascicularis*). 40th Annual Meeting of the Canadian Fertility & Andrology Society, St. John, NB, September 7 – 10, 1994.

13. Medical Research Council of Canada Studentship, 1988 – 1990.

14. Medical Sciences Programme Scholarship, 1987 – 1990.

15. Ontario Graduate Scholarship, 1985 – 1986; 1987 – 1988.

EMPLOYMENT EXPERIENCE:

1. Voluntary Clinical Professor, Department of Reproductive Medicine, University of California, San Diego, Health Sciences, San Diego, CA, June 2016 – present.

2. Research Director, Center for Endometriosis Research and Treatment (CERT), UC San Diego Health, San Diego, CA, September 2014 – present.

3. Adjunct Professor, Herbert Wertheim College of Medicine, Florida International University, Miami, FL, April 2013 – present.

4. Professor, Reproductive Biology Division, Department of Obstetrics & Gynecology, McMaster University, Hamilton, ON, June 2005 – present.

5. Affiliate Scientist, Institute of Population Health, University of Ottawa, Ottawa, ON, 2003 – 2014.

6. Director, Reproductive Biology Division, Department of Obstetrics & Gynecology, McMaster University, Hamilton, ON, 2002 – March 2010.

7. Coordinator, Resident Research Program, Department of Obstetrics & Gynecology, McMaster University, Hamilton, ON, 2003 – 2010.

8. Medical Director, Centre for Reproductive Care, Hamilton Health Sciences, Hamilton, ON, April 2005 – March 2008.

9. Adjunct Assistant Professor, Department of Obstetrics & Gynaecology, Foothills Hospital, University of Calgary, Calgary, AB, 1993 – 2008.

10. Associate Professor, Reproductive Biology Division, Department of Obstetrics & Gynecology, McMaster University, Hamilton, ON, June 2001 – June 2005.

11. Senior Science Advisor, Bureau of Chemical Hazards, Health Canada, Ottawa, ON, September 2000 – June 2001.

12. Adjunct Assistant Professor, Department of Obstetrics & Gynecology, McMaster University, Hamilton, ON, 1997 – 2001.

13. Associate Director/Director of Research, Center for Women's Health, Cedars-Sinai Medical Center, Los Angeles, CA, January 1999 – September 2000.

14. Acting Division Chief, Environmental & Occupational Toxicology Division, Environmental Health Directorate, Health Protection Branch, Health Canada, Ottawa, ON, May – October 1997; June 1998 – January 1999.

15. Head, Reproductive Toxicology Section, Environmental & Occupational Toxicology Division, Environmental Health Directorate, Health Protection Branch, Health Canada, Ottawa, ON, 1992 – June 1998.

16. Reproductive Toxicologist, Reproductive Toxicology Section, Environmental & Occupational Toxicology Division, Environmental Health Directorate, Health & Welfare Canada, Ottawa, ON, 1990 – 1992.

SCHOLARLY, AND PROFESSIONAL ACTIVITES:

i) **editorial boards** –
Reproductive Toxicology 2004 – 2008; 2015 – present.

BioMed Research International, 2014 – present.
Guest Editor, Obstetrics and Gynecology, 2014 – present.

Journal of Clinical Toxicology, 2013 – present.

Journal of Environmental & Analytical Toxicology, 2012 – present.

ISRN Toxicology, 2011 – present.

Journal of Toxicology and Environmental Health: B Critical Reviews, 2010 – present.

Journal of Applied Toxicology, 2008 – present.
Editor, Journal of Applied Toxicology, 2010 – present.

Immunology, Endocrine & Metabolic Agents in Medicinal Chemistry, 2009 – present.

Faculty of 1000, 2008 – present.

ii) **grant panels and committees** –

International Agency for Research on Cancer (IARC). Pentachlorophenol, Aldrin and Dieldrin; Volume 117. Lyon, France. October 1 – 12, 2016.

Canadian Institutes of Health Research (CIHR), College of Reviewers, 2015 – present.

Ontario Council on Graduate Studies, Ontario Women's Health Scholars Awards Selection Committee, elected membership, 2015 – present.

Canadian Breast Cancer Fund Grants Review Committee, 2014 – present.

Canadian Institutes of Health Research (CIHR), Operating Grant, Gender & Health Peer Review Committee, 2013 – present.

Development and Reproductive Toxicology (DART) Technical Committee, International Life Science Institute (ILSI), Health and Environmental Sciences Institute (HESI), Scientific Advisor, 2012 – present.

The Anti-NMDA Receptor Encephalitis Foundation, Inc, Board of Directors, 2012 – present.

Faculty of Health Sciences, McMaster University, Graduate Curriculum Committee, 2006 – present.

US-National Toxicology Program, Center for the Evaluation of Risk to Human Reproduction, Member of Expert Registry, 2005 – present.

CIHR Strategic Training Program in Tobacco Research (CIHR-STPTR), University of Waterloo, Mentor, 2005 – present.

College of Reviewers, Canada Research Chairs Program, Member, 2000 – present.

FIFRA Scientific Advisory Panel, US-Environmental Protection Agency (EPA), Ad hoc Member, 1999 – present.

CIHR Training Program in Reproduction, Early Development, and the Impact on Health (REDIH), Program Advisory Committee Member, 2009 – 2016.

CIHR Training Program in Reproduction, Early Development, and the Impact on Health (REDIH), Strategic Training Initiative in Health Research, Member and Mentor, 2009 – 2016.

CIHR Gender, Sex & Health, Committee for the Transitional Operating Grant, Member, 2013 – 2015.

National Institute of Environmental Health Sciences (NIEHS), National Toxicology Program (NTP) and the NTP Toxicology Branch, NTP Board of Scientific Counselors (BSC) meeting, Ad hoc Reviewer, December 9 – 10, 2014.

CIHR Clinical Investigation 'A', Grants Committee Panel, Member, 2013 – 2014.

Canadian Institutes of Health Research (CIHR) Catalyst Grant Committee, Genes and Chronic Disease, July 17 – 18, 2013.

National Sciences and Engineering Research Council (NSERC), Undergraduate Research Student Awards (URSA) Scholarship Review Committee, Member, March 13, 2013.

R. Samuel McLaughlin Centre for Population Health Risk Assessment, Affiliate, 2008 – 2013.

National Center for Environmental Assessment, US Environmental Protection Agency, Improving the risk assessment of persistent, bioaccumulative and toxic (PBT) chemicals in breast milk, Invited participant, Expert Workshop, October 24 – 26, 2012.

National Institute of Child Health and Human Development (NICHD) site visit, Division of Epidemiology, Statistics & Prevention Research, Expert Panel Member, October 22 – 24, 2012.

Council of Canadian Academies Panel, Integrating Emerging Technologies into Chemical Safety Assessment, Member, 2010 – 2012.

Council of Canadian Academies, Integrated Testing of Pesticides, Expert Panel Member, 2009 – 2012.

Centre for Disease Control, Infertility Research Working Group, Member, 2009 –2012.

Society for the Study of Reproduction, Committee on Reproduction and the Environment, Member, 2008 – 2012.

Institute of Medicine of the National Academies, "Veterans and Agent Orange: 8th Biennial Update", Expert Reviewer, 2011.

Strategic Training in Research in Reproductive Health Sciences (STIRRHS), University of Montréal, Mentor, 2006 – 2011.

ASRM ERSIG Advisory Board Member, June 2005 – 2011.

Canadian Fertility & Andrology Society, Strategies Planning Meeting, Member, November 19 –

20, 2010.

Health Canada, Assisted Human Reproduction Canada, The Environment and Reproductive Health: A Scientific Roundtable Steering Committee, Organizing Committee and Expert Panel Member, 2010.

National Institutes of Health/National Institute of Environmental Health Sciences – National Toxicology Program (NIH/NIEHS-NTP), Reproductive & developmental effects of soy products and genistein, Expert Panel Member, 2009 – 2010.

College of Physicians and Surgeons of Ontario, Obstetrics & Gynaecology Task Force - *In vitro* Fertilization, Mentor, 2008 – 2010.

Canadian Breast Cancer Fund Grants Review Committee, Member, 2007 – 2010.

CIHR Clinical Investigation 'A', Grants Committee Panel, Member, 2006 – 2010.

Environment and Health, Collaborations for Health, McMaster University, Co-theme Team Leader, 2005 – 2010.

US Environmental Protection Agency, the National Institute of Environmental Health Sciences, and the California Breast Cancer Research Program, Mammary Gland Evaluation and Risk Assessment Workshop, Invited Participant, November 16 – 17, 2009.

NIH Study Section, Integrative and Clinical Endocrinology and Reproduction Study Section, Member, October 5 – 6, 2009.

Health Canada and CIHR sponsored, Canadian Children's Environmental Health Research Workshop, Invited Participant and Session Chair, Ottawa, ON, February 9 – 11, 2009.

Preimplantation Genetic Diagnosis and Related Activities in Canada, Steering Committee Member, 2008 – 2009.

Assisted Human Reproduction Research Workshop: Developing a National Research Agenda, Scientific Planning Committee Member, Montreal, QC, October 15 – 16, 2008.

National Institute of Child Health and Human Development (NICHD) site visit, Team Member, September 24 – 26, 2008.

National Institute of Child Health and Human Development (NICHD), National Children's Study (NCS), Member of RFP Review Panel, April 2008.

CIHR Institute of Human Development and Child and Youth Health Workshop – Environmental Toxicology, Invited Participant, February 2008.

The Society of Obstetricians and Gynaecologists of Canada (SOGC), Research Committee, Member, 2003 – 2008.

Ontario Tobacco Research Unit, Small Grants Committee, Member, 2007.

National Institute of Child Health and Human Development (NICHD), Effects of Aspirin on Gestation and Reproduction Panel, Study Section Member, 2006.

US Environmental Protection Agency – Star program, Grant Review Panel Member, 2004; 2005.

National Institute of Child Health and Human Development (NICHD), Child Study; Study Section Member, 2005.

National Cancer Institute of Canada, Canadian Tobacco Control Research Initiative Peer Review Panel Member, 2004.

Toxic Substances Research Initiative, Health Canada, Healthy Environments and Consumer Safety Branch, 1999 – 2003.

- Chairman, Endocrine Disrupters Technical Review Committee, 1999 – 2000.

CIHR-Institute of Human Development, Child and Youth Health sponsored Pre and Post Implantation Consensus Workshop, Invited Participant, Niagara-on-the-Lake, ON, April 5 – 7, 2002.

Health Canada sponsored, Canadian Children's Environmental Health Research Workshop, Invited Participant and Session Chair, Ottawa, ON, March 17 – 19, 2002.

WHO/IPCS Steering Group on Endocrine Disruptors, 1998 – 2002.

- Global Inventory of Endocrine Disruptor Research.
- International Assessment of the State of Knowledge on Endocrine Disruptors.

Natural Sciences and Engineering Research Council (NSERC) Grant Selection, 1996 – 2001.

- Past Chair, 2001
- Committee Chair Person, 1999 – 2000.
- Committee Member, 1996 – 2000.

National Institutes of Environmental Health Safety, National Institutes of Health, and National Toxicology Program sponsored expert meeting on low-dose effects of endocrine disruptors,

Invited Panelist, October 2000.

US Environmental Protection Agency, Atrazine, Science Advisory Panel Member, June 2000.

SETAC-SOT co-sponsored Workshop on Environmental-Human Interconnections, Snowbird, UT, Invited Participant, June 10 – 15, 2000.

US Environmental Protection Agency – Endocrine Disruptor Screening and Testing Standardisation Committee, Mammalian Test Working Group Member, 1999 – 2000.

Joint US/EU Endocrine Disruptor Research, Expert Panel Member, 1999.

US-EPA ORD Strategy for Research on Environmental Risks to Children, Peer Reviewer, 1999.

Health Canada and Environment Canada, Interdepartmental Research Committee on Endocrine Disruptors, Co-Chair, 1998 – 1999.

Organisation for Economic Co-operation and Development (OECD), Member, 1997 – 1999.

- Working Group on Endocrine Disruptor Testing and Assessment, 1997 – 1999.
- Health Canada Endocrine Disruptor Committee, 1997 – 1999.
- National Co-ordinator, Test Guideline Program, 1997 – 1998.

National Sanitation Foundation - International, Health Effects Task Group, Member, 1996 – 1998.

Health Canada, Health Protection Branch, Animal Care Committee, Member, 1993 – 1998.

iii) **executive positions** –

Councillor, Society of Toxicology, Lake Ontario Regional Chapter, 2016 – present.

Board Member, The Anti-NMDA Receptor Encephalitis Foundation, Inc, 2012 – present.

Past President, Society of Toxicology, Lake Ontario Regional Chapter, 2014 – 2015.

Member, Membership Committee, Society for the Study of Reproduction, 2012 – 2015.

President, Society of Toxicology, Lake Ontario Regional Chapter, 2012 – 2014.

Member, Scientific Program Committee, Society for the Study of Reproduction, 2010 – 2012.

Chair, Core Committee, Society for the Study of Reproduction, 2010 – 2012.

Secretary Treasurer, Reproductive Developmental Toxicology Specialty Section, Society of

Toxicology, 2010 – 2012.

Member, Scientific Program Committee, Canadian Fertility & Andrology Society, 2006 – 2010.

Member of the Association of Professors of Obstetrics & Gynaecology, Science Committee, 2003 – 2009.

President, Canadian Fertility & Andrology Society, 2007 – 2008.

Chair, Scientific Program Committee, Canadian Fertility & Andrology Society, 2007 – 2008.

Vice-president, Canadian Fertility & Andrology Society, 2006 – 2007.

Chair, Science Panel, EM-COM web site, www.EMCOM.ca, 2002 – April 2005.

Member, Board of Directors, Infertility Awareness Association of Canada, 1998 – 2004.

Vice-president, President-elect, Society of Toxicology of Canada, 1998 – 1999.

Chair, Scientific Program Committee, Society of Toxicology of Canada, 1997.

Member of Scientific Program Committee, Society of Toxicology of Canada, 1996.

Member of Animal Care Committee, Society for the Study of Reproduction, 1995.

iv) **journal referee** –

Biology of Reproduction

British Journal of Obstetrics and Gynaecology

Comparative Biochemistry & Physiology

Critical Reviews in Toxicology
Endocrinology

Environmental Health Perspectives

Environmental Research

Environmental Toxicology

Fertility & Sterility

Food & Chemical Toxicology

Human Reproduction

Immunology, Endocrine & Metabolic Agents in Medicinal Chemistry

International Journal of Cell Biology

Journal of Applied Toxicology

Journal of Clinical Investigation

Journal of Toxicology & Environmental Health

Pesticide Biochemistry & Physiology

Placenta

Regulatory Toxicology and Pharmacology

Reproductive Biomedicine Online

Reproductive Toxicology

Systems Biology in Reproductive Medicine

The Journal of Urology

Toxicological Sciences

Toxicology and Industrial Health

v) **external grant review agencies –**
Canadian Institutes of Health Research

European Commission – Health Research

Natural Sciences and Engineering Research Council of Canada

Canadian Tobacco Control Research Initiative

Hospital for Sick Children, Toronto, ON

College of Reviewers for Canadian Research Chairs

Research Grants Council of Hong Kong

Puerto Rico Life Sciences Research Fund

Canterbury Medical Research Foundation

Netherlands Organization for Scientific Research - Earth and Life Sciences

AREAS OF RESEARCH INTEREST:

My research interests fall primarily into three categories as follows: (1) reproductive epidemiology and biomonitoring; (2) reproductive and development toxicity and carcinogenicity of environmental and dietary chemicals; and (3) the cellular and molecular mechanisms of endometriosis. My team of postdoctoral fellows and graduate students are assigned to each area carrying to better define the relationship between exposure to exogenous chemicals and adverse reproductive outcomes including estrogen related cancers.

COURSES TAUGHT:

i) **undergraduate –**
Pharmacology 4C03 Principles of Toxicology, Lectures on Endocrine Disruption

Undergraduate Medicine MF3 – Reproductive Biology

ii) **graduate –**

MS712 Reproductive Endocrinology

MS714 Industrial and Environmental Toxicology

MS720 Tobacco and Health: From Cells to Society

MS799 Independent Study in Reproductive Biology

SUPERVISORSHIPS:

i) **master –**

- Maria Haikalis, Identification of microRNA species in endometriosis. (Supervisor, 2015 – present).
- Eli Crapper, Characterization of a novel clinical marker of endometriosis. (Supervisor, 2015 – 2016). Medical School.
- Anne Doedée, Visiting Student (Netherlands), Neurotrophins and trks – novel reproductive tract proteins. (Supervisor, 2010). PDF, Germany.
- Alex Lagunov, Mechanisms regulating oocyte activation. (Supervisor, 2008 – 2010). Lab Director.
- Dana Anger, Tyrosine kinase receptor B expression in human endometrium. (Supervisor, 2005 – 2007). Cornerstone Research Group.
- Katie Stys, Mechanisms of AhR-ligand induced changes in inappropriate estrogen production in the endometrium. (Supervisor, 2003 – 2005). Health Law.
- Katie Edmunds, Proliferative effects of dietary isoflavones in the human endometrium. (Supervisor, 2002 – 2004). Health Law.
- Megan Miller, Effects of benzo[a]pyrene on matrix metalloproteinase expression and activity in breast cancer. (Supervisor, 2002 – 2004). High School Science Teacher.

ii) **doctoral –**

- Jocelyn Wessels, Brain derived neurotrophic factor (BDNF), a novel diagnostic marker for endometriosis. (Supervisor, 2011 – 2015). PDF, McMaster Unviersity.

- Mahmoud Aarabi, Queen's University, Characterization of post-acrosomal sheath protein PAWP. (Co-supervisor, 2008 – 2013). PDF, McGill, University.

- Anne Gannon , M.Sc., The effect of cigarette smoke constituents on ovarian follicle growth and apoptosis. (Supervisor, 2006 – 2013). PDF, Health Canada.

- Nakpangi A Johnson, Duquesne University – Does DDE shorten time to tumor formation in MMTV-neu mice? (Co-supervisor, 2008 –2011). Charles River Laboratories, USA.

- Mike Neal, M.Sc., The effect of environmental agents on follicle dynamics and oocyte quality. (Supervisor, 2002 – 2006) *incomplete.* Lab Directory, One Fertility.

- Ebrahim Nasir, Reproductive toxic effects of bilirubin on testicular function. (Co-supervisor, 2002 – 2003).

iii) **post-doctoral –**

- Hayley Furlong, Ph.D., Mitochondrial homeostatis and ovarian follicle development. (2014 – 2016).

- Marina Guerra, Ph.D., Ovarian toxicity of parabens. (2014 – 2015).

- Harry Peery, Ph.D., Mitochondrial oxidative stress following toxic insult. (2014 – 2015).

- Jean Clair Sadeu, Ph.D., Effect of environmental toxicants on folliculogenesis. (2009 – 2013).

- Heather Cameron, Ph.D., Effects of environmental toxicants on estrogen dependent mammary tumor development in mice. (2006 – 2008).

- Rocio Monroy, M.D., Cigarette smoking during pregnancy and glucose transport protein expression in the human placenta. (2006 – 2008).

- Gentao Liu, Ph.D., The reproductive effects of dietary galactose in the rat. (1999 – 2000).

- Jack Yang, Ph.D., The role of environmental pollutants in the pathophysiology of endometriosis in rodents and non-human primate models. (1996 – 1999).

- Michael Wade, Ph.D., The effect of environmentally relevant concentrations of priority contaminants on ovarian follicle differentiation, steroidogenesis and ovulation. (1996 – 1998).

- Daniel Cyr, Ph.D., The effects of methyl mercury on male reproduction. (1994 – 1995).

iv) **professional –**

- Andrea Mosher, M.D., Ph.D., OBS & GYN Resident research project, McMaster University, 2015 – present.
- Miguel Dominguez, D.V.M., Ph.D., Visiting Scientist, from Mexico, 2006 – 2009 & 2013.
- Sandra Gregorovich, M.D., Research Coordinator, 2009 – 2011.
- Heather Cameron, Ph.D., Research Associate, 2009 – 2010.
- Mehrnoosh Faghih, M.D., OBS & GYN Resident research project, McMaster University, 2006 – 2008.
- Myoung-seok Han, M.D., OBS & GYN, Visiting Scientist, from Korea, 2006 – 2007.
- Greg Athaide, M.D., OBS & GYN Resident research project, McMaster University, 2005 – 2006.
- Julie Francis, M.D., OBS & GYN Resident research project, McMaster University, 2004 – 2005.
- Pezhman Mirshokraei, D.V.M., Ph.D., Visiting Scientist, from Iran, 2003 – 2004.
- Alison Holloway, Ph.D., Cellular and molecular mechanisms of inappropriate estrogen production in endometriosis. (2001 – 2004).
- Anna Chomej, M.D., OBS & GYN Resident research project, McMaster University, 2003.

v) **supervisory committees –**

- Lucas Greville, 2017 – present, Ph.D. candidate, McMaster University.
- Tyler Pollock, 2014 – present, Ph.D. candidate, McMaster University.
- Evan Borman, 2013 – present, Ph.D. candidate, McMaster University.
- Jennifer Fazzari, 2012 – present, Ph.D. candidate, McMaster University.
- Michael Tsoulis, 2013 – 2015, M.Sc. candidate, McMaster University.

- Stephanie Zantinge, 2012 – 2014, M.Sc. candidate, McMaster University.
- Jonathan Lockwood, 2012 – 2013, M.Sc. candidate, McMaster University.
- Robert Berger, 2007 – 2010, M.Sc. candidate, McMaster University.
- Ayesha Khan, 2006 – 2009, Ph.D. candidate, McMaster University.
- Jenny Bruin, 2005 – 2009, Ph.D. candidate, McMaster University.
- Carolyn Cesta, 2007 – 2009, M.Sc. candidate, McMaster University.
- Jordan Shaw, 2007 – 2008, M.Sc. candidate, McMaster University.
- Rochelle Fernandez, 2004 – 2005, M.Sc. candidate, McMaster University.

vi) **others (summer/co-op students) –**

- Mona Kahlid, (January 2016 – present) Life Sciences, McMaster University. Project Title: Autophagy markers in ovarian follicular fluid.
- Allana Simon, (September 2015 – present) Health Sciences, McMaster University. Project Title: RANTES clinical marker of endometriosis, or not?
- Marina Bockaj, (September 2014 – present) Biomedical Engineering, McMaster University. Project Title: Miniaturized diagnostic test system for endometriosis.
- Nicholas Stalteri, (January 2016 – May 2016) Engineering, University of Waterloo. Project Title: Cigarette smoke-induced changes in autophagy gene expression and target tissues of the mouse.
- Yosef Ellenbogen, (September 2015 – July 2016) Health Sciences, McMaster University. Project Title: Fallopian tube endometriosis.
- Allegra Drumm, (September 2015 – September 2016) Biomedical Sciences, University of Guelph. Project Title: Interleukin 6 expression in the endometrium. Medical School.
- Thuy Linh Do, (September 2015 – May 2016) Health Sciences, McMaster University. Project Title: Toxicant-induced mitochondrial dysregulation in ovarian granulosa cells. Medical School.
- Aamer Somani, (January 2014 – September 2016) Biochemistry, McMaster University.

Project Title: Dietary chemical effects on endometrial epithelial cell aromatase activity. Medical School.

- Garima Aryal, (January 2014 – May 2016) Health Sciences, McMaster University. Project Title: Human developmental exposure to BPA: developmental effects.
- Sal Vivona, (September 2014 – 2015) Health Sciences, McMaster University. Project Title: Endometriosis markers. Medical School.
- Anna Parackal, (January 2014 – May 2015) Biopharmacology, McMaster University. Project Title: Effects of cigarette smoke on markers of autophagy in the mouse ovary. AstraZenica, Canada.
- Alia Tewari, (January – December 2014) Health Sciences, Western University, Project Title: Environmental obesogens, what's the skinny?
- Trevor Patch, (September 2013 – May 2014) Psychology, McMaster University. Project Title: Mechanisms of stress induced pregnancy loss in the mouse. Ph.D. student, University of Guelph.
- Piraveena Sivapatham, (September 2013 – May 2014) Life Sciences, McMaster University. Project Title: Connexin-26 expression in the endometrium and spontaneous abortion. BScN student, University of Toronto.
- Kabir Toor, (Summer 2013) Bachelor of Health Sciences, McMaster University. Project Title: Clinical markers in endometriosis. M.Sc. student, University of British Columbia.
- Alia Tewari, (Summer 2013) Bachelor of Health Sciences, University of Western Ontario. Project Title: Neurotrophins in endometriosis.
- Vanessa Kay, (Summer 2011; September 2011 – May 2013) Thesis student, McMaster University. Project Title: Measurement of urinary phthalate metabolites. M.D./Ph.D student, Queen's University.
- Vanessa Kay, (Summer 2010) Bachelor of Health Sciences, McMaster University. Project Title: Identification of a novel diagnostic marker for endometriosis.
- Vanessa Kay, (Summer 2009) Bachelor of Health Sciences, McMaster University. Project Title: Identification of a novel diagnostic marker for endometriosis.
- Natalie Cho, (Summer 2009) Bachelor of Health Sciences, McMaster University. Project

Title: Ovarian effects of Bisphenol A exposure. Resident of Anesthesiology, University of Ottawa.

- Mary Peric, (Summer 2009) Bachelor of Health Sciences, McMaster University. Project Title: Placenta change induced by in utero nicotine treatment in the rat. Naturopath.
- Melissa Coubrough, (Summer 2009) Midwifery, McMaster University. Project Title: Neurotrophic expression in breast cancer cell lines. Midwife.
- Otis Kryzanauskas, (Summer 2009) Midwifery, McMaster University. Project Title: Effects of maternal smoking on placenta glucose transport. Midwife.
- Mary Peric, (Summer 2008) Bachelor of Health Sciences, McMaster University. Project Title: Developmental effects of nicotine exposure in the rat.
- Vivian Ho, (Summer 2008) Bachelor of Health Sciences, McMaster University. Project Title: Developmental effects of nicotine exposure in the rat.
- Rami Elias, (Summer 2007) Bachelor of Health Sciences, McMaster University. Project Title: Neurotrophin and cognate receptor expression in endometriosis associated ovarian cancer.
- Mary Peric, (Summer 2007) Gr. 12 Hon. Student. Project Title: Brain derived neurotrophic factor and Tyrosine receptor kinase B expression in the reproductive tract of sexually mature BALB/c mice.
- Derek Chaves, (Fall 2004) Pharmacology & Toxicology Thesis Project, McMaster University. Project Title: Cyclooxygenase-II and matrix metalloproteinase expression in breast cancer cell lines. Architect.
- John Agzarian, (Summer 2004) Bachelor of Health Sciences, McMaster University. Project Title: Environmental toxicant mixture effects on thyroid gland morphology. (Funded by a scholarship from the Thyroid Foundation of Canada). Community Physician.
- Alex Petre, (Summer 2004) Bachelor of Science, McMaster University. Project Title: Toxicant induced changes in tissue remodelling enzyme expression in granulosa cells. (NSERC scholarship). M.D./Ph.D., University of Toronto.
- Dana Anger, (Summer 2004) Pharmacology & Toxicology Thesis Project, McMaster University. Project title: Mechanisms of methylchloranthene-induced changes in ovarian apoptosis. Cornerstone Research Group.

- Gareth Lim, (Summer 2003) Pharmacology & Toxicology Thesis Project, McMaster University. Project Title: Developmental toxicity of in utero exposure to drinking water disinfection by-products in the rat. PDF, University of British Columbia.

- Sarah Sinasac, (Summer 2002) Bachelor of Health Sciences, Medical Student, McMaster University. Project Title: Toxicant effects on granulosa cell steroidogenesis. Physician.

- Donna Grant, (Summer 1994) Biochemistry/Toxicology, University of Guelph. Project Title: Ovarian apoptosis in the PMSG primed immature rat ovary.

- Carmen Mertineit, (Winter 1992) Biology Thesis Project, McMaster University. Project Title: The effect of Hexachlorobenzene on the ovariectomized rat. Ph.D., Research Scientist, Astra Pharmaceuticals.

- Pete Ecclestone, (Winter 1991) Biology Thesis Project, McMaster University. Project Title: The effect of Lead intoxication on serum radioimmunoreactive vs. bioactive levels of pituitary gonadotropins and serum testosterone levels in the male cynomolgus monkey. Pharmaceutical sales.

- Greg Major, (Summer 1991) Biology, University of Western Ontario. Project Title: Modification of an enzyme fluorescent method for quantification of DNA in subcellular fractions. Orthopedic Surgeon, University of Colorado.

- Julie Pentick, (Fall 1990) Biochemistry, University of Waterloo. Project Title: Tissue distribution and subcellular localization of Hexachlorobenzene in the rat ovary. Contract researcher, Health Canada.

- Greg Major, (Summer 1990) Biology, University of Western Ontario. Project Title: Mating induced changes in the distribution of immunoreactive GnRH neural elements in the female rabbit.

vii) **graduate examining committees** –

- Tyler Pollock, Ph.D., Comprehensive Examiner, McMaster University, 2015.

- Evan Borman, Ph.D., Comprehensive Examiner, McMaster University, 2015.

- Jennifer Fazzari, M.Sc., Transfer Examiner, McMaster University, 2014.

- Stephanie Ondovcik, Ph.D., External Thesis Examiner, University of Toronto, 2013.

- Kristy Roth, Ph.D., Comprehensive Examiner, McMaster University, 2011.
- Jessica Kafka, Ph.D., Comprehensive Examiner, McMaster University, 2011.
- Robert Berger, M.Sc., Thesis Examiner, McMaster University, 2010.
- Ayesha Khan, Ph.D., Thesis Examiner, McMaster University, 2009.
- Jenny Bruin, Ph.D., Thesis Examiner, McMaster University, 2009.
- Arkadiusz (Eric) Hul, Ph.D., Thesis Examiner, McMaster University, 2009.
- Carolyn Cesta, M.Sc., Thesis Examiner, McMaster University, 2009.
- Jordan Shaw, M.Sc., Thesis Examiner, McMaster University, 2008.
- Navkiran Gill, Ph.D., Thesis Examiner, McMaster University, 2008.
- Anne Ellis, M.D., M.Sc., Thesis Examiner, McMaster University, 2008.
- Sudha Bhavanam, M.Sc., Thesis Examiner, McMaster University, 2008.
- Lorna Ryan, Ph.D., Thesis Examiner, McMaster University, 2007.
- Sherri Fernandez, M.Sc., Thesis Examiner, McMaster University, 2007.
- Alexandera Kollara, Ph.D., External Thesis Examiner, University of Toronto, 2006.
- Caleb Zavitz, M.Sc., Transfer Examiner, McMaster University, 2006.
- Rochelle Fernandez, M.Sc., Thesis Examiner, McMaster University, 2006.
- Tamara Lee Jocelyn, Ph.D., External Thesis Examiner, McGill University, 2005.
- Michael Cyr, M.Sc. Thesis Examiner, McMaster University, 2005.
- Erin McDonald, M.Sc., Thesis Examiner, McMaster University, 2005.
- Anthony Wood, Ph.D., External Thesis Examiner, University of Guelph, 2004.

- Julang Li, Ph.D., External Thesis Examiner, University of Ottawa, 1998.

RESEARCH FUNDING:

1. **Foster WG**, Leyland NA, Agarwal SK, Villeneuve P. Characterization of a novel clinical marker of endometriosis. CIHR Institute of Gender and Health $705,752. (***Awarded***). 2015 – 2020.

2. **Foster WG**, Allen-Vercoe E, Cyr D, Haddad S, Haines J, Hooks G, Kubwabo C, Langille M, Venners S. Developmental origins of toxicant-induced obesity and insulin resistance. Canadian Institutes of Health Research (CIHR) Team Grant: Developmental Origins of Health and Disease – Implications for Men, Women, Boys and Girls – LOI $10,000. (***Awarded***) 2015 – 2016.

3. **Foster WG** and Zhu, J. Mechanism(s) of cigarette smoke-induced ovarian follicle loss. Canadian Institutes of Health Research (CIHR) $873,835. (***Awarded***) 2011 – 2016.

4. Tayade C and **Foster WG**. A novel anti-angiogenic therapy for endometriosis. CIHR $483,410. (***Awarded***). 2011 – 2016.

5. **Foster WG**, Leyland NA, Agarwal SK, Villeneuve P. Characterization of a novel clinical marker of endometriosis. CIHR Institute of Gender and Health $100,000. (***Awarded***). 2014 – 2015.

6. Fraser W, Arbuckle T, **Foster WG**, et al., (9 co-applicants). Maternal-Infant Research on Environmental Chemicals – Child Development Plus (MIREC-CD+). Health Canada $149,165. (***Awarded***). 2013 – 2015.

7. Baltz JM, **Foster WG**, et al., (30 co-applicants). Training Program in Reproduction, Early Development, and the Impact on Health (REDIH). CIHR Human Development, Child and Youth Health $1,787,598. (***Awarded***) 2009 – 2015.

8. Cameron R, **Foster WG**, et al., (58 co-applicants). Population Intervention for Chronic Disease Prevention: A Pan-Canadian Program. CIHR $1,950,000. (***Awarded***) 2009 – 2014.

9. **Foster WG**. Evaluation of a dietary treatment for endometriosis. Concourse Health Sciences LLC $68,525 USD. (***Awarded***). 2012 – 2013.

10. **Foster WG**. Neurotrophins and Trks: Novel reproductive tract proteins. Natural Sciences and Engineering Research Center (NSERC) $185,000. (***Awarded***). 2008 – 2013.

11. Fraser W, Arbuckle T, **Foster WG**, et al., (9 co-applicants). Maternal-Infant Research on Environmental Chemicals : A National Profile of *In Utero* and Lactational Exposure to Environmental Contaminants. CIHR $1,248,126. (***Awarded***) 2006 – 2012.

12. Fraser W, Arbuckle T, **Foster WG,** et al., (9 co-applicants). Maternal-Infant Research on Environmental Chemicals – Infant Development (MIREC-ID). Health Canada $2,084,968. (***Awarded***). 2008 – 2011.

13. **Foster WG** and Cameron H. Dieldrin increases breast cancer metastasis via dysregulation of neurotrophin expression. CIHR $100,000. (***Awarded***) 2009 - 2010.

14. **Foster WG**, Yauk C, Quinn J, Robaire B, and McCarry B. Urban air particulate pollution & genetic instability. CIHR Team Grant LOI $10,000. (***Awarded***) 2009 - 2010.

15. **Foster WG**. Cellular and molecular mechanisms of toxicant-induced changes in ovarian follicular atresia. CIHR $616,408. (***Awarded***) 2006 – 2010.

16. **Foster WG**. Cellular and molecular mechanisms of toxicant-induced changes in follicular dynamics and ovarian regulation. Ontario Women's Health Council/CIHR Institute of Gender and Health Mid-career Award $375,000. (***Awarded***) 2005 – 2010.

17. Oko R and **Foster WG**. Improvement of ICSI treatment by co-injection of recombinant PAWP protein. CIHR RxND $100,000. (***Awarded***) 2007 – 2008.

18. **Foster WG**. Toxicant-induced resistance to ANOIKIS in estrogen sensitive target tissues. NSERC $35,136. (***Awarded***). 2007 – 2008.

19. **Foster W**, Holloway A, Krewski D, Kourti T. Surrogate biomarkers of *in utero* exposure to xenobiotics. American Chemistry Council $878,418. (***Awarded***) 2003 – 2007.

20. **Foster WG**. Toxicant induced tissue remodelling in estrogen sensitive target tissues. NSERC $144,000. (***Awarded***) 2003 – 2007.

21. **Foster W**, Holloway A, Krewski D, Muller W. Biomarkers of breast cancer. American Chemistry Council $941,751. (***Awarded***) 2002 – 2007.

22. Casper R and **Foster WG**. Identification of early pathogenetic events leading to endometriosis and discovery of novel therapeutic strategies. CIHR – Operating $415,794. (***Awarded***) 2003 – 2006.

23. **Foster WG**. AhR ligands and endometriosis: towards understanding their mechanism of

action. CIHR – Operating $345,601. (***Awarded***) 2002 – 2006.

24. **Foster WG**. Ontario Women's Health Council Career Award $100,000. (***Declined***) 2005.

25. **Foster W**, Holloway A. Effect of binary mixtures on estrogen sensitive target tissues. Canadian Network of Toxicology Centres $55,000. (***Awarded***) 2003 – 2004.

26. **Foster WG**. Hormonally active chemicals: cellular and molecular mechanisms of action. Canadian Foundation for Innovation-Ontario Innovation Trust (CFI-OIT) $422,096. (***Awarded***) 2002 – 2003.

27. Holloway A, and **Foster WG**. Effects of *in utero* chemical insult on postnatal health. Canadian Chlorine Coordinating Council $48,000. (***Awarded***) 2002 – 2003.

28. **Foster WG**, Hughes CL, and Chan S. Human developmental exposure to endocrine disruptors. New York Community Trust Fund $105,000 USD. (***Awarded***) 2002 – 2003.

29. **Foster WG**. Dietary factor modulation of endometrial tissue production of IL-6 and IL-6sR and angiogenic factors *in vitro*. Dow Chemical $25,000 USD. (***Awarded***) 2001 – 2002.

30. Davis V, **Foster WG**, and Hughes CL. Influence and localized DDT exposure on breast cancer. California Breast Cancer Research Program $305,989 USD. (***Awarded***) 2000 – 2002.

PEER REVIEWED PUBLICATIONS:

i) **books** –

1. Ritter L, Austin CP, Bend JR, Brunk CG, Caulfield T, Dellarco VL, Demers P, **Foster W**, Infante-Rivard C, Jumarie C, Kacew S, Kavlock RJ, Krewski D, Mezey PG, Shultz T. (2012) Integrating Emerging Technologies into Chemical Safety Assessment. Council of Canadian Academies. Ottawa, Canada.

ii) **contributions to books** –

1. Dominguez MA, Sadeu JC, Guerra MT, Furlong HC, Bains S, **Foster WG**. (2016) Chapter: Ovarian toxicity on environmental contaminants: 50 shades of grey. In: Translational Toxicology: Defining a new therapeutic discipline, Molecular Integrative Toxicology, Hughes C and Waters MD, Editors. Springer International Publishing, Switzerland.

2. Valez MP, Monnier P, **Foster WG**, Fraser WD. (2015) Chapter 7:The impact of phthalates on women's reproductive health: Current state-of-the-science and future directions. In: Our Chemical Selves – Gender, Toxics, and Environmental Health, Dayna Nadine Scott, Editor. UBC Press. Part III, Hormones as the 'Messengers of Gender'? 231-254.

3. Gannon AM, Sadeu JC, Agarwal SK, Hughes CL, **Foster WG**. (2013) Chapter 10:Cigarette smoking and ovarian function. In: Ovarian Toxicology, Patricia Hoyer, Editor. CRC Press. Part II, Ovotoxic Chemical Classes. 231-250.

4. Peery HE, Day GS, Doja A, Xia C, Fritzler M, **Foster W**. (2013) Chapter 129:Anti-NMDA receptor encephalitis in children: the disorder, its diagnosis, and treatment. In: Handbook of Clinical Neurology, Vol. 112, Pediatric Neurology Part II. 1229-1233.

5. Rier S and **Foster WG**. (2003) Environmental dioxins and endometriosis. Semin. Reprod. Med. 21(2):145-154.

6. **Foster W** and Hughes C. (2002) Chapter 2:Review of Normal Human Reproduction. In: Principles for Evaluating Human Reproductive Effects of Chemicals. International Programme on Chemical Safety of the World Health Organization.

7. Van Vugt DA, Krzemien A, **Foster W**, Lundhal S, Marcus S, Reid RL. (2000) Photodynamic endometrial ablation in non-human primates. In: Photomedicine in Gynecology and Reproduction. (P. Wyss, Tadir Y, Tromberg BJ, Haller U, eds.) Karger, Basal, Switzerland. 213-218.

iii) **journal articles** –

1. Guerra MT, Sanabria M, Leite GAA, Borges CS, Cucielo MS, Anselmo-Franci JA, **Foster WG**, Kempinas WG. (2016) Maternal exposure to butylparaben impairs testicular structure and sperm quality on male rats. Environ. Res. (*In-press*).

2. Yao C, **Foster WG**, Sadeu JC, Siddique S, Zhu J, Feng YL. (2016) Screening for DNA adducts in ovarian follicles exposed to benzo[a]pyrene and cigarette smoke condensate using liquid chromatography-tandem mass spectrometry. Sci. Total Environ. (*In-press*).

3. **Foster WG**, Evans JA, Little J, Arbour L, Moore A, Sauve R, León JA, Luo W. (2016) Human exposure to environmental contaminants and congenital anomalies: A critical review. Crit. Rev. Toxicol. (*In-press*).

4. Fisher M, Arbuckle TE, Liang CL, LeBlanc A, Gaudreau E, **Foster WG**, Haines D, Davis K, Fraser WD. (2016) Concentrations of persistent organic pollutants in maternal and cord blood from the maternal-infant research on environmental chemicals cohort study. Environ. Health. 15(1):59.

5. Guerra MT, Furlong HC, Kempinas WG, **Foster WG**. (2016) Effects of in vitro exposure to butylparaben and di-(2 ethylhexyl) phthalate, alone or in combination, on ovarian function. J. Appl. Toxicol. 26(9):1235-1245.

6. Furlong HC, Wessels JM, Guerra MT, Stämpfli MR, **Foster WG**. (2016) Hydroxychloroquine attenuates cigarette smoke induced autophagic signaling in the mouse ovary. Reprod. Toxicol. 61:105-113.

7. Kubwabo C, Kosarac I, **Foster WG**. (2016) Quantitative determination of nine urinary metabolites of organophosphate flame retardants using solid phase extraction and ultra performance liquid chromatography coupled to tandem mass spectrometry (UPLC-MS/MS). J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. 1014:24-30.

8. **Foster WG**. (2016) Diagnosing endometriosis: CA125 rules in, but not out. BJOG. 123(11):1769.

9. Perreault M, Xu VY, Hamilton S, Wright D, **Foster W**, Atkinson S. (2016) Validation of a food frequency questionnaire for bone nutrients in pregnant women. CJDPR. 77(3):133-139.

10. Stewart J, Breslin WJ, Beyer BK, Chadwick K, De Schaepdrijver L, Desai M, Enright B, **Foster W**, Hui JY, Moffat GJ, Tornesi B, Van Malderen K, Wiesner L, Chen CL. (2016) Birth control in clinical trials: Industry survey of current use practices, governance, and monitoring. Therapeutic Innovation & Regulatory Science. 50(2):155-168.

11. Wessels JM, Kay VR, Leyland NA, Agarwal SK, **Foster WG**. (2016) Assessing brain-derived neurotrophic factor as a novel clinical marker of endometriosis. Fertil. Steril. 105(1):119-128.

12. Agarwal SK, Daniels A, Drosman SR, Udoff L, Foster WG, Pike M, Spicer DV, Daniels JR. (2015) Treatment of endometriosis with the GnRHa deslorelin and add-back estradiol and supplementary testosterone. BioMed. Res. Intl. 2015:934164.

13. Furlong H, Stämpfli MR, Gannon AM, **Foster WG**. (2015) Cigarette smoke exposure triggers the autophagic cascade via activation of the AMPK pathway. Biol. Reprod. 93(4):1-10.

14. Agarwal SK, and **Foster WG**. (2015) Reduction in endometrioma size with three months aromatase inhibition and progestin add-back. Biomed. Res. Intl. (Obstet. & Gynecol.) Special Issue: The Impact of Endometriosis on the Health of Women. 2015:878517.

15. Hughes CL, **Foster WG**, Agarwal SK. (2015) The impact of endometriosis across the lifespan of women: Foreseeable research and therapeutic prospects. Biomed. Res. Intl. (Obstet. & Gynecol.) Special Issue: The Impact of Endometriosis on the Health of Women. 2015:158490.

16. Hughes CL, **Foster WG**, Agarwal SK, Mettler L. (2015) The impact of endometriosis on the health of women. Biomed. Res. Intl. (Obstet, & Gynecol.) Special Issue: The Impact of Endometriosis on the Health of Women. 2015:365951.

17. Borman ED, **Foster WG**, Greenacre MKE, Muir CC, deCatanzaro D. (2015) Stress lowers the threshold dose at which bisphenol A disrupts blastocyst implantation, in conjunction with decreased uterine closure and e-cadherin. Chem. Biol. Interac. 237:87-95.

18. Wessels JM, Leyland NA, Agarwal SK, **Foster WG**. (2015) Estrogen induced changes in uterine brain-derived neurotrophic factor and its receptors. Hum. Reprod. 30(4):925-936.

19. Adlard B, Davis K, Liang CL, Curren M, Rodriguez-Dozal S, Riojas-Rodriguez H, Hernández-Ávila M, **Foster** W, Needham L, Wong LY, Weber JP, Marro L, Leech T, Van Oostdam J. (2014) Persistent organic pollutants (POPs) and metals in primiparous women: A comparison from Canada and Mexico. Sci. Total Environ. 500-501:302-313.

20. Cruz G, **Foster W**, Paredest A, Yi KD, Uzumcu M. (2014) Long-term effects of early-life exposure to environmental oestrogens on ovarian function: Role of epigenetics. J. Neuroendocrinol. 26(9):613-624.

21. Lehmann GM, Verner MA, Luukinen B, Henning C, Assimon SA, LaKind JS, McLanahan ED, Phillips LJ, Verner MA, Davis MH, Powers CM, Erin HP, Haddad S, Longnecker MP, Poulsen MT, Farrer DG, Marchitti SA, Tan YM, Swartout JC, Sagiv SK, Welsh C, Campbell Jr. JL, **Foster WG**, Yang RSH, Fenton SE, Tornero-Valez R, Francis BM, Barnett JB, El-Masri HA, Simmons JE. (2014) Improving the risk assessment of lipophilic persistent environmental chemicals in breast milk. Crit. Rev. Toxicol. 44(7):600-617.

22. Kay VR, Bloom MS, **Foster WG**. (2014) Reproductive and developmental effects of phthalate diesters in males. Crit. Rev. Toxicol. 44(6):467-498.

23. Kubwabo C, Kosarac I, Lalonde K, **Foster WG**. (2014) Quantitative determination of free and total bisphenol A in human urine using labelled BPA-glucuronide and isotope dilution mass spectrometry. Anal. Bioanal. Chem. 406(18):4381-4392.

24. Arbuckle TE, Davis K, Marro L, Fisher M, Legrand M, LeBlanc A, Gaudreau E, **Foster WG**, Choeurng V, Fraser WD and the MIREC Study Group. (2014) Phthalate and bisphenol A exposure among pregnant women in Canada – Results from the MIREC study. Environ. Int. 68:55-65.

25. Lamb JC 4th, Boffetta P, **Foster WG**, Goodman JE, Hentz K, Rhomberg LR, Staveley J, Swaen G, Van Der Kraak G, Williams AL. (2014) Critical comments on the WHO-UNEP state of the science on endocrine disrupting chemicals – 2012. Regul. Toxicol. Pharmacol. 69(1):22-40.

26. Curren MS, Davis K, Liang CL, Adlard B, Thuppal V, Said F, **Foster WG**, Donaldson S, Van

Oostdam J. (2014) Comparing plasma concentrations of persistent organic pollutants in primiparous women from Northern and Southern Canada. Sci. Total Environ. 1:479-480.

27. Wessels J, Leyland N, **Foster WG**. (2014) The brain-uterus connection: Brain derived neurotrophic factor (BDNF) and its receptor (Ntrk2) are conserved in the mammalian uterus. PLoS One. 9(4):e94036.

28. Toor K, Wessels JM, Agarwal SK, Leyland NA, **Foster WG**. (2014) Clinical markers of endometriosis: Have we been too quick to judge? Med. Hypotheses. 82(4):493-501.

29. Rajabi N, Thorpe JB, **Foster WG**, deCatanzaro D. (2013) Novel male exposure reduces uterine e-cadherin, increases uttering luminal area, and diminishes progesterone levels while disrupting blastocyst implantation in inseminated mice. J. Steroid Biochem. Mol. Biol. 139:107-113.

30. Felker AM, Chen Z, **Foster WG**, Croy BA. (2013) Receptors for non-MHC ligands contribute to uterine Natural Killer cell activation during pregnancy in mice. Placenta. 34(9):757-764.

31. Sadeu JC and **Foster WG**. (2013) The cigarette smoke constituent benzo[a]pyrene disrupts metabolic enzyme, and apoptosis pathway member gene expression in ovarian follicles. Reprod. Toxicol. 40:52-59.

32. Mirshokraei P, Hassanpour H, Rahnama A, **Foster WG**. (2013) Gene expression of BDNF and its receptors, TrkB and p75 in the uterus and oviduct of pregnant and non-pregnant ewes. Res. Vet. Sci. 95(1):164-168.

33. Arbuckle TE, Fraser WD, Fisher M, Davis K, Liang CL, Lupien N, Bastien S, Velez MP, von DadelszenP, Hemmings DG, Wang J, Helewa M, Taback S, Sermer M, **Foster W**, Ross G, Fredette P, Smith G, Walker M, Shear R, Dodds L, Ettinger AS, Weber JP, D'Amour M, Legrand M, Kumarathasan P, Vincent R, Luo ZC, Platt RW, Mitchell G, Hidiroglou N, Cockell K, Villeneuve M, Rawn DFK, Dabeka R, Cao XL, Becalski A, Ratnayake N, Bondy G, Jin X, Wang Z, Tittlemier S, Julien P, Avard D, Weiler H, LeBlanc A, Muckle G, Boivin M, Dionne G, Ayotte P, Lanphear B, Séguin JR, Saint-Amour D, Dewailly É, Monnier P, Koren G, Ouellet E. (2013) Cohort Profile: The maternal-infant research on environmental chemicals research platform. Paediatr. Perinat. Epidemiol. 27(4):415-425.

34. Kay VR, Chambers C, **Foster WG**. (2013) Reproductive and developmental effects of phthalate diesters in females. Crit. Rev. Toxicol. 43(3):200-219.

35. Gannon AM, Stämpfli MR, **Foster WG**. (2013) Cigarette smoke exposure elicits increased autophagy and dysregulation of mitochondrial dynamics in murine granulosa cells. Biol.

Reprod. 88(3):63.

36. Siddique S, Sadeu JC, **Foster WG**, Feng YL, and Zhu J. (2013) *In vitro* exposure to cigarette smoke induces oxidative stress in follicular cells of F1 hybrid mice. J. Appl. Toxicol. 34(2):224-226.

37. **Foster WG**, Cheung AP, Davis K, Graves G, Jarrell J, Leblanc A, Liang CL, Leech T, Walker M, Weber JP, Van Oostdam J. (2012) Circulating metals and persistent organic pollutant concentrations in Canadian and non-Canadian born primiparous women from five Canadian centres: results of a pilot biomonitoring study. Sci. Total Environ. 435-436:326-336.

38. Peery HE, Day GS, Dunn S, Fritzler MJ, Prüss H, De Souza C, Doja A, Mossman K, Resch L, Xia C, Sakic B, Belbeck L, **Foster WG**. (2012) Anti-NMDA receptor encephalitis. The disorder, the diagnosis and the immunobiology. Autoimmun Rev. 11(12):863-872.

39. Khoufache K, Bondza PK, Harir N, Daris M, Leboeuf M, Mailloux J, Lemyre M, **Foster W,** Akoum A. (2012) Soluble human IL-1 receptor type 2 inhibits ectopic endometrial tissue implantation and growth: identification of a new potential target for endometriosis treatment. Am. J. Pathol. 181(4):1197-1205.

40. Rhomberg LR, Goodman JE, **Foster WG**, Borgert CJ, Van Der Kraak G, (2012) A critique of the European Commission document, "State of the Art Assessment of Endocrine Disrupters". Crit. Rev. Toxicol. 42(6):465-473.

41. Jarrell J, **Foster WG**, Kinniburgh DW. (2012) Phytoestrogens in human pregnancy. Obstet. Gynecol. Int. 2012:850313.

42. Khoufache K, Bazin S, Girard K, Guillemette J, Roy MC, Verreault JP, Al-Abed Y, **Foster W**, Akoum A. (2012) Macrophage migration inhibitory factor antagonist blocks the development of endometriosis in vivo. PLoS One. 7(5):e37264.

43. Clafshenkel WP, King TL, Kotlarczyk MP, Cline JM, **Foster WG,** Davis VL, Witt-Enderby PA. (2012) *Morinda cirtifolia* juice augments mammary gland differentiation and reduces mammary tumor growth in mice expressing the unactivated c-*erb*B2 transgene. Evid. Based Complement Alternat. Med. 2012:487423.

44. Kosarac I, Kubwabo C, Lalonde K, **Foster W**. (2012) A novel method for quantitative determination of free and conjugated bisphenol A in human maternal and umbilical cord blood serum using a two-step solid phase extraction and gas chromatography/tandem mass spectrometry. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 898:90-94.

45. Johnson NA, Ho A, Cline JM, Hughes CL, **Foster WG,** Davis VL. (2012) Accelerated mammary tumor onset in HER2/*Neu* mouse model exposed to DDT metabolites locally delivered to the mammary gland. Environ. Health Perspect. 120(8):1170-1176.

46. Sadeu JC, Doedée AM, Neal MS, Hughes EG, **Foster WG**. (2012) Neurotrophins (BDNF and NGF) in follicular fluid of women with different infertility diagnoses. Reprod. Biomed. Online 24(2):174-179.

47. Gannon AM, Stämpfli MR, **Foster WG**. (2012) Cigarette smoke exposure leads to follicle loss via an alternative ovarian cell death pathway in a mouse model. Toxicol. Sci. 125(1):274-284.

48. Nethery E, Wheeler AJ, Fisher M, Sjödin A, Li Z, Romanoff LC, **Foster W**, Arbuckle TE. (2012) Urinary polycyclic aromatic hydrocarbons as a biomarker of exposure to PAHs in air: a pilot study among pregnant women. J. Expo. Sci. Environ. Epidemiol. 22(1):70-81.

49. McCarver G, Bhatia J, Chambers C, Clarke R, Etzel R, **Foster W**, Hoyer P, Leeder JS, Peters JM, Rissman E, Rybak M, Sherman C, Toppair J, Turner K. (2011) NTP-CERHR expert panel report on the developmental toxicity of soy infant formula. Birth Defects Res. B Dev. Reprod. Toxicol. 92(5):421-468.

50. **Foster WG,** Gregorovich S, Morrison KM, Atkinson SA, Kubwabo C, Stewart B, Teo K. (2011) Human maternal and umbilical cord blood concentrations of polybrominated diphenyl ethers. Chemosphere. 84(10):1301-1309.

51. Sadeu JC and **Foster WG**. (2011) Cigarette smoke condensate exposure delays follicular development and function in a stage-dependent manner. Fertil. Steril. 95(7):2410-2417.

52. Lagunov A, Anzar, M, Sadeu JC, Kahn MI, Bruin JE, Woynillowicz AK, Buhr M, Holloway AC, **Foster WG**. (2011) Effect of *in utero* and lactational nicotine exposure on the male reproductive tract in peripubertal and adult rats. Reprod. Toxicol. 31(4):418-423.

53. Sadeu JC and **Foster WG**. (2011) Effect of *in vitro* exposure to benzo[a]pyrene, a component of cigarette smoke, on folliculogenesis, steroidogenesis and oocyte nuclear maturation. Reprod. Toxicol. 31(4):402-408.

54. Dominguez MA, Cho N, Zhang B, Neal MS, **Foster WG**. (2011) Brain-derived neurotrophic factor expression in granulosa lutein cells. Reprod. Biomed. Online. 22(1):17-24.

55. **Foster WG** and Hughes CL. (2011) Gene expression in oogenesis and implications for transgenerational effects of environmental toxicants. Biol. Reprod. 84(1):2-4.

56. Berger RG, **Foster WG**, deCatanzaro D. (2010) Bisphenol-A exposure during the period of blastocyst implantation alters uterine morphology and perturbs measures of estrogen and progesterone receptor expression in mice. Reprod. Toxicol. 30(3):393-400.

57. Sadeu JC, Hughes CL, Agarwal S, **Foster WG**. (2010) Alcohol, drugs, caffeine, tobacco, and environmental contaminant exposure: reproductive health consequences and clinical implications. Crit. Rev. Toxicol. 40(7):633-652.

58. Wright D, Gregorovich S, Hamilton S, Lukas P, **Foster WG**. (2010) Environmental chemical exposure research and characteristics of study subjects recruited from midwifery vs. hospital based clinics. CJMRP. 9(3):31-35.

59. Neal MS, Mulligan Tuttle AM, Casper RF, Lagunov A, **Foster WG**. (2010) Aryl hydrocarbon receptor antagonists attenuate the deleterious effects of benzo[a]pyrene on isolated rat follicle development. Reprod. Biomed. Online. 21(1):100-108.

60. **Foster WG**, Maharaj-Briceño S, Cyr DG. (2010) Dioxin-induced changes in epididymal sperm count and spermatogenesis. Environ. Health Perspect. 118(4):458-464. Reprinted with permission in: Cien Saude Colet. 16(6):2893-2905 and Review.

61. Morrison KM, Atkinson SA, Yusuf S, Bourgeois J, McDonald S, McQueen MJ, Persadie R, Hunter B, Pogue J, Teo K; FAMILY investigators. (Holloway A, **Foster W**, Steer P, Denburg J, Cyr M, Windsor S, Mohide P, Capes VS, Vaughn-Williams V, Gross J, Abdalla N, Sim C, Wright C, Sivaguranathan N, Singh P, Helden L, Beecroft ML.) (2009) The Family Atherosclerosis Monitoring In earLY life (FAMILY) study: rationale, design, and baseline data of a study examining the early determinants of atherosclerosis. Am. Heart J. 158(4):533-539.

62. **Foster WG**, Elias R, Faghih M, Dominguez MA, Elit L, Boutross-Tadross O. (2009) Immunohistochemical localization of tyrosine kinases A and B in endometriosis associated ovarian cancer. Histopathology. 54(7):907-912.

63. Tuttle AM, Stämpfli M, **Foster WG**. (2009) Cigarette smoke causes follicle loss in mice ovaries at concentrations representative of human exposure. Hum Reprod. 24(6):1452-1459.

64. Cameron HL and **Foster WG**. (2009) Developmental and lactational exposure to dieldrin alters mammary tumorigenesis in Her2/neu transgenic mice. PLoS One. 4(1):e4303.

65. Davis VL, Jayo MJ, Ho A, Kotlarczyk MP, Hardy ML, **Foster WG**, Hughes CL. (2008) Black cohosh increases metastatic mammary cancer in transgenic mice expressing c-erbB2. Cancer Res. 68(20):8377-8383.

66. Monroy R, Morrison K, Teo K, Atkinson S, Kubwabo C, Stewart B, **Foster WG**. (2008) Serum levels of perfluoroalkyl compounds in human maternal and umbilical cord blood samples. Environ. Res. 108(1):56-62.

67. **Foster WG,** Mirshokraei P, Holloway AC, Zhang B. (2008) Developmental and lactational exposure to environmentally relevant concentrations of dieldrin does not alter pregnancy outcome and mammary gland morphology in BALB/c mice. Environ. Res. 108(1):21-27.

68. **Foster WG.** (2008) Environmental estrogens and endocrine disruption: importance of comparative endocrinology. Endocrinology. 149(9):4267-4268.

69. Dominguez MA, Petre MA, Neal MS, **Foster WG**. (2008) Bisphenol A concentration-dependently increases human granulosa-lutein cell matrix metalloproteinase-9 (MMP-9) enzyme output. Reprod. Toxicol. 25(4):420-425.

70. **Foster WG.** (2008) Fetal and early postnatal environmental contaminant exposures and reproductive health effects in the female. Fertil. Steril. 89(2 Suppl):e53-54.

71. Phillips KP, **Foster WG**, Leiss W, Sahni V, Karyakina N, Turner MC, Kacew S, Krewski D. (2008) Assessing and managing risks arising from exposure to endocrine-active chemicals. J. Toxicol. Environ. Health B Crit. Rev. 11(3-4):351–372.

72. Phillips KP and **Foster WG**. (2008) Key developments in endocrine disrupter research and human health. J. Toxicol. Environ. Health B Crit. Rev. 11(3-4):322–344.

73. **Foster WG.** (2008) Endocrine toxicants including 2,3,7,8-terachlorodibenzo-p-dioxin (TCDD) and dioxin-like chemicals and endometriosis: is there a link? J. Toxicol. Environ. Health B Crit. Rev. 11(3-4):177–187.

74. **Foster WG**, Neal MS, Han MS, Dominguez MM. (2008) Environmental contaminants and human infertility: hypothesis or cause for concern? J. Toxicol. Environ. Health B Crit. Rev. 11(3-4):162–176.

75. **Foster WG** and Agzarian J. (2008) Toward less confusing terminology in endocrine disruptor research. J. Toxicol. Environ. Health B Crit. Rev. 11(3-4):152–161.

76. Phillips KP and **Foster WG**. (2008) Endocrine toxicants with emphasis on human health risks. J. Toxicol. Environ. Health B Crit. Rev. 11(3-4):149–151.

77. **Foster W**, Myllynen P, Winn LM, Ornoy A, Miller RK. (2008) Reactive Oxygen Species, Diabetes and Toxicity in the Placenta – A Workshop Report. Placenta. Suppl A, 29(22):S105-S107.

78. Cameron HL and **Foster WG**. (2008) Dieldrin promotes resistance of anoikis in breast cancer cells *in vitro*. Reprod. Toxicol. 25(2):256-262.

79. Neal MS, **Foster WG,** Younglai EV. (2008) The detrimental effects of smoking on female fertility and IVF success. Curr. Women's Hlth. Rev. 4:16-24.

80. Neal MS, Zhu J, **Foster WG**. (2008) Quantification of benzo[a]pyrene and other PAHs in the serum and follicular fluid of smokers versus non-smokers. Reprod. Toxicol. 25(1):100-106.

81. Anger DL and **Foster WG**. (2008) The link between environmental toxicant exposure and endometriosis. Front. Biosci. 13:1578-1593.

82. Holloway AC, Anger DA, Crankshaw DJ, Wu M, **Foster WG.** (2008) Atrazine-induced changes in aromatase activity in estrogen sensitive target tissues. J. Appl. Toxicol. 28(3):260-270.

83. Younglai EV, Wu YJ, **Foster WG.** (2007) Reproductive toxicology of environmental toxicants: emerging issues and concerns. Curr. Pharm Des. 13(29):3005-3019.

84. Anger DL, Zhang B, Boutross-Tadross O, **Foster WG**. (2007) Tyrosine receptor kinase B (TrkB) protein expression in the human endometrium. Endocrine. 31(2):167-173.

85. Agarwal SK, Estrada S, **Foster WG**, Wall LL, Brown D, Revis ES, Rodriguez S. (2007) What motivates women to take part in clinical and basic science endometriosis research? Bioethics. 21(5):263-269.

86. Neal MS, Zhu J, Holloway AC, **Foster WG.** (2007) Follicle growth is inhibited by benzo-[a]-pyrene, at concentrations representative of human exposure in an isolated rat follicle culture assay. Hum. Reprod. 22(4): 961-967.

87. **Foster WG** and Agzarian J. (2007) Reporting results of biomonitoring studies. Anal. Bioanal. Chem. 387(1):137-140.

88. Weselak AM, Arbuckle TE, **Foster WG**. (2007) Pesticide exposures and developmental outcomes: The epidemiological evidence. J. Toxicol. Environ. Health B Crit. Rev. 10(1-2):41-80.

89. Anger DL, Crankshaw DJ, **Foster WG.** (2006) Spontaneous appearance of uterine tumors in vehicle and 3-methylcholanthrene-treated Wistar rats. Repro. Toxicol. 22(4):760-764.

90. Younglai EV, Wu Y, **Foster WG**, Lobb DK, Price TM. (2006) Binding of progesterone to cell surfaces of human granulosa-lutein cells. J. Steroid Biochem. Mol. Biol. 101(1):61-67.

91. Wu Y, **Foster WG,** Younglai EV. (2006) Rapid effects of pesticides on human granulosa-lutein cells. Reproduction. 131(2):299-310.

92. **Foster WG**, Neal MS, YoungLai, EV. (2006) Ovarian toxicity of environmental toxicants. Immunol. Endocrinol. Metabolic Agents in Med. Chem. 6(1):37-43.

93. YoungLai EV, Wu YJ, **Foster WG.** (2006) Do pesticides have adverse effects on reproduction? Immunol. Endocrinol. Metabolic Agents in Med. Chem. 6(1):45-56.

94. Younglai EV, Wu Y, **Foster WG.** (2006) Rapid action of pesticides on cytosolic calcium concentrations in cultures of human umbilical vein endothelial cells. Repro. Toxicol. 21(3):271-279.

95. Liu G, Shi F, Blas-Machado U, Yu R, Davis VL, **Foster WG**, Magoffin DA, Hughes CL. (2006) Dietary galactose inhibits GDF-9 mediated follicular development in the rat ovary. Reprod. Toxicol. 21(1):26-33.

96. Holloway AC, Lim GE, Petrik JJ, **Foster WG**, Morrison KM, Gerstein HC. (2005) Fetal and neonatal exposure to nicotine in Wistar rats results in increased beta cell apoptosis at birth and postnatal endocrine and metabolic changes associated with type 2 diabetes. Diabetologia. 48(12):2661-2666.

97. **Foster WG**, Holloway AC, Hughes CL Jr. (2005) Dioxin-like activity and maternal thyroid hormone levels in second trimester maternal serum. Am. J. Obstet. Gyneol. 193(6):1900-1907.

98. **Foster WG**. (2005) Subclinical hypothyroidism increased the risk of placental abruption and poor neonatal outcomes. EBM. 10:153. (Commentary).

99. Yauk CL, Gingerich JD, Soper L, MacMahon A, **Foster WG**, Douglas GR. (2005) A *lacZ* transgenic mouse assay for the detection of mutations in follicular granulosa cells. Mutat. Res. 578(1-2):117-123.

100. Edmunds KM, Holloway AC, Crankshaw DJ, Agarwal SK, **Foster WG**. (2005) The effects of dietary phytoestrogens on aromatase activity in human endometrial stromal cells. Reprod. Nutr. Develop. 45(6):709-720.

101. Neal MS, Hughes EG, Holloway AC, **Foster WG**. (2005) Sidestream smoking is equally as damaging as mainstream smoking on IVF outcomes. Hum. Reprod. 20(9):2531-2535. (***Nominated for the 7th Annual Royan Award for one of the five best papers published in the field of reproductive medicine for the year 2005***).

102. Holloway AC, Stys KA, **Foster WG**. (2005) DDE-induced changes in aromatase activity in endometrial stromal cells in culture. Endocrine. 27(1):45-50.

103. Gao YJ, Holloway AC, Zheng ZH, Lim GE, Petrik JJ, **Foster WG**, Lee RM. (2005) Prenatal exposure to nicotine causes postnatal obesity and altered perivascular adipose tissue function. Obes. Res. 13(4):687-692.

104. Miller ME, Holloway AC, **Foster WG**. (2005) Benzo-[*a*]-pyrene increases invasion in MDA-MB-231 breast cancer cells via increased COX-II expression and prostaglandin E_2 (PGE_2) output. Clin. Exp. Metastasis. 22(2):149-156.

105. Liu G, Shi F, Blas-Machado U, Duong Q, Davis VL, **Foster WG**, Hughes CL. (2005) Ovarian effects of high lactose diet in the female rat. Reprod. Nutr. Dev. 45(2):185-192.

106. YoungLai EV, Holloway AC, **Foster WG**. (2005) Environmental and occupational factors affecting fertility and IVF success. Hum. Reprod. Update. 11(1):43-57.

107. Lim GE, Stals SI, Petrik JJ, **Foster WG**, Holloway AC. (2004) The effects of *in utero* and lactational exposure to chloroform on postnatal growth and glucose tolerance in male Wistar rats. Endocrine. 25(3):223-228.

108. Neal MS, Younglai EV, Holloway AC, **Foster WG**. (2004) Aromatase activity in granulosa cells as a predictor of pregnancy potential. Int. Congress Series. 1271:139-142.

109. **Foster WG**, Neal MS, YoungLai EV. (2004) Endocrine disrupters and ovarian function. Int. Congress Series. 1266:126-132.

110. Younglai EV, Holloway AC, Lim GE, **Foster WG**. (2004) Synergisticefects between FSH and 1,1-dichloro-2,2-bis(P-chlorophenyl)ethylene (*p,p'*-DDE) on human granulosa cell aromatase activity. Hum. Reprod. 19(5):1089-1093.

111. Younglai EV, Kwan TK, Kwan CY, Lobb DK, **Foster WG**. (2004) Dichlorodiphenylchloroethylene elevates cytosolic calcium concentrations and oscillations in primary cultures of human granulosa-lutein cells. Biol. Reprod. 70(6):1693-1700.

112. Hughes CL, Liu G, Beall S, **Foster WG**, Davis V. (2004) Effects of genistein or soy milk during late gestation and lactation on adult uterine organization in the rat. Exp. Biol. Med. 229(1):108-117.

113. **Foster WG**, Younglai EV, Bouttross-Tadross O, Hughes CL, Wade MG. (2004) Mammary gland morphology in Sprague-Dawley rats following treatment with an organochlorine

mixture *in utero* and neonatal genistein. Toxicol. Sci. 77(1):91-100.

114. Liu G, Hughes CL, Mathur R, **Foster WG**, Davis VL, Magoffin DA. (2003) Metabolic effects of dietary lactose in adult female rats. Reprod. Nutr. Dev. 43(6):567-576.

115. **Foster WG**. (2003) Environmental toxicants and human fertility. Minerva Ginecol. 55(5):451-457.

116. Rier S and **Foster WG**. (2003) Environmental dioxins and endometriosis. Semin. Reprod. Med. 21(2):145-154.

117. **Foster WG**. (2003) Do environmental contaminants adversely affect human reproductive physiology? J. Obstet. Gynaecol. Can. 25(1):33-44.

118. Rier S and **Foster WG**. (2002) Environmental dioxins and endometriosis. Toxicol. Sci. 70(2):161-170.

119. **Foster WG**, Hughes CL, Chan S, Platt L. (2002) Human developmental exposure to endocrine active compounds. Environ. Toxicol. Pharmacol. 12(2):75-81.

120. Younglai EV, **Foster WG**, Hughes EG, Trim K, Jarrell JF. (2002) Levels of environmental contaminants in human follicular fluid, serum, and seminal plasma of couples undergoing *in vitro* fertilization. Arch. Environ. Contam. Toxicol. 43(1):121-126.

121. Wade MG, Parent S, Finnson KW, **Foster W**, Younglai E, McMahon A, Cyr DG, Hughes C. (2002) Thyroid toxicity due to subchronic exposure to a complex mixture of 16 organochlorines, lead and cadmium. Toxicol. Sci. 67:207-218.

122. Wade MG, **Foster WG**, Younglai EV, McMahon A, Leingartner K, Yagminas A, Blakey D, Fournier M, Desaulniers D, Hughes CL. (2002) Effects of subchronic exposure to a complex mixture of persistent contaminants in male rats: systemic, immune and reproductive effects. Toxicol. Sci. 67(1):131-143.

123. **Foster WG** and Agarwal SK. (2002) Environmental contaminants and dietary factors in endometriosis. Ann. New York Acad. Sci. 955:213-229.

124. **Foster W**, Chan S, Platt L, Hughes C. (2002) Detection of dietary phytoestrogens in samples of second trimester human amniotic fluid. Tox. Lett. 129:199-205.

125. Harris R, **Foster W,** Surrey M, Agarwal SK. (2001) The association between right lower quadrant abdominal pain and appendiceal pathology in women with endometriosis. J. Am. Assoc. Gynecol. Laproscop. 8(4):536-541.

126. Surekha S, Farley A, Reid RL, **Foster WG**, Van Vugt DA. (2001) The Effect of 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin on Corticotrophin-Releasing Hormone, Arginine Vasopressin and Pro-opiomelanocortin Messenger Ribonucleic Acid Levels in the Hypothalamus of the Cynomolgus Monkey. Tox. Sci. 63:181-188.

127. Hughes Jr. CL and **Foster WG.** (2001) Environmental Estrogens and Anti-androgens in Women's Health. Menopausal Medicine. 9:7-12.

128. **Foster WG**. (2001) Endocrine Disruption and Human Reproductive Effects: An overview. Water Qual. J. Can. 36(2):253-271.

129. Hughes CL, **Foster W**, Chan S, Platt L, Thompson S, Hubbard S, DuBose A, Tyrey L. (2001) Extrapolation of rodent studies on amniotic fluid contaminants to human populations. Human Ecological Risk Assessment. 7:979-1002.

130. Pryor JL, Hughes C, **Foster WG**, Hales B, Robaire B. (2000) Critical windows of exposure for children's health: Reproductive system. Environ. Health Perspect. 108:491-503.

131. **Foster W**, Chan S, Platt L, Hughes C. (2000) Detection of endocrine disrupting chemicals in samples of second trimester human amniotic fluid. J. Clin. Endocrinol. Metab. 85:2954-2957.

132. Yang JZ, Agarwal SK, **Foster WG**. (2000) Subchronic exposure to 2,3,7,8-Tetrachlorodibenzo-*p*-Dioxin modulates the pathophysiology of endometriosis in the Cynomolgus monkey. Tox. Sci. 56:374-381.

133. Van Vugt DA, Krzemien A, Roy BN, Fletcher WA, **Foster W**, Lundahl S, Marcus SL, Reid RL. (2000) Photodynamic endometrial ablation in the nonhuman primate. J. Soc. Gynecol. Investig. 7:125-130.

134. **Foster WG**, Desaulniers D, Leingartner K, Wade MG, Poon R, Chu I. (1999) Reproductive effects of tris(4-chlorophenyl)methanol in the rat. Chemosphere. 39:709-724.

135. Desaulniers D, Leingartner K, Wade M, Fintelman E, Yagminas A, **Foster WG**. (1999) Effects of acute exposure to PCB 126 and 153 on anterior pituitary and thyroid hormones and FSH isoforms in adult Sprague Dawley male rats. Toxicol. Sci. 47: 158-169.

136. YoungLai EV, Collins JA, **Foster WG**. (1998) Canadian semen quality: An analysis of sperm density among eleven academic fertility centres. Fertil. Steril. 70(1):76-80.

137. Jarrell J, Gocmen A, **Foster W**, Brant R, Chan S, Sevcik M. (1998) Evaluation of reproductive outcomes in women inadvertently exposed to hexachlorobenzene in south-

eastern Turkey in the 1950s. Reproductive Toxicology. 12:469-476.

138. Desaulniers D, Leingartner K, Zacharewski T, **Foster WG**. (1998) Optimisation of an MCF7-E3 cell proliferation assay and effects on environmental pollutants and industrial chemicals. Toxicol. *In Vitro*. 12:409-422.

139. **Foster WG**. (1998) Endocrine Disrupters and development of the reproductive system in the fetus and children: Is there cause for concern? Can. J. Pub. Hlth. 89:S37-S41.

140. Todoroff EC, Sevcik M, Villeneuve DC, **Foster WG**, Jarrell JF. (1998) The effect of photomirex on the *in vitro* perfused ovary of the rat. Repro. Toxicol. 12:305-316.

141. Poon R, Chu I, Lecavalier P, Valli VE, **Foster W**, Gupta S, Thomas B. (1998) Effects of Antimony on rats following 90-day exposure via drinking water. Food Chem. Toxicol. 36:21-35.

142. Desaulniers D, Poon R, Phan W, Leingartner K, **Foster WG**, Chu I. (1998) Reproductive and Thyroid Hormone Levels in Rats Following 90-Day Dietary Exposure to CB (2,4,4'-Trichlotobiphenyl) or CB 77 (3,3',4,4'-tetrachlorobiphenyl). Toxicol. Indust. Health. 13:627-638.

143. **Foster WG**, Singh A, McMahon A, Rice DC. (1998) Chronic Lead-exposure effects in the Cynomolgus monkey (*Macaca fascicularis*) testis. Ultrastruct. Pathol. 22:63-71.

144. **Foster WG**, Ruka MP, Gareau P, Foster RA, Janzen EG. (1997) Histopathological characteristics of endometrial transplants in a new model of endometriosis. Can. J. Physiol. Pharmacol. 75:1188-1196.

145. Wade MG, Desaulniers D, Leingartnere K, **Foster WG**. (1997) Interactions between endosulfan and dieldrin on estrogen-mediated processes *in vitro* and *in vivo*. Reproductive Toxicol. 11:791-798.

146. Yang JZ and **Foster WG**. (1997) Continuous exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin inhibits the growth of surgically-induced endometriosis in the ovariectomized mouse treated with high dose estradiol. Toxicol. Industr. Health. 13:15-25.

147. Yang JZ, **Foster WG**. (1997) Stimulating effects of 4-Chlorodiphenyl Ether on surgically induced endometriosis in the mouse. Repro. Toxicol. 11:69-75.

148. **Foster WG**, McMahon A, Reed BL, Rice DC. (1996) Sperm chromatin structure is altered in Cynomolgus monkeys with environmentally relevant blood lead levels. Toxicol. Industr. Health. 12:723-735.

149. **Foster WG**, McMahon A, Rice DC. (1996) Subclinical changes in luteal function in cynomolgus monkey with moderate blood lead levels. J. Appl. Toxicol. 16:159-163.

150. Yang JZ, Van Vugt DA, Kennedy JC, **Foster WG**, Reid RL. (1996) Intrauterine 5-Aminolevulinic acid induces selective endometrial photosensitization in the rhesus and cynomolgus monkeys. J. Soc. Gynecol. Invest. 3:152-157.

151. **Foster WG**, Jarrell JF, YoungLai EV, Wade MG, Arnold DL, Jordan SA. (1996) An overview of some Reproductive Toxicology Studies: Conducted at Health Canada. Toxicol. Industr. Health. 12:447-459.

152. **Foster WG**. (1995) Reproductive Toxicology of Great Lakes Contaminants. Environ. Hlth. Perspec. 103(Suppl 9):63-69.

153. **Foster WG**, McMahon A, YoungLai EV, Hughes EG, Jarrell JF. (1995) Alterations in ovarian regulation following Hexachlorobenzene exposure during spontaneous cycles and ovulation induction in the cynomolgus monkey. Reprod. Toxicol. 9:541-548.

154. Bourque AC, Singh A, Lakanpal N, McMahon A, **Foster WG**. (1995) Ultrastructural changes in ovarian follicles of the monkey administered hexachlorobenzene. Am. J. Vet. Res. 56:1673-1677.

155. **Foster WG**, Mertineit C, McMahon A, Lecavalier PR. (1995) The effect of Hexachlorobenzene (HCB) on adrenal steroidogenesis in the ovariectomized female rat. J. Biochem. Toxicol. 10:129-135.

156. Yie S-M, Brown GM, Liu G-Y, Collins JA, Daya S, Hughes EG, **Foster WG**, YoungLai EV. (1995) Melatonin and steroids in human preovulatory follicular fluid: Seasonal variations and granulosa cell steroid production. Hum. Reprod. 10:50-55.

157. **Foster WG**, Jarrell JF, Hughes EG, YoungLai EV. (1994) The reproductive toxicology of chemotherapeutic agents and environmental toxins. Invited review: Current Trends in Exp. Endocrinol. 2:65-86.

158. Bourque A, Singh A, Dykeman A, McMahon A, **Foster W**. (1994) Hexachlorobenzene at low doses produces lesions in nonhuman primate ovary. Experientia. 50:A87.

159. Singh A, Cullen C, Dykeman A, Rice D, **Foster W**. (1993) Chronic lead exposure induces ultrastructural alterations in the monkey testis. J. Submicrosc. Cytol. Pathol. 25:479-486.

160. **Foster WG**, Jarrell JF, YoungLai EV. (1993) Developmental changes in the GnRH neuron of

the female rabbit: Effects of Tamoxifen citrate and Pregnant Mare Serum Gonadotropin. Can. J. Physiol. Pharmacol. 71:761-767.

161. **Foster WG**, McMahon A, YoungLai EV, Hughes EG, Rice DC. (1993) Reproductive endocrine effects of chronic Lead-exposure in the male cynomolgus monkey (*Macaca fascicularis*). Reprod. Toxicol. 7:203-209.

162. Cullen C, Singh A, Dykeman A, Rice D, **Foster W**. (1993) Chronic lead exposure induces ultrastructural alterations in the monkey seminal vesicle. J. Submicroscopic Cytol. Pathol. 25:127-135.

163. MacPhee IJ, Singh A, Wright GM, **Foster WG**, LeBlanc NN. (1993) Ultrastructure of granulosa lutein cells from rats fed hexachlorobenzene. Histol. Histopath. 8:35-40.

164. **Foster WG**, Pentick JA, McMahon A, Lecavalier PR. (1993) Body distribution and endocrine toxicity of Hexachlorobenzene (HCB) in the female rat. J. Appl. Toxicol. 13:79-83.

165. **Foster WG**, Stals SI, McMahon A. (1992) An ultrasound study of the effect of chronic Lead-exposure on endometrial cycle changes in the female cynomolgus monkey. Med. J. Primatol. 21:353-356.

166. YoungLai EV, Todoroff EC, **Foster WG**, Brown GM. (1992) Light-related pituitary response to gonadotropin-releasing hormone during sexual development in the female rabbit. Biol. Signals. 1:219-227.

167. **Foster WG**, Stals SI, McMahon A. (1992) A prospective analysis of endometrial cycle changes by ultrasound in the female cynomolgus monkey. Med. J. Primatol. 21:30-34.

168. Singh A, **Foster WG**, McMahon A, Rice DC, Villeneuve DC. (1992) Electron microscopy of seminal vesicles from monkeys exposed to Lead: A 9 year study. Experientia. 48:A5.

169. **Foster WG**. (1992) Reproductive toxicity of chronic Lead-exposure in the female cynomolgus monkey. Reprod. Toxicol. 6:123-131.

170. **Foster WG**, McMahon A, Villeneuve DC, Jarrell JF. (1992) Hexachlorobenzene (HCB) suppresses progesterone secretion during the luteal phase in the cynomolgus monkey. J. Appl. Toxicol. 12:13-17.

171. **Foster WG**, Pentick JA, McMahon A, Lecavalier PR. (1992) Ovarian toxicity of Hexachlorobenzene (HCB) in the superovulated rat. J. Biochem. Toxicol. 7:1-4.

172. **Foster WG** andYoungLai EV. (1991) An immunohistochemical study of the GnRH neuron

morphology and topography in the adult female rabbit Hypothalamus. Am. J. Anat. 191:293-300.

173. YoungLai EV, Thompson N, **Foster W**. (1990) Effects of steroid implants on the prepubertal increase in circulating gonadotropins and sexual receptivity in the female rabbit. J. Steroid. Biochem. 35(3-4):416-419.

174. **Foster WG**, Jarrell JF, Dolovich J, YoungLai EV. (1989) Immunoglobin-mediated hypersensitivity in response to chronic treatment with Gonadorelin-HCl (Factrel) in a female patient. Am. J. Obstet. Gynecol. 160(4):979-983.

175. YoungLai EV, Thompson N, **Foster W**. (1989) Effects of *In-Vivo* administration of GnRH on the release of gonadotropins in the female rabbit. J. Reprod. Fertil. 85(1):325-329.

176. **Foster WG**, Boyd WH. (1989) A light microscopic study of the hypophyseal angioarchitecture in the rabbit. Am. J. Anat. 184(3):205-211.

iv) **other, including proceedings of meetings –**

1. Bartell SM, Bois Fy, Calaf GM, Chiu WA, Connolly L, Demers PA, DeVito MJ, **Foster WG**. Friesen M, Fritschi L, Gibbons C, Hooth MJ, McLean DJ, Nishikawa A, Ross MK, Sergi CM, Umemura T, Yiin JH. (2016) Pentachlorophenol, Aldrin and Dieldrin; Volume 117. International Agency for Research on Cancer (IARC). Lyon, France. October 1 – 12, 2016.

2. **Foster WG**. (2015) Human exposure to environmental contaminants and congenital anomalies: a systematic review. Public Health Agency of Canada.

3. Lamb JC 4th, Boffetta P, **Foster WG**, Goodman JE, Hentz K, Rhomberg LR, Staveley J, Swaen G, Van Der Kraak G, Williams AL. (2014) Critical review of WHO-UNEP state of the science of endocrine disrupting chemicals – 2012.

4. Assimon SA, Barnett J, Campbell J, Davis M, El-Masri H, Farrer D, Fenton S, Foster P, **Foster W**, Francis B, Haddad S, Karmaus W, Knadle S, Lakind J, Lehmann G, Longnecker M, Marchitti S, McLanahan E, Poulsen M, Rogan W, Sagiv S, Simmons JE, Swartout J, Tornero-Velez R, Verner M, Welsh C, Yang R. (2013) Improving the risk assessment of persistent, bioaccumulative, and toxic chemicals in breast milk. Workshop summary report. National Center for Environmental Assessment, US Environmental Protection Agency. *ofmpub.epa.gov/eims/eimscomm.getfile?p_download_id=516257*

5. **Foster WG**, Elliott SJ, Eyles JD, Gregorovich S, Kalsi I, Siu S, Crosse E, van Zandvoort M, Reffle J, Turner W, Pollett GL. (2011) Results of a polychlorinated byphenyl (PCB) blood survey in former transformer workers and Pottersburg Creek residents. Ontario Ministry of

Health and Long-Term Care. Report to Middlesex London Health Unit.

6. **Foster WG**. (2011) Tri-national Biomonitoring Study: I. Assessments of persistent pollutants and selected metals in the blood of first birth mothers in southern Canada and Mexico and in women of reproductive age in the United States. Commission for Environmental Cooperation. Ottawa, ON.

7. McCarver G, Bhatia J, Chambers C, Clarke R, Etzel R, **Foster W**, Hoyer P, Leeder JS, Peters J, Rissman E, Rybak M, Sherman C, Toppair J, Turner K. (2010) NTP Final CERHR Expert Panel Report on Soy Infant Formula. NIEHS-NIH, US Dept. of Health & Human Services.

8. Kubwabo C, Gregorovich S, Monroy R, Morrison K, Atkinson S, Stewart B, Teo K, **Foster WG**. (2009) Determination of polybrominated diphenyl ethers in human maternal serum and cord blood samples using accelerated solvent extraction and GC/EI-MS/MS. 29th International Symposium on Halogenated Persistent Organic Pollutants (Dioxin 2009), Beijing, China. August 23 – 28, 2009.

9. Vélez MP, Monnier P, **Foster WG**, Fraser WD. (2008) The impact of phthalates on women's reproductive health: State-of-the science and future directions. National Network on Environment and Women's Health.

10. **Foster WG** and Rousseaux C. (1995) The reproductive toxicology of great lakes contaminants. In: Proceedings of the State of the Lakes Environment Conference.

11. **Foster WG**, Singh A, Rice DC, McMahon A. (1991) Reproductive effects of chronic Lead-exposure in the male cynomolgus monkey (*Macaca fascicularis*). Symposium On Lead In Adults, Durham, NC. December 9 – 11, 1991.

12. Inskip MJ, Yagminas A, Franklin CA, **Foster W**, Wandelmaier F, Haines D, Blenkinsop J. (1991) Maternal-fetal transfer of Lead in a non-human primate *Macaca fascicularis*: Preliminary studies using stable isotope tracers. In: Proceedings of International Conference on Heavy Metals in the Environment. CEP Consultants.

NON-PEER REVIEWED PUBLICATIONS:

i) **journal articles** –

1. **Foster WG** and Moore E. (2008) Chemical exposures and infertility. J. Infertility Awareness Association of Canada.

2. Neal MS and **Foster WG.** (2005) Applications for *in vitro* follicle culture assays. Fertility World. 3:10-11.

3. **Foster WG**. (2004) Chemical exposures and human fertility. Fertility Magazine.

4. YoungLai EV and **Foster WG**. (2004) Dichlorodiphenylchloroethylene and human fertility. A.R.T. & Science. 3(3):6-8.

5. **Foster WG** and Beecroft ML. (2003) Chemical exposures and human fertility. Infertility Awareness 4:30-31.

6. Hughes CL Jr. and **Foster WG.** (2001) Some potential health effects of endocrine-disrupting chemicals across the lifespan of adult women. ASRM Menopausal Medicine 9(2):7-12.

7. Safe S, **Foster W**, Lamb J, Newbold R, Van Der Kraak G. (2000) Estrogenicity and Endocrine Disruption. CAST. 16:1-16.

8. Yang JZ and **Foster WG**. (1997) Causes of endometriosis: Do environmental contaminants play a role? Infertility Awareness 13:10-13.

ARTICLES IN PREPARATION:

1. Furlong HC, Stampfli MR, Gannon AM, **Foster WG**. (2016) Is cigarette smoke induced autophagy unique to the ovary – A tissue comparison study.

2. Aryal G, Kubwabo C, **Foster WG**. (2016) Free Bisphenol A (BPA) concentrations in maternal and cord blood.

UNPUBLISHED DOCUMENTS:

1. Provisional Patent No. 61889085: An assay for inflammatory disease progression and response to treatment.

2. Provisional Patent No. 796896: Method to predict pregnancy potential of an oocyte.

3. International PCT Application No. PCT/CA2007/002114: Trkβ: a diagnostic tool in endometriosis.

PAPERS GIVEN AT SCIENTIFIC MEETINGS:

i) **invited presentations** –

1. **Foster WG**. Life-style and environmental effects on ovarian function: Myths or real concerns? 62nd Annual Meeting of the Canadian Fertility & Andrology Society. Toronto, ON. September 22 – 24, 2016.

2. **Foster WG**. Environmental toxicant exposure and dysregulation of ovarian function. Environment & Reproductive Science Summit 2016 – Environmental, Nutritional, and Genetic Factors Affecting Reproduction. Sponsored by the American Society for Reproductive Medicine, the Society for the Study of Reproduction and the Society of Reproductive Biologists and Technologists. The ONLY Canadian invited speaker. Dallas/Fort Worth TX. March 5 – 6, 2016.

3. **Foster WG**. Environmental toxicant exposure induced dysregulation of ovarian function. 48th Annual Meeting of the Society for the Study of Reproduction. San Juan, Puerto Rico. June 18 – 22, 2015.

4. **Foster WG**. Toxicant-induced changes in ovarian function and follicle development: From clinic to bench and back. Texas A & M University. College Station, TX. February 2 – 3, 2015.

5. **Foster WG**. Endometriosis: clinical and experimental aspects. Female Reproductive Tract session. International Conference on Reproductive Biology and Toxicology, Auditorium of the University of São Paulo Law School (FDRP). Ribeirão Preto, São Paulo, Brazil. November 10 – 11, 2014.

6. **Foster WG**. Research: making tomorrow better today. Round table discussion, Translation of Basic Science to Real-World Practice. International Conference on Reproductive Biology and Toxicology, Auditorium of the University of São Paulo Law School (FDRP). Ribeirão Preto, São Paulo, Brazil. November 10 – 11, 2014.

7. **Foster WG**. The Jekyll and Hyde tale of neurotrophin expression in the endometrium. Escola Paulista de Medicina, Universidade Federal de São Paulo, Centro de Pesquisa em Urologia, Ribeirão Preto, São Paulo, Brazil. November 10 – 11, 2014.

8. **Foster WG.** Clinical markers of endometriosis. University of São Paulo, São Paulo, Brazil. November 5, 2014.

9. **Foster WG**. Cigarette smoke exposure induced dysregulation of mitochondrial homeostasis in granulosa cells. University State de São Paulo (UNESP), Botucatu, Brazil. November 6, 2014.

10. **Foster WG**. The role of phthalates in obesity: what's the skinny? City Wide Endocrine Rounds, University of Toronto. October 2014.

11. **Foster WG.** Publish or Perish: The pitfalls and tricks to getting your scientific article published. XIII Workshop da Pós-Graduação, Publicação, Pesquisa e Ensino, Salão Nobre da FMB, UNESP, Botucatu/ São Paulo, Brazil. June 4 – 7, 2014.

12. **Foster WG.** Endometriosis: Animal models and the role of toxicants. XIII Workshop da Pós-Graduação, Publicação, Pesquisa e Ensino, Salão Nobre da FMB, UNESP, Botucatu/ São Paulo, Brazil. June 4 – 7, 2014.

13. **Foster WG.** The ovarian effects of environmental toxicants: Clinical implications. The Society of Obstetricians and Gynaecologists of Canada, Ontario CME Program, Toronto, ON. November 28 - 30, 2013.

14. **Foster WG**. Characterization of a novel clinical marker of endometriosis. Grand Rounds, Department of Obstetrics & Gynecology, McMaster University, Hamilton, ON. September 4, 2013.

15. **Foster WG.** Exposure to environmental toxicants and ovarian dysfunction. International Workshop in Neuroendocrinology – Brazilian International Symposium on Integrative Neuroendocrinology. Dourado, Brazil. August 4 – 7, 2013.

16. **Foster WG.** Mechanisms of cigarette smoke induced ovarian follicle loss: The bumpy road to discovery. Graduate Program on General and Applied Biology, Institute of Biosciences of Botucatu, Universidade Estadual Paulista (unesp), Brazil. July 29 – August 2, 2013.

17. **Foster WG**. Clinical markers of endometriosis: what's new? Rounds, Department of Obstetrics & Gynecology, McMaster University. Hamilton, ON. May 29, 2013.

18. **Foster WG**. Strengths and limitations of *in vitro* test methods for reproductive toxicology. 52nd Annual Meeting of the Society of Toxicology. San Antonio, TX. March 10 – 14, 2013.

19. **Foster WG**. Women's reproductive health. Canadian Memorial Chiropractic College (CMCC). Toronto, ON. March 1, 2013.

20. **Foster WG**. Environmental toxicants and reproductive health. Centre INRS-Institut Armand Frappier. Laval, QC. December 7, 2012.

21. **Foster WG.** The ovarian toxic effects of cigarette smoke exposure. Florida International University. Miami, FL. November 29, 2012.

22. **Foster WG.** The ovarian effects of environmental toxicants and clinical implications. 58th Annual Meeting of the Canadian Fertility & Andrology Society. Ottawa, ON. September 6 – 9, 2012.

23. **Foster WG.** Effects of endocrine disruption on reproductive function in the female. Center for Research in Biology of Reproduction (CRBR) Laval University, QC. December 15, 2011.

24. **Foster WG.** Mechanisms of cigarette smoke-induced sub-optimal ovarian follicle development and atresia. Department of Veterinary Medicine and Biomedical Sciences, Texas A&M University. College Station, TX. September 30, 2011.

25. **Foster WG.** Windows of susceptibility: improving understanding of physiological and exposure differences. International Council of Chemical Associations Long-Range Research Initiative (ICCA-LRI) & Health Canada Workshop; Advancing Exposure Science to Improve Chemical Safety. Quebec City, QC. June 22 – 23, 2011.

26. **Foster WG**. Mechanisms of environmental toxicant induced ovarian follicle loss. Department of Reproductive Medicine, Seminars in Reproductive Science and Medicine, University of California San Diego. April 5, 2011.

27. **Foster WG.** The environment and its' implication to cancer and other diseases. The Kiwanis Club of Oakville, monthly meeting. Oakville, ON. November 15, 2010.

28. **Foster WG.** Regulating estrogen production and metabolism in estrogen-dependent diseases. Citywide Rounds. University of Toronto, Mount Sinai Hospital. Toronto, ON. November 27, 2009.

29. **Foster WG.** Bisphenol A (BPA): science, policy options and risk communication. Improving the Public Communications of Chemical-related Health Risks Workshop. University of Ottawa, ON. September 30 – October 1, 2009.

30. **Foster WG**. Current and emerging issues in reproductive medicine and the placenta biology. 2009 Human Placenta Workshop, Opening Address. Queen's University. Kingston, ON. July 19, 2009.

31. **Foster WG.** Exposure to environmental toxicants and consequences for adverse health effects. Bay Area Restoration Council's 17th Annual Community Workshop – "Looking Beyond 2015". Parks Canada Discovery Centre. Hamilton, ON. April 25, 2009.

32. **Foster WG.** Effects of environmental contaminants on human reproductive health. Hamilton's 3rd Annual Health Research in the City. Hamilton, ON. February 11, 2009.

33. **Foster WG.** Breast Cancer and Environmental Toxicants: New Approaches to Animal Studies in Research, Environment & Health Seminar Series, Centre for Environment. University of Toronto, ON. January 29, 2009.

34. **Foster WG.** Impact of cigarette smoke and its constituents on ovarian function, Public Health Symposium on Infertility. Centers for Disease Control and Prevention. Atlanta, GA. September 15 – 17, 2008.

35. **Foster WG.** The animal evidence: Critical effects and dose-response. Dioxin 2008. 28th International Symposium in Halogenated Persistent Organic Pollutants (POPs). Birmingham, England, UK. August 17 – 22, 2008.

36. **Foster WG.** Bisphenol A: a reproductive hazard but what are the public health implications? Ontario Public Health Unit and Medical Officers of Health Webinar. June 24, 2008.

37. **Foster WG**. Hormone mimics and human health. The 5th PCB Workshop. New Knowledge Gained from Old Pollutants. Iowa City, IA. May 18 – 22, 2008.

38. **Foster WG.** Environmental hazards posed by chemicals and the problems associated with trying to understand the true risk. The Probus Club of Hamilton Mountain, meeting. Hamilton, ON. May 1, 2008.

39. **Foster WG**, Neal MS, Mulligan Tuttle A, Dominguez MA. Impact of environmental factors on ovarian function. Preconception Care Research: Improving Birth Outcomes and Reproductive Health Workshop. Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), National Institutes of Health (NIH), U.S. Department of Health and Human Services. Potomac, MD. April 14 – 15, 2008.

40. **Foster WG.** Reproductive toxicity of environmental toxicants: state of the science and data gaps. CIHR Institute of Human Development and Child and Youth Health Workshop. Montreal, QC. February 11, 2008.

41. **Foster WG.** Reproductive toxicology: Potential hazards and risks to human reproductive health from environmental contaminants. San Diego Reproductive Endocrinology Society. San Diego, CA. November 15, 2007.

42. **Foster WG.** Reproductive toxicity of environmental contaminants: Potential hazards vs. risk to human reproductive health. Reproductive Medicine Grand Rounds. University of California. San Diego, CA. November 14, 2007.

43. **Foster WG.** Reproductive effects of endocrine disrupting compounds. Seminars in Reproductive Science and Medicine. University of California. San Diego, CA. November 14, 2007.

44. **Foster WG.** Cellular and molecular mechanisms of cigarette smoke-induced reproductive

toxicity. 2nd Sino-Canada Workshop on Reproductive Medicine. Ottawa, ON. October 19 – 21, 2007.

45. **Foster WG.** Emerging issues in Reproductive Environmental Toxicant Research. UCSF-CHE Summit on Environmental Challenges to Reproductive Health and Fertility. University of California at San Francisco, CA. January 28 – 30, 2007.

46. **Foster WG**. Endocrine Disruption. National Policy Consultation Series on Children's Health and Environment. Ottawa, ON. January 23 – 24, 2007.

47. **Foster WG.** Resistance to Anoikis in estrogen dependent disease. National Institutes of Child Health and Human Development, Division of Epidemiology, Statistics, & Prevention Research. Rockville, MD. January 11, 2007.

48. **Foster WG.** Anoikis in estrogen sensitive target tissues and disease. Department of Anatomy and Physiology, Queen's University. Kingston, ON. December 7, 2006.

49. **Foster WG.** Reproductive effects of endocrine toxicants: From the lab to the clinic. 52nd Annual Meeting of the Canadian Fertility & Andrology Society. Ottawa, ON. November 15 – 18, 2006.

50. **Foster WG.** Environmental toxicants and ovarian function. American Society of Reproductive Immunology. Nashville, TN. June 15 – 17, 2006.

51. **Foster WG.** Ovarian cancer: Mechanisms of action for environmental toxicants and dietary chemicals. National Toxicology Program sponsored Hormonally-induced reproductive tumors: Relevance of rodent bioassays workshop. Raleigh, NC. May 22 – 24, 2006.

52. **Foster WG.** Reproductive effects of environmental endocrine toxicants. First Sino-Canada Bilateral Workshop on Reproductive Health Research. Beijing, China. November 15 – 18, 2005.

53. **Foster WG.** 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) induced increase in endometrial aromatase activity is mediated through altered PGE2 metabolism. Institute of Gender and Health Annual Meeting. Vancouver, BC. November 8, 2005.

54. **Foster WG.** Endocrine toxicants and laboratory evidence. Endocrine Toxicants: Ecological and population health impacts. University of Ottawa. Ottawa, ON. October 28, 2005.

55. **Foster WG.** Relative effects of environmental toxicants on aromatase expression and activity in human endometrial stromal cell cultures. NIEHS sponsored workshop on Atrazine. Iowa City, IA. April 25 – 26, 2005.

56. **Foster WG.** Adverse effects of environmental toxicants on female reproduction. Society of Toxicology of Canada, 37th Annual Symposium. Montreal, QC. December 6 – 7, 2004.

57. **Foster WG**. Current context and need for basic research on environmental effects on reproduction. CIHR Workshop, 50th Annual Meeting of the Canadian Fertility & Andrology Society. Jasper, AB. November 24 – 28, 2004.

58. **Foster WG.** Endocrine disrupters and ovarian function. International Federation of Fertility Societies. Montreal, QC. May 23 – 28, 2004.

59. **Foster WG.** Environmental toxicants and breast cancer. Juravinski Cancer Centre. Hamilton, ON. January 2004.

60. **Foster WG.** *In utero* exposure to endocrine toxicants and dietary phytoestrogens. Linus Pauling Institute, Oregon State University. Corvallis, OR. December 18, 2003.

61. **Foster WG.** Are dioxins involved in the pathogenesis of endometriosis? Department of Pharmacology & Toxicology, and the Department of Obstetrics & Gynecology, Queen's University. Kingston, ON. October 30, 2003.

62. **Foster WG**. Environmental toxicants: ovulation and endometriosis. World Congress on Endometriosis. San Diego, CA. February 24 – 27, 2002.

63. **Foster WG.** and YoungLai EV. Presence of environmental contaminants in human follicular fluid, serum, and seminal plasma of couples undergoing *in vitro* fertilization. 28th Aquatic Toxicology Workshop. Winnipeg, MB. September 30 – October 3, 2001.

64. **Foster W.** Chan S. Platt L. and Hughes C. Human developmental exposure to endocrine active compounds. Key Note lecture, 5th International Symposium on Biological Monitoring. Banff, AB. September 18 – 21, 2001.

65. **Foster W.** Human exposure and potential health effects of endocrine disrupting compounds. Symposium # 2. Effects of Pollutants on Humans. Canadian Federation of Biological Societies, 44th Annual Meeting. Ottawa, ON. June 20 – 23, 2001.

66. **Foster WG** and Agarwal SK. Environmental contaminants and dietary factors in endometriosis. Endometriosis: Emerging Research and Intervention Strategies. National Institute of Child Health and Human Development (NICHD). Bethesda, MD. April 9 – 10, 2001.

67. **Foster WG**. Effects of environmental contaminants on ectopic endometrium in animal

models of endometriosis. University of Laval. Quebec City, QC. November 18, 2000.

68. **Foster WG**. Environmental contaminants - effects upon human reproductive physiology. Grand Rounds. Cedars-Sinai Medical Center. Los Angeles, CA. August 16, 2000.

69. **Foster WG**, Chan S, Platt L, Hughes CL. Detection of organochlorine compounds and phytoestrogens in second trimester human amniotic fluid. Int. Assoc. Great Lakes Res. Cornwall, ON. May 22 – 26, 2000.

70. **Foster WG**. Effects of environmental contaminants on ectopic endometrium in animal models of endometriosis. University of Western Ontario. London, ON. April 20, 2000.

71. **Foster WG**, Hughes CL. An overview of endocrine disruption and human health. Endocrine Disruption Workshop: Establishing a National Science Agenda on the scientific assessment of endocrine disrupting substances. Huntsville, ON. February 17 – 19, 2000.

72. **Foster WG**. Environmental contaminants and dietary factors: Consequence for human health. University of Ottawa. Ottawa, ON. November 18, 1999.

73. **Foster WG**, Hughes CL, Platt L, Chan S. Detection of Endocrine Disrupting Chemicals and Dietary Factors in Samples of Second Trimester Human Amniotic Fluid. Environmental Hormones: Past, Present, Future. Tulane University. New Orleans, LA. October 18 – 20, 1999.

74. **Foster WG**, EU/US Transatlantic Co-operation in Human Environmental Health; Expert Panel Meeting on Opportunities for Collaborative EU/US Research Programmes. Ispra, Italy. April 1999.

75. **Foster WG**. 1999 Special Meeting of the Toxicology Forum: Dose-Response Considerations for Potential Endocrine Active Substances. Washington, DC. April 1999.

76. **Foster WG**. Endocrine Disruptors and Human Health Effects: Health Canada Perspective. Southern Ontario Reproductive Biology Workshop. University of Waterloo. Waterloo, ON. May 8, 1998.

77. **Foster WG**. Endocrine Disruptors Panel Member. Health Conference. Montreal, QC. May 15 – 17, 1997.

78. **Foster WG**. Human health effects of endocrine modulating substances: fact or fiction? Grand Rounds, Department of Obstetrics & Gynecology, McMaster University. Hamilton, ON. May 21, 1997.

79. **Foster WG**. Endocrine Disruptors: Potential Risks to Children. What on Earth Conference sponsored by the Canadian Institute of Child Health. Ottawa, ON. May 26 – 27, 1997.

80. **Foster WG**. White House expert committee on endocrine disrupters. Washington, DC. January 27, 1997.

81. **Foster WG**. Environmental Exposures and Human Reproduction: Women. Reproductive Health and The Environment Symposium. Toronto Department of Health. Toronto, ON. May 1996.

82. **Foster WG**. Endocrine Disruptors Chemicals: The Human Connection. DOE Workshop on Endocrine Disrupter Issue. Hull, QC. May 1996.

83. **Foster WG**. Endocrine Disruptors: Human Health Effects. Canadian Chemical Producers Association. Ottawa, ON. June 1996.

84. **Foster WG**. Animal models in reproductive biology and toxicology. Canadian Association for Laboratory Animal Science. Ottawa, ON. June 2, 1994.

85. **Foster WG**. Steroid actions on Endometriotic Tissue. Canadian Workshop on Human Reproduction and Reproductive Biology "From Bench to Bedside". Miami, FL. April 25 – 29, 1994.

86. **Foster WG**. Environmental contaminants and adverse reproductive health outcomes: Current status and future directions. University of Laval. Quebec City, QC. March 25, 1994.
87. **Foster WG**. Biomarkers in Reproduction: Effects in females. IPCS/Australian workshop on Biomarkers. Adelaide, Australia. October 11 – 15, 1993.

88. **Foster WG**, McMahon A, Rice D. The reproductive effects of chronic Lead exposure male cynomolgus monkey. Loebb Research Institute, Ottawa Civic Hospital, University of Ottawa. Ottawa, ON. January 1993.

89. **Foster WG**, McMahon A, Rice D. The reproductive effects of chronic Lead exposure male cynomolgus monkey. Canadian Wildlife Service. Hull, QC. January 1993.

90. **Foster WG**. Environmental Contaminants and Reproduction. Toxicology Research Division Seminar Series. Ottawa, ON. January 1993.

91. **Foster WG**. Session Moderator: Ovarian Function. 12th Annual Ottawa Reproductive Biology Workshop and 1993 Southern Ontario Reproductive Biology Meeting. May 25 – 26, 1993.

92. **Foster WG**. Environmental Contaminants and Reproduction. Ontario Farm Women's

Association. Guelph, ON. June 1993.

93. **Foster WG**, McMahon A, Rice D. Reproductive toxicity of chronic Lead-exposure in the female and male cynomolgus monkey. Symposium On Lead In Adults. Durham, NC. December 9 – 11, 1991.

ii) Contributions to peer reviewed presentations –

1. Haikalis ME, Wessels JM, Leyland NA, **Foster WG**. Insights into microRNA expression and use as diagnostic markers for endometriosis. 62nd Annual Meeting of the Canadian Fertility & Andrology Society. Toronto, ON. September 22 – 24, 2016.

2. Furlong H, Stämpfli MR, Gannon A, **Foster WG**. Identification of microRNAs as biomarkers for toxic insult in the ovary. 49th Annual meeting of the Society for the Study of Reproduction. San Diego, CA. July 16 – 20, 2016.

3. Haikalis ME, Wessels JM, Leyland NA, **Foster WG**. Preliminary insights on microRNA expression and use as diagnostic markers for endometriosis. 49th Annual Meeting of the Society for the Study of Reproduction. San Diego, CA. July 16 – 20, 2016.

4. Furlong H and **Foster WG**. Novel microRNAs as biomarkers for toxic insult in the ovary. Annual Meeting for the European Society of Human Reproduction and Embryology (ESHRE). Helsinki, Finland. July 3 – 6, 2016.

5. Furlong H and **Foster WG**. Environmental toxicant exposure and dysregulation of ovarian function. The Reproductive Health Summit. London, UK. April 19 – 21, 2016.

6. Guerra MT, Sanabria M, Leite GA, Borges C, Anselmo-Franchi JA, **Foster WG**, Kempinas WG. Maternal exposure to butylparaben impairs testicular structure and function on adult male Wistar rats. 48th Annual Meeting of the Society for the Study of Reproduction. San Juan, Puerto Rico. June 18 – 22, 2015.

7. Wessels JM, Somani A, **Foster WG**. Regulation of brain-derived neurotrophic factor by estrogen, GABA, and melatonin in endometrial epithelial cells. 48th Annual Meeting of the Society for the Study of Reproduction. San Juan, Puerto Rico. June 18 – 22, 2015.

8. Furlong H, Gannon AM, Stämpfli M, **Foster W**. Autophagic cascade triggered by cigarette smoke-induced activation of the AMPK pathway. 48th Annual Meeting of the Society for the Study of Reproduction. San Juan, Puerto Rico. June 18 – 22, 2015.

9. Furlong HC, Stampfli MR, Gannon AM, Foster WG, Cigarette smoke exposure triggers the

autophagic cascade via activation of the AMPK pathway. CIHR Training Program in Reproduction, Early Development and the Impact on Health (REDIH) meeting. Montreal, QC. June 2015.

10. Kosarac I, Kubwabo C, **Foster W**. Determination of urinary metabolites of organophosphate flame retardants using ultra performance liquid chromatography (UPLC) tandem mass spectrometry (MS/MS). 63rd Annual Conference of the American Society for Mass Spectrometry. St. Louis, MO. May 31 – June 4, 2015.

11. Furlong HC, Stampfli MR, Gannon AM, **Foster WG**. Cigarette smoke exposure triggers the autophagic cascade via activation of the AMPK pathway. Research Plenary, Department of Obstetrics & Gynecology, Faculty of Health Sciences, McMaster University. Hamilton, ON. May 2015.

12. Furlong HC, Stampfli MR, Gannon AM, **Foster WG**. Cigarette smoke exposure triggers the autophagic cascade via activation of the AMPK pathway. 54th Annual Meeting of the Society of Toxicology. San Diego, CA. March 22 – 26, 2015.

13. Furlong HC, Stampfli MR, Gannon AM, **Foster WG**. Cigarette smoke exposure triggers the autophagic cascade via activation of the AMPK pathway. CIHR Training Program in Reproduction, Early Development and the Impact on Health (REDIH) meeting. Ottawa, ON. December 2014.

14. Wessels JM, Leyland NA, Agarwal SK, **Foster WG**. Good proteins gone bad: brain-derived neurotrophic factor and its receptors in endometriosis implants and the role of estrogen. Oral presentation at the 70th Annual Meeting of the American Society for Reproductive Medicine. Honolulu, HI. October 18 – 22, 2014.

15. Wessels JM, Leyland NA, Agarwal SK, **Foster WG**. Estrogen regulation of brain-derived neurotrophic factor in the uterus and the link to endometriosis. 60th Annual Meeting of the Canadian Fertility & Andrology Society. Quebec City, QC. September 11 – 14, 2014.

16. Curren MS, Davis K, Liang CL, Adlard B, **Foster WG,** Donaldson SG, Kandola K, Brewster J, Potyrala M, Van Oostam J. Comparison of persistent pollutants (POPs) and metals in primiparous women from Canada and Mexico. 26th Annual Conference of the International Society for Environmental Epidemiology. Seattle, WA. August 24 – 28, 2014.

17. deCatanzaro D, Thorp J, Rajabi N, Partch T, **Foster W**. Novel male exposure and predator stress similarly disrupt blastocyst implantation while suppressing uterine closure and levels of uterine e-cadherin in mice. International Congress of Neuroendocrinology. Sydney, Australia. August 19, 2014.

18. Wessels JM, Leyland NA, Agarwal SK, **Foster WG**. (2015) The Jekyll and Hyde of estrogen induced changes in uterine brain-derived neurotrophic factor and its receptors. 47th Annual Meeting of the Society for the Study of Reproduction. Grand Rapids, MI. July 19 – 23, 2014.

19. Agarwal SK and **Foster WG**. Medical shrinkage of endometriomas with aromatase inhibition and progestin add-back. 69th Annual Meeting of the American Society of Reproductive Medicine. Boston, MA. October 12 – 17, 2013.

20. Wessels JM, Leyland NA, Agarwal SK, Murji A, **Foster WG**. Can brain-derived neurotrophic factor be a clinical marker for endometriosis? Oral presentation at the 69th Annual Meeting of the American Society of Reproductive Medicine. Boston, MA. October 12 – 17, 2013.

21. Wessels J, Leyland N, **Foster WG**. The brain-uterus connection: Uterine expression of brain-derived neurotrophic factor (BDNF) and its receptor vary over the estrous cycle. 46th Annual Meeting of the Society for the Study of Reproduction. Palais des congress de Montréal. Montreal, QC. July 22 – 26, 2013.

22. Curren MS, Liang CL, Davis K, Thuppal V, Said F, Adlard B, Donaldson S, Kandola K, Brewster J, **Foster WG**, Van Oostdam J. Examination of contaminant exposures for populations from northern and southern Canada. Environmental Health 2013, Boston, MA. March, 2013.

23. Gannon AM, Stämpfli MR, **Foster WG**. Dysregulation of mitochondrial dynamics and activation of the autophagy cascade occur in a mouse model of cigarette smoke-induced ovarian follicle loss. 68th Annual Meeting of the American Society of Reproductive Medicine. San Diego, CA. October 20 – 24, 2012.

24. Sadeu JC and **Foster WG**. Mechanism of benzo[a]pyrene-induced inhibition of follicle growth and dysfunction. 68th Annual Meeting of the American Society of Reproductive Medicine. San Diego, CA. October 20 – 24, 2012.

25. Sadeu JC and **Foster WG**. In vitro exposure to benzo[a]pyrene alters the expression of factors controlling follicle growth. 58th Annual Meeting of the Canadian Fertility & Andrology Society. Ottawa, ON. September 6 – 9, 2012.

26. Gannon AM, Stämpfli MR, **Foster WG**. Cigarette smoke exposure triggers dysregulation of mitochondrial dynamics, leading to autophagy-mediated ovarian follicle loss in a mouse model. 45th Annual Meeting of the Society for the Study of Reproduction. Pennsylvania State University. State College, PA. August 12 – 15, 2012.

27. Wessels J and **Foster WG**. Uterine expression of brain-derived neurotrophic factor (BDNF) and its receptor during the estrous cycle and menstrual cycle. 45th Annual Meeting of the

Society for the Study of Reproduction. Pennsylvania State University. State College, PA. August 12 – 15, 2012.

28. Sadeu JC, Doedée AM, Neal M, Hughes EG, **Foster WG**. Neurotrophins (BDNF and NGF) in ovarian follicular fluid of women with different infertility diagnoses. 57th Annual Meeting of the Canadian Fertility & Andrology Society. Toronto, ON. September 21 – September 24, 2011.

29. Gannon AM, Stämpfli MR, **Foster WG**. Cigarette smoke exposure triggers autophagy-mediated ovarian follicle loss in a mouse model. 44th Annual Meeting of the Society for the Study of Reproduction. Oregon Convention Center. Portland, OR. July 31 – August 4, 2011.

30. Sadeu JC, Doedée AM, **Foster WG**. Localization of ovarian neurotrophins (BDNF, NT-4/5 & NGF) and their receptors (Trk B & p^{75NTR}): Role in follicle growth. 2011 Annual Ottawa Reproductive Biology Workshop. Ottawa, ON. June 17, 2011.

31. Cameron H and **Foster WG.** Endocrine toxicants promote resistance to anoikis and invasiveness of breast cancer cells *in vitro*. 50th Anniversary Annual Meeting & ToxExpo, Society of Toxicology. Washington, DC. March 6 – 10, 2011.

32. Mulligan Tuttle AM, Stämpfli M, **Foster WG**. Cigarette smoke exposure results in significant follicle loss via an alternative ovarian cell death pathway. 50th Anniversary Annual Meeting & ToxExpo, Society of Toxicology. Washington, DC. March 6 – 10, 2011.

33. Sadeu JC and **Foster WG.** Cigarette smoke condensate inhibits follicular development, oocyte maturation and dysregulates steroids synthesis *in vitro*: Implications for human fecundity. 50th Anniversary Annual Meeting & ToxExpo, Society of Toxicology. Washington, DC. March 6 – 10, 2011.

34. Johnson NA, Meng WS, Witt-Enderby PA, **Foster WG,** Davis VL. Localized exposure to DDT congeners influence mammary gene expression. 50th Anniversary Annual Meeting & ToxExpo, Society of Toxicology. Washington, DC. March 6 – 10, 2011.

35. Sadeu JC and **Foster WG**. Cigarette smoke condensate (CSC) inhibits follicular development, oocyte nuclear maturation and disrupts progesterone synthesis *in vitro*: Implications for human fecundity. 56th Annual Meeting of the Canadian Fertility & Andrology Society. Vancouver, BC. September 29 – October 2, 2010.

36. Doedée A, Sadeu JC, **Foster WG**. The effects of bisphenol A (BPA) on the expression of neurotrophins (NTs) and neurotrophin receptors (NTRs) during *in vitro* follicle growth. 56th Annual Meeting of the Canadian Fertility & Andrology Society. Vancouver, BC. September 29 – October 2, 2010.

37. Mulligan Tuttle A and **Foster WG**. Cigarette smoke exposure results in significant follicle loss and decreased pro-survival expression in an *in vivo* mouse model. 56th Annual Meeting of the Canadian Fertility & Andrology Society. Vancouver, BC. September 29 – October 2, 2010.

38. Mallach G, Davidson A, Arbuckle T, Nethery E, Van Ryswk K, You H, Fisher M, **Foster W**, Moore E, Ripley D, Wheeler AJ. Assessing the Value of Including GPS in Personal Exposure Monitoring. ISES-ISEE 2010 Joint Conference of International Society of Exposure Science & International Society for Environmental Epidemiology. Seoul, Korea. August 28 – September 1, 2010.

39. Sadeu JC and **Foster WG.** Benzo[a]pyrene (B[a]P)-treatment at concentrations representative of human exposure attenuates ovarian follicle development and survival. 49th Annual Meeting of the Society of Toxicology. Salt Lake City, UT. March 7 – 11, 2010.

40. Lagunov A, Sadeu JC, Bruin JE, Woynillowicz AK, Anzar M, Khan MIR, Buhr M, Holloway AC, **Foster WG**. Effect of *in utero* and lactational nicotine exposure on the male reproductive tract in pubertal and adult rats. 55th Annual Meeting of the Canadian Fertility & Andrology Society. Montreal, QC. November 18 – 21, 2009.

41. Neal MS and **Foster WG**. Aryl hydrocarbon receptor (AhR) antagonists attenuates the deleterious effect of benzo[a]pyrene on isolated rat follicle growth *in vitro*. 54th Annual Meeting of the Canadian Fertility & Andrology Society. Calgary, AB. November 26 – 29, 2008.

42. Dominguez MA, Zhang B, **Foster WG**. BDNF and Trk B expression in the mouse ovary. 54th Annual Meeting of the Canadian Fertility & Andrology Society. Calgary, AB. November 26 – 29, 2008.

43. Mulligan Tuttle A, Stämpfli M, **Foster WG**. *In vivo* and *in vitro* follicle loss caused by cigarette smoke and benzo[a]pyrene exposure at physiologically relevant concentrations. 54th Annual Meeting of the Canadian Fertility & Andrology Society. Calgary, AB. November 26 – 29, 2008.

44. Boutross-Tadross O, Faghih M, Elias R, Elit L, **Foster WG**. Immunolocalization of tyrosine kinase receptor B (TrkB) expression in endometriosis associated ovarian cancer (EAOC) cells. 3rd Intercontinental Congress of Pathology. Barcelona, Spain. May 17 – 22, 2008.

45. Mulligan Tuttle A and **Foster WG**. Cigarette smoke and benzo[a]pyrene cause follicle loss *in vivo* and *in vitro* at physiologically relevant concentrations. 47Th Annual Meeting of the Society of Toxicology. Seattle, WA. March 16 – 20, 2008.

46. Monroy R, Bourgeois J, Shaw D, Morrison K, Atkinson S, Teo K, **Foster WG**. Effects of cigarette smoking in pregnancy on the placenta vasculosyncytial membrane thickness. 47th Annual Meeting of the Society of Toxicology. Seattle, WA. March 16 – 20, 2008.

47. Mulligan Tuttle A and **Foster WG.** Cigarette smoke causes follicle loss *in vivo* at physiologically relevant concentrations. 4th Annual Invitational Symposium for Research to Inform tobacco Control. A pre-conference symposium at the Society for Research on Nicotine and Tobacco Annual Conference. Portland, OR. February 2008.

48. Davis VL, Johnson NA, Jayo MJ, Hughes CL, **Foster WG**. Localized exposure to *p,p'* DDE accelerates mammary tumor development in MMTV-*neu* transgenic mice. Future Research on Endocrine Disruption: Translation of Basic and Animal Research to Understand Human Disease. Durham, NC. August 27 – 29, 2007.

49. Monroy R, Bourgeois J, Shaw D, Morrison K, Teo K, Atkinson S, **Foster WG**. Effects of maternal smoking on the placenta vasculosyncytial membrane thickness. 13th International Federation of Placenta Associations. Kingston, ON. August 17 – 22, 2007. Poster of Mention and Y. W. Loke New Investigator Award.

50. Cameron HL and **Foster WG.** The organochlorine pesticide dieldrin increases resistance to anoikis and invasiveness of breast cancer cells *in vitro*. American Association for Cancer Research Edward A. Smuckler Memorial Pathobiology of Cancer Workshop, Snowmass Village, CO. July 15 – 22, 2007.

51. Faghih M, Elias R, Boutross-Tadross O, Elit L, **Foster WG**. Immunolocalization of tyrosine kinase receptor B (TrkB) expression in endometriosis associated ovarian cancer (EAOC) cells. Society of Gynaecologists of Canada Annual Clinical Meeting. June 23, 2007.

52. Cameron HL and **Foster WG.** Developmental and lactational exposure to environmentally relevant concentrations of dieldrin in neu/ErB2 transgenic mice. American Association for Cancer Research special conference. Albuquerque, NM. May 30 – June 2, 2007.

53. Neal M, Holloway AC, **Foster WG**. Ovotoxicity of benzo[a]pyrene and the ovarian protective effects of aryl hydrocarbon receptor antagonist. 52nd Annual Meeting of the Canadian Fertility & Andrology Society Meeting. Ottawa, ON. November 15 – 18, 2006. Organon Canada Ltd. Ontario Region Student/Resident Award.

54. **Foster WG**, Monroy R, Morrison K, Teo K, Atkinson S, Kubwabo C. Serum Levels of perfluorinated compounds in human maternal and umbilical cord blood samples. International Council of Chemical Associations. Minneapolis, MN. July 26 – 27, 2006.

55. Neal MS, Petrik J, **Foster WG**, Holloway AC. *In utero* and lactational exposure to nicotine: ovarian effects. Conjoint American Society for Reproductive Medicine and Canadian Fertility & Andrology Society Meeting. Montreal, QC. October 16 – 19, 2005. Best Basic Science Paper and Alpha Award.

56. Neal MS, Zhu J, **Foster WG**. Quantification of benzo-[*a*]-pyrene (B[*a*]P) in serum and follicular fluid and its effects on follicle growth in an isolated follicle culture assay. Conjoint American Society for Reproductive Medicine and Canadian Fertility & Andrology Society Meeting. Montreal, QC. October 16 – 19, 2005.

57. Stys KA, Zhang B, Holloway AC, **Foster WG**. Dioxin-induced changes in endometrial aromatase expression and activity. Society for the Study of Reproduction. Quebec City, QC. July 24 – 27, 2005.

58. Neal MS, Lim GE, YoungLai EV, Daya S, Holloway AC, **Foster WG**. Aromatase activity in granulosa cells as a predictor of pregnancy. International Federation of Fertility Societies. Montreal, QC. May 23 – 28, 2004.

59. **Foster WG**, Agarwal SK, Holloway AC. 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) induced increased endometrial aromatase activity is mediated through altered PGE_2 metabolism. 51st Annual Meeting of the Society for Gynecologic Investigation. Huston, TX. March 24 – 27, 2004.

60. Edmunds KE, Holloway AC, Beecroft ML, Daya SH, Crankshaw DJ, **Foster WG**. The effects of dietary phytoestrogens on aromatase activity in endometrial stromal cell cultures. 51st Annual Meeting of the Society for Gynecologic Investigation. Houston TX. March 24 – 27, 2004.

61. Lim GE, Stals SI, **Foster WG**, Petrik JJ, Holloway AC. Fetal exposure to water disinfection byproducts alters postnatal growth and glucose homeostasis. 51st Annual Meeting of the Society for Gynecologic Investigation. Houston TX. March 24 – 27, 2004.

62. **Foster WG**, Wade MG, Hughes CL, YoungLai EV. Developmental exposure to a complex mixture of environmental toxicants interacts with postnatal genistein to induce changes in reproductive development of female Sprague Dawley rats. 43rd Annual Meeting of the Society of Toxicology. Baltimore, MD. March 21 – 25, 2004.

63. YoungLai E, Holloway A, Lim G, Neal M, **Foster W.** Synergistic effects of FSH & DDE on human granulosa cell aromatase. 49th Annual Meeting of the Canadian Fertility & Andrology Society. Victoria, BC. November 5 – 8, 2003.

64. YoungLai E, Kakuda N, Neal M, **Foster W**, Buhr M. Additive effects of progesterone &

DDE on calcium flux in human sperm. 49th Annual Meeting of the Canadian Fertility & Andrology Society. Victoria, BC. November 5 – 8, 2003.

65. Neal M, Holloway A, Hughes E, **Foster W**. Effect of sidestream and mainstream smoking on IVF outcomes. 49th Annual Meeting of the Canadian Fertility & Andrology Society. Victoria, BC. November 5 – 8, 2003.

66. Chomej A, Holloway A, **Foster W**. Effects of Bisphenol A on tissue remodeling enzymes in human luteinized granulosa cells *in vitro*. 49th Annual Meeting of the Canadian Fertility & Andrology Society. Victoria, BC. November 5 – 8, 2003.

67. Miller M, Holloway AC, **Foster WG**. Benzo-a-pyrene alters invasion in human breast cancer cell lines BT-474 and MDA-MB-231 through altered prostaglandin E_2 (PGE_2) metabolism. 2nd Annual AACR International Conference, Frontiers in Cancer Prevention Research. Phoenix, AZ. October 26 – 30, 2003.

68. Holloway A, Beecroft ML, Sinasac S, YoungLai E, Daya S, Edmunds K, **Foster WG**. Environmental Toxicant Induced Changes in Aromatase Activity in Estrogen Sensitive Target Tissues. Society for the Study of Reproduction, 36th Annual Meeting. Cincinnati, OH. July 19 – 22, 2003.

69. WWW.EMCOM.CA - A risk communication vehicle about endocrine disruption. Phillips KP, Aronson KJ, Brunet P, **Foster WG**, Kacew S, Leiss W, Mehta M, Poirier R, Salem T, Van Der Kraak G, Wade MG, Walker M, Wigle D, Krewski D. Society for the Study of Reproduction, 36th Annual Meeting. Cincinnati, OH. July 19 – 22, 2003.

70. Davis VL, Jayo MJ, Hardy ML, Ho A, Shaikh F, Lee H, **Foster WG**, Hughes CL. Effects of black cohosh on mammary tumor development and progression in MMTV-*neu* transgenic mice. American Association of Cancer Researchers. Washington, DC. July 11 – 14, 2003.

71. Davis VL, **Foster W**, Jayo M, Ho A, Shaikh F, Lee H, Hughes C. Influence of Locally Stored DDT Metabolites on Mammary Cancer Development induced by the *neu* proto-oncogene. American Association of Cancer Researchers. Washington, DC. July 11 – 14, 2003.

72. Hughes CL, Davis V, Shaikh F, Villegas M, Ho A, **Foster WG.** The effects of *in utero* and Lactational Exposure to Genistein or Daidzein on reproductive development in FVB/N mice and occurrence of mammary tumors in MMTV-*neu* transgenic mice. US-EPA Sponsored Endocrine Disruptors Workshop. Research Triangle Park, NC. October 29 – 31, 2002.

73. **Foster W**, Sinasac S, Holloway A. Histopathological changes in endometrial stroma and epithelium of 2,3,7,8-tettachlorodibenzo-*p*-dioxin (TCDD) treated monkeys. 48th Annual Meeting of the Canadian Fertility & Andrology Society. Charlevoix, QC. September 25 – 28,

2002.

74. **Foster WG**, Daya SH, Holloway AC. Inappropriate estrogen production induced by environmental toxicants. 48th Annual Meeting of the Canadian Fertility & Andrology Society. Charlevoix, QC. September 25 – 28, 2002.

75. Davis VL, Villegas M, Shaikh F, Ho A, **Foster WG**, Hughes CL. The Effects of *in Utero* and Lactational Exposure to the Soy Isoflavones, Genistein and Daidzein, on Reproductive Development of Male and Female Mice. Society for Gynecologic Investigation. Los Angeles, CA. March 20 – 24, 2002.

76. Helliwell J, Agarwal SK, **Foster WG**. Endometriosis and rheumatological disease: An observational study using diagnostic criteria from the American Board or Rheumatology. Society of Gynecologic Investigation. Los Angeles, CA. March 20 – 24, 2002.

77. Rodriquez S, Estrada S, **Foster W,** Agarwal SK. What motivates women to take part in clinical and basic science endometriosis research? World Congress on Endometriosis. San Diego, CA. February 24 – 27, 2002.

78. Agarwal SK, Estrada S, **Foster WG**. Pilot study evaluating the efficacy of delorelin with add-back low-dose sex steroids for the treatment of pelvic pain secondary to laparoscopically confirmed endometriosis. World Congress on Endometriosis. San Diego, CA. February 24 – 27, 2002.

79. **Foster WG**. Yang JZ. Fournier M. Immunological effects of 2,3,7,8-tetrachlorodiebnzo-p-dioxin (TCDD) in monkeys with endometriosis. 47th Annual Meeting of the Canadian Fertility & Andrology Society. Whistler, BC. October 3 – 6, 2001.

80. YoungLai EV. Wade MG. **Foster WG**, Hughes CL. Endocrine disrupting effects of a mixture of environmentally relevant pollutants in the pregnant female. 47th Annual Meeting of the Canadian Fertility & Andrology Society. Whistler, BC. October 3 – 6, 2001.

81. Wade MG, **Foster WG**, YoungLai EV, Desaulniers D, Hughes CL. Thyroid and reproductive effects of *in utero*/lactational exposure to a mixture of persistent environmental contaminants. Canadian Federation of Biological Societies, 44th Annual Meeting. Ottawa, ON, June 20 – 23, 2001.

82. Agarwal SK, Surrey MW, **Foster WG**, Harris R. The occurrence of and screening for appendiceal disease in women with endometriosis and right lower quadrant pain. 49th Annual Meeting of the Pacific Coast Reproductive Society. Rancho Mirage, CA. April 2001.

83. **Foster WG**, Chan S, Platt L, Hughes C. Detection of phytoestrogens in samples of second

trimester human amniotic fluid. 46th Annual Meeting of the Canadian Fertility & Andrology Society. St. John's, NL. September 13 – 16, 2000.

84. Liu G, Blas-Machado U, Duong Q, Lee H, **Foster WG**, Davis VL, Magoffin DA, Hughes CL. Long-term exposure to high lactose diet fails to induce ovarian dysfunction in Long Evans rats. Society for the Study of Reproduction. Madison, WI. July 15 – 18, 2000.

85. Shridhar S, Farley A, Reid RL, **Foster W**, Van Vugt DA. Tetrachlorodibenzo-p-dioxin (TCDD) on gene expression of Corticotropin Releasing Hormone, Arginine Vasopressin, and Pro-opiomelanocortin in the hypothalamus of non-human primates. 82nd Annual Meeting of the Endocrine Society. Toronto, ON. June 21 – 24, 2000.

86. YoungLai EV, **Foster WG**, Wade MG, Hughes CL. Effect of a mixture of environmental contaminants on the male rat. American Society of Andrology. Boston, MA. April 7 – 11, 2000.

87. Hughes C, **Foster W**, Platt L, Chan S, Thompson S, Hubbard S, DuBose A, Tyrey L. Midgestation intrauterine exposure of the human fetus to dietary isoflavones in North America: How does this exposure compare to animal studies in late gestation and lactation that alter developmental endpoints. Third International Symposium on The Role of Soy in Preventing and Treating Chronic Disease. Washington, DC. October 31 – November 3, 1999.

88. **Foster WG**, Chan S, Platt L, Hughes C. *In utero* exposure of the human fetus to xenobiotics endocrine disrupting chemicals: Detection of organochlorine compounds in samples of second trimester human amniotic fluid. 81st Endocrine Society Annual Meeting. San Diego, CA. June 12 – 15, 1999.

89. Douglas GR, Gingerich JD, Soper LM, McMahon A, **Foster WG**. Gene mutations in follicular granulosa cells of super-ovulated lacZ transgenic mice. Environmental Mutagen Society Meeting. Washington, DC. March 27 – April 1, 1999.

90. Desaulniers D, Leingartner K, Cooke GM, Wade M, Yang J, Yagminas A, **Foster WG**. Effects of a "Human-milk PCB-DDT-DDE" mixture on MCF-7 cell proliferation *in vitro* and on nitrosomethylurea-induced mammary tumours in the rat. Society of Toxicology of Canada. Montreal, QC. December 2 – 3, 1998.

91. Ruka MP, Gareau PJ, Janzen EG, **Foster WG**. Use of MRI T1/T2 weighted images to evaluate the lesions in surgically induced endometriosis in a mouse model. Royal College of Physicians and Surgeons Annual Meeting. Toronto, ON. September 24 – 27, 1998.

92. Desaulniers D, Leingartner K, Fintelman E, Wade M, Yagminas A, **Foster WG**. Relative sensitivity of several endocrine biomarkers following acute exposure to PCB-126 and 153 in

adult Sprague Dawley male rats. Health Conference '97. Montreal, QC. May 12 – 15, 1997.

93. Buttar HS, Moffatt JH, McMahon A, **Foster WG**. Sperm motility analysis, sperm chromatin strucutre assay and serum hormone assessment in alph-chlorohydrin treated rats. Society of Teratology. Palm Beach, FL. June 21 – 27, 1997.

94. Van Vugt DA, Kizemian A, Roy BN, **Foster W**, Lundahl S, Marcus S, Reid RL. Photodynamic endometrial ablation in non-human primates. 42nd Annual Meeting of the Canadian Fertility & Andrology Society. Lake Louise, AB. November 20 – 23, 1996.

95. Yang JZ, **Foster WG**. Repeated exposure to TCDD inhibits growth of surgically induced endometriosis in the mouse. Society for the Study of Reproduction. London ON. 1996.

96. Van Vugt DA, Kizemian A, Roy BN, **Foster W**, Lundahl S, Marcus S, Reid RL. Photodynamic endometrial ablation in rhesus monkeys: 22nd Annual Meeting of the Society of Obstetrics & Gynaecology Canada. 1996.

97. **Foster WG**, Desaulniers D, Chu I, Poon R. Reproductive effects of tris (4-chlorophenyl) methanol in the male rat. 41st Annual Meeting of the Canadian Fertility & Andrology Society. Montebello, QC. September 19 – 23, 1995.

98. Jarrell JF, Gocmen A, **Foster WG**. A long-term review of the human reproductive outcomes of inadvertent exposure to hexachlorobenzene. 41st Annual Meeting of the Canadian Fertility & Andrology Society. Montebello, QC. September 19 – 23, 1995.

99. Yang JZ, Van Vugt DA, Reid RL, **Foster WG**. Intrauterine 5-aminolevulinic acid induces selective endometrial photosensitization in the Rhesus and Cynomolgus monkeys. 41st Annual Meeting of the Canadian Fertility & Andrology Society. Montebello, QC. September 19 – 23, 1995.

100. YoungLai EV, **Foster WG**, Jarrell JF. Relationship between environmental contaminants in human reproductive fluids and results of *in vitro* fertilization. 41st Annual Meeting of the Canadian Fertility & Andrology Society. Montebello, QC. September 19 – 23, 1995.

101. Ruka MP, McCutcheon JL, Gareau PJ, **Foster WG**, Janzen EG. Surgically-induced model of endometriosis in B6C3F1 mice. Royal College of Physicians & Surgeons of Canada, 64th Annual Meeting. Montreal QC, September 13 - 17, 1995.

102. Ruka MP, McCutcheon JL, Gareau PJ, **Foster WG**, Janzen EG. Evaluation of estrogen/progesterone treatment on the functional integrity of the endometrium temporally exposed to warm ischemia prior to its auto-transplantation. Royal College of Physicians & Surgeons of Canada, 64th Annual Meeting. Montreal, QC. September 13 – 17, 1995.

103. **Foster WG**, Rice DC, McMahon A, Reed BL. Altered semen quality in Cynomolgus monkeys with occupationally relevant circulating concentrations of lead. Society for the Study of Reproduction. Davis, CA. July 9 – 12, 1995.

104. Desaulniers DM, Leingartner K, Stoddart K, **Foster WG**. Development of an *in vitro* bioassay to assess the estrogenic potential of xenobiotics using cloned and wild types MCF-7 cells. Society for the Study of Reproduction. Davis, CA. July 9 – 12, 1995.

105. Desaulniers D, Leingartner K, **Foster WG**. Comparison of proliferation indices derived from 3H-thymidine incorporation and the metabolic reduction of alamar blue (Tm) in estradiol-stimulated human breast cancer cells. 14th Annual Ottawa Reproductive Biology Workshop. Ottawa, ON. June 2 – 3, 1995.

106. **Foster WG**, Jarrell JF, YoungLai EV. Preliminary results of a survey of contaminants present in serum and follicular fluid of women participating in fertility clinics in six Canadian cities. Southern Ontario Reproductive Biology Workshop. Kingston, ON. May 5, 1995.

107. **Foster WG**, McMahon A, Reed BL, Rice DC. Flow cytometric and conventional analysis of semen from chronically lead exposed Cynomolgus monkeys (Macaca fasicularis). Southern Ontario Reproductive Biology Workshop. Kingston, ON. May 5, 1995.

108. YoungLai EV, Baillie J, Yie S-M, Hughes EG, Collins JA, **Foster WG**. Acrosin activity in frozen sperm samples does not correlate with *in vitro* fertilization of human oocytes. International Federation of Gynaecology & Obstetrics (FIGO). Montreal, QC. September 24 – 30, 1994.

109. **Foster WG**, Rice DC, McMahon A. Suppression of luteal function in the chronically lead exposed cynomolgus monkey (*Macaca fascicularis*). 40th Annual Meeting of the Canadian Fertility & Andrology Society. St. John, NB. September 7 – 10, 1994.

110. **Foster WG**, Yagminas A, McMahon A. Reverse phase HPLC demonstration of cortisol, corticosterone, progesterone, and 17 a-hydroxyprogesterone in rat serum. 40th Annual Meeting of the Canadian Fertility & Andrology Society. St. John, NB. September 7 – 10, 1994.

111. YoungLai EV, Yie S-M, Hughes EG, Collins JA, **Foster WG**. Seasonal variation in hormones in human follicular fluid and granulosa cell steroidogenesis. 40th Annual Meeting of the Canadian Fertility & Andrology Society. St. John, NB. September 7 – 10, 1994.

112. **Foster WG**, McMahon A, YoungLai EV, Hughes EG. Ovarian toxicity of hexachlorobenzene in the cynomolgus monkey (*Macaca fascicularis*). Society for the Study of Reproduction,

27th Annual Meeting. Ann Arbor, MI. August 24 – 27, 1994.

113. **Foster WG**. Alterations in ovarian function following exposure to hexachlorobenzene (HCB) in the rodent and primate model. 13th Annual Ottawa Reproductive Biology Workshop. Ottawa, ON. June 8, 1994.

114. **Foster WG**, Rice DC, McMahon A. Adverse effects of low circulating lead levels on menstrual cycle characteristics in the monkey. Southern Ontario Reproductive Biology Workshop (Tuck et al.), Hamilton, ON. May 6, 1994.

115. **Foster WG**. Steroid actions on endometriotic tissue. The Canadian Workshop on Human Reproduction and Reproductive Biology. Miami, FL. April 25 – 29, 1994.

116. Singh A, Bourque A, Lakhanpal N, McMahon A, **Foster WG**. Electron microscopy of ovary from the monkey administered hexachlorobenzne. Society of Toxicology. Dallas, TX. March 12 – 18, 1994.

117. Bourque A, Singh A, Dykeman A, McMahon A, **Foster W**. Hexachlorobenzene at low doses produces lesions in nonhuman primate ovary. 26th Annual Meeting of the USGEB/USSBE. University of Bern, Switzerland. March 17 – 18, 1994.

118. **Foster WG**, McMahon A, YoungLai EV, Hughes EG. Ovarian toxicity of hexachlorobenzene in the cynomolgus monkey (*Macaca fascicularis*). Montreal area Reproductive and Developmental Biologists First Annual Research Day. Montreal, QC. November 15, 1993.

119. **Foster WG**. Biomarkers in Reproduction - Effects in females. IPCS/Australian Workshop on Biomarkers. Adelaide, Australia. October 11 – 15, 1993.

120. Cullen C, Singh A, **Foster WG**. Electron microscopy of monkey seminal vesicle. 159th National Meeting of the American Association for the Advancement of Science. Boston, MA. February 11 – 16, 1993.

121. **Foster WG**, McMahon A. Effects of organochlorine contaminants on steroidogenesis in the rat. 12th Annual Ottawa Reproductive Biology Workshop and 1993 Southern Ontario Reproductive Biology Meeting (Tuck et al.). May 25 – 26, 1993.

122. Singh A, Dykeman A, Rice D, **Foster WG**. Electron microscopy of testis from the monkey fed Lead: A 9-year study. American Association of Veterinary Anatomists. July 15 – 18, 1993.

123. **Foster WG**, McMahon A, YoungLai EV, Hughes EG, Rice DC. Reproductive endocrine effects of chronic Lead exposure in the male cynomolgus monkey (*Macaca fascicularis*). 38th

Annual Meeting of the Canadian Fertility & Andrology Society. November 27, 1992.

124. Todoroff EC, Sevcik M, Brännstrom M, Janson PO, **Foster WG**, Villeneuve DC, Jarrell JF. The effect of photomirex on the *in vitro* perfused ovary of the rat. 38th Annual Meeting of the Canadian Fertility & Andrology Society. November 27, 1992.

125. Reid RL, Kennedy JC, Van Vugt DA, Yang JZ, Fletcher A, **Foster W**. Evidence in the human and the non human primate for aminolevulinic acid (ALA) induced selective endometrial photosensitization: A potential agent for photodynamic ablation of the endometrium. Society for Gynecologic Investigation. March 31 – April 3, 1993.

126. Singh A, **Foster WG**, Dykeman A, Villeneuve DC. Hexachlorobenzene toxicity in the rat ovary II. Ultrastructure induced by medium (10 mg/kg) dose exposure. Proceedings of the 50th Annual Meeting, Electron Microscopy Society of America. August 16 – 21, 1992.

127. Singh A, **Foster WG**, Arendz J. Chronic Lead exposure induces ultrastructural alterations in the monkey seminal vesicle. C.V.M.A. July 5 – 8, 1992.

128. Singh A, **Foster W**, McMahon A, Villeneuve DC. Lead-induced alterations in the testis of monkeys: An ultrastructural study. Annual Meeting, Society of Toxicology. February 23 – 27, 1992.

129. Singh A, **Foster WG**, McMahon A, Rice DC, Villeneuve DC. Electron microscopy of seminal vesicles from monkeys exposed to Lead: A 9-year study. 24th Annual Meeting, USGEB/USSBE. March 19 – 20, 1992.

130. Singh A, **Foster W**, Villeneuve D. Hexachlorobenzene-induced alterations in the ovary of rats. 105th Annual Meeting, American Association of Veterinary Anatomists. March 1992.

131. **Foster WG**, McMahon A, Pentick JA. Hexachlorobenzene (HCB) augments circulating progesterone concentration in the female rat. 34th Annual Meeting Canadian Federation of Biological Societies. (Abstract) p. 95. 1991.

132. McMahon A, **Foster WG**, Pentick JA, Lecavalier PR. Tissue distribution and ovarian subcellular localization of Hexachlorobenzene (HCB) in the female rat. 34th Annual Meeting Canadian Federation of Biological Societies. (Abstract) p. 95. 1991.

133. **Foster WG**, McMahon A, Villeneuve DC, Jarrell JF. Hexachlorobenzene (HCB) suppresses progesterone secretion during the luteal phase in the female cynomolgus monkey. 73rd Annual Meeting of The Endocrine Society. (Abstract) P. 118. 1991.

134. **Foster WG**, Jarrell JF, Younglai EV. Sexual maturation in the female rabbit: Effect of

Tamoxifen and Pregnant Mare Serum. 32nd Annual Meeting Canadian Federation of Biological Societies. (Abstract) p. 35. 1989.

135. **Foster WG**, Jarrell JF, Younglai EV. Light and ultrastructural characterisation of gonadotropin hormone-releasing hormone (GnRH) neurones in the rabbit. 31st Annual Meeting Canadian Federation of Biological Societies. (Abstract) p. 114. 1988.

136. **Foster WG**, Jarrell JF, Dolovich J, Younglai EV. IgE mediated hypersensitivity in response to chronic treatment with Gonadorelin-HC_l (Factrel) in a female patient. 44th Annual Meeting Society of Obstetrics & Gynecology Canada. (Abstract) p. 105. 1988.

KNOWLEDGE TRANSLATION:

i) **Consulting – Government & Industry**

2014 – present — The Endometriosis Association, USA., Scientific Advisory Panel. *The Endometriosis Association Advisors are selected for their expertise on specific scientific and/or clinical aspects of endometriosis. They advise the Association on important matters related to endometriosis, review research proposals the Association might fund, provide a sounding board for Association board and staff, and help promote the Association's mission.*

2012 – present — Development and Reproductive Toxicology (DART), International Life Science Institute (ILSI), Health and Environmental Sciences Institute (HESI), Washington DC. *Elected to serve as one of only two academic science advisors to the International Life Sciences Institute Developmental and Reproductive Toxicology Panel. As a panel member participate in setting panel objectives and budget, contribute to panel sponsored research projects, and authoring technical reports.*

2014 – 2015 — Health Canada, Public Health Agency of Canada (PHAC), Ottawa ON. *I was invited to carry out a systematic review of the literature relating exposure to environmental contaminants and development of congenital anomalies. This report was prepared under contract to PHAC and was used to guide PHAC decisions on inclusion of occupational groups and chemical exposure priorities for inclusion in a new Canadian congenital anomalies surveillance program. Results of this contract have also been accepted for publication in the peer reviewed scientific literature.*

2013 – 2014 — Exponent, Inc., Alexandria VA. *Provided expert technical advice for inclusion in government submission on the relevance of exposure to hormonally active chemicals in adverse human health outcomes.*

2011	Institute of Medicine of the National Academies, "Veterans and Agent Orange: 8th Biennial Update". *By invitation provided expert technical advice on the reproductive health implications associated with exposure to Agent Orange.*
2008	Government of Canada: Risk Management of Toxic Substances - Canadian Environment Network, Sound Management of toxic substances. *Invited to provide expert technical advice for chemical substance prioritization, methods for hazard identification, and risk assessment to protect the health of Canadians.*

ii) **Reports** –
Reviewer for Ecojustice, Toronto ON, Special review of Atrazine: Proposed Decision for Consultation, Health Canada, January 12, 2016.

iii) **Media** –

Magazine	Scientia	Searching for new biomarkers of endometriosis http://www.scientiapublications.com/professor-warren-foster-science-diffusion/ May/June 2016.
Newspaper	The Hamilton Spectator	How worried should we be about bisphenol A? April 15, 2016.
Press Release	Science Media Centre	Expert reaction to study on endocrine disrupting chemicals and onset of menopause in women. January 28, 2015.
Radio	CHML AM900 – The Scott Thompson Show	Bisphenol A (BPA) and its' association with exposure and risk for fertility problems, diabetes, and other adverse health conditions. January 16, 2015.
Press Release	Science Media Centre	Expert reaction to phthalates in pregnancy and asthma in children. September 17, 2014.
Public Lecture	McMaster University	Characterization of a novel clinical marker of endometriosis. September 4, 2013.
Press	Science Media Centre	BPA, brain and behavior – experts. May 27, 2013.

Release		
Press Release	Science Media Centre	Expert reaction to research into urinary bisphenol A and arterial stenosis. August 15, 2012.
Newsletter	Plastics Europe	Questioning the latest BPA "scare story". May 16, 2012.
Newsletter	Food Production – Rory Harrington	Study conclusion over bisphenol A breast cancer hazard challenged by scientists. May 9, 2012.
Newsletter	Hamilton Health Sciences	Enlightening girls who 'light up' about reproductive health. September 1, 2011.
Television	BBC News, Europe	EU bans bisphenol A chemical from babies' bottles. November 25, 2010.
Public Lecture	Kiwanis Club of Oakville Ontario Canada	The environment and its' implication to cancer and other diseases. November 15, 2010.
Newspaper	CBC News: The Associated Press	Higher BPA levels in urine tied to fewer sperm. October 28, 2010.
Television	CTV News: The Canadian Press	Higher BPS levels linked to lower semen quality. October 28, 2010.
Press Release	Science Media Centre	Should bisphenol A be banned? December 4, 2009.
Newspaper	The Hamilton Spectator	Researchers tap local mothers-to-be. May 21, 2009.
Newspaper	The Canadian Press	Lead levels in blood plummet after lead in gas, paints phases out. November 19, 2008.
Magazine	Today's Parent – Wendy Haaf	Scent of a baby – The flap over phthalates. December 1, 2008.
Newspaper	The Hamilton Spectator	Lead is common, but high levels pose risks. November 11, 2008.

Television	CTV News: The Canadian Press	Toxic shower curtains? Study overblown, experts say. June 12, 2008.
Public Lecture	The Probus Club of Hamilton Mountain	Environmental hazards posed by chemicals and the problems associated with trying to understand the true risk. May 1, 2008.
Radio	WAMU 88.5 – The Dianne Rehm Show, Washington DC	Bisphenol A exposure and human health concerns. April 29, 2008.
Newspaper	The Canadian Press – Anne-Marie Tobin	Sperm under siege: Though a man makes 100 million sperm every day, a variety of chemicals and workplace conditions put reproduction at risk. April 22, 2008.
Television	CTV News: The Canadian Press	Environment affects how well men's 'boys' can swim. April 21, 2008.
Magazine	Backpacker	Bottle Blues: Canada calls Nalgene plastic toxic. April 16, 2008.
Newspaper	The New York Times – Ian Austen	Canada likely to label plastic ingredient toxic. April 16, 2008.
Newspaper	The News, Healthy Living, Ancaster/ Dundas Edition	Toxic baby bottle scare or simply scary science? February 15, 2008.
Public Lecture	The Hamilton Spectator Auditorium, Science in the City	Hormone mimics and human health. May 9, 2006.

ADMINISTRATIVE RESPONSIBILITIES:

i) **Local** – McMaster University/Hamilton Health Sciences

1. Representative, Animal Research Ethics Board, Faculty of Health Sciences, 2016 – present.
2. Member, Tenure & Promotion Committee, Department of Obstetrics & Gynecology, 2008 – present.
3. Member, Animal Advisory Committee, Faculty of Health Sciences, 2007 – present.

4. Member, Graduate Curriculum Committee, Faculty of Health Sciences, 2006 – present.
5. Member, Undergraduate Hearing Committee, Faculty of Health Sciences, 2005 – present.
6. Member, Finance Management Committee, Department of Obstetrics & Gynecology, 2004 – present.
7. Member, Resident Program Evaluation Committee, Department of Obstetrics & Gynecology, 2003 – present.
8. Member, Postgraduate Education Committee, Department of Obstetrics & Gynecology, 2003 – present.
9. Member, Search Committee for Chair, Department of Obstetrics & Gynecology, 2009 – 2010.
10. Member, Reproductive Endocrinology and Infertility Fellowship Committee, Department of Obstetrics & Gynecology, 2008 – 2010.
11. Member, Undergraduate Student Appeals Committee, Faculty of Health Sciences, 2005 – 2010.
12. Co-theme Team Leader, Environment & Health, Collaborations for Health, 2005 – 2010.
13. Director, Reproductive Biology Division, Department of Obstetrics & Gynecology, 2002 – 2010.
14. Coordinator, Resident Research Program, Department of Obstetrics & Gynecology, 2002 – 2010.
15. Member, Research Advisory Committee, Hamilton Health Sciences, 2006 – 2009.
16. Member, Search Committee for Associate Dean, Graduate Studies, 2007 – 2008.
17. Medical Director, Centre for Reproductive Care, Hamilton Health Sciences, 2005 – 2008.
18. Member, BioSafety Committee, Faculty of Health Sciences, 2003 – 2005.

ii) **National -**

1. Member, Grants Review Committee, Canadian Breast Cancer Fund, 2014 – present.
2. Member, Operating Grant Review Panel, Gender & Health Peer Review Committee, Canadian Institutes of Health Research (CIHR), 2013 – present.
3. Board Member, The Anti-NMDA Receptor Encephalitis Foundation, Inc, 2012 – present.
4. Mentor, CIHR Strategic Training Program in Tobacco Research (CIHR-STPTR), University of Waterloo, 2005 – present.
5. Member of Expert Registry, US-National Toxicology Program, Center for the Evaluation of Risk to Human Reproduction, 2005 – present.
6. Member, Research Advisory Board of the Reproductive and Developmental Origins of Health, Disability and Disease, Queen's University, 2010 – 2016.
7. Member and Mentor, CIHR Training Program in Reproduction, Early Development, and the Impact on Health (REDIH), University of Ottawa, 2009 – 2016.
8. Member, Gender, Sex & Health, Committee for the Transitional Operating Grant, CIHR, 2013 – 2015.
9. Member, Clinical Investigation 'A', Grants Committee Panel, CIHR, 2013 – 2014.
10. Member, Catalyst Grant Committee, Genes and Chronic Disease, CIHR, 2013.
11. Affiliate, R. Samuel McLaughlin Centre for Population Health Risk Assessment, 2008 – 2013.

12. Panel Member, Integrating Emerging Technologies into Chemical Safety Assessment, Council of Canadian Academies, 2010 – 2012.
13. Expert Panel Member, Integrated Testing of Pesticides, Council of Canadian Academies, 2009 – 2012.
14. Mentor, Strategic Training in Research in Reproductive Health Sciences (STIRRHS), University of Montréal, 2006 – 2011.
15. Expert Panel Member, Assisted Human Reproduction Canada, The Environment and Reproductive Health: A Scientific Roundtable Steering Committee, Health Canada, 2010.
16. Organizing Committee Member, Assisted Human Reproduction Canada, The Environment and Reproductive Health: A Scientific Roundtable Steering Committee, Health Canada, 2010.
17. Member, Obstetrics & Gynecology Task Force - In vitro Fertilization, College of Physicians and Surgeons of Ontario, 2008 – 2010.
18. Member, Grants Review Committee, Canadian Breast Cancer Fund, 2007 – 2010.
19. Member, Clinical Investigation 'A', Grants Committee Panel, CIHR, 2006 – 2010.
20. Member, Scientific Program Committee, Canadian Fertility & Andrology Society, 2006 – 2010.
21. Invited Participant and Session Chair, Canadian Children's Environmental Health Research Workshop. Health Canada and CIHR sponsored, Ottawa, ON, February 9 – 11, 2009.
22. Member, Science Committee, Association of Professors of Obstetrics & Gynaecology, 2003 – 2009.
23. Member, Assisted Human Reproduction Canada/CIHR Sponsored Workshop Scientific Organizing Committee, 2008.
24. President, Canadian Fertility & Andrology Society, 2007 – 2008.
25. Chair, Scientific Program Committee, Canadian Fertility & Andrology Society, 2007 – 2008.
26. Member, Research Committee, The Society of Obstetricians and Gynaecologists of Canada (SOGC), 2003 – 2008.
27. Chair, Science Panel, EM-COM web site, www.EMCOM.ca, 2002 – April 2005.
28. Member, Board of Directors, Infertility Awareness Association of Canada, 1998 – 2004.
29. Toxic Substances Research Initiative, Healthy Environments and Consumer Safety Branch, Health Canada, 1999 – 2003.
30. Invited Participant and Session Chair, Canadian Children's Environmental Health Research Workshop. Health Canada sponsored, Ottawa, ON, March 17 – 19, 2002.
31. Chairman, Endocrine Disrupters Technical Review Committee, 1999 – 2000.
32. Vice-president, President-elect, Society of Toxicology of Canada, 1998 – 1999.
33. Co-chair, Interdepartmental Research Committee on Endocrine Disruptors, Health Canada,1998 – 1999.
34. Member, Endocrine Disruptor Committee, Organisation for Economic Co-operation and Development (OECD), Health Canada, 1997 – 1999.
35. Chairperson, Committee on Endocrine Disruptors, Health Protection Branch, Health Canada, 1994 – 1999.
36. Member, Animal Care Committee, Health Protection Branch, Health Canada, 1993 – 1998.

iii) **International -**

1. Councillor, Lake Ontario Regional Chapter, Society of Toxicology, 2016 – present.
2. Science Advisory Board Member, Endometriosis Association, 2014 – present.
3. Scientific Advisor, Development and Reproductive Toxicology (DART) Technical Committee, International Life Science Institute (ILSI), Health and Environmental Sciences Institute (HESI), 2012 – present.
4. Member, Center for Scientific Review (CSR), US Department of Health & Human Services, National Institutes of Health (NIH), 2010 – present.
5. Ad hoc Member, FIFRA Scientific Advisory Panel, US-Environmental Protection Agency (EPA), 1999 – present.
6. Past President, Lake Ontario Regional Chapter, Society of Toxicology, 2014 – 2015.
7. Member, Membership Committee, Society for the Study of Reproduction, 2012 – 2015.
8. President, Lake Ontario Regional Chapter, Society of Toxicology, 2012 – 2014.
9. Ad hoc Reviewer, National Institute of Environmental Health Sciences (NIEHS), National Toxicology Program (NTP) and the NTP Toxicology Branch, NTP Board of Scientific Counselors (BSC) meeting, December 9 – 10, 2014.
10. Invited participant, National Center for Environmental Assessment, US Environmental Protection Agency, Expert Workshop: Improving the Risk Assessment of persistent, bioaccumulative and toxic (PBT) chemicals in breast milk, October 24 – 26, 2012.
11. Expert Panel Member, site visit, Division of Epidemiology, Statistics & Prevention Research, National Institute of Child Health and Human Development (NICHD), October 22 – 24, 2012.
12. Secretary Treasurer, Reproductive Developmental Toxicology Specialty Section, Society of Toxicology, 2010 – 2012.
13. Chair, Core Committee, Society for the Study of Reproduction, 2010 – 2012.
14. Committee on Reproduction and the Environment, Society for the Study of Reproduction, 2008 – 2012.
15. Expert Panel Member, Reproductive & developmental effects of soy products and genistein, National Institutes of Health/National Institute of Environmental Health Sciences – National Toxicology Program (NIH/NIEHS-NTP), 2009 – 2010.
16. Invited Participant, Mammary Gland Evaluation and Risk Assessment Round Robin Workshop, Oakland, CA, November 16 – 17, 2009.
17. Member, NIH Study Section, Integrative and Clinical Endocrinology and Reproduction Study Section, October 5 – 6, 2009.
18. Team Member, National Institute of Child Health and Human Development (NICHD) site visit, September 24 – 26, 2008.
19. Member, WHO/IPCS Steering Group on Endocrine Disrupters, 1998 – 2002.
    - Global Inventory of Endocrine Disrupter Research.
    - International Assessment of the State of Knowledge on Endocrine Disrupters.
20. Expert Panel Member, Joint US/EU Endocrine Disruptor Research, 1999.
21. Member, Organisation for Economic Co-operation and Development (OECD), Working Group on Endocrine Disrupter Testing and Assessment, 1997 – 1999.
    - National Co-ordinator, Test Guideline Program – OECD, 1997 – 1998.

22. Member, National Sanitation Foundation - International, Health Effects Task Group, 1996 – 1998.

8.0 Statement of compensation:

I have been asked by the defendant's lawyer to provide an expert opinion on the carcinogenicity of glyphosate in rodents. I am compensated at a rate of $250 hour, $400 per hour for deposition, and a trial rate of $400 per day plus expenses. I have not testified as an expert witness in the past four years.

Hollingsworth

Ranjit S. Dhindsa
dir 202 898 5806
rdhindsa@hollingsworthllp.com

August 18, 2016

PRIVILEGED AND CONFIDENTIAL

VIA ELECTRONIC MAIL

Dr. Warren G. Foster
McMaster University
Department of Obstetrics & Gynecology
1280 Main Street West, HSC-3N52D
Hamilton, Ontario, Canada L8S 4K1

Re: Monsanto Roundup® Litigation

Dear Dr. Foster:

This letter confirms that Hollingsworth LLP ("HLLP"), on behalf of Monsanto Company ("Monsanto"), has retained you to provide expert consulting services to HLLP, for the purpose of assisting HLLP in representing Monsanto in connection with potential and/or actual litigation against Monsanto involving injuries allegedly caused by Roundup® and/or glyphosate ("the Litigation"). You acknowledge that you have received, and/or likely will receive, confidential information from HLLP and that you likely will generate work product (orally and/or in writing) to assist us in representing Monsanto in the Litigation. You agree that you will maintain all information exchanged between HLLP and you (whether orally or in writing) as strictly confidential and privileged, unless we inform you, at some time in the future, that certain information needs to be disclosed in the Litigation. You also agree to maintain the fact that you have been retained by HLLP as strictly confidential and privileged, unless we inform you, at some time in the future, that your identity as HLLP's expert has been disclosed in the Litigation. Furthermore, you agree to not do any consulting or other work for any other corporation, law firm, or person with respect to any actual or potential legal claims involving Roundup® and/or glyphosate. You will be compensated at your standard hourly rate for time spent working with HLLP on the Litigation, namely $250.00 per hour.

Hollingsworth LLP

If you agree to these terms, please sign the letter below and send it back to me. We look forward to working with you.

Sincerely,

Ranjit S. Dhindsa

SEEN AND AGREED:

Dr. Warren G. Foster

TECH TOX CONSULTING

[redacted]
Phone:
Fax: [redacted]

Invoice

Attention: Ranjit S. Dhindsa
c/o Hollingsworth LLP
1350 I Street NW
WASHINGTON DC 20005

Date: October 17, 2016

Invoice #: 201610001

Terms: Net 30 Days

HST #: [redacted]

SERVICES:

Professional services rendered in connection with
Glyphosate litigation

1	22 07 2016	Initial telephone meeting	15 min	62.50
2	26 07 2016	Telephone meeting Ron and Erica	1 hr	250.00
3	02 08 2016	Review of IARC document General + Mice pp. 1-35	1 hr 30 min	375.00
4	03 08 2016	FTF Meeting in Hamilton (Ron & Erica)	4 hrs	1,000.00
5	09 08 2016	Review of IARC document (Rat +) pp. 35-92	2 hrs	500.00
6	09 08 2016	List of participants	NC	NC
7	09 08 2016	Complaint Case No. 2:16-cv-5035	45 min	187.50
8	09 08 2016	EPA Report of the CARC	2 hrs	500.00
9	10 08 2016	Williams et al., 2000 ppp. 117-139	2 hrs	500.00
10	11 08 2016	Williams et al., 2000 ppp. 139-165	2 hrs	500.00
11	15 08 2016	Greim et al., 2015	1 hr 30 min	375.00
12	15 08 2016	EFSA correspondence chain	1 hr 45 min	437.50
13	18 08 2016	Data reports tabs 1-8	1 hr 15 min	312.50
14	18 08 2016	Chruscielska et al. 2000	30 min	125.00
15	18 08 2016	Data reports tabs 10-14	30 min	125.00
16	18 08 2016	Pub Med lierature search	1 hr	250.00
17	18 08 2016	Monsanto and EPA correspondence 1983-1984	1 hr	250.00
18	18 08 2016	Monsanto and EPA correspondence 1984-1985	1 hr	250.00

19	18 08 2016	Monsanto and EPA correspondence 1986-1991	1 hr 45 min	437.50
20	18 08 2016	Monsanto and EPA correspondence 1991-1993	1 hr	250.00
21	21 08 2016	G1993 Glyphosate: Toxicology Chapter for RED BC D183195	1 hr 30 min	375.00
22	22 08 2016	FTF Meeting in Hamilton (Ron)	4 hr 30 min	1,125.00
23	15/10/2016	EPA Evaluation of Glyphosate	4 hr 30 min	1,125.00
24	15/10/2016	IARC Guidelines for Conducting a review	2 hrs	500.00

Services Fee	$9,812.50	
13% HST	$1,275.63	
TOTAL	**$11,088.13**	**USD**

Please make cheque payable to:

Dr. Warren G. Foster

[redacted]

TECH TOX CONSULTING

[redacted]
Phone:
Fax: [redacted]

Invoice

Attention: Ranjit S. Dhindsa
c/o Hollingsworth LLP
1350 I Street NW
WASHINGTON DC 20005

Date: January 03, 2017

Invoice #: 201701001

Terms: Net 30 Days

HST #: [redacted]

SERVICES:

Professional services rendered in connection with
Glyphosate litigation

1	19-10-2016	IARC Guidelines for Conducting a review – pages 111-227	3 hr	750.00
2	21-10-2016	1981 Rat Studies - EPL Reprot and EPA Registration	2 h 30 min	625.00
3	21-10-2016	BFR Report	2 h 30 min	625.00
4	21-10-2016	Drafting of report – outline	1 hr	250.00
5	24-10-2016	Air fare for Flight to Washington, meals & mileage		1,804.47
6	24-10-2016	Travel and Meeting in Washington with legal team	14 hr	3,500.00
7	24-10-2016	Drafting of report – outline	1 hr	250.00
8	25-10-2016	Literature search – review	1 hr	250.00
9	26-10-2016	BFR Report	1 h 30 min	375.00
10	2016-10-27	Review of peer reviewed literature from search	1 h 30 min	375.00
11	27-10-2016	1990 Rat studies	3 hr	750.00
12	31-10-2016	Preparing draft report	2 hr	500.00
13	03-11-2016	Drafting of report	3 hr	750.00
14	08-11-2016	Drafting of report	2 hr	500.00
15	09-11-2016	Drafting of report	2 hr	500.00
16	10-11-2016	Writing report	4 hr	1,000.00
17	11-11-2016	Writing report	4 hr	1,000.00
18	14-11-2016	Writing report and editing	2 hr	500.00

19	15-11-2016	Report editing	1 hr	250.00
20	21-11-2016	Seralini papers	2 hr	500.00
21	23-11-2016	F2F meeting at McMaster with Ron	4 hr 30 min	1,125.00
22	08-12-2016	Review of Dr. Portier's comments	1 hr	250.00
23	15-12-2016	Revision to expert report	4 hr	1,000.00
24	20-12-2016	F2F meeting at McMaster with Ron – FL	4 hr 30 min	1,125.00

Services Fee	18,554.47
13% HST	2,412.08
TOTAL	**20,966.55 USD**

Please make cheque payable to:

Dr. Warren G. Foster

[redacted]

TECH TOX CONSULTING

Invoice

Date: June 1, 2017
Invoice #: 201706001

To: Hollingsworth LLP
c/o Ranjit S. Dhindsa
1350 I Street NW
WASHINGTON DC 20005

Terms: Net 30 Days

HST #: [redacted]

Services:

Professional services rendered in connection with Glyphosate litigation

Item	Date	Description		Line Total
1	2017-01-26	Revisions to expert report	1 hr 30 min	$ 375.00
2	2017-01-27	Review Bucher, 2000 paper	0 hr 30 min	$ 125.00
3	2017-01-27	Review of OECD TG 451/453 & cancer background	2 hr 0 min	$ 500.00
4	2017-01-29	Summarize key elements of carcinogenicity	2 hr 0 min	$ 500.00
5	2017-01-30	Revisions to animal studies	5 hr 0 min	$ 1,250.00
6	2017-01-31	Revised Exhibits	2 hr 0 min	$ 500.00
7	2017-02-01	Preparation for F2F meeting	1 hr 30 min	$ 375.00
8	2017-02-01	F2F meeting with Ron at McMaster	4 hr 30 min	$ 1,125.00
9	2017-02-06	Review Greim et al., 2015	1 hr 0 min	$ 250.00
10	2017-02-13	Editing report	2 hr 30 min	$ 625.00
11	2017-03-07	F2F meeting with Ron at McMaster	3 hr 0 min	$ 750.00
12	2017-03-21	Meeting preparation	1 hr 0 min	$ 250.00
13	2017-03-22	F2F meeting with Ron at McMaster	4 hr 30 min	$ 1,125.00
14	2017-04-06	Edit report	5 hr 30 min	$ 1,375.00
15	2017-04-11	F2F with Ron at McMaster & prep for meeting	5 hr 0 min	$ 1,250.00
16	2017-04-25	Reading Gold et al., 1987; 1989; Eustis et al., 1994, Smith et al., 2016	4 hr 0 min	$ 1,000.00

Invoice

TECH TOX CONSULTING

Date: June 1, 2017
Invoice #: 201706001

Item	Date	Description		Line Total
17	2017-04-26	Final Report of the Dec. 2016 FIFRA meeting	5 hr 45 min	$ 1,437.50
18	2017-04-27	Boorman et al., 1994	hr 30 min	$ 125.00
19	2017-04-27	Revise and send report to Ron	2 hr 0 min	$ 500.00
20	2017-05-03	F2F meeting with Ron at McMaster	3 hr min	$ 750.00
21	2017-05-04	F2F meeting with Ron at McMaster	3 hr min	$ 750.00
22	2017-05-06	Printing, review of Expert witness reports, & comments	2 hr min	$ 500.00
23	2017-05-07	Review of Dr. Neugut & Chris Portier reports	6 hr 30 min	$ 1,625.00
24	2017-05-08	Complete Portier review & teleconference with Ron and John	4 hr 30 min	$ 1,125.00
25	2017-05-14	Revisions to expert report (1:00 - 3:30)	2 hr 30 min	$ 625.00
26	2017-05-15	F2F meeting with Ron & John at McMaster (9:00 - 6:00)	9 hr 0 min	$ 2,250.00
27	2017-05-16	F2F meeting with Ron & John at McMaster (8:30 - 10:30)	2 hr 0 min	$ 500.00
28	2017-05-18	Review Dr. Jameson's expert report	2 hr 30 min	$ 625.00
29	2017-05-19	Teleconference with Ron & John	0 hr 30 min	$ 125.00
30	2017-05-21	Revisions to expert witness report with rebuttal to Portier et al.,	4 hr 30 min	$ 1,125.00
31	2017-05-24	FTF meeting with Ron at McMaster	7 hr 0 min	$ 1,750.00
32	2017-05-25	FTF meeting with Ron at McMaster	6 hr 0 min	$ 1,500.00
33	2017-05-30	Literature search, reading and drafting of ADME section	4 hr 30 min	$ 1,125.00
			Subtotal	27,812.50
			Sales Tax	$ 3,615.63
			Total	31,428.13

Make all checks payable to : **TECH TOX CONSULTING**

[redacted]

Phone: [redacted]

Spontaneous Neoplastic Lesions in the Crl:CD-1®(ICR)BR Mouse

March, 2000

Information Prepared by
Mary L. A. Giknis, Ph.D.
Charles B. Clifford, D.V.M., Ph.D.

CHARLES RIVER
LABORATORIES

INTRODUCTION

The data presented in these tables was gathered from 51 toxicology studies of at least 78 weeks duration. All studies were performed in the United States or Europe by contract laboratories or industrial toxicology facilities.

PURPOSE

The purpose of this compilation is to offer the study director, reviewing toxicologist and/or study pathologist some reported incidences of neoplasms in Crl:CD-1 (ICR)BR mice, maintained as control animals, in studies of 78-104 weeks duration. Diagnoses in this compilation are intentionally grouped in a manner to provide the user with a range of reported incidences of similar types of lesions. This compilation is not intended in any way to propose a system of standardized nomenclature nor does it separately include each and every variant of the lesion.

COMMON STUDY PARAMETERS

The 51 studies included in this publication were initiated between January 1987 and December of 1996 in seven different laboratories. All studies used male and/or female Crl:CD-1® (ICR)BR mice from three different Charles River Laboratories production sites: Raleigh, North Carolina; Kingston, New York and Portage Michigan.

The mice in these studies were from control groups of dietary or gavage studies and were approximately 4-7 weeks of age at study initiation. Some groups were untreated while others received the study vehicle, all served as control groups.

The mice included in this publication were generally singly housed in hanging wire mesh cages, fed a diet of Purina 5002 Certified Rodent Chow and had free access to water. The animal rooms were generally maintained at average temperatures of 72 +/- 5° Fahrenheit with an average relative humidity of 30-70%. A 12hr/12hr light/dark cycle was employed in all studies. Since these studies were conducted in different facilities over a period of several years, there was some variation in environmental conditions. The overall environmental conditions were not considered by those performing the studies to have had any effect on the quality or integrity of the studies. Information on the health monitoring, other than that associated with pathological examination conducted in accordance with scheduled or moribund sacrifices, was not available.

DATA SETS PRESENTED

Survival data are presented by study as the actual number surviving to terminal sacrifice and as a percent survival at terminal sacrifice, Tables1 and 2. The survival data are also presented in graphic form, Graphs 1 and 2. Survival data were not available for all studies at the time of publication. Only those studies for which data were available are represented on the graphs.

The overall incidences of all neoplastic lesions observed in any organ are reported and summarized by sex, Tables 3 and 4. These data also include neoplastic lesions from mice that died or were found moribund and killed prior to terminal sacrifice. It does not include information from mice that were killed at any interim sacrifice. Due to the apparent diversity in terminology and the variability among studies in the incidence of

Minimum and Maximum Percent Found (Minimum and Maximum % Found)

The range reported is the lowest and highest percent incidence for each lesion from the studies where the diagnosis was made. Therefore, if a study did not include a particular diagnosis, it was excluded from these calculations. The minimum and maximum percent found values should be considered in conjunction with the Number of Studies Using the Diagnosis.

The individual study percentages, Minimum % Found and Maximum % Found, were calculated by dividing the number of times each diagnosis was made by the total number of organs examined in each study and then multiplying the resultant value by 100 to express it as a percent. Values are expressed to the second decimal place.

ADDITIONAL INFORMATION

If additional information is desired regarding the conduct of these studies or the incidence of a particular neoplasm please contact Mary Giknis through Charles River Laboratories, or via e-mail at MLAGIKNIS@att.net.

SYNONYMS

Synonymous terms or diagnoses were frequently encountered in different studies and were combined under a single, often broad diagnosis, which was considered to be the primary diagnosis. Although some effort was made to use currently acceptable terms, it is beyond the scope of this publication to propose a system of preferred diagnoses. The synonyms which were included in the various diagnoses are presented in the synonym list which follows. Where possible, terminology is consistent with the classification system proposed by the Society of Toxicologic Pathologists.

Skin:
Nerve Sheath Tumor = Schwannoma

Testis:
Sertoli Cell Tumor, Benign = Sertoliform Adenoma

Uterus:
Endometrium, Adenocarcinoma = Endometrial Carcimoma
Endometrial Stromal Sarcoma = Endometrial Sarcoma

Whole Body/Multiple Organ:
Lymphoma, Malignant = Lymphosarcoma
Mast Cell Tumor = Mastocytoma

ABBREVIATIONS

NR = Not Recorded or not available at the time of publication.

Table 1: Summary of Individual Study Information and Survival/Males

Study Identification	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Study Initiation Date	1987	1988	1988	1988	1988	1988	1989	1989	1989	1990	1990	1990	1990	1991	1991	1991
Total Number on Study	53	47	50	49	50	59	50	60	50	48	50	50	69	50	59	60
Number Surviving to Termination	NR	40	NR	NR	31	NR	NR	45	NR	NR	NR	47	NR	31	38	NR
% Survival		85.11			62.00			75.00				94.0		62.0	64.4	
Study Duration in Weeks	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78

Study Identification	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Study Initiation Date	1992	1992	1992	1992	1993	1993	1993	1993	1994	1995	1989	1992	1990	1991	1991	1993
Total Number on Study	50	50	50	50	50	50	60	50	50	70	50	49	60	70	65	60
Number Surviving to Termination	35	39	NR	NR	NR	NR	NR	NR	NR	NR	30	39	11	40	NR	22
% Survival	70.0	78.0									60.0	79.6	18.3	57.1		36.7
Study Duration in Weeks	78	78	78	78	78	78	78	78	78	78	97	100	104	104	104	104

Study Identification	33	34	35	36	37	38	39	40	41	42	43	44	45	46
Study Initiation Date	1993	1993	1993	1993	1994	1994	1994	1995	1995	1995	1995	1996	1996	1996
Total Number on Study	70	50	65	50	50	65	65	60	60	60	80	50	50	116
Number Surviving to Termination	17	18	24	NR	22	43	28	26	29	NR	NR	27	22	NR
% Survival	24.3	36.0	36.9		44.0	66.2	43.1	43.3	48.3			54.0	44.0	
Study Duration in Weeks	104	104	104	104	104	104	104	104	104	104	104	104	104	104

Graph 1: Male Survival

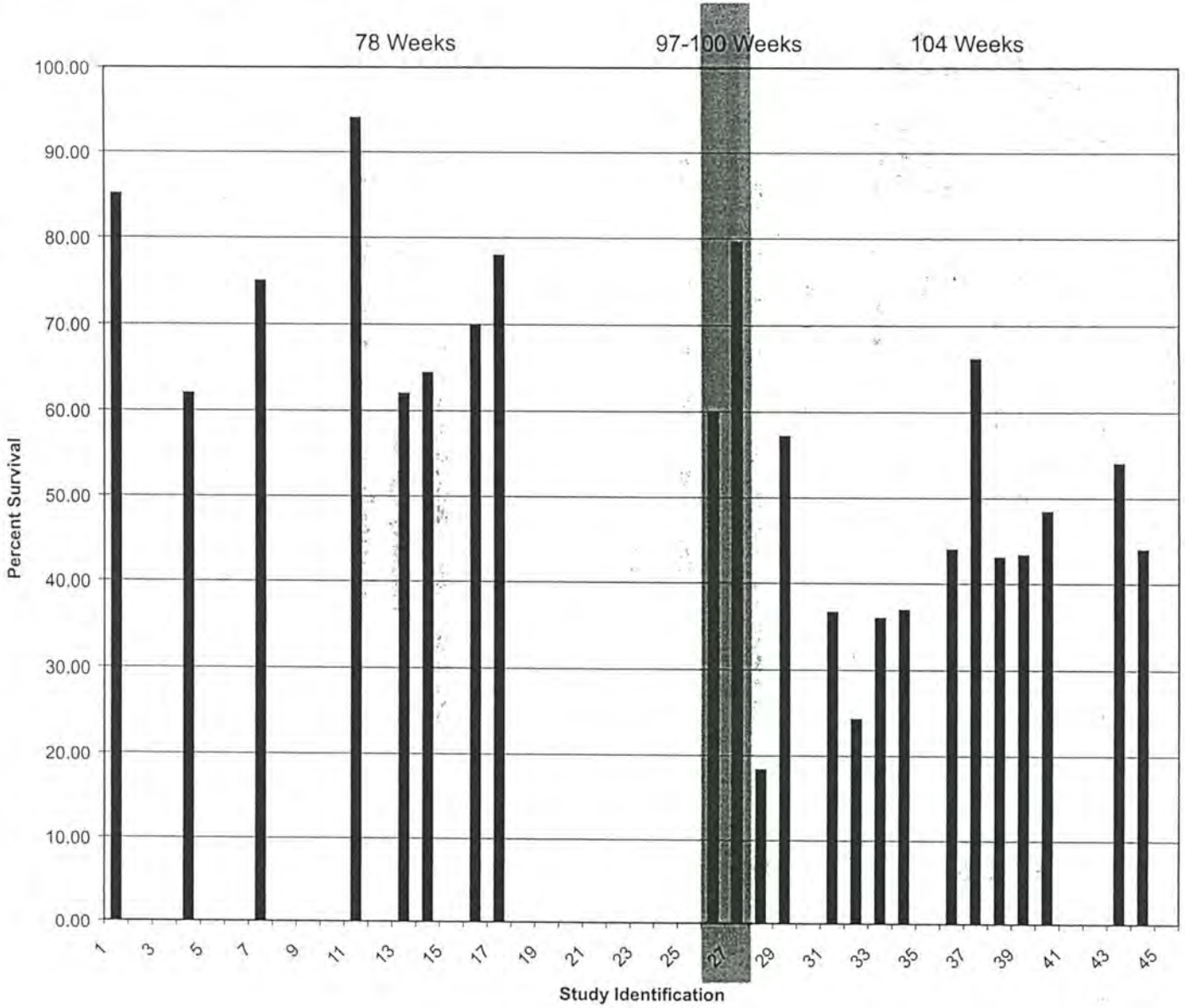


Table 3: Neoplasms/Males

		TOTAL		# STUDIES		
		# ORGANS	PERCENT	USING THIS	MINIMUM	MAXIMUM
LOCATION AND TUMOR	# STUDIES	# LESIONS	OF TOTAL	DIAGNOSIS	% FOUND	%FOUND
DIGESTIVE SYSTEM						
ORAL CAVITY	46	2577				
SALIVARY GLAND	46	2577				
STOMACH	46	2546				
Nonglandular Mucosa/Squamous Cell Papilloma		3	0.12	3	1.67	1.72
Adenocarcinoma		1	0.04	1	1.79	1.79
SMALL INTESTINE	46	2455				
Adenoma		1	0.04	1	1.72	1.72
Adenocarcinoma		5	0.20	4	1.67	2.90
LARGE INTESTINE/CECUM/ANUS	46	2482				
Adenocarcinoma		3	0.12	2	1.43	4.08
LIVER	46	2571				
Hepatocellular Adenoma		269	10.46	44	2.86	28.00
Hepatocellular Carcinoma		136	5.29	39	1.54	16.00
Hemangioma		9	0.35	7	1.54	4.00
Hemangiosarcoma		29	1.13	15	1.11	5.00
GALL BLADDER	46	2257				
Adenoma		3	0.13	3	1.69	2.00
Papilloma		6	0.27	3	2.08	5.00
PERITONEUM	46	2577				
Fibrosarcoma		1	0.04	1	1.69	1.69
Lipoma		2	0.08	2	1.43	2.00

		TOTAL		# STUDIES		
		# ORGANS	PERCENT	USING THIS	MINIMUM	MAXIMUM
LOCATION AND TUMOR	# STUDIES	# LESIONS	OF TOTAL	DIAGNOSIS	% FOUND	%FOUND
SKIN						
SKIN	46	2552				
Papilloma/Squamous Cell Papilloma		4	0.16	4	1.47	2.00
Trichoepithelioma, Benign		1	0.04	1	2.63	2.63
SKIN, cont'd						
Chondroma		1	0.04	1	1.67	1.67
Fibroma		2	0.08	2	2.00	2.08
Fibrosarcoma		2	0.08	2	1.54	2.00
Hemangioma		1	0.04	1	1.54	1.54
Hemangiosarcoma		4	0.16	4	1.43	1.67
Leiomyosarcoma		1	0.04	1	1.43	1.43
Mast Cell Tumor		1	0.04	1	1.54	1.54
Nerve Sheath Tumor, Benign		1	0.04	1	1.67	1.67
Nerve Sheath Tumor, Malignant		3	0.12	3	1.43	2.00
Sarcoma		1	0.04	1	2.00	2.00
Neurofibroma		1	0.04	1	2.00	2.00
ENDOCRINE SYSTEM						
ADRENAL	46	2526				
Cortex, Adenoma		30	1.19	17	1.56	7.14
Cortex, Carcinoma		1	0.04	1	2.00	2.00
Pheochromocytoma, Benign		11	0.44	7	1.11	5.00
Spindle Cell Tumor, Benign		6	0.24	4	1.56	4.00
PANCREAS	46	2559				
Islet Cell, Adenoma		4	0.16	3	1.54	2.00
Hemangiosarcoma		1	0.04	1	1.69	1.69
PITUITARY	46	2504				
Adenoma		6	0.24	5	1.45	3.23
Carcinoma		1	0.04	1	2.04	2.04
Pars Intermedia, Adenoma		1	0.04	1	2.00	2.00
THYROID	46	2524				
C-Cell, Adenoma		1	0.04	1	2.00	2.00
Follicular Cell, Adenoma		12	0.48	12	1.11	2.00
Follicular Cell, Carcinoma		1	0.04	1	2.00	2.00

		TOTAL		# STUDIES		
		# ORGANS	PERCENT	USING THIS	MINIMUM	MAXIMUM
LOCATION AND TUMOR	**# STUDIES**	**# LESIONS**	**OF TOTAL**	**DIAGNOSIS**	**% FOUND**	**%FOUND**
Lymphoma, Malignant		4	0.16	1	8.00	8.00
THYMUS	46	2037				
Lymphoma, Malignant		7	0.34	1	14.89	14.89
LYMPH NODES	46	2504				
Hemangioma		3	0.12	3	1.43	2.04
Hemangiosarcoma		2	0.08	2	2.00	2.00
Lymphoma, Malignant		3	0.12	1	6.00	6.00
WHOLE BODY/MULTIPLE ORGAN	46	2565				
Lymphoma, Malignant		105	4.09	33	1.45	21.67
Lymphoma, Lymphocytic		11	0.43	8	1.69	4.08
Leukemia, Granulocytic		6	0.23	6	1.43	2.04
Leukemia, Lymphocytic		3	0.12	2	2.00	3.33
Hemangiosarcoma		29	1.13	8	1.67	12.00
Histiocytic Sarcoma		35	1.36	19	1.11	8.00
Mast Cell Tumor, Malignant		4	0.16	3	1.43	2.00
SPECIAL SENSES						
EYE	46	2539				
Harderian Gland, Adenoma		120	4.73	31	1.67	14.00
Harderian Gland, Adenocarcinoma		11	0.43	7	1.43	8.33
EAR	46	2575				
Pinna, Hemangioma		1	0.04	1	1.67	1.67
Pinna, Papilloma		1	0.04	1	1.67	1.67

		TOTAL		# STUDIES		
		# ORGANS	PERCENT	USING THIS	MINIMUM	MAXIMUM
	# STUDIES	# LESIONS	OF TOTAL	DIAGNOSIS	% FOUND	%FOUND
RESPIRATORY SYSTEM						
NASAL CAVITY	48	2781				
LUNG	48	2773				
Adenoma, Alveolar/Bronchiolar		236	8.51	43	1.67	26.67
Adenocarcinoma, Alveolar/Bronchiolar		113	4.08	35	0.77	18.37
Mesothelioma, Benign		1	0.04	1	1.67	1.67
UROGENITAL SYSTEM						
KIDNEY	48	2857				
Adenoma/Tubular Adenoma		1	0.04	1	2.00	2.00
Adenocarcinoma/Tubular Adenocarcinoma		1	0.04	1	2.00	2.00
Transitional Cell Carcinoma		1	0.04	1	2.00	2.00
URINARY BLADDER	48	2718				
Transitional Cell Carcinoma		1	0.04	1	2.17	2.17
Leiomyosarcoma		4	0.15	4	1.75	2.44
Undifferentiated Sarcoma, Malignant		1	0.04	1	2.00	2.00
OVARY	48	2735				
Cystadenoma		18	0.66	12	1.54	6.00
Granulosa Cell Tumor, Benign		6	0.22	6	1.47	2.08
Tubular Adenoma		22	0.80	13	1.43	8.16
Luteal Cell Tumor, Benign		6	0.22	5	1.47	4.00
Luteal Cell Tumor, Malignant		1	0.04	1	1.11	1.11
Sertoliform Adenoma		2	0.07	2	2.00	2.04
Theca Cell Tumor, Benign		6	0.22	6	0.77	2.04
Theca Cell Tumor, Malignant		1	0.04	1	2.04	2.04
Hemangioma		8	0.29	7	1.11	2.90
Hemangiosarcoma		2	0.07	2	1.75	2.00
Leiomyoma		4	0.15	4	1.69	2.13
Oviduct, Fibroma		2	0.07	2	0.77	2.04
UTERUS	48	2812				
Endometrium, Adenoma		3	0.11	3	1.54	2.00
Endometrium, Adenocarcinoma		11	0.39	7	0.86	4.00
Endometrial Stromal Polyp		146	5.19	35	1.67	17.14

		TOTAL		# STUDIES		
		# ORGANS	PERCENT	USING THIS	MINIMUM	MAXIMUM
	# STUDIES	# LESIONS	OF TOTAL	DIAGNOSIS	% FOUND	%FOUND
SKIN						
SKIN	48	2803				
Basal Cell Tumor, Benign		1	0.04	1	1.67	1.67
Basal Cell Carcinoma		1	0.04	1	2.00	2.00
Squamous Cell Papilloma		4	0.14	4	1.43	2.00
Squamous Cell Carcinoma		8	0.29	7	1.43	3.33
Fibrosarcoma		10	0.36	8	1.54	4.29
Leiomyosarcoma		1	0.04	1	2.00	2.00
SKIN, cont'd.						
Liposarcoma		2	0.07	1	4.00	4.00
Rhabdomyosarcoma		1	0.04	1	1.54	1.54
Sarcoma		3	0.11	3	1.43	1.67
Nerve Sheath Tumor, Malignant		14	0.50	3	1.67	14.00
MAMMARY GLAND	48	2573				
Adenoma		2	0.08	2	2.04	2.63
Adenocarcinoma		42	1.63	22	0.78	8.33
Adenoacanthoma		1	0.04	1	1.79	1.79
Adenoacanthoma, Malignant		5	0.19	3	2.08	3.85
Fibrosarcoma		3	0.12	2	2.04	2.35
ENDOCRINE SYSTEM						
ADRENAL	48	2797				
Cortex, Adenoma		7	0.25	5	0.78	3.08
Cortex, Adenocarcinoma		1	0.04	1	2.00	2.00
Pheochromocytoma, Benign		8	0.29	5	0.78	5.00
Pheochromocytoma, Malignant		1	0.04	1	1.96	1.96
Spindle Cell Tumor, Benign		7	0.25	5	1.54	4.00
PANCREAS	48	2774				
Acinar Cell Adenoma		2	0.07	2	1.54	2.00
Islet Cell, Adenoma		6	0.22	6	1.54	2.08
PITUITARY	48	2697				
Adenoma		55	2.04	27	0.78	14.29
Carcinoma		1	0.04	1	1.69	1.69
Pars Intermedia, Adenoma		1	0.04	1	1.45	1.45

		TOTAL		# STUDIES		
		# ORGANS	PERCENT	USING THIS	MINIMUM	MAXIMUM
	# STUDIES	# LESIONS	OF TOTAL	DIAGNOSIS	% FOUND	%FOUND
BLOOD VESSEL	48	2533				
HEMATOPOIETIC/LYMPHOID SYSTEM						
BONE MARROW	48	2817				
Fibrosarcoma		1	0.04	1	1.54	1.54
Plasmacytoma		1	0.04	1	2.04	2.04
Hemangiosarcoma		2	0.07	2	1.67	1.69
SPLEEN	48	2772				
Hemangioma		2	0.07	2	1.69	2.00
SPLEEN, cont'd.						
Hemangiosarcoma		12	0.43	11	1.43	3.85
Leiomyosarcoma		1	0.04	1	2.00	2.00
THYMUS	48	2404				
Thymoma, Malignant		2	0.08	2	1.49	2.00
Lymphoma, Thymic		1	0.04	1	1.89	1.89
LYMPH NODES	48	2742				
Hemangioma		5	0.18	4	1.43	4.17
WHOLE BODY/MULTIPLE ORGAN	48	2822				
Lymphoma, Malignant		274	9.71	41	1.67	50.00
Lymphoma, Lymphocytic		30	1.06	4	2.00	27.45
Fibrous Histiocytoma		1	0.04	1	2.00	2.00
Histiocytic Sarcoma		111	3.93	31	1.67	18.33
Lymphoma, Histiocytic		10	0.35	4	2.08	6.38
Leukemia, Lymphocytic		6	0.21	2	1.54	8.62
Leukemia, Granulocytic		7	0.25	5	0.77	4.08
Mast Cell Tumor, Malignant		1	0.04	1	2.00	2.00
Hemangioma		4	0.14	3	1.43	2.67
Hemangiosarcoma		25	0.89	9	1.67	12.00

Table 5: Incidence of Neoplasms by Study for Selected Organs/Males

Study Identification	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
LIVER	53	47	50	49	50	59	50	60	50	47	50	50	68	50	59	60	50	50	50	49	50	50	60
Hepatocellular Adenoma	4	7	5	7	3	7	3	3	2	2	5	2	11	3	9	3	12	6	5	2	3	3	3
Hepatocellular Carcinoma	4	6	1	1	2		1	4	1	2	1	1	6	2	2	2	3	4		4	2	2	1
Hemangioma										1							1						
Hemangiosarcoma	2						2				2			1		1	2		2				
LUNG	53	47	50	49	50	58	50	60	50	48	50	50	69	50	59	60	50	50	50	49	50	50	60
Adenoma, Alveolar/Bronchiolar	6	9	9	10	1	5	6	6		10	3	8	15	6	8	3	13	7	1	2	2	2	4
Adenocarcinoma, Alveolar/Bronchiolar	1	1	3	1	4	2	5	3		3			16		1		3	5	6	3	3	3	4
Hemangiosarcoma																							
WHOLE BODY/MULTIPLE ORGAN	53	47	50	49	50	59	50	60	50	46	50	50	69	50	59	60	50	50	50	50	50	50	50
Lymphoma, Malignant	2	2	1	4	1	3	1	2	2	2		1	1	1	2		7		1		1		
Lymphoma, Lymphocytic												2	2	1	1							1	
Leukemia, Granulocytic																							1
Leukemia, Lymphocytic								2															1
Hemangiosarcoma																							
Histiocytic Sarcoma							1	1					2		1					1			
Mast Cell Tumor, Malignant	1						1																

Table 6: Incidence of Neoplasms by Study for Selected Organs/Females

Study Identification	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
LIVER	52	49	50	47	49	60	50	57	49	47	50	49	70	48	59	60	50	49	49	50	50	59	50	50
Hepatocellular Adenoma					1	1	1											1	1	1				
Hepatocellular Carcinoma			1										1			2		1	1		1			
Undifferentiated Carcinoma																								
Hemangioma																							1	
Hemangiosarcoma	2					1					1	1					1						1	
LUNG	52	49	50	48	49	60	50	57	50	48	50	49	70	49	59	60	50	50	50	50	50	59	50	50
Adenoma, Alveolar/Bronchiolar	3	6	6	5	2	2	5	6		3	5	5	11	3	6	5	8	3	2	2	2	2	6	2
Adenocarcinoma, Alveolar/Bronchiolar		3	4	1	3	2	4	2	4	5		1	7		2		3	6	1	2		5	1	1
Mesothelioma, Benign																								
WHOLE BODY/MULTIPLE ORGAN	52	49	50	48	50	60	50	58	50	47	50	49	70	49	59	60	50	50	50	50	50	59	50	50
Lymphoma, Malignant	2	2	7	6	1	5	7	10	2	5	4	2		3		6	4	3	1	3	3	9		5
Lymphoma, Lymphocytic													9										1	
Fibrous Histiocytoma											1													
Histiocytic Sarcoma		1						1					2	2	1	2	2		2	1	1	3	1	
Lymphoma, Histiocytic			3	1						3														
Leukemia, Lymphocytic								5																
Leukemia, Granulocytic														2									2	
Mast Cell Tumor, Malignant																								
Hemangioma													1											
Hemangiosarcoma													2											

Regulatory Toxicology and Pharmacology 69 (2014) 22–40

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Critical comments on the WHO-UNEP State of the Science of Endocrine Disrupting Chemicals – 2012

James C. Lamb IV [a,*], Paolo Boffetta [b], Warren G. Foster [c], Julie E. Goodman [d], Karyn L. Hentz [a], Lorenz R. Rhomberg [d], Jane Staveley [a], Gerard Swaen [e], Glen Van Der Kraak [f], Amy L. Williams [a]

[a] Exponent, 1800 Diagonal Road, Suite #500, Alexandria, VA 22314, USA
[b] The Tisch Cancer Institute and Institute for Translational Epidemiology, Icahn School of Medicine at Mount Sinai, 17 East 102 Street Floor West Tower, 5th Floor Room 5-142, New York, NY 10029, USA
[c] Department of Obstetrics & Gynecology, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada
[d] Gradient, 20 University Road, Cambridge, MA 02138, USA
[e] Exponent, The Lenz, 1st Floor, Hornbeam Park, Harrogate HG2 8RE, UK
[f] Department of Integrative Biology, University of Guelph, Guelph, ON N1G 2W1, Canada

ARTICLE INFO

Article history:
Received 4 December 2013
Available online 13 February 2014

Keywords:
Endocrine-disrupting chemicals (EDCs)
Weight of evidence
Causation
State of the science
Disease trends
Low-dose effects
Non-monotonic dose response (NMDR)

ABSTRACT

Early in 2013, the World Health Organization (WHO) released a 2012 update to the 2002 *State of the Science of Endocrine Disrupting Chemicals*. Several significant concerns have been identified that raise questions about conclusions reached in this report regarding endocrine disruption. First, the report is not a state-of-the-science review and does not follow the 2002 WHO recommended weight-of-evidence approach. Second, endocrine disruption is often presumed to occur based on exposure or a potential mechanism despite a lack of evidence to show that chemicals are causally established as endocrine disruptors. Additionally, causation is often inferred by the presentation of a series of unrelated facts, which collectively do not demonstrate causation. Third, trends in disease incidence or prevalence are discussed without regard to known causes or risk factors; endocrine disruption is implicated as the reason for such trends in the absence of evidence. Fourth, dose and potency are ignored for most chemicals discussed. Finally, controversial topics (i.e., low dose effects, non-monotonic dose response) are presented in a one-sided manner and these topics are important to understanding endocrine disruption. Overall, the 2012 report does not provide a balanced perspective, nor does it accurately reflect the state of the science on endocrine disruption.

1. Introduction

In 2002, the World Health Organization (WHO), in collaboration with the International Programme on Chemical Safety (IPCS), produced the *Global Assessment of the State-of-the-Science of Endocrine Disruptors* (WHO-IPCS, 2002). In the intervening ten years, interest in the question of endocrine disruption as a possible environmental issue has only increased and a substantial quantity of research related to endocrine disruption has been conducted. Consequently, a more current state-of-the-science review was warranted and the WHO, in collaboration with the United Nations Environment Programme (UNEP), published what is presented as an "update" to the 2002 report: *State of the Science of Endocrine Disrupting Chemicals – 2012* (WHO-UNEP, 2012a) and a companion report: *State of the Science of Endocrine Disrupting Chemicals 2012 Summary for Decision-Makers* (WHO-UNEP, 2012b).

The WHO-IPCS 2002 report was not an assessment of particular agents or risks, but set out to summarize the prevailing state of scientific knowledge – what was known, what was uncertain, and what the prospects were for resolving the uncertainties with

Abbreviations: ADDM, Autism and Developmental Disabilities Monitoring Network; ADHD, attention deficit hyperactivity disorder; AhR, aryl hydrocarbon receptor; AHS, Agricultural Health Study; ASD, autism spectrum disorder; BPA, bisphenol A; CDC, Centers for Disease Control and Prevention; DDD, dichlorodiphenyldichloroethane; DDE, dichlorodiphenyldichloroethylene; DDT, dichlorodiphenyltrichloroethane; DES, diethylstilbestrol; DNA, deoxyribonucleic acid; DNT, developmental neurotoxicity; DTU, Technology University of Denmark; EDCs, endocrine disrupting chemicals; EFSA, European Food Safety Authority; GLP, Good Laboratory Practices; GRADE, Grades of Recommendation Assessment Development and Evaluation; IARC, International Agency for Research on Cancer; IPCS, International Programme on Chemical Safety; MOA, mode of action; NMDR, non-monotonic dose response; NOAEL, no observed adverse effect level; NRC, National Research Council; OECD, Organization for Economic Cooperation and Development; PCB, polychlorinated biphenyls; POP, persistent organic pollutants; TBT, tributyl tin; TCDD, tetrachlorodibenzo-p-dioxin; TDI, tolerable daily intake; TPT, triphenyl tin; TRH, thyrotropin-releasing hormone; UNEP, United Nations Environment Programme; USEPA, United States Environmental Protection Agency; USFDA, United States Food and Drug Administration; WHO, World Health Organization.

* Corresponding author. Fax: +1 571 227 7299.
E-mail addresses: jlamb@exponent.com (J.C. Lamb IV), paolo.boffetta@gmail.com (P. Boffetta), foster@mcmaster.ca (W.G. Foster), jgoodman@gradientcorp.com (J.E. Goodman), khentz@exponent.com (K.L. Hentz), lrhomberg@gradientcorp.com (L.R. Rhomberg), jstaveley@exponent.com (J. Staveley), gswaen@exponent.com (G. Swaen), gvanderk@uoguelph.ca (G. Van Der Kraak), awilliams@exponent.com (A.L. Williams).

http://dx.doi.org/10.1016/j.yrtph.2014.02.002
0273-2300

further research and data collection regarding endocrine disruption. The report described patterns in natural human and animal populations that were considered possible manifestations of endocrine disruption and assessed the basis for evaluating whether these patterns should be regarded as real and robust, whether explanations for them other than endocrine disruption could be possible, and what might be the state of toxicological evidence for attributing them to an interference with endocrine-mediated control by environmental chemicals at prevailing environmental concentrations. Issues under debate were described forthrightly along with the nature and extent of evidence available to support the differing points of view. Importantly, the assessment was notable not only for its product, but also its process. A large and widely representative set of international experts, including those with a variety of views, articulated and employed a weight-of-evidence methodology to integrate various kinds and lines of evidence and to gauge how, and how well, the collective evidence supported conclusions, which were then extensively reviewed. The aim was to be appropriately circumspect, yet earnestly probing – that is, neither to be alarmist, focusing only on feared possibilities, nor to be complacent and dismissive of concerns that had yet to be adequately supported scientifically. The WHO-IPCS 2002 assessment largely succeeded in these aims and it won wide acceptance and respect as an objective picture of what science had to say (and the limits as to what current knowledge allowed it to say) about the possibilities, prevalence, and magnitude of impacts of environmental chemicals on natural populations through interaction with endocrine systems.

Unfortunately, the 2012 report falls well short of the standard set by the earlier 2002 assessment, both in its openness and objectiveness of process and as a substantial evaluation of current scientific knowledge and thinking on the issues. Whereas the 2002 assessment was produced by consensus among a large set of scientists spanning the range of views on the matter, the 2012 report was produced by a more limited set of authors. The 2002 report articulated and used a weight-of-evidence evaluation process and, while the 2012 report criticizes that process, it does not replace it with anything else, relying instead on an unexplained "best professional judgment." The 2002 report attempted to integrate information on exposure, toxicological testing (including dose-dependence of effects), the ability of putative disruptors to interfere with endocrine-mediated control, and patterns of appearance of possibly endocrine-related effects in populations. In contrast, the 2012 report discusses each of these elements independently and specifically declines to consider how these aspects can be brought together to assess whether there are real and current endocrine disruption problems or how well an integrated view of the scientific evidence can answer that question.

The present paper identifies several concerns regarding the WHO-UNEP 2012 report. Namely, the report fails to present an objective assessment of the current state of the science of endocrine disruption and does not, in fact, serve to update the 2002 assessment. Instead, the 2012 report seeks to replace the earlier assessment with a much less thoroughly reasoned evaluation that stresses possibilities of concern rather than an assessment of evidence about whether those possibilities result in real human health or environmental problems. An underlying concern with the report is the presentation of evidence in a manner that infers that the information demonstrates endocrine disruption without full consideration of alternative explanations for the observed effects. This is partially achieved by the imprecise use of key terms or concepts. For example, throughout the report there are sections titled "Epidemiological evidence for EDCs [endocrine disrupting chemicals] causing [insert health effect under discussion, e.g., early puberty]." This title gives the reader an impression that evidence will be presented on chemicals that cause that particular effect, when these sections should have more appropriately been characterized as a discussion of EDCs associated with these effects. Section 2.4 provides other examples and more detail on the use of inference to imply rather than show that EDCs are causally associated with certain effects. The following Table 1 provides a summary of key terms, with their definitions, as used in this paper.

The observations, comments and criticisms of the WHO-UNEP 2012 report are provided with further discussion and examples below. The aim of this critique is not to reevaluate these points, nor to conduct a comprehensive assessment of the issue of endocrine disruption or the 2012 report. Rather this critique illustrates, with specific examples, where the 2012 report has made statements that claim or imply a finding about endocrine disruption as a cause of actual effects, but have not been supported by a balanced and thorough evaluation of the pertinent evidence. New data and new understanding of the endocrine activity of chemicals have been developed since 2002 and using this information to build on the 2002 analysis is worthwhile – but, the WHO-UNEP 2012 report does not achieve this goal. This paper focuses on the limitations of the WHO-UNEP 2012 report and outlines specific concerns that include: the inconsistency between the *Summary for Decision-Makers* and the main report; the lack of a transparent and systematic framework for identifying, reviewing, and evaluating data; the failure to update the 2002 WHO-IPCS report as stated; the informal approach to assessing causation from endocrine disrupting chemicals (EDCs); the reliance on disease trends to suggest associations with EDCs; and ignoring the role of exposure, dose, and potency in endocrine disruption. Each of these key concerns is presented below in the Discussion section and includes specific examples of limitations in the WHO-UNEP 2012 report.

2. Discussion

2.1. Companion report: summary for decision-makers

In addition to the State of the Science of Endocrine Disrupting Chemicals – 2012 main report, a second publication, *State of the Science of Endocrine Disrupting Chemicals 2012 Summary for Decision-Makers* was simultaneously released (WHO-UNEP, 2012b). The relationship between the 2012 main report and the *Summary for Decision-Makers* is confusing at best. Based on the title of this document, one might presume that this document is a summary of – or at least based on the analysis of – the main report. But a closer look reveals that the Summary is actually characterized as "another product" of the process. In some cases, the Summary does present an overview of key findings from the main report, but there are many parts of the Summary which include conclusions and assertions not reflected in the main report. Indeed, some conclusions are matters not mentioned at all in the main report. It is very important to draw this distinction and make clear that the Summary for Decision-Makers is not truly representative of the content of the main report. Thus, there are even more shortcomings in the Summary for Decision-Makers than in the 2012 report itself.

The *Summary for Decision-Makers* presents a broader scope in discussions and the statements are presented as more definitive conclusions compared to those in the main report. For example, a list of diseases are presented in Figure 5 of the Summary and described as being induced by endocrine disrupting chemicals (EDCs); however, no references are provided in the Summary to support these inferences and insufficient data are presented in the main report itself to show that these diseases are in fact induced or caused by any EDC. The lack of references in the Summary or even cross-references to particular sections in the main report makes it difficult for any reader to find the basis for many of these

Table 1
Key terms and definitions.

Key terms	Definition
Endocrine disruptor	"An endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub) populations" (WHO-IPCS, 2002)
Adverse effect	"Change in morphology, physiology, growth, development, reproduction or life span of an organism, system, or (sub) population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influence." (IPCS, 2004)
Association	Statistical relationship between two or more events, characteristics, or other variables
Causation	To say that an agent causes an adverse effect means that the agent interacts with an organism to produce changes that lead to adverse effects that would not have occurred had the agent not been present
	Bradford Hill (1965) suggested several factors be considered to differentiate a causal association from a non-causal association. Based on a modification of the Bradford Hill criteria, WHO-IPCS (2002) developed causation criteria in a "structured format for assessing postulated relationships between altered health outcomes and exposure to EDCs [endocrine disrupting chemicals]."
Non-monotonic dose response	"non-[m]onotonic . . . dose response curves are often referred to as U-shaped (with maximal responses observed at low and high doses) or inverted U-shaped (with maximal responses observed at intermediate doses)." (WHO-UNEP, 2012a,b, p. 8)
	"non-[m]onotonic dose responses (NMDRs) – measured biological effects with dose response curves that contain a point of inflection where the slope of the curve changes sign at one or more points within the tested range." (USEPA, 2013)
Threshold	Dose or exposure concentration of an agent below which a stated effect is not observed or expected to occur (USEPA, 2014; IPCS, 2004)
Potency	The power of an agent to produce a desired effect (Dorlands Medical Dictionary)
	Together, affinity and efficacy determine the potency of a ligand (an agent) to activate specific hormone receptors and to elicit specific cellular responses in target tissues (Borgert et al., 2013)

general statements or to gauge their degree of scientific support. Efforts to simplify the information for decision makers or lay people have resulted in a failure to appropriately characterize or present the existing data gaps and uncertainties. Consequently, the Summary overstated the strength of its conclusions. For example, the Summary stated that EDC exposures are linked with a variety of neurobehavioral diseases or disorders, of which dyslexia is included, but in the main report, the only data provided on dyslexia relate to trends in prevalence and no information is given on an association, let alone causation, with any EDC. Additionally, the Summary approached the state of the science from the same perspective as the main report, and consequently, has many of the same limitations related to: the lack of a framework for evaluating data and conducting a weight of evidence review, the use of inference to imply causal relationships between chemicals and diseases, the reliance on disease trends to suggest associations with EDCs, and ignoring the role of dose and potency in endocrine disruption.

2.2. State of the science?

How the WHO-UNEP 2012 report is to be considered an assessment of the state of the science is unclear to the reader. The report never defines what might be meant by "state of the science" nor discusses what such an assessment should cover and characterize. A state-of-the-science review should have a defined scope with a systematic approach to the collection and review of data, and a clear methodology for the integration and assessment of these data. Several factors need to be considered in the process of integrating and interpreting data, particularly when evaluating the relevance of experimental animal studies to human health or wildlife. To that end, one of the key concerns is how dose–responses observed in experimental animal studies compare to the exposures potentially experienced by humans or wildlife. Other factors that also should be considered in the integration of data include: the quality of the available data, the consistency of the results, the presence of bias and confounding factors that may influence the findings, and the identification of data gaps. When interpreting data, the complete spectrum of findings should be considered and any controversy or debate on the issues at hand should be presented. As discussed in greater detail below, given the undefined scope and lack of a structured methodology for integrating and assessing the weight of evidence in the WHO-UNEP 2012 report, the state of the science was not evaluated in a consistent, objective manner. In some cases, the weaknesses of the data are acknowledged, such as the lack of data on EDC exposure and ovarian cancer; in other cases, important studies are ignored that counter those cited in the report.

The literature published on the potential endocrine activity of specific chemicals is extensive and beyond the scope of either the 2002 or 2012 reports. Neither report could be expected to undertake complete reviews for even a small subset of chemicals, but a systematic methodology would have ensured that a representative spectrum of the available literature was captured in the review. Publications selected for discussion in the report were not systematically assessed for their quality; nor were the results placed in a broader perspective. The only description of the literature retrieval process indicates that emphasis was placed on literature published after 2000 through March 2012. Despite the claimed focus on literature available in the last ten years or so, a substantial number of citations relied upon in the 2012 report are ones that were previously cited in the 2002 report or were published prior to the turn of the century. The conclusion that the WHO-UNEP 2012 report is not an actual update to WHO-IPCS 2002 report, but rather, a reworking of that earlier report is further supported by the 2012 report's strong reliance on older citations. This is discussed more fully in Section 2.3 of this paper.

A particular concern regarding the data collection process is that there seems to be a strong preference toward citing studies that report an association with exposure and omitting those studies that do not support such associations. For example, studies that showed some association between levels of pesticides in the environment and adverse effects on frogs are cited (e.g., Hayes et al., 2003; McDaniel et al., 2008), while studies to the contrary (e.g., Du Preez et al., 2009; Kloas et al., 2009; Murphy et al., 2006; Skelly et al., 2010; Smith et al., 2005; Spolyarich et al., 2011) are not mentioned. The same is true of the experimental studies cited for thyroid effects from polychlorinated biphenyls (PCBs). While twelve studies were cited in the discussion, half of these studies were from one research laboratory and thus do not represent independent verification of findings in separate laboratories. Other studies investigating PCB effects on the thyroid were not included in the review (e.g., Kato et al., 2004, 2010; Martin and Klaassen, 2010). Another example

of the selective citation of literature in the 2012 report is the failure to cite many of the experimental animal studies on bisphenol A (BPA) that were conducted with larger numbers of animals and dose groups under Good Laboratory Practices (GLP) (e.g., Stump et al., 2010; Tyl et al., 2002, 2008). As recommended by Conrad and Becker (2011a,b), all well conducted laboratory studies, both GLP and non-GLP, should be considered in a review in order to provide a comprehensive understanding of the mode of action (MOA), hazards, and risks of a chemical. It is not expected that all of the available data could be critically reviewed in the state-of-the-science report, but greater objectivity and balance would have helped the reader better understand the controversies and elucidate the key uncertainties that exist in the field of endocrine disruption.

A discussion of divergent results in the literature is essential to determine why different results were generated, particularly since one of the hallmarks of science is reproducibility. If only one of several laboratories is able to generate a particular outcome, it suggests that the outcome is not very robust or that there may be something unique in how the results were obtained and thus they may not be reliable. An additional reason for discussing divergent results is to make the reader aware of controversy that is fundamental to the scientific process, rather than give a false sense of agreement. Only an informed individual can make decisions about the impact of data discrepancies in the use of these data. Finally, by identifying studies with divergent findings, important gaps in understanding are characterized and potential research priorities can be better formulated. Thus, the failure to identify studies with different results is potentially misleading, contrary to best scientific practices, and does not facilitate sound policy decisions.

An additional concern regarding the specific citations referenced in the WHO-UNEP 2012 report are the discrepancies that exist between the findings reported in the publications and how these findings are described in the WHO-UNEP report. For example, it is stated in the WHO-UNEP report that "[a]rsenic exposure is strongly associated with prostate cancer" and two citations are provided (Benbrahim-Tallaa and Waalkes, 2008; Schuhmacher-Wolz et al., 2009 in WHO-UNEP, 2012a, p. 131). One of these citations, Schuhmacher-Wolz et al. (2009), merely references the other. The statement in the WHO-UNEP 2012 report does not reflect the outstanding questions about potential prostate cancer risk at low environmental exposures. The association between arsenic in drinking water and prostate cancer is primarily based on the results of ecological studies in Taiwan, which are not sufficient for assessing causation. The potential MOA is not defined and is only speculated to be endocrine-mediated. Furthermore, the International Agency for Research on Cancer (IARC, 2004), which reviewed the same epidemiology studies as Benbrahim-Tallaa and Waalkes (2008), concluded that there was sufficient evidence for arsenic in drinking water to cause skin cancer, lung cancer, and bladder cancer, but no conclusions were reached based on prostate cancer. Thus, the statement in the WHO-UNEP 2012 report that arsenic is "strongly associated with prostate cancer" is not consistent with IARC's more comprehensive review of the same literature.

Later in the discussion of mechanisms for prostate cancer, Soto et al. (1995) is cited as "good evidence that the organochlorine pesticides [are] shown to be associated with increased prostate cancer risks." In fact, the publication of Soto et al. (1995) relates to an *in vitro* assay and does not deal with the risk of prostate cancer from pesticides. Similarly, Yolton et al. (2011) is cited as showing that concentrations of BPA and phthalates in maternal urine during early pregnancy were associated with higher hyperactivity and aggression in 2-year old girls, but not in boys (WHO-UNEP, 2012a, pp. 114–115). This appears to be in error since the reference cited did not investigate hyperactivity or aggression. However, an earlier publication from the same research group, Braun et al. (2009), reported an association between mean urinary BPA concentrations and behavior scores in female children – but no assessment of phthalates was conducted in this earlier study. Finally, Figure 2.31 (WHO-UNEP, 2012a, p. 182) presents data on concentrations of DDE (dichlorodiphenyldichloroethylene, a metabolite of dichlorodiphenyltrichloroethane [DDT]) in osprey eggs, with a straight line drawn through the points to depict a clear decline over time. This figure is reported to be "based on data from Henny et al. (2010)." The original paper, however, does not draw a best-fit line through these values, but rather, the data are presented in a bar graph with no statistically significant differences for the time periods: 1981–82, 1993, 1998, and 2000–01. In other words, there is no real trend for these four periods; only the values in 2006 and 2008 show a statistically significant decrease in DDE concentrations in osprey eggs. These are just a few examples of the discrepancies found in the WHO-UNEP 2012 report; they do not reflect all of the citations that have been mischaracterized or mistakenly referenced. In some cases, the discrepancies may only be mistakes in referencing the correct citation, but in others, there appears to be a tendency to exaggerate the findings or conclusions of the original authors.

Finally, a process for assessing the quality or reliability of the studies considered for review in the WHO-UNEP 2012 report is not described. In fact, the quality of the underlying studies does not appear to have been evaluated at all. This is of particular concern because not all studies should be given the same weight. For example, epidemiology studies that employ weak research designs (e.g., ecological and cross-sectional study designs) or include small sample sizes should not be given as much weight as studies with stronger designs (e.g., case-control and cohort studies) and larger sample sizes. Also, comparing *in vitro* exposures to relevant *in vivo* exposures is fraught with difficulties. *In vitro* studies can be relevant for investigating MOA and potential for endocrine activity, but cannot provide useful information on dose–response, do not take into account the disposition of a chemical in the body (its absorption, distribution, metabolism and elimination), and fail to account for homeostasis or other pathways and processes that respond to certain MOAs. Various methods exist to evaluate and weigh the quality and reliability of studies included in a review, such as the systematic approach for evaluating toxicology and ecotoxicology data described by Klimisch et al. (1997), which is also employed by the Organization for Economic Cooperation and Development (OECD) and relied upon by the OECD's 34 member countries in the investigation of high production volume chemicals (OECD, 2005). Yet inexplicably, neither this approach nor an alternative approach for evaluating the quality of the data reviewed was applied in the WHO-UNEP 2012 review.

Most importantly, the WHO-UNEP 2012 report did not adopt a weight-of-evidence approach for the evaluation of data on endocrine disruption. In the 2002 review, an objective and transparent framework was developed for assessing the relationship between potential EDCs and health outcomes, which the 2012 report does not apply in its evaluation and interpretation of data. Further, no alternative approach was proposed or applied to evaluate the weight of evidence. However, an overall evaluation of the data must have been conducted because conclusions at the end of each section discussing health effects are titled: *Strength of Evidence*. No description is provided regarding how this "strength of evidence" was determined, although "best professional judgment" (WHO-UNEP, 2012a, p. 19) was said to be applied. Specific concerns about the lack of a structured approach for assessing causation in the 2012 report are further discussed in Section 2.4 of this paper.

In the WHO-UNEP 2012 report, evidence for endocrine disruption in humans and wildlife is presented as narrative reviews of the data. These assessments were described as being founded on an aggregation of the information related to biological plausibility, relevant exposures, consistency of the data across species, and dose–response and temporality (WHO-UNEP, 2012a, p. 19).

Despite this description, there appears to be little, if any, integration of the data. For example, biological plausibility is often cited in the 2012 report as the basis of concern for causality, but the evidence presented generally is limited to data on the role of endogenous hormones and not based on mechanistic data, potency or actual exposures for any of the chemicals of potential concern. Exposure data are clearly not integrated with the rest of the data, as exposure information is provided separately in Chapter 3 of the WHO-UNEP 2012 report. Species concordance is frequently used to try and bridge experimental animal or wildlife data with human data on particular observed effects, but there is a lack of integration of the data with exposure information or mechanisms. Potential species differences do not seem to have been considered, and in some cases, these differences may be critical to the interpretation of the data (an example is provided regarding thyroid effects in Section 2.6.2). Although the report purports to incorporate dose–response in its evaluation of the evidence, it tends to be ignored when specific chemicals or adverse outcomes[1] are discussed; this issue is discussed in greater detail in Section 2.6 below.

Other factors not specifically mentioned in the WHO-UNEP 2012 report that should have been considered in the integration and interpretation of data include the reproducibility of the data and consistency of data across different lines of evidence (epidemiology, *in vivo* and *in vitro* data), data gaps, and the existence of controversy or differences in interpretation of study findings. As mentioned previously, the reproducibility of a finding in different studies, different research labs, or in different study populations was not addressed in the report. Although it is noted in a number of places where significant gaps still exist in the data, these data gaps and their implications are not always carried forward into the report conclusions. Further, no specific research recommendations are provided based on such data gaps in the 2012 report, which is an important and useful element for developing data in the future.

Other recently published reports have done a better job of providing state-of-the-science reviews on controversial issues in science. For example, the U.S. Environmental Protection Agency (USEPA) recently released a draft review on the scientific issues surrounding the phenomenon of a non-monotonic dose–response (NMDR) (USEPA, 2013). USEPA took a methodical and even-handed approach to its review. The scope of the report was provided, clearly stating the scientific questions that were to be addressed. USEPA also clearly described and respected the limits of the report; for example, they acknowledged that it was not a comprehensive treatise. In a succinct manner, USEPA described both the supporting and conflicting evidence for NMDRs. As part of this assessment, the uncertainties associated with the interpretation of the data were considered. Overall, the USEPA draft report on NMDRs is a good example of a well-conducted, state-of-the-science review and illustrates how controversial issues, such as NMDRs, can be approached in an objective manner. Similarly, the European Food Safety Authority (EFSA) reviewed a developmental neurotoxicity (DNT) study in rats and other recent scientific literature in context of the risk assessment for BPA (EFSA, 2010), a compound mentioned in the WHO-UNEP 2012 report as being an EDC. Although EFSA's review had a narrow scope, specific parameters were provided for the inclusion of studies in the assessment. EFSA also described its reliance on clear "quality criteria" to assess the strengths and weaknesses of the studies it reviewed. Those criteria addressed issues of study design, conduct, recordkeeping, and interpretation. The EFSA review provides an example in which a comprehensive evaluation of the literature considers explanations for toxicity beyond those related to endocrine disruption; a feature which is not generally considered in the WHO-UNEP 2012 report. A final example is the independent review by the National Research Council (NRC) on the adverse health effects associated with perchlorate ingestion (NRC, 2005) and USEPA's draft risk assessment for perchlorate. The scope of the panel's review was outlined with specific issues to be addressed including whether or not USEPA "considered all relevant literature (both supporting and non-supporting), consistently critiqued that literature, and then used appropriate scientific studies to develop its health risk assessment" (p. 30). Thus, the full spectrum of evidence was considered in the NRC's review and a systematic approach was used to evaluate the data. The NRC report is a comprehensive, systematic review utilizing a weight-of-evidence approach to investigate a potential causal relationship and considered dose–response in evaluating potential human health effects.

Overall, the WHO-UNEP 2012 report does not meet the expectations of a state-of-the-science review. In contrast to the above examples of state-of-the-science assessments conducted and based on objective reviews of the literature, using sound methodology and defined goals, the WHO-UNEP 2012 report did not provide a summary of the state of the science for endocrine disruption. A number of limitations regarding the lack of a defined scope for the review, the absence of a process for identification, integration, and interpretation of data, and the lack of a structure for evaluating the weight of the evidence calls into question the conclusions reached in the report. The WHO-UNEP 2012 report can, at most, be characterized as a discussion of some of the scientific issues that should be considered when discussing endocrine disruption, but it is not, by any means, a true summary of the state of the science.

2.3. Not an update of 2002 state-of-the-science report

Although the WHO-UNEP 2012 report is described as an update of the WHO-IPCS 2002 *Global Assessment of the State-of-the-Science of Endocrine Disruptors* (WHO-IPCS, 2002), it does not achieve this goal. An update relies and builds upon the information reviewed and assessments of data from an earlier document. The WHO-UNEP 2012 report reviews much of the same information cited in the 2002 report by frequently citing literature from the year 2001 or earlier. Although there is a section in each subchapter on specific health effects that is titled *Scientific progress since 2002*, these are only brief sketches that are not sufficiently scientifically robust to achieve the objective of fully revisiting the state of scientific understanding. Often, the information contained therein is limited to a series of bullets that lack citations in support of the statements made. In some cases, the support for these bullets in the text of the chapter is scant, and it generally does not reflect the scientific consensus or weight of evidence available on the topic at hand. Importantly, there is no consideration of whether new information changed the state of understanding since the earlier 2002 report. The WHO-UNEP 2012 report does not provide any follow-up on the research recommendations presented in the WHO-IPCS 2002 report, nor are any new research recommendations presented. Notably, there are no references to the conclusions reached in the WHO-IPCS 2002 report; nor does the WHO-UNEP 2012 report provide any support or explanation when its conclusions differ from the 2002 report. Thus, it is unclear in what way the WHO-UNEP 2012 report can be considered an "update."

Perhaps the most significant factor that belies the characterization of the WHO-UNEP 2012 report as an update is the failure to use the framework and causal criteria for assessing endocrine disruptors originally proposed in the WHO-IPCS 2002 report. The framework put forward in the WHO-IPCS 2002 report was based on established and broadly accepted scientific methods for the evaluation of data to determine whether associations could be considered causally related – not just chance findings or the result of bias. Not only does the WHO-UNEP 2012 report not rely on the

[1] When referring to adverse outcomes in this paper, this includes adverse health effects in humans, experimental animals, and wildlife.

weight-of-evidence framework developed in the 2002 report, it also does not provide an appropriate alternative framework by which the data could be assessed to determine causal relationships. Instead of using clear and objective principles as the foundation for the report, "best professional judgment" was used (WHO-UNEP, 2012a, p. 19) without fully describing any criteria by which such judgment was to be applied. Consequently, the WHO-UNEP 2012 report is not an update but, rather, a selective re-evaluation of information largely included in the 2002 report.

In order to better understand why and how the WHO-UNEP 2012 report is not an update of the WHO-IPCS 2002 report, selected examples are provided below of adverse outcomes reviewed in both documents. These examples compare the data relied upon, how the available evidence was characterized, the state of the science, and conclusions reached about the weight of evidence in the respective reports.

2.3.1. Sperm/semen quality

Sperm or semen quality was evaluated in both state-of-the-science reviews. The WHO-IPCS 2002 report concluded that a global trend for declining semen quality was not supported by the existing data. This conclusion was based on a broad review of studies investigating sperm counts, from the first study suggesting a decline (Nelson and Bunge, 1974), and included the first meta-analysis (Carlsen et al., 1992), longitudinal retrospective studies in single centers (e.g., Auger et al., 1995), and broader investigations around the world (e.g., Auger and Jouannet, 1997; Jørgensen et al., 2001; Swan et al., 1997; Younglai et al., 1998). The review in the 2002 report described the limitations and biases of the various studies. Concerns included the use of retrospective study designs, evaluation of semen or sperm samples from men that may not have been representative of the general population (e.g., patients at infertility clinics), differences in methods for recruiting study subjects, variability in analytical methods, and lack of control or consideration of the other factors that are known to impact sperm quality (e.g., age, sexual abstinence). The lack of data on specific chemical exposures raised questions about assessing the strength of the association. The WHO-IPCS 2002 report acknowledged that, while it was biologically plausible and some experimental evidence was available to support EDCs affecting sperm quality, the "lack of any demonstration to date of an endocrine-disrupting mechanism for other chemical exposures indicates the need for more studies before firm conclusions can be drawn" (WHO-IPCS, 2002, p. 56). In its evaluation of the strength of evidence for: *Semen Quality and Testis Function in Humans* (WHO-IPCS, 2002, p. 124), the 2002 report concluded that the overall strength of evidence was weak based on an assessment of temporality, strength of association, consistency, and biological plausibility.

The WHO-UNEP 2012 report described various studies of declining semen quality, also starting with the first meta-analysis by Carlsen et al. (1992) and cites many of the same references relied on in the WHO-IPCS 2002 report. The discussion is not limited to or focused on the newest studies on sperm or semen quality, nor is it a comprehensive review of all of the data. Although several prospective studies of the general population were mentioned as being conducted by Nordic, Baltic, German, Spanish and Japanese researchers, the only citations provided are two studies of Finnish and Danish men (Jørgensen et al., 2011, 2012). A retrospective analysis of French men with total infertile partners that were subjects in assisted reproductive technology was also cited (Rolland et al., 2013). However, several other studies published since the WHO-IPCS 2002 report that do not show a decrease or an increase in sperm counts are not cited (e.g., Axelsson et al., 2011; Costello et al., 2002; Elia et al., 2012; Marimuthu et al., 2003; Pal et al., 2006). In a recent commentary from Bonde et al. (2011), it is shown that sperm cell counts in Danish military draftees have remained stable and, in fact, suggest higher counts in the last four years of the study between 2007 and 2010. This contribution to the extensive and ongoing discussion about sperm cell count trends is not cited at all. The selective citation of literature and the failure to include many studies that do not support a decline in sperm counts indicates that the 2012 report provided an unbalanced review of the literature.

Unlike the WHO-IPCS, 2002 report, there is no mention of the limitations and potential biases of any of the studies reviewed, despite the fact that all of the limitations noted in the 2002 report continue to apply. The new, prospective studies highlighted in the 2012 report also have low study participation rates: 13.4% in the Finnish study (Jørgensen et al., 2011) and 24% in the Danish study (Jørgensen et al., 2012). Such low participation rates raise concerns about whether the men in these studies were truly representative of the general population. Furthermore, there is only a brief mention that the issue of declining sperm counts remains controversial. Questions about the implications of conflicting findings related to sperm quality are not carried over into the conclusions on male reproductive health. Recent reviews of this issue continue to characterize the reports of declining sperm counts as controversial because of the differences seen geographically and temporally (Fisch and Braun, 2013; Sharpe, 2010). These reviews identify a number of factors that may account for these differences that are not related to endocrine disruption, such as lab techniques, sexual behavior resulting in differences in abstinence, lifestyle factors (e.g., obesity, drug use), and genetic variations.

The WHO-UNEP 2012 report presents limited data regarding specific EDCs and the potential for affecting sperm quality. Epidemiology studies are characterized as showing weak associations with EDCs, which is based on single citations for most of the chemicals of possible concern identified in the report. Single publications are not sufficient to assess the weight of evidence (or to provide one's "professional judgment") of a potential association between an exposure and an adverse effect. Furthermore, the 2012 report does not evaluate the limitations of these studies; nor does it consider experimental animal studies in conjunction with epidemiology to assess the hypothesis that EDCs could cause male reproductive disorders, including effects on sperm or semen quality. The 2012 report relies only on "suboptimal or poor semen quality in large proportions (20–40%) of men in countries in which this has been studied" and that there is "some evidence for a declining semen quality," thereby suggesting that these perceived trends are the consequence of exposure to endocrine disruptors with no strong evidence to support this claim.

Despite acknowledging that the epidemiology data only show weak associations for a decline in sperm quality related to specific EDCs, the conclusions in the 2012 report for male reproductive health focus on the observation of decreased sperm counts, ignoring the variability in sperm quality reported around the world and the well founded scientific questions that have been raised about this issue. Based on the evidence presented in the WHO-UNEP 2012 report, it does not appear that the evidence for changes in sperm quality differ from that reported in the WHO-IPCS 2002 report. Therefore, the reason for the discrepancy between the conclusions of the two reports is difficult to explain. When an objective, structured, and transparent weight-of-the-evidence analysis reaches one conclusion and a subjective analysis concludes the opposite, logic dictates the driving force for such a difference stems from the methodology and bias inherent in a subjective analysis.

2.3.2. Adrenal disorders

As noted in the WHO-UNEP 2012 report, adrenal dysfunction was not discussed in detail in the WHO-IPCS 2002 report because the "available research to date [was] very limited" (WHO-IPCS, 2002, p. 86). The observation of severe adrenocortical hyperplasia in the Baltic ringed and gray seal were mentioned and it was stated in the 2002 report that although adrenal effects in wildlife were

associated with dichlorodiphenyldichloroethane (DDD), DDT, and PCBs, the involvement of these compounds in the cause of these disorders was uncertain.

The WHO-UNEP 2012 report contains a specific section discussing adrenal disorders, which seems appropriate given the lack of systematic review in 2002. The information presented in the WHO-UNEP 2012 report focuses heavily on data and literature from 2002 or earlier, particularly regarding effects in wildlife. This suggests that the state of knowledge regarding endocrine disruption related to perturbations of the adrenal gland has not changed significantly since 2002. No new data are provided on the observation of adrenocortical hyperplasia in Baltic seals – the one new article, Lind et al. (2003), addresses bone mineral density in Baltic grey seals, not adrenocortical hyperplasia. Significantly, there is a change in the conclusions about the weight of evidence regarding adrenocortical hyperplasia in wildlife and various compounds, despite the lack of new information. In 2002, this was considered an uncertain association, but in 2012, it was concluded that this was a causal relationship. This change in the interpretation of the data should have been highlighted and the evidence to support this change should have been stated explicitly. Additional concerns about the weight of evidence regarding the causal relationship between exposure to EDCs and the observation of adrenocortical hyperplasia in Baltic seals are discussed later in this paper (see Section 2.4.1). It is unclear how a more definitive conclusion could be reached regarding adrenocortical hyperplasia based on essentially the same data reviewed in 2002, which suggests that the WHO-UNEP 2012 report is not an update, but a re-interpretation, of the same data. Alternatively, the authors of the WHO-UNEP 2012 report may view the 2002 report as flawed. If this is the case, then the reasons behind the divergent opinions should have been explicitly stated.

2.3.3. Endometriosis

In the WHO-IPCS 2002 report, endometriosis was reviewed as a reproductive outcome possibly associated with EDCs. Some epidemiology studies of PCBs and dioxins indicating associations with endometriosis were contrasted with other studies that failed to observe an association. Experimental animal evidence was presented for 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and the data were judged to be conflicting – some studies indicated that TCDD may have a role in the development of endometriosis, but other studies did not support such an association. The 2002 report mentioned the criticisms of the studies of TCDD in rhesus monkeys, including factors potentially confounding the results. The question of relevance of the rodent studies for humans was also raised given the high doses required to induce endometriosis in rodents. Regarding possible MOAs, the roles of estrogen and progesterone were discussed in the context of known disease development and the aryl hydrocarbon receptor (AhR) was reviewed in relation to TCDD exposure and endometriosis. The 2002 report concluded:

> "Relative to the hypothesis of an association between a stressor and an outcome, evidence is judged to be weak because of conflicting data from humans and animals, lack of association in women exposed to high amounts of TCDD, and antiestrogenic effects of TCDD. In humans, occurrence of endometriosis shows dependency on estrogen–progesterone balance, suggesting that an EDC-related mechanism may be possible." (WHO-IPCS, 2002, p. 125)

The WHO-UNEP 2012 discussion on endometriosis is not an update of the WHO-IPCS 2002 report as it mainly consists of a re-review of the information evaluated in 2002. A majority of the literature cited in the WHO-UNEP 2012 report linking TCDD to endometriosis was cited in the WHO-IPCS 2002 report, but the 2012 report failed to note that many of these studies (e.g., Bruner-Tran et al., 1999; Rier et al., 2001) employed high doses of TCDD, which were considered of questionable relevance to endometriosis in humans in the 2002 report. Much of the PCB literature was also reviewed and cited in the earlier WHO-IPCS 2002 report. Although the WHO-UNEP 2012 report cites nine publications as finding a "relationship between circulating phthalate (and phthalate esters) and endometriosis" (p. 44), none of these publications mention phthalates. Four other publications are mentioned regarding studies on phthalate ester metabolites, but the findings are inconsistent for various phthalate esters, which is acknowledged in the WHO-UNEP 2012 report. It is interesting to note that, in the discussion of phthalate esters and endometriosis, there is no mention of a study by Itoh et al. (2009) that did not observe an association between measures of phthalate metabolites in urine and endometriosis. Regarding a potential MOA for endometriosis, general hormonal influences on the disease are discussed, but little information is reviewed on specific EDCs. The WHO-UNEP 2012 report stated that epigenetic changes have been reported to be involved in endometriosis – particularly those induced by *in utero* exposure. However, only two review articles are cited (Cakmak and Taylor, 2010; Guo, 2009), neither of which address causal relationships for specific chemicals. It is not clear what, if anything, has changed regarding the state of the science in terms of cause and effect for endocrine disruption and endometriosis since the WHO-IPCS 2002 report.

The examples above illustrate how the WHO-UNEP 2012 report is not an update of the WHO-IPCS 2002 report. In some cases, the 2012 report reaches conclusions that conflict with those reached in the earlier report, despite the lack of new information to support a change in the weight of evidence. In other cases, the 2012 report simply presents a re-evaluation of the studies reviewed in 2002. The fact that the 2012 report often reaches more definitive conclusions based on the same data emphasizes the use of non-transparent, subjective decision making for evaluating potential causal relationships compared to the earlier 2002 report. In light of this, although the 2012 report is stated to be an update, in actuality, the 2012 report is a revised review of the state of the science presented in 2002 and does not build upon what was previously done.

2.4. Inference, not causation

Causation is a critical element in the definition of an endocrine disruptor and consideration of causal relationships should have been given more attention in the WHO-UNEP 2012 report. Instead, the 2012 report tended to focus on only part of the definition of an endocrine disruptor – the potential to alter some aspect of endocrine function. Consequently, information is presented or discussed on possible or potential endocrine disruption without considering whether alterations to the endocrine system caused any adverse effects. The failure to differentiate between the *potential* for endocrine system interaction of some sort and actual disruption of control of physiology or development as a result of such interactions misleads the reader about the weight of evidence for particular disruptive effects. This could result in inappropriate regulatory actions and research priorities when additional research still might be needed to establish these causal relationships. Any state-of-the science review should be a balanced and objective review of all of the available literature with identification of data gaps along with clear statements of the conclusions supported by the data and the inferences that cannot be supported. Only through such an approach can all parties, in particular, stakeholders without a strong scientific background, be confident in the evidence-based decisions that arise from such a report.

There have been many definitions for an endocrine disruptor over the years and the WHO-UNEP 2012 report stated it relies on the definition in general use today; this same definition was used in the WHO-IPCS 2002 report. Integral to this definition are three important components:

- The substance must act through an endocrine MOA that alters function of the endocrine system;
- The substance must cause an adverse health effect; and

- That adverse effect must be causally related to and occur as a consequence of the altered endocrine function.

All three of these components are necessary to demonstrate that a chemical is an endocrine disruptor. This requires the differentiation between endocrine-mediated effects from other known MOAs and the linking of an observation of an adverse effect to that endocrine MOA. Various chemicals or classes of chemicals were presented in the WHO-UNEP 2012 report as though they were EDCs, but the report does not provide any information to show that the effects of these chemicals are, in fact, the result of alterations in endocrine function. In the 2012 report, the interference with endocrine function by a chemical is often considered sufficient evidence for endocrine disruption, when in fact this only demonstrates that a chemical has the ability to interact with the endocrine system (Tinwell et al., 2013) and the *potential* for endocrine disruption.

The WHO-UNEP 2012 report frequently relies solely on the observation of adverse effects in endocrine organs or the existence of possible endocrine MOAs. The report often reaches conclusions based on only one or two elements of the definition for an endocrine disruptor. It is the third and essential component of the definition that leads to the determination of endocrine disruption. In order to reach the conclusion that a chemical is an endocrine disruptor, one needs to have a systematic method for assessing causation. Various methods have been developed to assess causal relationships; the most commonly referenced approach relates to the criteria outlined by Sir Austin Bradford Hill (Hill 1965). The WHO-IPCS 2002 framework was proposed based on Bradford Hill's criteria to assess relationships between exposures to potential EDCs and altered health outcomes (see WHO-IPCS, 2002, Chapter 7). Several factors were specifically considered important in establishing the overall weight of evidence for a causal relationship: temporality, strength of the association, biological gradient (dose response), consistency of the observations, biological plausibility, and evidence of recovery. A number of illustrative examples were presented to demonstrate how the framework would work to evaluate hypotheses that particular EDCs cause specific adverse outcomes.

The WHO-UNEP 2012 report did not adopt the WHO-IPCS 2002 framework for assessing causation. In fact, the authors criticize the earlier approach because it failed to distinguish between the quality of evidence and strength of the recommendations, as recommended in the Grades of Recommendation Assessment, Development and Evaluation (GRADE) scheme used in clinical medicine (GRADE, 2011). Despite the reference to this alternative approach, the 2012 report did not apply the GRADE scheme (which, because it was derived for clinical medicine, would not be appropriate in any case for the evaluation of toxicology data and studies in wildlife). Although there is substantial discussion of weight of evidence, no systematic approach is described or adopted to assess the weight of evidence for causation; instead, "best professional judgment" is the basis for making expert assessments of the data and aggregated data are purported to be presented on trends, biological plausibility, relevant exposures, consistency across species, dose–response, and temporality (WHO-UNEP, 2012a, p. 19). However, not all of these categories are captured in the discussion of all endpoints of concern; in particular, few data are presented on dose–response. Although a section, titled *Strength of evidence*, is included in the discussion of health effects and conclusion-like statements are provided for some adverse health outcomes, in many cases, these conclusions do not reflect the uncertainties or limitations that were described in the main text. More importantly, these statements do not represent the totality of the evidence, often ignoring contrary study results that were inappropriately excluded. Furthermore, little effort was made to synthesize the findings across the categories in the report – from temporal trends to exposure and biological plausibility – in order to provide a complete picture of the state of science for MOA, dose response, and adverse health effects. The use of the narrative approach in the WHO-UNEP 2012 report allowed for a selective presentation of information without a critical review of the data. In the absence of a formal assessment of causation, subjective inference is relied onto suggest causation.

For each of the specific health or environmental adverse outcomes discussed, information is presented in the 2012 report for a sequence of topics in such a way to suggest that they were related when, often, there is no connection at all. Each subchapter discusses the trends in the subject disease or health outcome first, followed by a suggestion that exposure to environmental chemicals contributes to these trends. In some cases, limited or no data are provided to support environmental chemical exposures as contributing to the trends and often, other causes of the trends are ignored or dismissed. For example, the report stated that there is a rising trend for breast cancer and notes that this trend cannot be explained by improved diagnosis or changes in risk factors, including genetic factors (WHO-UNEP, 2012a, p. 126). This statement about the trends is followed by the comment that twin studies have highlighted the importance of environmental factors. The juxtaposition of these two sentences gives the reader the impression that the rising trend in breast cancer must be a consequence of these environmental chemical exposures. However, the term "environment" as used in twin studies encompasses all modifiable (i.e., non-genetic) factors, not just environmental chemical exposures. Several publications have concluded that the observed increase in breast cancer incidence in some countries can be explained by the introduction and promotion of mammography and breast cancer screening (Glass et al., 2007; Séradour et al., 2009; Weedon-Fekjær et al., 2012). In addition, in several countries around the world, the trend for breast cancer has been in decline since 2002, which has been attributed to the dramatic reduction in hormone replacement therapy in post-menopausal women (Glass et al., 2007; IOM, 2012; LeClère et al., 2013; Séradour et al., 2009; Weedon-Fekjær et al., 2012). Therefore, the implication that the trends are due to environmental chemical exposures is misleading when evidence for some of the changes in breast cancer trends point to improved diagnostic tests or to changes in recommended post-menopausal therapies. In the IOM (2012) review of the state of the science regarding environmental risk factors for breast cancer, the factors with the clearest evidence included: hormone therapy products, oral contraceptives, being overweight or obese, alcohol consumption and ionizing radiation; the evidence for exposure to industrial chemicals was considered to be limited and in some cases only suggest a possible association for an increased risk of breast cancer. Further discussion of the issues associated with disease trends as presented in the WHO-UNEP 2012 report is described in more detail in Section 2.5.

Another part of the discussion in the 2012 report for each adverse outcome included a description of the normal role of endogenous hormones. This information is useful in understanding normal physiology, basic mechanisms and showing the potential biological plausibility for endocrine disruption; however, it does not demonstrate that environmental chemicals are acting in the same manner as endogenous hormones. Evidence is presented from epidemiology and experimental animal studies for particular chemicals associated with various health outcomes, but rarely is there any discussion of specific mechanisms for the highlighted chemicals. For example, androgens and estrogens are mentioned as playing a role in normal prostate development; however, an androgen- or estrogen-mediated MOA cannot be described for any chemical mentioned in this section. In fact, the report notes that "the precise mechanisms by which the chemicals related to prostate cancer induce the carcinogenic process remain to be

resolved" (p. 131). Overall, the reader is left with the impression that environmental chemicals could cause disease via the same mechanisms as endogenous hormones without any understanding of the normal feedback mechanisms that exist for homeostasis as well as differences in potency and dose–response (discussed in more detail in Section 2.6). Moreover, a failure to recognize the complexity of hormone–receptor interaction and activation is lacking. Specifically, despite a short discussion at the beginning of the WHO-UNEP 2012 report, the roles of co-activators, repressors, and transcription factor interactions, and receptor cross-talk are completely ignored throughout the rest of the report. Recent studies also point to competition for transcription factors (Kollara and Brown, 2006); yet these issues are completely overlooked.

Another shortcoming of the WHO-UNEP 2012 report that affects the cohesive evaluation of causation is the fact that little, if any, discussion of exposure occurred in the report in the context of specific effects or related to specific hormonal MOAs. Generally, when exposures were mentioned, a reference is made to Chapter 3 of the 2012 report. In Chapter 3, only general information is presented on major classes of chemicals that the WHO-UNEP 2012 report described as known or potential endocrine disruptors. No quantitative information on exposure is presented for individual chemicals, and potential human exposures are not considered in context of dose–response data. The segregation of exposure data from the information on chemicals of concern, potential adverse effects associated with that chemical, and biological plausibility or possible MOAs for endocrine disruption make the assessment of the causal relationship between exposure and effects impossible.

The lack of a systematic approach to assess causation for specific chemicals and associated health outcomes resulted in conclusions that were predisposed to the identification of potential EDCs. The selective citation of literature without discussion of contradictory studies and the failure to consider alternative causes of reported effects gives the reader the impression that the weight of evidence is stronger than is justified by the available scientific data. This calls into question the integrity of decisions at all levels of the 2012 report. Specific examples are provided below that further demonstrate the issues with the evaluation of causation in the WHO-UNEP 2012 report, with an emphasis on highlighting the key factors that are typically used in a causation assessment.

2.4.1. Adrenocortical hyperplasia in seals

The WHO-UNEP 2012 report concluded that there was sufficient evidence to demonstrate that a mixture of PCBs and DDT caused adrenocortical hyperplasia and a Cushing-like condition in Baltic seals. The report attributes recoveries in seal populations to the "drastic reduction of DDT and PCBs in Baltic biota" (p. 149). This conclusion is not supported by the discussion in the main text of the report; at best, it is based on limited data. Additionally, the report ignores conflicting data and fails to consider alternative causes for the adrenocortical hyperplasia.

The data on adrenocortical hyperplasia in seals or other aquatic mammals are limited and inconsistent. Although the WHO-UNEP 2012 report mentions that adrenocortical hyperplasia has not been reported in seal populations outside the Baltic Sea, it does not discuss evidence from Great Britain that reported contrasting results for other marine mammals (Kuiken et al., 1993); nor does the report address alternative causes for these observations. For example, Kuiken et al. (1993) measured the concentrations of chlorinated hydrocarbons in the carcasses of harbor porpoises and found that adrenocortical hyperplasia was not associated with increased levels of these chemicals, but rather associated with chronic stressors causing their death (e.g., malnutrition, prolonged illness). Another study that is not mentioned, Clark et al. (2006), also found a significantly higher adrenal gland mass in Atlantic bottlenose dolphins that were chronically stressed compared to those that were acutely stressed, suggesting that other factors may be involved in the observation of adrenocortical hyperplasia in seals. Lair et al. (1997) suggested that the adrenal hyperplasia seen in beluga whales may be part of the normal aging process. Although these latter studies are on cetaceans, rather than pinnipeds, they point to alternative causes of adrenocortical hyperplasia in marine mammals and should have been considered in the data review and assessment of causation.

The report also fails to address data inconsistencies for the specific persistent organic pollutants (POPs) mentioned, including the differences observed between experimental animal studies and reports in wildlife. For example, in two-year chronic toxicity studies of PCBs conducted in female Sprague–Dawley rats (NTP, 2006, 2010), increased adrenocortical atrophy was reported, which is in contrast to the findings of adrenocortical hyperplasia found in Baltic seals. Cancer bioassays for several commercial PCB mixtures (i.e., Aroclors) do not report adrenal effects in rats exposed to these PCBs for two years in the diet (Mayes et al., 1998), which further calls into question the identification of PCBs as the cause of adrenocortical hyperplasia in seals. While the degree of similarity between adrenal glands of seals and rats is unknown, it is expected that similar, rather than opposite, effects from the same chemical are more likely in mammals. The WHO-UNEP 2012 report acknowledges in the introduction to the section on adrenocortical hyperplasia that a DDT metabolite showed degeneration and necrosis in the adrenal cortex of laboratory mice – a finding inconsistent with hyperplasia – but this is not mentioned further.

No information on plasma cortisol levels is available for the Baltic Sea seals, as the WHO-UNEP 2012 report notes, which are critical for differentiating a stress response from a direct toxic effect on the hypothalamus–pituitary-adrenal axis as the cause of the observed hyperplasia (Harvey and Sutcliffe, 2010). Although the WHO-UNEP 2012 report recognizes the possible role of stress and aging of wildlife in the development of adrenal hyperplasia in the main text, when the strength of evidence is described, these factors are ignored. The 2012 report does not consider alternative causes for the observed species recovery, and stress can plausibly explain these observations given other changes occurring during the same time period in the Baltic Sea (e.g., reductions in nutrient inputs, eutrophication, oxygen deficiency, and oil discharges) (HELCOM, 2012).

It is presumed that because the effect occurred in an endocrine organ, this must be the result of endocrine disruption, but no data are provided to show that these effects are the result of an endocrine MOA. Given the limited data available on the observation of adrenocortical hyperplasia, inconsistent findings in experimental animal studies, conflicting data in other wildlife species and other regions, and stress as a plausible alternative cause for these observations, it is questionable that there is sufficient evidence to demonstrate that these compounds caused the adrenocortical hyperplasia observed in the Baltic seals.

2.4.2. Prostate cancer

Prostate cancer is included in the discussion of various hormonal cancers in the WHO-UNEP 2012 report and it is stated that there is sufficient evidence for a link between pesticide exposures and prostate cancer (WHO-UNEP, 2012a, p. 130). However, the report does not reach any conclusions or make any statements about the evidence for an endocrine-mediated MOA. The mere association between pesticides and prostate cancer is insufficient to demonstrate causation. More importantly, this link has not been shown to be attributable to an alteration in endocrine function. Moreover, based on the data presented in main text of the report, from an objective view, there is not sufficient evidence to conclude that a link between pesticide exposures in general and prostate cancer

even exists. The report mentions that individual pesticides have been reported to be associated with prostate cancer; however, these data are not consistent. For example, the report mentions that oxychlordane was linked with an increased risk for prostate cancer based on Ritchie et al. (2003). Another study cited in the report (Hardell et al., 2006) that did not observe an association between oxychlordane and prostate cancer is ignored. Other biomonitoring studies that examined oxychlordane levels are not cited in the 2012 report, including two studies that failed to observe an association with prostate cancer (Aronson et al., 2010; Sawada et al., 2010) and one that did (Xu et al., 2010). Thus, the evidence from biomonitoring studies for an association between oxychlordane and prostate cancer is inconsistent. In these same studies, similar results were seen for other organochlorine pesticides where only one or two statistically significant associations were reported for any individual pesticide. Although other epidemiology studies are cited in the WHO-UNEP 2012 report, these studies did not directly measure exposure through analysis of blood, fat, or urine, and therefore, are considered to be more susceptible to bias and should be given less weight in an overall assessment of the evidence. For example, the Agricultural Health Study (AHS) (Alavanja et al., 2003; Koutros et al., 2010) obtained exposure information based on a questionnaire that collected data on duration and frequency of pesticide use, which is a less reliable measure of exposure compared to biomonitoring studies.

It is interesting to note that the two meta-analyses of pesticide applicators (van Maele-Fabry and Willems, 2004) and pesticide manufacturers (van Maele-Fabry et al., 2006) characterized the weight of evidence for pesticide exposure for these workers as weak (rate ratios of less than two). A general limitation with the epidemiology studies on pesticide exposure and cancer is the use of multiple comparisons to assess risks for many different types of pesticides and various cancer endpoints, which increases the likelihood that a statistically significant association will be detected based on chance alone. In addition, the long latency between initiation and detection of the cancer make it very difficult to identify relevant exposures. Overall, the data for pesticides as a broad category are weak and those for individual pesticides are limited with regard to an association with prostate cancer, let alone to causation.

While it has been speculated that hormones play a role in the development of prostate cancer given the involvement of sex steroids in the development of the prostate, it is recognized that many other factors may be involved in the etiology of prostate cancer. These factors include age, family history (genetics), race, dietary fat, and other dietary factors (NCI, 2013). As the WHO-UNEP 2012 report notes, the mechanism by which pesticides could induce prostate cancer is currently unknown. Therefore, while it may be biologically plausible for endocrine disruption to be contributing to prostate cancer, insufficient data are available to show that pesticides are involved in the induction of prostate cancer by an endocrine-mediated MOA.

Based on the issues discussed above regarding the strength of the association, lack of consistency among studies of pesticide workers and manufacturers and lack of evidence for an endocrine MOA, the overall weight of evidence for pesticides causing prostate cancer through endocrine disruption should have been determined to be either inconclusive or weak, not "sufficient" as determined in the WHO-UNEP 2012 report.

As the above examples illustrate, the WHO-UNEP 2012 report presented information on chemicals and various adverse outcomes, but whether these exposures actually cause these effects was not established scientifically. Several factors for establishing causation, such as demonstrating exposure to the chemical, dose–response, and consistency in the data, were frequently ignored. Most critically, the lack of a formal framework or standardized approach to evaluate the data on specific chemicals and the potential causal association with adverse outcomes via an endocrine-mediated MOA is a significant shortcoming in the WHO-UNEP 2012 review.

2.5. What do temporal trends in diseases show?

The WHO-UNEP 2012 report indicates that the high incidence and increasing trends of many endocrine-related disorders in humans is one of "three strands of evidence [that] fuel concerns over endocrine disruptors" (p. vii). The report indicates that "worldwide, there has been a failure to adequately address the underlying environmental causes of trends in endocrine diseases and disorders" (WHO-UNEP, 2012a, p. ix).

It should be noted that, in some cases, the WHO-UNEP 2012 report does put trends in perspective. For example, the report indicates that elevated BPA levels in the body could be a result of polycystic ovarian syndrome and not the other way around. In other cases, however, alternative explanations are either not discussed, or when they are discussed, they are dismissed in favor of endocrine disruption as an explanation without a sufficient evaluation of the science. For example, as noted earlier, the report briefly acknowledges, but appears to dismiss, the role of alternative factors in reported trends for breast cancer or sperm/semen quality.

The 2012 report concluded that because the increase in disease trends has occurred primarily over the last few decades, these trends cannot be entirely attributable to genetic causes. It is implied that if these trends are not the result of genetic heritability, then the only other explanation is environmental exposure to chemicals. However, environmental factors go well beyond chemicals, and cover a multitude of characteristics in human populations including: diet, exercise, lifestyle factors, infectious agents, and even drug use; for wildlife, these include factors related to habitat, food supply, disease, predation, and competition – factors which can be completely unrelated to environmental chemical exposures. The WHO-UNEP 2012 report does not acknowledge that, much of the time, the environmental causes of the diseases being discussed are not chemical exposures.

There are many factors that can influence the appearance of an increasing trend (either temporally or geographically) in disease incidence or prevalence. For human health considerations, these include changes in diagnostic criteria, screening, medical interventions, and treatment. Other life style trends are important; for example, giving birth at an older age can have a substantial impact on the incidence of birth defects and congenital abnormalities. Another significant trend is the obesity epidemic and being overweight has been found to increase the risk of male infertility (Hammoud et al., 2008). In addition, the report sometimes described trends based on a compilation of data from different sources such that, what appears to be a trend actually may be a reflection of different data collection methods.

Through selective citation in the WHO-UNEP 2012 report, an impression is created that certain diseases have an increasing incidence or prevalence. However, publications that are not cited in the WHO-UNEP 2012 report often provide opposing evidence that the occurrence of the disease of interest is not on the rise. For example, in the 2012 report, several papers are cited to indicate that the prevalence of hypospadias is increasing. Fisch et al. (2010) was not cited, who stated that "[a] review of the epidemiologic data on this issue amassed to date clearly demonstrates that the bulk of evidence refutes claims for an increase in hypospadias rates." The postulated decreasing trends in semen quality have also been highly contested by other scientists, as already discussed above.

In the end, even if an environmental exposure and a health outcome trend are related spatially or temporally in an ecological epidemiology study, all one can be sure of is that a statistical correlation exists. Such study designs cannot determine whether people with health effects are the same people with a particular exposure in these studies. In addition, it generally cannot be known if the exposure occurred before the health outcome (or vice versa), whether they each have a common cause, or whether they are completely independent. As a result, these analyses are the weakest form of scientific evidence for evaluating causation.

The reasoning presented in the WHO-UNEP report that the rising disease trends must be associated with exposure to EDCs becomes even more questionable given that exposure to most of the compounds named in the report have not increased over the last twenty or thirty years, but rather, have decreased. Concentrations of DDE in human milk in Germany have been reported to be reduced by approximately 90% from 1984 to 2001 (Wilhelm et al., 2007). Another biomonitoring study in Germany clearly shows that human exposure to phthalates has declined over time, with the exception of some new compounds that were recently introduced to the market (Wittassek et al., 2007). Historical biomonitoring data from a number of countries including the Czech Republic, Norway, and the U.S. indicates that human exposure to compounds like persistent chlorinated pollutants has decreased in the last two decades (Cerná et al., 2012; Ferriby et al., 2007; Nøst et al., 2013). These trends in declining human exposures over the last two decades in the Western world contradict the hypothesis presented in the WHO-UNEP 2012 report that a continuing rise in disease trends, if any, can be related to EDC exposure in humans.

Overall, any claims of trends indicating an endocrine cause must be supported by a systematic review of all relevant data regarding disease trends, exposure trends, and alternative explanations for observed statistical correlations. Below, two examples are discussed in which the WHO-UNEP 2012 report described trends of increasing endocrine-related disorders and concluded they are due to environmental EDCs, without considering whether the weight of evidence supports such conclusions or alternative explanations are more likely. Note that these examples are not weight-of-evidence analyses, but the identification of factors and limitations related to the discussion of health trends in the WHO-UNEP 2012 report and how these factors may lead to erroneous conclusions.

2.5.1. Autism spectrum disorders (ASDs)

Autism spectrum disorders (ASDs) – which include autistic disorder, Asperger's syndrome, and pervasive developmental disorders not otherwise specified – are developmental disabilities that are diagnosed based on behavioral symptoms and failure to reach certain developmental milestones (CDC, 2012). ASDs are characterized by communication and socialization problems, as well as atypical behaviors and interests. ASD symptoms are usually apparent before the age of three and can vary in severity and presentation. While IQ decrement can co-occur with ASDs, they are not associated with ASDs *per se* (CDC, 2012).

The WHO-UNEP 2012 report claims that "the increase in autism spectrum disorders is indisputable" (p. 109) and there is "sufficient evidence to conclude that a number of factors, including environmental, contribute to the increases in autism spectrum disorders" (p. 119). The first claim is based on two studies, one published in 1976 and the other in 2007 (Rice, 2007; Wing et al., 1976). More current analyses were not cited; there is a considerable body of other literature evaluating ASD prevalence that the WHO-UNEP 2012 report did not consider.

At least some of the increase in ASD prevalence is due to changes in diagnostic criteria, better diagnostic techniques and increased case ascertainment (CDC, 2012). For example, when the Centers for Disease Control and Prevention (CDC) created new diagnostic standards for its Autism and Developmental Disabilities Monitoring Network (ADDM), there was a significant increase in ASD prevalence compared to the prevalence estimated using older standards (Rice et al., 2012). Also, a recent reanalysis of older studies found ASD prevalence to be consistent with current reports when the data were analyzed according to contemporary diagnostic criteria, indicating that ASD was likely underestimated in earlier studies (Duchan and Patel, 2012). Increases in ASD prevalence may also be partly attributed to diagnostic substitution, as children who would have been diagnosed with learning disabilities or mental retardation in the past are currently diagnosed with ASD (reviewed by Fombonne et al., 2009). ASD diagnosis relies on behavioral identification, which leaves room for wide variation in clinical judgment and is influenced by differing cultural and social norms worldwide (Elsabbagh et al., 2012). Even within a culture, there is evidence of low inter-evaluator agreement about diagnoses, although it is not clear in which direction it would influence prevalence measures (Duchan and Patel, 2012). Finally, the success of national awareness efforts and the growth of alternative treatments may also contribute to a perceived increase in ASD prevalence, with more children being tested, diagnosed, and treated (Duchan and Patel, 2012). Therefore, it is not clear that there is a true increase in ASDs.

Regardless of temporal trends, the WHO-UNEP 2012 report presents no evidence that environmental factors, much less EDCs, contribute to ASDs. The report summary stated that "insufficiency of thyroxine during pregnancy is also associated with reduced intelligence quotient, ADHD and even autism in children" (p. xii). Yet, the WHO-UNEP 2012 report does not provide a reference to support this claim, nor discuss any other factor(s) that may contribute to ASDs. There is no clear etiology for autism and no known MOA for it being induced, thus the current science cannot substantiate endocrine disruption as causal. Many potential non-EDC risk factors have been studied, including genetics, older paternal age, sex, prenatal nutrition, and *in utero* exposure to antidepressants and pain-killers (e.g., Duchan and Patel, 2012; Gentile et al., 2013; Guinchat et al., 2012; Kinast et al., 2013; Schmidt et al., 2011). Although the weight of the evidence supporting these associations varies, the available data indicate that a good deal of research on possible causes of ASDs is not considered in the WHO-UNEP 2012 report.

Overall, the two references (Rice, 2007; Wing et al., 1976) regarding endocrine disruption and trends reported for ASDs on which the WHO-UNEP 2012 report relies are not representative of the literature as a whole. Thus, the report's conclusions that there are actual increases in the spectrum of autism-related disorders and that these increases are due to endocrine disruption are not supported based on the current state of the science.

2.5.2. Wildlife population declines

The WHO-UNEP 2012 report stated that the evidence for "endocrine disrupting POPs such as PCBs and organochlorines" (p. 186) as causes of wildlife population declines has increased since 2002 due to observed increases in the populations since restrictions on the use of these chemicals. The logic presented – that as chemical exposures increased, populations declined and, conversely, as chemicals were removed from the market and as exposures declined, populations recovered, would be reasonable if: (1) the chemical exposures are documented; (2) the levels of exposure occurring are sufficient to impact the organisms; (3) the organism-level impacts are manifested in population-level impacts; and (4) other possible causes for population changes are adequately considered. The WHO-UNEP 2012 report falls short in demonstrating the linkages that would be required to make a case based on all of these points.

The two most prominent examples cited in the WHO-UNEP 2012 report are links between DDT and bird populations and between tributyl tin (TBT) and snail populations. For the latter example, the report cites publications by Jörundsdóttir et al. (2005) and Morton (2009). Jörundsdóttir et al. (2005) observed reductions in the levels of imposex in the dogwhelk (*Nucilla lapillus*) in Iceland, mainly near small harbors with no change seen in larger harbors. As no measurements were made of TBT concentrations, Jörundsdóttir et al. (2005) stated that the continued impacts in the large harbors are "presumably associated" with continued use of TBT paints on larger vessels. Morton (2009) documented a 20-fold increase in the population of *N. lapillus* on the southeastern coast of England during the period May 2004–August 2008, which coincided with the period over which TBT was banned as an antifoulant paint globally. Morton (2009) stated that, "due to the lack of confirmatory chemical data, the changes in population size, structure, and reproduction herein reported upon for *N. lapillus* cannot be correlated positively with changes in ambient TBT levels." The WHO-UNEP 2012 report discusses recovery in the abundance of North Sea brown shrimp, although there is no known mechanism of endocrine disruption by TBT in crustaceans. Verhaegen et al. (2012), as cited in the WHO-UNEP 2012 report, state that the inability to demonstrate an "unarguable causative link" between decreased organotin concentrations and recovery of the shrimp stock is due to the lack of data on both exposure and effects in these organisms. None of these weaknesses in the conclusions of the cited studies are mentioned in the WHO-UNEP 2012 report, and therefore, the observed trends of imposex and the recovery of snail and shrimp populations cannot be definitively attributed to TBT, much less endocrine disruption.

Although the ability of organotins to cause masculinization of female gastropods (including the development of imposex) is probably the most-recognized EDC effect in wildlife over the past 30+ years, the vast majority of field studies on this phenomenon do not include chemical analyses of body burdens (Titley-O'Neal et al., 2011). This review article by Titley-O'Neal et al. (2011) is cited in the WHO-UNEP 2012 report, but a number of interesting points from the review were not mentioned. For example, the WHO-UNEP 2012 report does not note the lack of agreement among researchers on the mechanism for induction of effects, the observation of imposex prior to the use of TBT, the natural occurrence of imposex in some species, the lack of sensitivity of a number of species to TBT-induced imposex, and the fact that female masculinization by TBT or triphenyl tin (TPT) has been confirmed in the laboratory in only a small fraction of species affected (7.5% or 20 species confirmed out of 268 total species examined). Thus, the statement in the WHO-UNEP 2012 report that the "temporal relationship between a measure of exposure and population parameters" for TBT as an example of the "best evidence of a relationship between EDCs and wildlife populations," does not reflect the uncertainties in the available information, including the studies cited in the report.

The WHO-UNEP 2012 report acknowledges the difficulty in making the link between declines/recoveries in wildlife populations and EDCs, stating that many factors may be responsible. These factors may include food, habitat, competition, predation, overall environmental quality, climate change, and human activities (e.g., harvesting, traffic, noise). Regardless, the report emphasizes chemicals, specifically EDCs, as the main causative factor. Even in the case of TBT and *N. lapillus*, which is arguably the best known example of EDC effects on wildlife, there are other factors that impact the distribution and abundance of this gastropod. This species is sensitive to changes in nutrient levels, substrate loss, toxic algal blooms, and oil spills (Bryan, 1968; Gibbs et al., 1999; Robertson, 1991). For the DDT example cited in the WHO-UNEP report, a number of confounding factors are likely to have affected the recovery of osprey populations, as discussed by Henny et al. (2010), who stated that "expansion of suitable habitat (reservoirs) and enhanced use of artificial nest sites confounds a simple conclusion that recent population increases were solely a recovery from earlier contaminant exposure," especially in the western United States. The WHO-UNEP 2012 report concluded that the "strength of the evidence linking EDC exposure to most wildlife population declines is insufficient," then goes onto make the statement that "an endocrine mechanism for wildlife declines is probable but not conclusive" (p. 186). It would be more appropriate to conclude that the evidence for an endocrine mechanism is hypothetical, particularly given the fact that for the two best known examples for wildlife declines, DDT and TBT, an endocrine mechanism, while possible, is only one of many potential factors that may be contributing to the observed population dynamics.

2.6. Importance of dose–response and potency

In Chapter 1 of the WHO-UNEP 2012 report, an effort is made to describe the endocrine system in general – the glands involved, hormones produced, molecular mechanisms involved in mediating responses, and the physiological processes that are regulated by this system. A few of the feedback mechanisms that are an integral part of this system (e.g., how insulin secretion is affected by changes in blood glucose levels) are also mentioned. These various negative feedback loops are important in regulating the production and release of hormones, the expression of various hormone receptors, and generally maintaining homeostasis (i.e., a stable internal environment). What the WHO-UNEP 2012 report fails to fully discuss, however, is the fact that the endocrine system is specifically designed to respond to environmental fluctuations and such homeostatic responses generally are considered normal, adaptive, and necessary as long as they are transient and within the normal homeostatic range (Goodman et al., 2010; Rhomberg et al., 2012). In fact, the responsive nature of the endocrine system is essential to health as seen in the hormonal changes that occur when a woman becomes pregnant. In other words, not all modulations of endocrine function are necessarily adverse. Based on this fact, it can be generally accepted that endocrine activity observed through *in vitro* testing (or even some *in vivo* assays) is not sufficient to classify a substance as an endocrine disruptor if these tests do not address whether the alterations cause actual harm in a whole organism or its offspring. Rather, such a substance may be considered endocrine-active only (EFSA, 2013a); without a clear indication of consequent adversity in a living organism, the substance does not reach the level of an endocrine *disruptor*. Nevertheless, the WHO-UNEP 2012 report often presented evidence of *in vitro* or *in vivo* endocrine modulation (rather than adversity) as support for certain substances being classified as EDCs. For example, in the section on adrenal disorders, only *in vitro* data are presented as evidence of potential endocrine disruption in humans; no epidemiologic evidence of adrenal effects in people is available. Given the fact that a large proportion of the cited *in vitro* studies relate to alterations in gene expression, the data presented fall short of demonstrating an adverse effect and are not sufficient for establishing endocrine disruption. Despite the lack of robust evidence for adrenal disorders in humans as a result of exposure to environmental chemicals, the WHO-UNEP 2012 report identified the adrenal cortex as "the most commonly affected and vulnerable endocrine organ in toxicology" (p. 148).

The WHO-UNEP 2012 report contends that hormonal dose–response curves are non-monotonic and states that NMDRs for EDCs are to be expected. As support for NMDRs, the WHO-UNEP 2012 report cited a recent review by Vandenberg et al. (2012) in which numerous examples were presented of EDCs that exhibit these types of behaviors. This review has been duly criticized, however,

for its selective dismissal of studies that do not show these effects and the general acceptance of those studies that do show these types of responses without any type of critical evaluation of study quality; the inclusion of studies that do not address adverse effects, but rather, transient, adaptive responses; and a failure to consider whether the doses examined in these studies are of any relevance to human exposure levels (Rhomberg and Goodman, 2012).

The Danish Centre on Endocrine Disrupters, in its examination of the evidence presented by Vandenberg et al. (2012) for NMDRs, noted that the majority of these data were from *in vitro* studies and inappropriately included findings for which the U-shaped or inverted U-shaped curves were the product of general toxicity (DTU Food, 2013). It was concluded that 45% of the *in vitro* examples cited by Vandenberg et al. (2012) were the result of cytotoxicity and thus were not examples of true NMDRs. Of the remaining examples cited, approximately one-third were judged to be false and another one-third were considered questionable. Further, only 5 of the 34 *in vivo* examples cited by Vandenberg et al. (2012) were considered to show "clear evidence" of NMDRs. In other words, while examples of NMDRs do exist, they are not as common as Vandenberg et al. (2012) suggests. Recently, USEPA also conducted an expert review of the experimental evidence for NMDRs (USEPA, 2013). In this draft report, EPA noted that such responses are not uncommon in *in vitro* studies and often relate to "lower-order biological endpoints" rather than apical endpoints. However, "[t]here is currently no reproducible evidence that the early key events involved in the expression of NMDRs that are identified at low doses are predictive of adverse outcomes that may be seen in humans or wildlife populations for estrogen, androgen or thyroid endpoints" (USEPA, 2013, p. 8). USEPA concluded that, while NMDRs for adverse effects have been occasionally seen in intact organisms, NMDRs are relatively uncommon. Further, such dose–response curves – when observed – typically occur at high doses, well above the NOAELs identified in standard testing paradigms. In summary, the limited available evidence for low dose effects and NMDRs does not preclude the need to consider dose in assessing the potential hazards of chemicals to the endocrine system.

The WHO-UNEP 2012 report also asserts the potential for EDCs to act at very low doses (i.e., doses below the no observed adverse effect level [NOAEL] or doses below a dose that is environmentally relevant to humans). The implication is that no threshold can be characterized for adverse effects; therefore, any exposure poses some kind of concern, though the magnitude of the effect, or even its direction, let alone its adversity, is not considered relevant from this perspective. However, there are many reasons to question the assertion of effects at very low doses. First, as noted by EFSA (2010), studies of low-dose effects often suffer from various methodological shortcomings (including the use of small numbers of animals and single doses). Additionally, their findings are of questionable toxicological relevance and frequently cannot be replicated in subsequent, more robust studies (EFSA, 2010).

Second, the WHO-UNEP 2012 report generally does not address the fact that many EDCs have much lower potency than endogenous hormones (Nohynek et al., 2013; Sharpe, 2003). In fact, the report claimed that in the diethylstilbestrol (DES) case study, while other chemicals may be less potent than DES, these effects are "equally undesirable when the exposure occurs in early development where potency seems less important" (WHO-UNEP, 2012a, p. 25). In other words, the report suggests that at vulnerable developmental stages, potency may not be very relevant. This statement confuses the issue of potency and sensitivity at different life stages. Although certain substances may be more or less potent depending on the particular life stage at which exposure occurs, potency is always important – no matter the developmental window. The WHO-UNEP 2012 report rightly stated that hormone potency and receptor affinity are not the same things (p. 12). It is further suggested that potency depends on many different factors, but it is vague as to what those factors might be (other than receptor abundance). At the receptor level, potency is determined by both the affinity of a substance to bind to a receptor site as well as the efficacy with which that substance activates the receptor (Borgert et al., 2013). Endogenous hormones have both strong affinity for their receptor sites as well as high efficacy for activation of these receptors; thus, hormones generally are highly potent for modulating endocrine function. Exogenous chemicals, on the other hand, are rarely as potent as hormones, either due to reduced affinity, reduced efficacy, or both (e.g., Gaido et al., 1997; Nilsson, 2000). Although the WHO-UNEP 2012 report claims that "very low concentrations of environmental endocrine disruptors could add to the endogenous hormone effect to produce a response that is much greater than would be predicted based on the hormone alone" (p.8), this theory fails to consider the existence of biological thresholds. As described by Borgert et al. (2013), given the lower potency of most exogenous chemicals, the additional presence of these chemicals will not significantly alter hormone receptor occupancy; thus, a biological threshold for potency exists.

At the whole organism level, potency relates to the ability of a substance to produce a biological effect and may be substantially different from the potency measured with *in vitro* assays (EFSA, 2013a). In considering potential EDCs, therefore, potency should not be determined based on the results of *in vitro* studies. Rather, EFSA recommends that potency should be based on the ability of a substance to produce an adverse health effect *in vivo* (EFSA, 2013a). This ability will depend on not just a substance's potency at its receptor site, but also on its disposition in the body, the timing of exposure (i.e., the particular life stage of development) and the dose and duration of exposure. Therefore, dose remains an important factor in assessing the potency of potential EDCs to cause adverse health effects.

Endocrine disruption is generally posited throughout the WHO-UNEP 2012 report in terms of "adverse" outcomes, yet the report fails to provide a concrete definition for what may be considered an adverse response. In particular, the 2012 report did not adopt the IPCS (2004) definition of an adverse health effect; "change in morphology, physiology, growth, development, reproduction or life span of an organism, system, or (sub) population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influence." The need to clearly delineate adverse from adaptive responses, particularly when considering results from *in vitro* assays, was addressed and possible definitions for these terms proposed in a recent workshop (Keller et al., 2012). Certain endocrine-mediated adverse effects – such as cancer or reproductive disorders – clearly can be judged as detrimental. For other endpoints – such as alterations in hormone levels – it is more difficult to delineate an adaptive response that is within the limits of homeostasis from one that has gone beyond those limits for a sufficient period of time and, therefore, capable of causing an adverse effect. This is particularly true when hormone levels are only measured at a single time point shortly after exposure with no indications of whether the response is transient or more permanent. Thus, the mere presence of a change does not necessarily mean that the outcome is adverse. For example, in the discussion on thyroid-related disorders and diseases (WHO-UNEP, 2012a, p. 97), the report suggests that chemicals can interfere with thyroid hormone signaling without affecting serum hormone levels, but it is stated that methods to evaluate this are not yet available. Changes in thyroid hormone signaling alone cannot be characterized as adverse without evidence to show that these changes lead to impairment in function (Blanco and Kim, 2006). Furthermore, the failure to observe changes in serum thyroid hormone levels would indicate a lack of consequence from the change in signaling.

In its opinion on the scientific criteria for the identification of endocrine disruptors, EFSA specifically discussed a "threshold of adversity," noting that toxicologically relevant responses occur only when the degree of endocrine modulation elicited is beyond that which could be counteracted through homeostatic mechanisms (EFSA, 2013a). These thresholds are similar to those that exist for responses measured from other physiological systems operating within the body. Further, their importance in characterizing endocrine disruption has been emphasized by toxicologists concerned about the European Commission's recommendations for the regulation of EDCs (Dietrich et al., 2013). In contrast, the WHO-UNEP 2012 report proposes that thresholds for endocrine disruption "should not be assumed" (p. 19) and exposures to endocrine-active substances – no matter the level – will add to the already present hormone levels in the body and thus alter endocrine function in a threshold-independent manner. However, using mathematical calculations within a systems biology construct, Borgert et al. (2013) have posited that the endocrine system is able to discriminate potent hormonal signals from the "background noise" of other endogenous molecules, making the system relatively robust in its responses and resistant to spurious interferences by substances with lower potency.

Because thresholds exist, not only for inducing an endocrine response, but also for moving beyond adaptive modulation toward adversity, it is important to understand at what doses the observed responses occur and how these doses compare to the levels at which people or wildlife are typically exposed. In the discussions in Chapter 2 of the report regarding various adverse outcomes, however, the WHO-UNEP 2012 report often fails to mention the doses at which findings are observed. It is important to note that doses administered in experimental animal studies are often orders of magnitude above those to which people or wildlife are generally exposed. Frequently, if one delves deeper into the scientific literature as seen in the examples below, it is apparent that the doses associated with the reported findings are extremely high, well above those to which people or wildlife may be typically exposed.

It is unclear why the WHO-UNEP 2012 report fails to consider dose in the discussions of evidence for endocrine disruption in humans and wildlife. In the beginning of Chapter 2 of the report, it is noted that the focus was on the "identification of the characteristics of the hazards posed by endocrine disruptors rather than risk assessment," (WHO-UNEP, 2012a, p. 23) because accurate risk assessments are difficult in light of limited human exposure data and the combined effects of mixtures. However, the report often draws conclusions that appear to go beyond a simple assessment of potential hazard. For example, the report concluded that environmental exposures play a role in the observed increased incidences of hormonal cancers – rather than saying that these exposures have been associated with the cancers (WHO-UNEP, 2012a,b, p. 137). Similarly, the report states that adrenal changes seen in Baltic seals were caused by exposure to DDT, PCBs, and their metabolites – instead of saying that they have been associated with these exposures (WHO-UNEP, 2012a, p. 147). Further, at the end of Chapter 1, the WHO-UNEP 2012 report stated that "best professional judgment was used to make expert assessments of the data linking exposure to chemicals with each disease/dysfunction," (p. 19) and relevant exposures and dose–responses were considered. Therefore, although dose and exposure were specified as important factors in the evaluation of endocrine disruption, in reality, this does not appear to have been the case.

To illustrate, a few selected examples are discussed below. These are not isolated examples, but rather, representative of how the lack of the consideration of dose in the WHO-UNEP 2012 report leads to a false impression that humans are at risk of endocrine effects from their daily exposures to chemicals.

2.6.1. DES or genistein and endometrial cancer

In the discussion of animal studies of EDCs and endometrial cancer, the WHO-UNEP 2012 report cites the study of Kabbarah et al. (2005) as showing "[g]reater than 90% of CD-1 pups neonatally exposed to DES or the phytoestrogen genistein develop endometrial cancer by 18 months of age whilst C57Bl/6 mice are resistant" (WHO-UNEP, 2012a, p. 130). In this study, both DES and genistein were injected subcutaneously into the pups (a route of administration not relevant to environmental exposures for humans) at doses of 1 and 50 mg/kg/day, respectively, on postnatal days 1–5. The DES dose is over 1000-fold higher than the typical estrogen dose that women receive from low-dose daily oral contraceptive pills (Kripke, 2005). Further, the daily intake of genistein in Japanese subjects, a population known to have high intake of soy isoflavones, has been shown to be <1 mg/kg/day (Nakamura et al., 2000; Wakai et al., 1999), at least 50-fold lower than the dose administered to mice in Kabbarah et al. (2005). Thus, the doses of DES and genistein used in this study are well beyond those to which people would be typically exposed. Another issue is that this study was actually conducted in knockout mice with a genetic predisposition for DNA repair errors – a fact that is not mentioned in the WHO-UNEP 2012 report and that further brings the human relevance of the findings into question.

2.6.2. PCBs and neurodevelopmental effects

In another example, in the discussion on neurodevelopmental disorders, the WHO-UNEP 2012 report cites three studies as consistent evidence that "PCB exposures decrease serum thyroid hormone levels," (p. 113) with no mention of study details. However, the doses at which effects were observed in those studies are extremely high. In Goldey et al. (1995) and Zoeller et al. (2000), rats were exposed to 1, 4, or 8 mg/kg/day of Arochlor 1254 on gestational days 6–21. Although circulating T4 levels were reduced on postnatal days 1–30, they recovered by postnatal day 45, indicating a transient effect. More importantly, Goldey et al. (1995) reported that pup mortality was 20% and 50% in the 4 and 8 mg/kg/day dose groups, respectively, indicating that these doses were extremely high. In the third study (Bastomsky, 1974), adult rats were injected with an even higher dose of 25 mg/kg/day of Arochlor 1254 for 4 days. In contrast, the mean intake of PCBs from consumption of the French diet was recently estimated at 2.71 ng/kg/day for adults and 3.77 ng/kg/day for children (Sirot et al., 2012), while that from consumption of the Japanese diet was estimated at 1.45–2.08 pg/kg/day (Nakatani et al., 2011). Thus, human exposures to PCBs in the diet are over 300,000 times lower than the doses used in the experimental studies (on a human equivalent dose basis) cited in the WHO-UNEP 2012 report as consistent evidence of effects of PCBs on thyroid function. This dramatic difference in dose calls into question the relevance of these studies to human environmental exposures.

It should be further noted that, although the thyroid develops and functions in a manner generally similar between rodents and humans, differences exist that make neonatal rats more susceptible and less capable of compensating for possible alterations in function than humans. For example, the human fetal pituitary can respond to thyrotropin-releasing hormone (TRH) as early as gestation week 25 and thyroid-stimulating hormone reaches peak serum levels somewhere around this same time, while the hypothalamic-pituitary-thyroid axis in rats does not respond to TRH signals until a couple of weeks after birth (Howdeshell, 2002). Further, free thyroid hormone levels can be maintained during pregnancy in humans via increased peripheral metabolism and enhanced thyroid hormone binding to serum proteins (Ahmed et al., 2008; Howdeshell, 2002). Consequently, children with congenital hypothyroidism may be born with low-normal concentrations of

thyroid hormone due to compensation by the maternal system (Ahmed et al., 2008).

2.6.3. BPA and adverse effects

Throughout the WHO-UNEP 2012 report, BPA is mentioned as being responsible for a variety of adverse findings in rodents, including fibroid development in mice and rats (p. 42); defeminization and other alterations in social behaviors in female rats (p. 115); altered mammary gland development leading to increased tumor induction (p. 128); endometriosis in offspring of exposed mice (p. 130); and modified immune responses in mice (p. 169), to name a few. In all of these cases, the doses of BPA associated with these findings are not reported. Further, in Chapter 3 of the WHO-UNEP 2012 report, the various ways in which people may be exposed to BPA are emphasized (e.g., in the call-out box on page 196 under the heading called "origin and use"). The report also notes that BPA is found in virtually all people (WHO-UNEP, 2012a, p. 225), but no information is provided on the magnitude of exposures, the biological concentrations that have been measured in people, or whether these would be sufficient to cause adverse effects. The implication is that people are at risk of adverse health effects because they are exposed to BPA. However, a number of recent weight-of-evidence evaluations have been conducted to assess the potential risks to humans from BPA exposure (EFSA, 2013b; Goodman et al., 2006, 2009; Hengstler et al., 2011; Teeguarden and Hanson-Drury, 2013). These reviews document that some BPA results reported in investigatory experiments have not been replicated in subsequent studies and that many studies have used non-oral exposure routes that bypass first-pass liver metabolism and thus are not relevant to human oral exposures. The majority of BPA studies have been conducted at doses well above those to which humans are generally exposed and human exposures are generally well below the current BPA tolerable daily intake (TDI) of 0.05 mg/kg/day derived from two- and three-generation reproductive studies in rodents. More specifically, daily BPA exposures were recently estimated by EFSA (2013b) to be ≤857 ng/kg/day for toddlers and ≤495 ng/kg/day for infants 1–5 days of age; these values are 50- to 100-fold lower, respectively, than the BPA TDI value of 0.05 mg/kg/day. Daily BPA exposures were also estimated by the U.S. Food and Drug Administration (USFDA) to be 100–200 ng/kg/day for children and adults and 200–400 ng/kg/day for infants (USFDA, 2009); these values are even lower than those estimated by EFSA. In other words, the implication of human health risks from BPA exposure raised in the WHO-UNEP 2012 report is unfounded when the data are considered in the context of actual doses administered and concentrations to which people are typically exposed.

These examples demonstrate that dose was not considered in the discussion of experimental animal studies in the WHO-UNEP 2012 report. In all of the illustrated cases, potential human exposures are orders of magnitude lower than those administered in toxicology studies. In addition, some of these studies utilized routes of exposure (i.e., injection) that bypass normal metabolism and elimination of chemicals and therefore, are not relevant for assessing potential environmental exposure to humans. Further, the WHO-UNEP 2012 report did not address other dose-related issues such as thresholds for effects and potency, when discussing specific chemicals.

3. Conclusions

The WHO-UNEP 2012 state-of-the-science report on endocrine disruptors is purported to be an update of the WHO-IPCS 2002 state-of-the-science report – however, it is neither a state-of-the-science review, nor is it an update of the 2002 report.

The 2012 report cannot be characterized as a state-of-the-science review because it lacks several key features for this type of assessment including: the lack of a defined scope for the review, the absence of a process for identification, integration, and interpretation of data, the lack of a structure for evaluating individual studies for relevance and reliability, and an objective method for evaluating the weight of the evidence. These deficiencies undermine the conclusions reached in the report. The WHO-UNEP 2012 report can be more appropriately characterized as a selected discussion of aspects of science that should be considered when discussing endocrine disruption, but it is not a summary of the current state of the science.

Neither can the WHO-UNEP 2012 report be considered an update to the WHO-IPCS 2002 report because it does not build on and modify the earlier analysis, giving reasons and support for the changes in the state of the science. A true update to the earlier report would cite the 2002 conclusions, articulate what data, findings, or new understanding since 2002 should be considered and evaluate how and whether the 2002 conclusions need to be modified in light of the newer information. In addition, the WHO-UNEP 2012 report does not address research recommendations from the earlier report. In some cases, the 2012 report reviews the same data from the 2002 report, but reaches conclusions that conflict with those of the earlier report, despite the lack of new information to support a change in the weight of evidence. The fact that the 2012 report reaches more definitive conclusions based on the same data emphasizes a reliance on subjective decision making and less stringent criteria for evaluating potential causal relationships compared to the earlier 2002 report. Although the WHO-UNEP 2012 report is stated to be an update, this report in actuality is a revised review of the state of the science that does not build upon what was previously done and disregards the WHO-IPCS 2002 proposed framework for causation in favor of "best professional judgment" on these matters.

A key concern with the WHO-UNEP 2012 report is the use of subjective inference instead of a formal framework to assess the potential role of causation for endocrine disruption. The report adopted a narrative approach for the data review that does not represent a weight-of-evidence assessment. Rather than demonstrate causation, the report relies on inference to suggest that exposures to chemicals and adverse outcomes are related. The WHO-UNEP 2012 report presented information on chemicals and various adverse outcomes, but whether the exposure causes these effects was not determined or adequately considered in an objective, transparent, and scientific manner. Several key factors for establishing causation, such as a demonstrated exposure to the chemical, dose–response, and consistency in the data were frequently ignored. For example, temporal trends in human diseases or wildlife populations are presented without consideration of alternative explanations for these trends (especially diagnostic criteria and reporting changes). Exposures to chemicals considered to have the potential for endocrine disruption exist and are suggested as contributing to the observed trends, but there is little consideration of whether these exposures are sufficient to explain the alleged effects and whether the patterns of exposure are congruent with the trends. Above all, the lack of a framework to collectively evaluate, in an objective and comprehensive manner, the data on specific chemicals and the alleged adverse outcomes is a significant shortcoming in the WHO-UNEP 2012 review.

The WHO UNEP 2012 report fails to fully address a number of critical factors that must be considered when defining EDCs; specifically dose, dose–response, potency, and adversity. First, the substance must be shown to cause an adverse effect in an intact organism, their progeny or (sub)populations; therefore, *in vitro* data alone are insufficient for classifying a compound as an endocrine disruptor. Further, the observed effect must be shown to go

beyond adaptive modulation of endocrine function; that is, it must result in an adverse outcome. Second, thresholds exist for inducing such adverse effects. The 2012 report does not give appropriate consideration to thresholds and dose–response for adverse effects. Despite the chosen definition of endocrine disruption as producing adverse changes, the report treats all effects as evidence of disruption and makes an *a priori* rejection of thresholds. Where examples from animal testing are discussed there is little consideration of dose–response, when, in fact, only some doses of some compounds can cause endocrine disruption in the laboratory. The WHO-UNEP 2012 report also does not address the fact that most EDCs have much lower potency than endogenous hormones and potency is important regardless of the life stage at which exposure occurs. Finally, consideration of low dose effects and NMDRs do not exclude the need to consider potency and exposure.

It is also important to emphasize that the *Summary for Decision-Makers*, while implied by the title to be a synopsis of the main report, it is not truly representative of the main report. In many cases, the *Summary for Decision-Makers* compounds the limitations of the main report by making statements without supporting references and providing more definitive conclusions. Consequently, this companion report cannot be considered a Summary nor should it be relied onto make decisions regarding the regulation of endocrine disruptors.

Overall, the WHO-UNEP 2012 report on endocrine disruptors fails to achieve its objectives as an updated state-of-the-science review on endocrine disrupting chemicals, and therefore, should not be used to support evidence-based decisions. The scientific literature on endocrine disruption is voluminous, complex, heavily nuanced and covers multiple disciplines, and has certainly expanded greatly since the WHO-IPCS 2002 report – thus, warranting a true update of the state of the science. Science and policy decisions should not rely on professional opinion alone; rather, efforts to advance our understanding of the potential impacts of endocrine disrupting chemicals on human health and wildlife need to be based on objective and systematic reviews that transparently capture the best available science and rely on explicit criteria for the evaluation of the evidence. Therefore, a balanced and objective review of all of the available literature with clearly stated objectives and limitations is essential. Moreover, any updated review of the literature must acknowledge controversies regarding the issues, identify data gaps, and provide clear statements of conclusions that are supported by the data. Only through such an approach can all parties, regardless of position on the issues, be confident in evidence-based decisions that arise from such a report.

Conflict of interest

The employment affiliations of the authors are as shown on the cover page. James Lamb, Karyn Hentz, Jane Staveley, Gerard Swaen, and Amy Williams are employees of Exponent. Exponent is a publically traded engineering and scientific consulting firm that provides expertise in toxicology, ecological toxicology, and epidemiology. James Lamb has been a testifying expert in litigation involving several chemicals discussed in this review, including pesticides, PCBs and dioxins. Lorenz Rhomberg and Julie Goodman are employees of Gradient, a private consulting firm that provides services to both private and public organizations on toxicological and human health risk assessment issues. Paolo Boffetta is on the faculty of the Icahn School of Medicine at Mount Sinai and, as an independent consultant, has provided technical advice to industry and government on cancer epidemiology and has been a testifying expert in litigation involving PCBs. Warren Foster is on the faculty of McMaster University and, as an independent consultant, has provided technical advice to industry and government organizations on the health impacts associated with exposure to environmental chemicals. Glen Van Der Kraak has served as a consultant to industry regarding endocrine-related effects on wildlife. This review has been conducted with funding support from several sponsors: American Chemistry Council (ACC), CropLife America (CLA), CropLife Canada (CLS), CropLife International (CLI), European Chemical Industry Council (Cefic), and European Crop Protection Association (ECPA). These sponsors were provided an opportunity to review a draft of the paper and offer comments for consideration by the authors. The authors have sole responsibility for the content and the writing of the paper. The interpretations and views expressed in the paper are not necessarily those of the sponsors, or the authors' employers or clients. Both Glen Van Der Kraak and Warren Foster were members of the Steering committee for the WHO-IPCS 2002 Global Assessment of the State of the Science of Endocrine Disruptors; Glen Van Der Kraak also served as one of the four editors for this report.

References

Ahmed, O.M., El-Gareib, A.W., El-Bakry, A.M., Abd El-Tawab, S.M., Ahmed, R.G., 2008. Thyroid hormones states and brain development interactions. Int. J. Dev. Neurosci. 26, 147–209.

Alavanja, M.C., Samanic, C., Dosemeci, M., Lubin, J., Tarone, R., Lynch, C.F., Knott, C., Thomas, K., Hoppin, J.A., Barker, J., Coble, J., Sandler, D.P., Blair, A., 2003. Use of agricultural pesticides and prostate cancer risk in the Agricultural Health Study cohort. Am. J. Epidemiol. 157, 800–814.

Aronson, K.J., Wilson, J.W., Hamel, M., Diarsvitri, W., Fan, W., Woolcott, C., Heaton, J.P., Nickel, J.C., Macneily, A., Morales, A., 2010. Plasma organochlorine levels and prostate cancer risk. J. Expo. Sci. Environ. Epidemiol. 20, 434–445.

Auger, J., Jouannet, P., 1997. Evidence for regional differences of semen quality among fertile French men. Fédération Française des Centres d'Etude et de Conservation des Oeufs et du Sperme humains. Hum. Reprod. 12, 740–745.

Auger, J., Kunstmann, J.M., Czyglik, F., Jouannet, P., 1995. Decline in semen quality among fertile men in Paris during the past 20 years. N. Engl. J. Med. 332, 281–285.

Axelsson, J., Rylander, L., Rignell-Hydbom, A., Giwercman, A., 2011. No secular trend over the last decade in sperm counts among Swedish men from the general population. Hum. Reprod. 26, 1012–1016.

Bastomsky, C.H., 1974. Effects of a polychlorinated biphenyl mixture (Aroclor 1254) and DDT on biliary thyroxine excretion in rats. Endocrinology 95, 1150–1155.

Benbrahim-Tallaa, L., Waalkes, M.P., 2008. Inorganic arsenic and human prostate cancer. Environ. Health Perspect. 116 (2), 158–164.

Bianco, A.C., Kim, B.W., 2006. Deiodinases: implications of the local control of thyroid hormone action. J. Clin. Invest. 116 (10), 2571–2579.

Bonde, J.T., Ramlau-Hansen, C.H., Olsen, J., 2011. Trends in sperm counts: the saga continues. Epidemiology 22, 617–619.

Borgert, C.J., Baker, S.P., Matthews, J.C., 2013. Potency matters: thresholds govern endocrine activity. Regul. Toxicol. Pharmacol. 67, 83–88.

Braun, J.M., Yolton, K., Dietrich, K.N., Hornung, R., Ye, X., Calafat, A.M., Lanphear, B.P., 2009. Prenatal bisphenol A exposure and early childhood behavior. Environ. Health Perspect. 117, 1945–1952.

Brunet-Tran, K.L., Rier, S.E., Eisenberg, E., Osteen, K.G., 1999. The potential role of environmental toxins in the pathophysiology of endometriosis. Gynecol. Obstet. Invest. 1, 45–56.

Bryan, G.W., 1968. The effect of oil-spill removers ('detergents') on the gastropod *Nucella lapillus* on a rocky shore and in the laboratory. J. Mar. Biol. Ass. UK 48, 1067–1092.

Cakmak, H., Taylor, H.S., 2010. Molecular mechanisms of treatment resistance in endometriosis: the role of progesterone-hox gene interactions. Semin. Reprod. Med. 28, 68–74.

Carlsen, E., Giwercman, A., Keiding, N., Skakkebaek, N.E., 1992. Evidence for decreasing quality of semen during past 50 years. BMJ 305, 609–613.

Centers for Disease Control and Prevention (CDC), 2012. Autism and Developmental Disabilities Monitoring Network surveillance year 2008 principal investigators. Prevalence of autism spectrum disorders — Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. Morb. Mortal. Wkly. Rep. Surveill. Summ. 61, 1–19.

Černá, M., Krsková, A., Cejchanová, M., Spěváčková, V., 2012. Human biomonitoring in the Czech Republic: an overview. Int. J. Hyg. Environ. Health 215, 109–119.

Clark, E.S., Cowan, D.F., Pfeifer, D.C., 2008. Morphological changes in the Atlantic Bottlenose Dolphin (*Tursiops truncatus*) adrenal gland associated with chronic stress. J. Comp. Pathol. 138 (4), 208–216.

Conrad Jr., J.W., Becker, R.A., 2011a. Enhancing credibility of chemical safety studies: emerging consensus on key assessment criteria. Environ. Health Perspect. 119, 737–743.

Conrad Jr., J.W., Becker, R.A., 2011b. Chemical safety studies: Conrad and Becker respond. Environ. Health Perspect. 119, a508–a509.

Costello, M.F., Sjoblom, P., Haddad, Y., Steigrad, S.J., Bosch, E.G., 2002. No decline in semen quality among potential sperm donors in Sydney, Australia, between 1983 and 2001. J. Assist. Reprod. Genet. 19, 284–290.

Dietrich, R., von Aulock, S., Marquardt, H., Blaauboer, B., Dekant, W., Hengstler, J., Kehrer, J., Collier, A., Batta Gori, G., Pelkonen, O., Nijkamp, F.P., Lang, F., Stemmer, K., Li, A., Savolainen, K., Hayes, A.W., Gooderham, N., Harvey, A., 2013. Scientifically unfounded precaution drives European Commission's recommendations on EDC regulation, while defying common sense, well-established science and risk assessment principles. Toxicol. in Vitro (13), 00166-5, pii: S0887-2333.

Danish Centre on Endocrine Disrupters (DTU Food), 2013. Input for the REACH-review in 2013 on endocrine disrupters (tærskelværdi-projekt, j.nr. MST-621-00050). Final report. Technical University of Denmark (DTU), National Institute of Food, Division of Toxicology and Risk Assessment.

Du Preez, L., Kunene, N., Hanner, R., Giesy, J., Solomon, K., Hosmer, A., Van Der Kraak, G., 2009. Population-specific incidence of testicular ovarian follicles in *Xenopus laevis* from South Africa: a potential issue in endocrine testing. Aquat. Toxicol. 95, 10–16.

Duchan, E., Patel, D.R., 2012. Epidemiology of autism spectrum disorders. Pediatr. Clin. North Am. 59, 27–43.

European Food Safety Authority (EFSA), 2010. Scientific report of the endocrine active substances task force. EFSA J. 8, 1932.

European Food Safety Authority (EFSA), 2013a. Scientific opinion on the hazard assessment of endocrine disruptors: scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment. EFSA J. 11, 3132.

European Food Safety Authority (EFSA), 2013b. DRAFT scientific opinion on the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs – part: exposure assessment. *EFSA J.*

Elia, J., Imbrogno, N., Delfino, M., Rossi, T., Mazzilli, R., Nofroni, I., Toscano, V., Mazzilli, F., 2012. Comparative study of seminal parameters between samples collected in 1992 and samples collected in 2010. Arch. Ital. Urol. Androl. 84, 26–31.

Elsabbagh, M., Divan, G., Koh, Y.J., Kim, Y.S., Kauchali, S., Marcin, C., Montiel-Nava, C., Patel, V., Paula, C.S., Wang, C., Yasamy, M.T., Fombonne, E., 2012. Global prevalence of autism and other pervasive developmental disorders. Autism Res. 5, 160–179.

Ferriby, L.L., Knutsen, J.S., Harris, M., Unice, K.M., Scott, P., Nony, P., Haws, L.C., Paustenbach, D., 2007. Evaluation of PCDD/F and dioxin-like PCB serum concentration data from the 2001–2002 National Health and Nutrition Examination Survey of the United States population. J. Expo. Sci. Environ. Epidemiol. 17 (4), 358–371.

Fisch, H., Braun, S.R., 2013. Trends in global semen parameter values. Asian J. Androl. 15, 169–173.

Fisch, H., Hyun, G., Hensle, T.W., 2010. Rising hypospadias rates: disproving a myth. J. Pediatr. Urol. 6, 37–39.

Fombonne, E., Quirke, S., Hagen, A., 2009. Prevalence and interpretation of recent trends in rates of pervasive developmental disorders. McGill J. Med. 12, 99–107.

Gaido, K.W., Leonard, L.S., Lovell, S., Gould, J.C., Babai, D., Portier, C.J., McDonnell, D.P., 1997. Evaluation of chemicals with endocrine modulating activity in a yeast-based steroid hormone receptor gene transcription assay. Toxicol. Appl. Pharmacol. 143 (1), 205–212.

Gentile, I., Zappulo, E., Militerni, R., Pascotto, A., Borgia, G., Bravaccio, C., 2013. Etiopathogenesis of autism spectrum disorders: fitting the pieces of the puzzle together. Med. Hypotheses 81, 26–35.

Gibbs, P.E., Green, J.C., Pascoe, P.C., 1999. A massive summer kill of the dog-whelk, *Nucella lapillus*, on the north Cornwall coast in 1995: Freak or forerunner? J. Mar. Biol. Ass. UK 79, 103–109.

Glass, A.G., Lacey Jr., J.V., Carreón, J.D., Hoover, R.N., 2007. Breast cancer incidence, 1980–2006: combined roles of menopausal hormone therapy, screening mammography, and estrogen receptor status. J. Natl. Cancer Inst. 99, 1152–1161.

Goldey, E.S., Kehn, L.S., Lau, C., Rehnberg, G.L., Crofton, K.M., 1995. Developmental exposure to polychlorinated-biphenyls (Aroclor-1254) reduced circulating thyroid-hormone concentrations and causes hearing deficits in rats. Toxicol. Appl. Pharmacol. 135, 77–88.

Goodman, J.E., McConnell, E.E., Sipes, I.G., Witorsch, R.J., Slayton, T.M., Yu, C.J., Lewis, A.S., Rhomberg, L.R., 2006. An updated weight of the evidence evaluation of reproductive and developmental effects of low doses of bisphenol A. Crit. Rev. Toxicol. 36, 387–457.

Goodman, J.E., Witorsch, R.J., McConnell, E.E., Sipes, I.G., Slayton, T.M., Yu, C.J., Franz, A.M., Rhomberg, L.R., 2009. Weight-of-evidence evaluation of reproductive and developmental effects of low doses of bisphenol A. Crit. Rev. Toxicol. 39, 1–75.

Goodman, J.E., Dodge, D.G., Bailey, L.A., 2010. A framework for assessing causality and adverse effects in humans with a case study of sulfur dioxide. Regul. Toxicol. Pharmacol. 58, 308–322.

GRADE, 2011. GRADE working group. Frequently asked questions. Available at: <http://www.gradeworkinggroup.org/FAQ/index.htm> (last accessed 08.11.13).

Gumlchar, V., Thorsen, P., Laurent, C., Cans, C., Bodeau, N., Cohen, D., 2012. Pre- peri- and neonatal risk factors for autism. Acta Obstet. Gynecol. Scand. 91, 287–300.

Gun, S.W., 2009. Epigenetics of endometriosis. Mol. Hum. Reprod. 15, 587–607.

Hammond, A.O., Wilde, N., Gibson, M., Parks, A., Carrell, D.T., Meikle, A.W., 2008. Male obesity and alteration in sperm parameters. Fertil. Steril. 90, 2221–2225.

Hardell, L., Andersson, S.O., Carlberg, M., Bohr, L., van Bavel, B., Lindström, G., Björnfoth, H., Ginman, C., 2006. Adipose tissue concentrations of persistent organic pollutants and the risk of prostate cancer. J. Occup. Environ. Med. 48, 700–707.

Harvey, P.W., Sutcliffe, C., 2010. Adrenocortical hypertrophy: establishing cause and toxicological significance. J. Appl. Toxicol. 30, 617–626.

Hayes, T., Haston, K., Tsui, M., Hoang, A., Haeffele, C., Vonk, A., 2003. Atrazine-induced hermaphroditism at 0.1 ppb in American leopard frogs (*Rana pipiens*): laboratory and field evidence. Environ. Health Perspect. 111, 568–575.

Helsinki Commission (HELCOM), 2012. Baltic sea environment fact sheets for 2012. Helsinki Commission, Baltic Marine Environment Protection Commission.

Hengstler, J.G., Foth, H., Gebel, T., Kramer, P.J., Lilienblum, W., Schweinfurth, H., Völkel, W., Wollin, K.M., Gundert-Remy, U., 2011. Critical evaluation of key evidence on the human health hazards of exposure to bisphenol A. Crit. Rev. Toxicol. 41 (4), 263–291.

Henny, C.J., Grove, R.A., Kaiser, J.L., Johnson, B.L., 2010. North American osprey populations and contaminants: historic and contemporary perspectives. J. Toxicol. Environ. Health B Crit. Rev. 13, 579–603.

Hill, A.B., 1965. The environment and disease: association or causation? Proc R Soc Med. 58, 295–300.

Howdeshell, K.L., 2002. A model of the development of the brain as a construct of the thyroid system. Environ. Health Persp. 100, 337–348.

International Agency for Research on Cancer (IARC), 2004. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 84: Some Drinking-water Disinfectants and Contaminants, including Arsenic. World Health Organization, Geneva, pp. 41–287.

Institute of Medicine (IOM), 2012. Breast Cancer and the Environment: A Life Course Approach. Committee on Breast Cancer and the Environment: The Scientific Evidence, Research Methodology, and Future Directions. The National Academies Press, Washington, DC.

International Programme on Chemical Safety (IPCS), 2004. IPCS risk assessment terminology. Harmonization Project Document No. 1., World Health Organization, Geneva.

Itoh, H., Iwasaki, M., Hanaoka, T., Sasaki, H., Tanaka, T., Tsugane, S., 2009. Urinary phthalate monoesters and endometriosis in infertile Japanese women. Sci. Total Environ. 408 (1), 37–42.

Jørgensen, N., Andersen, A.G., Eustache, F., Irvine, D.S., Suominen, J., Petersen, J.H., Andersen, A.N., Auger, J., Cawood, E.H., Horte, A., Jensen, T.K., Jouannet, P., Keiding, N., Vierula, M., Toppari, J., Skakkebaek, N.E., 2001. Regional differences in semen quality in Europe. Hum. Reprod. 16 (5), 1012–1019.

Jørgensen, N., Vierula, M., Jacobsen, R., Pukkala, E., Perheentupa, A., Virtanen, H.E., Skakkebaek, N.E., Toppari, J., 2011. Recent adverse trends in semen quality and testis cancer incidence among Finnish men. Int. J. Androl. 34 (4 Pt 2), e37–48.

Jørgensen, N., Joensen, U.N., Jensen, T.K., Jensen, M.B., Almstrup, K., Olesen, I.A., Juul, A., Andersson, A.M., Carlsen, E., Petersen, J.H., Toppari, J., Skakkebæk, N.E., 2012. Human semen quality in the new millennium: a prospective cross-sectional population-based study of 4867 men. BMJ Open 2 (4).

Jörundsdóttir, K., Svavarsson, K., Leung, K.M.Y., 2005. Imposex levels in the dog whelk *Nucella lapillus* (L.) – continuing improvement at high latitudes. Mar. Pollut. Bull. 51 (8–12), 744–749.

Kabbarah, O., Sotelo, A.K., Mallon, M.A., Winkeler, E.L., Fan, M.Y., Pfeifer, J.D., Shibata, D., Gutmann, D.H., Goodfellow, P.J., 2005. Diethylstilbestrol effects and lymphomagenesis in Mlh1-deficient mice. Int. J. Cancer 115 (6), 666–669.

Kato, Y., Ikushiro, S., Haraguchi, K., Yamazaki, T., Ito, Y., Suzuki, H., Kimura, R., Yamada, S., Inoue, T., Degawa, M., 2004. A possible mechanism for decrease in serum thyroxine level by polychlorinated biphenyls in Wistar and Gunn rats. Toxicol. Sci. 81 (2), 309–315.

Kato, Y., Haraguchi, K., Ito, Y., Fujii, A., Yamazaki, T., Endo, T., Kaga, N., Yamada, S., Degawa, M., 2010. Polychlorinated biphenyl-mediated decrease in serum thyroxine level in rodents. Drug Metab. Dispos. 38 (4), 697–704.

Keller, D.A., Juberg, D.R., Catlin, N., Farland, W.H., Hess, F.G., Wolf, D.C., Doerrer, N.G., 2012. Identification and characterization of adverse effects in 21st century toxicology. Toxicol. Sci. 126 (2), 291–297.

Kinast, K., Peeters, D., Kolk, S.M., Schubert, D., Homberg, J.R., 2013. Genetic and pharmacological manipulations of the serotonergic system in early life: neurodevelopmental underpinnings of autism-related behavior. Front. Cell. Neurosci. 7, 72.

Klimisch, H.J., Andreae, M., Tillmann, U., 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul. Toxicol. Pharmacol. 25 (1), 1–5.

Kloas, W., Lutz, I., Springer, T., Krueger, H., Wolf, J., Holden, L., Hosmer, A., 2009. Does atrazine influence larval development and sexual differentiation in *Xenopus laevis*? Toxicol. Sci. 107 (2), 376–384. http://dx.doi.org/10.1093/toxsci/kfn232.

Kollara, A., Brown, T.J., 2006. Functional interaction of nuclear receptor coactivator 4 with aryl hydrocarbon receptor. Biochem. Biophys. Res. Commun. 346 (2), 526–534.

Koutros, S., Alavanja, M.C., Lubin, J.H., Sandler, D.P., Hoppin, J.A., Lynch, C.F., Knott, C., Blair, A., Freeman, L.E., 2010. An update of cancer incidence in the Agricultural Health Study. J. Occup. Environ. Med. 52 (11), 1098–1105.

Kripke, C., 2005. Lower- vs higher-dose estrogen for contraception. Am. Fam. Physician 72 (7), 1224–1225.

Kuiken, T., Hofle, U., Bennett, P.M.T., Allchin, C.R., Kirkwood, J.K., Baker, J.R., Appleby, E.C., Lockyer, C.H., Walton, M.J., Sheldrick, M.C., 1993. Adrenocortical hyperplasia, disease and chlorinated hydrocarbons in the harbour porpoise (*Phocoena phocoena*). Mar. Pollut. Bull. 26 (8), 400–406.

Lair, S., Béland, P., De Guise, S., Martineau, D., 1997. Adrenal hyperplastic and degenerative changes in beluga whales. J. Wildl. Dis. 33 (3), 430–437.

Leclère, B., Molinié, F., Tyrtaire, B., Stracci, F., Daubisse-Marliac, L., Colonna, M., 2013. Trends in incidence of breast cancer among women under 40 in seven European countries: a GRELL cooperative study. Cancer Epidemiol. 37 (5), 544–549.

Lind, P.M., Bergman, A., Olsson, M., Orberg, J., 2003. Bone mineral density in male Baltic grey seal (*Halichoerus grypus*). AMBIO 32 (6), 385–388.

Marimuthu, P., Kapilashrami, M.C., Misro, M.M., Singh, G., 2003. Evaluation of trend in semen analysis for 11 years in subjects attending a fertility clinic in India. Asian J. Androl. 5 (3), 221–225.

Martin, L., Klaassen, C.D., 2010. Differential effects of polychlorinated biphenyl congeners on serum thyroid hormone levels in rats. Toxicol. Sci. 117 (1), 36–44.

Mayes, B.A., McConnell, E.E., Neal, B.H., Brunner, M.J., Hamilton, S.B., Sullivan, T.M., Peters, A.C., Ryan, M.J., Toft, J.D., Singer, A.W., Brown Jr., J.F., Mennon, R.G., Moore, J.A., 1998. Comparative carcinogenicity in Sprague–Dawley rats of the polychlorinated biphenyl mixtures Aroclors 1016, 1242, 1254, and 1260. Toxicol. Sci. 41 (1), 62–76.

McDaniel, T.V., Martin, P.A., Struger, J., Sherry, J., Marvin, C.H., McMaster, M.E., Clarence, S., Tetreault, G., 2008. Potential endocrine disruption of sexual development in free ranging male northern leopard frogs (*Rana pipiens*) and green frogs (*Rana clamitans*) from areas of intensive row crop agriculture. Aquat. Toxicol. 88 (4), 230–242.

Morton, B., 2009. Recovery from imposex by a population of the dogwhelk, *Nucella lapillus* (Gastropoda: Caenogastropoda), on the southeastern coast of England since May 2004: a 52-month study. Mar. Pollut. Bull. 58 (10), 1530–1538.

Murphy, M.B., Hecker, M., Coady, K.K., Tompsett, A.R., Du Preez, L.H., Everson, G.J., Solomon, K.R., Carr, J.A., Smith, E.E., Kendall, R.J., Van Der Kraak, G.J., Giesy, J.P., 2006. Atrazine concentrations, gonadal gross morphology and histology in *ranid* frogs collected in Michigan agricultural areas. Aquat. Toxicol. 76, 230–245.

Nakamura, Y., Tsuji, S., Tonegai, Y., 2000. Determination of the levels of isoflavonoids in soybeans and soy-derived foods and estimation of isoflavonoids in the Japanese daily intake. J. AOAC Int. 83 (3), 635–650.

Nakatani, T., Yamamoto, A., Ogaki, S., 2011. A survey of dietary intake of polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans, and dioxin-like coplanar polychlorinated biphenyls from food during 2000–2002 in Osaka City, Japan. Arch. Environ. Contam. Toxicol. 60 (3), 543–555.

National Cancer Institute (NCI), 2013. Prostate Cancer Prevention (PDQ®). Health Professional Version. <http://www.cancer.gov/cancertopics/pdq/prevention/prostate/healthprofessional/page3/AllPages> (accessed 08.11.13).

Nelson, C.M., Bunge, R.G., 1974. Semen analysis: evidence for changing parameters of male fertility potential. Fertil. Steril. 25 (6), 503–507.

Nilsson, R., 2000. Endocrine modulators in the food chain and environment. Toxicol. Pathol. 28 (3), 420–431.

Nohynek, G.J., Borgert, C.J., Dietrich, D., Rozman, K.K., 2013. Endocrine disruption: fact or urban legend? Toxicol. Lett. (13), 01365–01369. pii: S0378-4274.

Nost, T.H., Breivik, K., Fuskevag, O.M., Nieboer, E., Odland, J.O., Sandanger, T.M., 2013. Persistent organic pollutants in Norwegian men from 1979 to 2007: intraindividual changes, age-period-cohort effects, and model predictions. Environ. Health Perspect. 121 (11-12), 1292–1298.

National Research Council (NRC), 2005. Health Implications of perchlorate ingestion. Committee to Assess the Health Implications of Perchlorate Ingestion. The National Academies Press, Washington, DC.

National Toxicology Program (NTP), 2006. Toxicology and carcinogenesis studies of 3,3′,4,4′,5-pentachlorobiphenyl (PCB 126) (CAS No. 57465-28-8) in female Harlan Sprague-Dawley rats (gavage studies). Natl Toxicol Program Tech Rep Ser., May;(529):4-168.

National Toxicology Program (NTP), 2010. Toxicology and carcinogenesis studies of 2,3′,4,4′5-pentachlorobiphenyl (PCB 118) (CAS No. 31508-00-6) in female Harlan Sprague-Dawley rats (gavage studies) Natl Toxicol Program Tech Rep Ser., Nov;(559):1-174.

Organization for Economic Cooperation and Development (OECD), 2005. Manual for investigation of HPV chemicals. Chapter 3: data evaluation. OECD Cooperative Chemicals Assessment Programme (CoCAP).

Pal, P.C., Rajalakshmi, M., Manocha, M., Sharma, R.S., Mittal, S., Rao, D.N., 2006. Semen quality and sperm functional parameters in fertile Indian men. Andrologia 38 (1), 20–25.

Rhomberg, L.R., Goodman, J.E., 2012. Low-dose effects and nonmonotonic dose-responses of endocrine disrupting chemicals: Has the case been made? Regul. Toxicol. Pharmacol. 64, 130–133.

Rhomberg, L.R., Goodman, J.E., Foster, W.G., Borgert, C.J., Van Der Kraak, G., 2012. A critique of the European Commission document, "State of the Art Assessment of Endocrine Disruptors". Crit. Rev. Toxicol. 42 (6), 465–473.

Rice, C.D., 2007. Prevalence of autism spectrum disorders–Autism and Developmental Disabilities Monitoring Network, six sites, United States, 2000. Morb. Mortal. Wkly. Rep. Surveill. Summ. 56 (1), 1–11.

Rice, C.E., Rosanoff, M., Dawson, G., Durkin, M.S., Croen, L.A., Singer, A., Yeargin-Allsopp, M., 2012. Evaluating changes in the prevalence of the autism spectrum disorders (ASDs). Public Health Rev. 34 (2), 1–22.

Rier, S.E., Turner, W.E., Martin, D.C., Morris, R., Lucier, G.W., Clark, G.C., 2001. Serum levels of TCDD and dioxin-like chemicals in Rhesus monkeys chronically exposed to dioxin: correlation of increased serum PCB levels with endometriosis. Toxicol. Sci. 59 (1), 147–159.

Ritchie, J.M., Vial, S.L., Fuortes, L.J., Guo, H., Reedy, V.E., Smith, E.M., 2003. Organochlorines and risk of prostate cancer. J. Occup. Environ. Med. 45 (7), 697–702.

Robertson, A., 1981. Effects of a toxic bloom of *Chrysochromulina polylepis* on the common dog-whelk, *Nucella lapillus* on the Swedish west coast. J. Mar. Biol. Assoc. UK 71, 569–578.

Rolland, M., Le Moal, J., Wagner, V., Royère, D., De Mouzon, J., 2013. Decline in semen concentration and morphology in a sample of 26,609 men close to general population between 1989 and 2005 in France. Hum. Reprod. 28 (2), 462–470.

Sawada, N., Iwasaki, M., Inoue, M., Itoh, H., Sasazuki, S., Yamaji, T., Shimazu, T., Tsugane, S., 2010. Plasma organochlorines and subsequent risk of prostate cancer in Japanese men: a nested case-control study. Environ. Health Perspect. 118 (5), 659–665.

Schmidt, R.J., Hansen, R.L., Hartiala, J., Allayee, H., Schmidt, L.C., Tancredi, D.J., Tassone, F., Hertz-Picciotto, I., 2011. Prenatal vitamins, one-carbon metabolism gene variants, and risk for autism. Epidemiology 22 (4), 476–485.

Schuhmacher-Wolz, U., Dieter, H.H., Klein, D., Schneider, K., 2009. Oral exposure to inorganic arsenic: evaluation of its carcinogenic and non-carcinogenic effects. Crit. Rev. Toxicol. 39, 271–298.

Séradour, B., Allemand, H., Weill, A., Ricordeau, P., 2009. Changes by age in breast cancer incidence, mammography screening and hormone therapy use in France from 2000 to 2006. Bull. Cancer 96 (4), E1–E6.

Sharpe, R.M., 2003. The 'oestrogen hypothesis' – where do we stand now? Int. J. Androl. 26, 2–15.

Sharpe, R.M., 2010. Environmental/lifestyle effects on spermatogenesis. Philos. Trans. R. Soc. London, Ser. B. 365 (1546), 1697–1712.

Sirot, V., Tard, A., Venisseau, A., Brosseaud, A., Marchand, P., Le Bizec, B., Leblanc, J.C., 2012. Dietary exposure to polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans and polychlorinated biphenyls of the French population: Results of the second French Total Diet Study. Chemosphere 88 (4), 492–500.

Skelly, D.K., Bolden, S.R., Dion, K.B., 2010. Intersex frogs concentrated in suburban and urban landscapes. EcoHealth 7, 374–379.

Smith, E.E., Du Preez, L.H., Gentles, B.A., Solomon, K.R., Tandler, B., Carr, J.A., Van Der Kraak, G.J., Kendall, R.J., Giesy, J.P., Gross, T.S., 2005. Assessment of laryngeal muscle and testicular cell types in *Xenopus laevis* (Anura: Pipidae) inhabiting maize and non-maize growing areas of South Africa. Afr. J. Herpetol. 54, 69–76.

Soto, A.M., Lin, T.M., Sakabe, K., Olea, N., Damassa, D.A., Sonnenschein, C., 1995. Variants of the human prostate LNCaP cell line as a tool to study discrete components of the androgen-mediated proliferative response. Oncol. Res. 7, 545–558.

Spolyarich, N., Hyne, R.V., Wilson, S.P., Palmer, C.G., Byrne, M., 2011. Morphological abnormalities in frogs from a rice-growing region in NSW, Australia, with investigations into pesticide exposure. Environ. Monit. Assess. 173, 397–407.

Stump, D.G., Beck, M.J., Radovsky, A., Garman, R.H., Freshwater, L., Sheets, L.P., Marty, M.S., Waechter, J.M., Dimond, S.S., Van Miller, J.P., Shiotsuka, R.N., Beyer, D., Chappelle, A.H., Hentges, S.G., 2010. Developmental neurotoxicity study of dietary bisphenol A in Sprague–Dawley rats. Toxicol. Sci. 115, 167–182.

Swan, S.H., Elkin, E.P., Fenster, L., 1997. Have sperm densities declined? A reanalysis of global trend data. Environ. Health Perspect. 105 (11), 1228–1232.

Tveguarden, J.G., Hanson-Drury, S., 2013. A systematic review of bisphenol A "low dose" studies in the context of human exposure: a case for establishing standards for reporting "low dose" effects of chemicals. Food Chem. Toxicol. (13), 00459-6. pii: S0278-6915.

Tinwell, H., Colombel, S., Blanck, O., Bars, R., 2013. The screening of everyday life chemicals in validated assays targeting the pituitary-gonadal axis. Regul. Toxicol. Pharmacol. 66 (2), 184–196.

Titley-O'Neal, C.P., Munkittrick, K.R., Macdonald, B.A., 2011. The effects of organotin on female gastropods. J. Environ. Monit. 13 (9), 2360–2388.

Tyl, R.W., Myers, C.B., Marr, M.C., Thomas, B.F., Keimowitz, A.R., Brine, D.R., Veselica, M.M., Fail, P.A., Chang, T.Y., Seely, J.C., Joiner, R.L., Butala, J.H., Dimond, S.S., Cagen, S.Z., Shiotsuka, R.N., Stropp, G.D., Waechter, J.M., 2002. Three-generation reproductive toxicity study of dietary bisphenol A in CD Sprague-Dawley rats. Toxicol. Sci. 68 (1), 121–146.

Tyl, R.W., Myers, C.B., Marr, M.C., Sloan, C.S., Castillo, N.P., Veselica, M.M., Seely, J.C., Dimond, S.S., Van Miller, J.P., Shiotsuka, R.N., Beyer, D., Hentges, S.G., Waechter Jr., J.M., 2008. Two-generation reproductive toxicity study of dietary bisphenol A in CD-1 (Swiss) mice. Toxicol. Sci. 104 (2), 362–384.

US Environmental Protection Agency (USEPA), 2013. State of the science evaluation: nonmonotonic dose responses as they apply to estrogen, androgen, and thyroid pathways and EPA testing and assessment procedures. Draft. Available at: <http://epa.gov/ncct/download_files/edr/NMDR.pdf> (last accessed 08.11.13).

US Environmental Protection Agency (USEPA), 2014. Vocabulary catalog list detail – integrated risk information system (IRIS) Glossary. Available at: <http://ofmpub.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=IRIS%20Glossary> (last accessed 24.01.14).

US Food and Drug Administration (USFDA), 2009. Memorandum to M. Hepp (Division of Food Contact Notifications, Regulatory Group 1) Re: Exposure to bisphenol A (BPA) for infants, toddlers and adults from the consumption of infant formula, toddler food and adult (canned) food. FDA-2010-N-0100-0001. Available at: <http://www.regulations.gov/#!documentDetail;D=FDA-2010-N-0100-0009> (last accessed 08.11.13).

van Maele-Fabry, G., Willems, J.L., 2004. Prostate cancer among pesticide applicators: a meta-analysis. Int. Arch. Occup. Environ. Health 77 (8), 559–570.

van Maele-Fabry, G., Libotte, V., Willems, J., Lison, D., 2006. Review and meta-analysis of risk estimates for prostate cancer in pesticide manufacturing workers. Cancer Causes Control 17 (4), 353–373.

Vandenberg, L.N., Colborn, T., Hayes, T.B., Heindel, J.J., Jacobs Jr., D.R., Lee, D.H., Shioda, T., Soto, A.M., vom Saal, F.S., Welshons, W.V., Zoeller, R.T., Myers, J.P., 2012. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. Endocr. Rev. 33 (3), 378–455.

Verhaegen, Y., Monteyne, E., Neudecker, T., Tulp, I., Smagghe, G., Cooreman, K., Roose, P., Parmentier, K., 2012. Organotins in North Sea brown shrimp (*Crangon crangon* L.) after implementation of the TBT ban. Chemosphere 86 (10), 979–984.

Wakai, K., Egami, I., Kato, K., Kawamura, T., Tamakoshi, A., Lin, Y., Nakamura, T., Wada, M., Ohno, Y., 1999. Dietary intake and sources of isoflavones among Japanese. Nutr. Cancer 33 (2), 139–145.

Weedon-Fekjær, H., Bakken, K., Vatten, L.J., Tretli, S., 2012. Understanding recent trends in incidence of invasive breast cancer in Norway: age-period-cohort analysis based on registry data on mammography screening and hormone treatment use. BMJ 30 (344), e299.

World Health Organization, International Programme on Chemical Safety (WHO-IPCS), 2002. In: Damstra, T., Barlow, S., Bergman, A., Kavlock, R., Van Der Kraak, G. (Eds.), Global Assessment of the State-of-the-Science of Endocrine Disrupters. World Health Organization, Geneva.

World Health Organization, 2012a. United Nations Environment Programme (WHO-UNEP), In: Bergman, A., Heindel, J.J., Jobling, S., Kidd, K.A., Zoeller, R.T. (Eds.), State of the Science of Endocrine Disrupting Chemicals. Available at: <http://www.who.int/ceh/publications/endocrine/en/index.html> (last accessed 08.11.13).

World Health Organization, 2012b. United Nations Environment Programme (WHO-UNEP), In: Bergman, A., Heindel, J.J., Jobling, S., Kidd, K.A., Zoeller, R.T. (Eds.), State of the Science of Endocrine Disrupting Chemicals 2012 Summary for Decision-Makers. Available at: <http://www.who.int/ceh/publications/endocrine/en/index.html> (last accessed 08.11.13).

Wilhelm, M., Ewers, U., Wittsiepe, J., Fürst, P., Hölzer, J., Eberwein, G., Angerer, J., Marczynski, B., Ranft, U., 2007. Human biomonitoring studies in North Rhine-Westphalia, Germany. Int. J. Hyg. Environ. Health 210 (3–4), 307–318.

Wing, L., Yeates, S.R., Brierley, L.M., Gould, J., 1976. The prevalence of early childhood autism: comparison of administrative and epidemiological studies. Psychol. Med. 6 (1), 89–100.

Wittassek, M., Wiesmüller, G.A., Koch, H.M., Eckard, R., Dobler, L., Müller, J., Angerer, J., Schlüter, C., 2007. Internal phthalate exposure over the last two decades – a retrospective human biomonitoring study. Int. J. Hyg. Environ. Health 210, 319–333.

Xu, X., Dailey, A.B., Talbott, E.O., Ilacqua, V.A., Kearney, G., Asal, N.R., 2010. Associations of serum concentrations of organochlorine pesticides with breast cancer and prostate cancer in U.S. adults. Environ. Health Perspect. 118 (1), 60–66.

Yolton, K., Xu, Y., Strauss, D., Altaye, M., Calafat, A.M., Khoury, J., 2011. Prenatal exposure to bisphenol A and phthalates and infant neurobehavior. Neurotoxicol. Teratol. 33, 558–566.

Younglai, E.V., Collins, J.A., Foster, W.G., 1998. Canadian semen quality: an analysis of sperm density among eleven academic fertility centers. Fertil. Steril. 70 (1), 76–80.

Zoeller, R.T., Dowling, A.L., Vas, A.A., 2000. Developmental exposure to polychlorinated biphenyls exerts thyroid hormone-like effects on the expression of RC3/neurogranin and myelin basic protein messenger ribonucleic acids in the developing rat brain. Endocrinology 141 (1), 181–189.

Phase 1 — High level summary

The high level summary will be developed based on the outline shared with you on March 17, 2013. This summary will be a broad-based response to the UNEP/WHO 2012 report and will focus on key, significant concerns. This document could be used to develop talking points for each of the industry associations. The sponsors have indicated that the intended audiences are policy makers.

The high level summary will cover these areas:

1. Selective citing of the literature
2. Lack of use of an appropriate Weight of Evidence framework for establishing cause and effect in relation to linking the chemical evidence to the endpoint
3. Lack of discussion around other causative factors – lifestyle, diet etc.
4. Weak evidence around whether or not the effects being cited are actually endocrine effects per se.
5. Lack of evidence / weak evidence from epidemiological studies – including consideration of other literature which take a bigger picture view – such as these trends over time may be driven by better identification/detection of these illnesses; that lifespans are actually improving; and that some of these trends need to be contextualized over a longer period, such as I.Q.

The draft report will be collaboratively prepared by Exponent and Gradient staff including: Julie Goodman, Karyn Hentz, Jim Lamb, Lorenz Rhomberg, and Amy Williams. This phase will require familiarization with the UNEP/WHO 2012 report, as well as the WHO 2002 report. Our preliminary estimate of the labor costs to prepare the high level summary is ~~$45,000.~~ The proposed dates for deliverables are outlined below. **[Note – we'd like to keep cost for this activity at around 20-25K]**

Phase 2 — A detailed review of the UNEP/WHO document

As requested, the UNEP/WHO 2012 State of the Science reports will be evaluated for completeness, objectivity, consistency and accuracy. We shall also determine whether the main report followed standard scientific methods for conducting weight-of-evidence (WoE) reviews, including the use of transparent procedures for literature reviews, consistent data evaluation procedures, and a transparent framework for integrating study results to arrive at determinations of associations and cause and effect conclusions. In addition, we will address the fact that the 2012 report appears to have abandoned of the WHO/IPCS weight-of-evidence methodology. Finally, we will review the evidence provided in the report on endocrine disruption to determine 1) whether it was scientifically justifiable for the authors to use their opinion or judgment in lieu of a formal WoE framework, and 2) the impact of use of opinion/judgment in lieu of a transparent WoE framework. The sponsors have indicated that the intended audiences are regulatory agencies, scientists in these agencies, non-regulatory scientific policy venues such as SAICM and general scientific and science policy communities.

Several allegations have been identified in these reports that will be addressed in our critical review and include:

- The lack of threshold for endocrine disrupting chemicals (EDC) and effects of EDC's below accepted thresholds for adverse effects;
- Commentary regarding the incidence rates of adverse health effects in humans and wildlife

and overview of other factors suggested to be involved their etiology;

- Commentary regarding the diet and lifestyle drivers for the adverse effects in humans and overview of other factors which may be involved in their development;
- Windows of developmental sensitivity where dose and potency are purportedly no longer critical for EDC's harmful effects;
- Non-monotonic dose response curves for EDC's; and
- The linkage between chemical exposure and adverse health effects in humans.

We will review all claims of this sort in the main report and evaluate the extent to which they are solidly based on a balanced and objective evaluation of available published literature. In particular, we will highlight points of speculation and conjecture that are not substantiated by the totality of the scientific evidence.

The draft detailed critical review will be collaboratively prepared by the Exponent and Gradient staff. This draft will then be shared with select Experts for review and comments as co-authors to the document. The following collaborators are being considered for inclusion in the critical review: Paolo Boffetta of Mount Sinai School of Medicine, NY; Chris Borgert; Neil Carmichael; Warren Foster of McMaster University; Richard Sharpe of University of Edinburgh; Gerard Swaen at Dow (Europe); and Glen VanDerKraak of the University of Guelph. Exponent and Gradient staff will incorporate comments and suggestions from the Experts to the extent feasible, but retain the final rights regarding the text of the report. The critical review will draw from work previously conducted by these experts, including the publications cited below. While the critical review will specifically address the issue raised in the WHO/UNEP report, much of the content of the critiecal review will be applicable to the broader issues and therefore will be exportable, with modification as appropriate, into other science and science policy venues. The draft will then be provided to the Sponsors for their review and comments before finalizing.

The preliminary cost estimate for labor by the Exponent and Gradient staff to prepare the detailed critical review is $150,000. In addition, we assume that a $4000-$5000 honorarium or fee would be appropriate for each of the Experts for an additional cost of $28,000 to $35,000 if all considered Experts participate. The total labor cost for this Phase is preliminarily estimated to be $178,000 to $185,000. We will coordinate with the Sponsors for the retrieval of literature that are not already available in the Exponent or Gradient libraries to control costs. The retrieval of literature by us will be passed on directly, without mark-up. The proposed dates for deliverables are outlined below.

Phase 3 — Will not be part of this contract at this time. Decision on Phase 3 will be deferred until completion of Phase 1 and Phase 2. Publication of a commentary in Regulatory Toxicology and Pharmacology

After the final report is completed, the report will be posted on Sponsor's web sites and distributed by Sponsors. A short commentary or editorial will be prepared in parallel for publication in the journal *Regulatory Toxicology and Pharmacology*, but will not be submitted to the Sponsor until the full critical review (Phase 2) is complete. The commentary/editorial would cite the full report and provide a URL to the report. The sponsors have indicated that the intended audiences are regulatory agencies, scientists in these agencies, non-regulatory scientific policy venues such as SAICM and

general scientific and science policy communities. Sponsors desire to have the overall findings of the in-depth review / critique become part of the peer reviewed scientific literature, so that the findings are more readily citable and have increased credibility.

The draft commentary/editorial will be collaboratively prepared by Exponent and Gradient staff and reviewed by the same Experts who co-authored the detailed critical review. It is estimated that the labor costs for preparation of the commentary/editorial will be $35,000. It is assumed that the honorarium or fees for Experts from Phase 2 will include the review of the manuscript. We will also arrange and pay for the editorial/commentary to be an open access publication, which will be billed as a direct cost without mark-up. The proposed dates for deliverables are outlined below.

Publications:
Lamb JL, Hentz KL. Chapter 3. Impact of endocrine disruptor policies and regulations. In: Endocrine Toxicology, Volume 27 in Target Organ Toxicology. Eldridge JS, Stevens JT (eds), New York: Informa Healthcare, 2010.

Lamb JC, Neal BH, Ginevan ME, Bus JS, Mahlburg WM. Herbicide effects on embryo implantation and litter size. Environ Health Perspect 2003; 111(9):A450.

Lamb JC. Current issues in toxicity testing. New endpoints in reproductive toxicology and their relevance to risk assessment. pp. p. 98–99. In: An Evaluation and Interpretation of Reproductive Endpoints for Human Health Risk Assessment. ILSI Press, 1998.

Lamb JC. Introduction to exposures to endocrine–active substances in the human diet. pp. 131–134. In: Human Diet and Endocrine Modulation, Estrogenic and Androgenic Effects. ILSI Press, 1998.

Lamb JC. New endpoints in reproductive toxicology and their relevance to risk assessment. Chapter III–Current Issues in Toxicity Testing. In: Evaluation and Interpretation of Reproductive Endpoints for Human Health Risk Assessment, 1998.

Lamb JC. Testing chemicals for endocrine–modulating activity and using the test data in risk assessments. pp. 220–228. In: Human Diet and Endocrine Modulation, Estrogenic and Androgenic Effects. ILSI Press, 1998.

Low-dose effects and nonmonotonic dose-responses of endocrine disrupting chemicals: has the case been made?
Rhomberg LR, Goodman JE. Regul Toxicol Pharmacol. 2012 Oct;64(1):130-3. doi: 10.1016/j.yrtph.2012.06.015. Epub 2012 Jun 27.

A critique of the European Commission document, "State of the Art Assessment of Endocrine Disrupters". Rhomberg LR, Goodman JE, Foster WG, Borgert CJ, Van Der Kraak G. Crit Rev Toxicol. 2012 Jul;42(6):465-73. doi: 10.3109/10408444.2012.690367. Epub 2012 May 26. Review.

+++
ENVIGO

SPL 0041-0216
24 July 2008

Envigo Research Limited
Shardlow Business Park
Shardlow
Derbyshire
DE72 2GD
UK

Summary Histopathology Data for 18 Month Control Study
Safepharm Laboratories Internal Project 0041-0216

The attached document was prepared to summarize the histopathology of neoplasms observed in an internal project (number 0041-0216) conducted by Safepharm Laboratories Ltd.

Safepharm Laboratories Ltd. was acquired initially by Harlan Laboratories Ltd. and subsequently Huntingdon Life Sciences. The company now operates as part of Envigo.

I can confirm that I was the Study Director for the project (number 0041-0216) and was responsible for the presentation of this summary.

Eric Wood
Manager Safety Assessment
Envigo

SPL 0041-0216
24 July 2008

Observations on the Development of Spontaneous Neoplasms in Male and in Female Crl: CD-1 (ICR) BR Strain Mice Following 18-Months on Control Diet.

Brooks PN, Dunster JS, and Wood E
Safepharm Laboratories Ltd, PO Box 45, Derby DE1 2BT
(Internal Publication)

SUMMARY

Male and female Crl: CD-1 (ICR) BR strain mice were maintained for a period of 18 months on standard laboratory diet at the Safepharm Laboratories Limited animal facility. Food and water were provided *ad libitum*. All mice found dead or killed *in extremis* during the course of the investigation together with those killed at termination were subjected to a full necropsy examination and histopathological evaluation. All neoplastic lesions were identified and a tabulation of spontaneous tumour incidences compiled.

METHODS

50 male and 50 female mice of the Crl: CD-1 (ICR) BR strain were supplied by Charles River (UK) Limited, Margate, Kent. Mice were aged 6-8 weeks at the start of the study and the weight variation did not exceed ± 20% of the mean weight for either sex. Animals were maintained at 21 ± 2° C with a target humidity of 55 ±15%. Lighting was 12 hours of continuous artificial light in each 24 hour period and there were at least 15 air changes per hour. Mice were housed in groups of no more than 3 by sex in polypropylene cages with wood flake bedding.

Rodent 5LF2 (Certified) diet (BCM IPS Limited, London, UK) with batch analysis, and tap water were provided *ad libitum*, and both the diet and the drinking water were routinely analysed to ensure the absence of any contaminant that could reasonably be anticipated to influence tumour development. Animals were provided with wooden chew blocks and cardboard tunnels both of which were routinely sampled and analysed for any contaminants having the potential to influence the purpose or integrity of the investigation. The study was conducted according to Safepharm policy on animal welfare and to the requirements of the United Kingdom's Animals (Scientific Procedures) Act 1986 in order to cause the minimum of suffering or distress to animals.

Mice were inspected twice daily, early and late during the working periods, for mortality or signs of ill-health and clinical signs were observed daily throughout the duration of the study. Detailed clinical examination was performed once a week and the findings recorded. From week 48 onward animals were palpated and the location, size and description of any masses recorded. Bodyweights were recorded on Day 1 of the study, at four-weekly intervals thereafter, and at termination.

A blood smear was prepared for all animals during the final week of the study (Week 78) and differential white cell counts performed. Post mortem studies were undertaken on all animal found dead or killed *in extremis* and on those killed at termination of the study. Mice killed *in extremis* and those killed at termination were euthanased by exsanguination following intravenous overdose of sodium pentobarbitone. A full external and internal necropsy examination was performed on all animals and a comprehensive list of organs and tissues taken or sampled into 10% neutral buffered formalin.

Tissue were routinely processed to paraffin wax, sectioned, and stained with haematoxylin and eosin. All tissues were examined microscopically and all neoplastic lesions reported. Non-neoplastic conditions were observed but not reported since they were not the primary objective of this investigation.

RESULTS

This investigation is concerned only with the development of neoplastic lesions. Non-neoplastic conditions were observed in most organs and tissues and these were entirely consistent with age-related degenerative or atrophic changes seen commonly in ageing or geriatric mice. Secondary changes were also seen in mice that developed severely debilitating neoplastic pathology. No unexpected non-neoplastic changes were encountered.

Of the 50 male and 50 female mice examined none was lost to autolysis or cannibalism. A total of 51 tumours was seen, 23 of these occurring in male mice and 28 in females. 12 (24%) male mice and 12 (24%) female mice died before scheduled termination. 20 (40%) male mice and 21 (42%) female mice had at least one neoplasm and 3 (6%) male mice and 6 (12%) female mice developed multiple neoplasms at different sites. 6 premature death male mice and 8 premature death female had tumours. 6 (12%) male mice and 6 (12%) female mice developed malignant lymphoma and one (2%) female mouse was diagnosed with myeloid leukaemia. Unusually, histiocytic sarcoma was not reported for any female mice in this study although this tumour type is relatively commonly seen in females, particularly in the uterus, of this mouse strain. No testicular or ovarian tumours were observed.

A breakdown of tumour incidence by type including all tumours encountered in the study is given in Table 1.

SUMMARY AND CONCLUSION:

Care needs to be exercised when interpreting the incidences of neoplastic lesions from mouse oncogenicity studies employing group size of 50 per sex because appreciable differences can be encountered in the prevalence of commonly reported tumours, even though a group size of 50 animals per sex is considered to provide a strong statistical basis for evaluation. The absence of a particular tumour type in control mice can suggest an effect of treatment when that tumour is present among high dose animals. However, the 'effect' in such cases can be a consequence of an absence among control animals and not a presence among treated mice. The compilation of background data on the prevalence of neoplastic conditions among control mice is thus of great value in assessing the true incidence of conditions, but the incidences between batches of animals, and thus the variability of distribution, is of as much, of not more, value than the overall background prevalence.

This investigation provided incidences of neoplastic lesions that developed in male and in female mice maintained on a standard rodent diet during a period of 18 months. Although the group size was relatively small many of the tumour types commonly encountered in laboratory mice were observed and these can be compared with incidences from other batches of mice maintained on the same diet for the same duration at the Safepharm animal facility.

Table 1

Incidence and % Incidence of Neoplastic Lesions by Tissue for Terminal Kill and Interim Death Animals Combined

Number of Mice	Males 50		Females 50	
CONDITION	**n**	**%**	**n**	**%**
ADRENAL GLAND				
Cortical adenoma b	4	8	0	0
BONE				
Osteoma b	0		1	2
HARDERIAN GLAND				
Adenoma b	4	8	2	4
GASTRO-INTESTINAL TUMOUR				
Stomach adenocarcinoma m	1	2	0	0
LIVER				
Hepatocellular carcinoma m	3	6	0	0
Haemangiosarcoma m	1	2	0	0
LUNG				
Adenoma b	0	0	1	2
Adenocarcinoma m	2	4	1	2
Combined	2	4	2	4
MAMMARY GLAND				
Adenocarcinoma m	0	0	3	6
Adenosquamous carcinoma m	0	0	1	2
PITUITARY				
Adenoma b	0	0	1	2
SKIN/SUBCUTIS				
Fibrosarcoma m	1	2	0	0
OVARY				
Luteoma b			1	2
UTERUS				
Endometrial stromal polyp b			4	8
Leiomyoma b			3	6
Stromal sarcoma m			2	4
Haemangiosarcoma m			1	2
THYROID				
Adenoma b	1	2	0	0
LYMPHOID/HAEMOPOIETIC				
Myeloid leukaemia m	0	0	1	2
Malignant lymphoma m	6	12	6	12
*Histiocytic sarcoma m	0	0	0	0
Combined	6	12	7	14
OVERALL TUMOUR INCIDENCE				
Primary benign tumours	9	18	10	20
Primary malignant tumours	14	28	14	28
Multiple benign tumours	0	0	0	0
Multiple malignant tumours	1	2	2	4
Multiple any tumours	3	6	4	8

* Histiocytic sarcomas are not generally regarded as lymphoid in origin but are usefully included here

Critical Reviews in Toxicology

http://informahealthcare.com/txc
ISSN: 1040-8444 (print), 1547-6898 (electronic)

Crit Rev Toxicol, 2015; 45(3): 185–208
 DOI: 10.3109/10408444.2014.1003423

informa healthcare

REVIEW ARTICLE

Evaluation of carcinogenic potential of the herbicide glyphosate, drawing on tumor incidence data from fourteen chronic/carcinogenicity rodent studies

Helmut Greim[1], David Saltmiras[2,6], Volker Mostert[4,5], and Christian Strupp[3,6]

[1]Technical University Munich, Arcisstr. 21, 80333 Munich, Germany, [2]Monsanto Company, 800 North Lindbergh Blvd., 63167 St. Louis, MO, USA, [3]ADAMA MAH BV Amsterdam NL Schaffhausen Branch, Spitalstrasse 5, 8200 Schaffhausen, Switzerland, [4]Knoell Consult GmbH, Dynamostr. 19, 68165 Mannheim, Germany, [5]Extera, Nelly-Sachs-Str. 37, 40764 Langenfeld, Germany, and [6]Glyphosate Task Force, http://www.glyphosatetaskforce.org/

Abstract

Glyphosate, an herbicidal derivative of the amino acid glycine, was introduced to agriculture in the 1970s. Glyphosate targets and blocks a plant metabolic pathway not found in animals, the shikimate pathway, required for the synthesis of aromatic amino acids in plants. After almost forty years of commercial use, and multiple regulatory approvals including toxicology evaluations, literature reviews, and numerous human health risk assessments, the clear and consistent conclusions are that glyphosate is of low toxicological concern, and no concerns exist with respect to glyphosate use and cancer in humans. This manuscript discusses the basis for these conclusions. Most toxicological studies informing regulatory evaluations are of commercial interest and are proprietary in nature. Given the widespread attention to this molecule, the authors gained access to carcinogenicity data submitted to regulatory agencies and present overviews of each study, followed by a weight of evidence evaluation of tumor incidence data. Fourteen carcinogenicity studies (nine rat and five mouse) are evaluated for their individual reliability, and select neoplasms are identified for further evaluation across the data base. The original tumor incidence data from study reports are presented in the online data supplement. There was no evidence of a carcinogenic effect related to glyphosate treatment. The lack of a plausible mechanism, along with published epidemiology studies, which fail to demonstrate clear, statistically significant, unbiased and non-confounded associations between glyphosate and cancer of any single etiology, and a compelling weight of evidence, support the conclusion that glyphosate does not present concern with respect to carcinogenic potential in humans.

Keywords

amino acid, carcinogenicity, epidemiology, glyphosate, herbicide, mouse, neoplasm, phosphonomethylglycine, Roundup, rat, regulatory, tumor

History

Received 6 November 2014
Revised 19 December 2014
Accepted 28 December 2014
Published online 24 February 2015

Table of Contents

Address for correspondence: David Saltmiras, Monsanto Company, 800 North Lindbergh Blvd., 63167 St. Louis, MO, USA. Tel: +1 (314) 684-8856. E-mail: david.a.saltmiras@monsanto.com

Introduction

Glyphosate (Figure 1), an aminophosphonic analog of the natural amino acid glycine, is widely used as an herbicide for the control of annual and perennial grasses and broad-leaved weeds. Glyphosate inhibits 5-enolpyruvateshikimate-3-phosphate synthase (EPSPS), an enzyme of the aromatic acid biosynthesis pathway, which is not present in the animal kingdom. Glyphosate-based herbicide formulations (GBFs) were introduced in 1974 and are formulated with

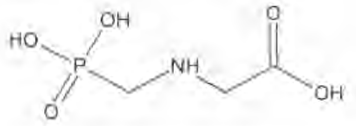


Figure 1. Structure of glyphosate acid.

sodium-, potassium-, ammonium- and isopropyl ammonium-salt forms of the active ingredient. The bulk-manufactured active herbicide glyphosate has the synonyms glyphosate technical acid, technical grade glyphosate and glyphosate acid.

The economic importance of glyphosate for growers is high. It has been estimated that a hypothetical ban of glyphosate would lead to decreases in the production of wheat, fodder, maize and oilseeds, by 4.3–7.1%, with the result of an estimated annual welfare loss of 1.4 billion USD to society in the European Union alone (Schmitz and Harvert 2012). Furthermore, glyphosate plays an important role in integrated pest management strategies, and affords the environmental benefit of substantially reduced soil erosion resulting from of no-till and reduced-till agriculture.

The long-term toxicity and carcinogenicity of glyphosate has been investigated by multiple entities including academia, registrants, and regulatory authorities, and the data generated have been evaluated in support of herbicide regulatory approvals in many world regions including the USA (US EPA 1993) and the European Union (EC 2002), and several scheduled reevaluations are currently ongoing in the USA, Canada, Japan and Europe (Germany Rapporteur Member State 2015a), with imminent conclusions.

Studies of appropriate scientific quality are the basis for regulatory decision making. Mandatory testing guidelines (TGs) exist for toxicological studies submitted for regulatory review of active substances for plant protection in many regions of the world. Such TGs have been released, *inter alia*, by the United States Environmental Protection Agency (US EPA 2012), the European Union (EU 2008), the Japanese Ministry of Agriculture, Forestry and Fisheries (JMAFF 2000), and the Organization of Economic Co-operation and Development (OECD 2012b). These TGs set quality standards for each type of study by giving guidance regarding test species, strains, and number of animals to be used, the choice of dosing, exposure duration, and parameters to be measured and observed, as well as for the reporting of results. Due to the lack of effective legal and regulatory provisions for the sharing of vertebrate study data in the past, and to guarantee the safety of technical glyphosate obtained from different processes of synthesis, several manufacturers of glyphosate had to initiate toxicological testing programs of their own. Occasionally, regulatory studies had to be repeated to reflect major changes in the underlying TG. In the case of glyphosate, this has given rise to a multitude of studies for the same toxicological endpoints, leading to the availability of an extraordinarily robust scientific study database that can be considered unique among pesticides, industrial chemicals, and pharmaceuticals. Such a remarkable volume of studies addressing the same endpoints, conducted over the last 40 years by several independent companies and laboratories while toxicology test guidelines have evolved, warrants investigation for consistency, reliability, and application to their intended purpose; identifying potential human health hazards and setting appropriate endpoints for human health risk assessment. Studies conducted with equivalent test substances using the same TG are readily comparable and can be evaluated by regulators following standardized schemes. Minor differences in the findings reported by such repetitive studies are attributable to statistical chance, natural biological variability, type of basal diet, rate of feed consumption, animal strain differences, choice of dose levels, inter-strain genetic drift over time due to varying vendor breeding practices, changes in animal care and husbandry practices across laboratories over the years, inter-laboratory variations in clinical measurements, and differences between individual pathologist evaluation and interpretation of tissue specimens.

Glyphosate is under significant political pressure due to its widespread use, particularly in association with use on genetically modified crops. One focus area of contention has been the human safety of glyphosate, which has been repeatedly challenged by interest groups via the media, as well as select research publications in the scientific literature (Antoniou et al. 2012, Aris and Leblanc 2011, Aris and Paris 2010, Benachour and Seralini 2009, Gasnier et al. 2010, Paganelli et al. 2010, Romano et al. 2012, Romano et al. 2010). To that end, one specific publication by Seralini et al. (2012, retracted) drew significant criticism from both the toxicology and broader scientific communities (Barale-Thomas 2013, Berry 2013, de Souza and Oda 2013, Grunewald and Bury 2013, Hammond et al. 2013, Langridge 2013, Le Tien and Le Huy 2013, Ollivier 2013, Panchin 2013, Sanders et al. 2013, Schorsch 2013, Tester 2013, Trewavas 2013, Tribe 2013). After a special review of the investigators' raw data by a mutually agreed-upon expert panel, the manuscript was retracted by *Food and Chemical Toxicology* (FCT), for reasons of inconclusive data and unreliable conclusions (Hayes 2014). The Editor of the *International Journal of Toxicology* highlighted this manuscript as an example of possible failure of the peer review process in a well-respected toxicology journal with an editorial board of well-known and respected toxicologists (Brock 2014). The manuscript was later republished without peer-review in an open access journal (Seralini et al. 2014), but will not be addressed in this data evaluation due to the inappropriate study design, insufficient reporting of tumor incidence data, and the lack of a data supplementary to the manuscript.

The chronic/carcinogenicity studies discussed in this paper have been submitted to and evaluated by a variety of agencies over time, including the World Health Organization (WHO/FAO 2004b, WHO/FAO 2004a), the United States Environmental Protection Agency (US EPA 1993), the European Rapporteur Member State Germany for the initial glyphosate Annex I listing (EC 2002) and the recent European re-evaluation (Germany Rapporteur Member State 2015a), as well as the ongoing reevaluations in the USA, Canada and Japan. These regulatory bodies, drawing upon internal and/or external expertise, have consistently concluded that glyphosate is devoid of carcinogenic risk to humans.

The purpose of this article is to provide the broader scientific community with insight into this large body of carcinogenicity data on glyphosate, originally generated for

DOI: 10.3109/10408444.2014.1003423

regulatory purposes. Each study discussed in this review has been assigned a reliability score in Tables 3–19, following the Klimisch scoring system (Klimisch et al. 1997). In this system, a score of 1 is assigned to studies that are fully reliable based on compliance with Good Laboratory Practice (GLP) and adherence to appropriate study guidelines. A score of 2 is appropriate if some guideline requirements are not met, but if these deficiencies do not negatively affect the validity of the study for its regulatory purpose. Studies with a reliability of 3 employ a test design that is not fit for the scientific purpose of the study, due to significant scientific flaws, or the objective of the study not covering the regulatory endpoints, or both. Such studies can provide supplemental information but do not allow a stand-alone appraisal of a regulatory endpoint. No studies were assigned a reliability of 4, since each report contained sufficient information to judge the validity of the study.

This manuscript presents the robust glyphosate carcinogenicity data generated by industry. Study summaries will focus on carcinogenicity evaluation, to allow third parties the opportunity to independently evaluate the carcinogenicity data presented alongside other relevant data on carcinogenicity, i.e. genotoxicity testing and epidemiology, and facilitate a multidisciplinary carcinogenicity assessment as proposed in the literature, by recognized experts in the fields of toxicology and human health risk assessment (Adami et al. 2011).

Absorption, distribution, metabolism and excretion of glyphosate

A number of absorption, distribution, metabolism, and excretion studies (ADME) have been conducted on glyphosate for evaluation in regulatory submissions (EC 2002, US EPA 1993, WHO/FAO 2004a) and also by academic institutions (Anadon et al. 2009). Glyphosate consistently demonstrates low gastrointestinal absorption (20–40%). Its metabolism is very limited, whereby only small quantities of a single metabolite, aminomethylphosphonic acid (AMPA), are eliminated in feces. AMPA is likely produced by the limited metabolism of glyphosate by the gastrointestinal microflora, rather than via mammalian metabolism. Glyphosate is structurally akin to a phase II metabolite, a glycine-conjugate of methyl phosphonate, and thus avails itself to rapid urinary excretion. Systemic elimination is biphasic, with alpha-phase half-lives in the range of 6–14 h (Anadon et al. 2009, WHO/FAO 2004a).

Toxicological properties of glyphosate

Table 1 contains a short overview of toxicological endpoints of glyphosate that have been published in the List of Endpoints identified for glyphosate by the Rapporteur in the European Union under Regulation 1107/2009 (Germany Rapporteur Member State 2015c). Glyphosate is of low acute toxicity via all routes of exposure. Glyphosate's active ingredient, an organic acid, has an irritating effect on mucosa which is evidenced by eye irritation and effects on oral and gastrointestinal mucosa; final formulated products contain more neutral pH salt forms, as reflected in the tabulated eye irritation data reported in Table 11, on page 109 of the 2004 JMPR Toxicological Evaluation (WHO/FAO 2004a). Glyphosate is not mutagenic, not neurotoxic, and has no effect on pre-natal development and fertility at doses not exceeding the maximum tolerated dose (MTD).

Genotoxicity

Very recently, a review of the vast body of genotoxicity studies on glyphosate and GBFs has been published (Kier and Kirkland 2013), including an online data supplement presenting detailed data from 66 separate *in vitro* and *in vivo* genotoxicity assays. The authors incorporated these studies and published genotoxicity data into a weight-of-evidence analysis. The vast majority (over 98%) of the available bacterial reversion and *in vivo* mammalian micronucleus and chromosomal aberration assays were negative. Negative results for *in vitro* gene mutation and a large majority of negative results for clastogenic effect assays in mammalian cells support the conclusion that glyphosate is not genotoxic for these endpoints in mammalian test systems. DNA damage effects are reported in some instances for glyphosate at high or toxic dose levels. The compelling weight of evidence is that glyphosate and typical GBFs are negative in core assays, indicating that the reported high-dose effects are secondary to toxicity and are not due to DNA-reactive mechanisms. Mixed results were observed for micronucleus assays in non-mammalian systems and DNA damage assays of GBFs. These effects of GBFs may also be associated with surfactants present in the formulated products. Kier and Kirkland conclude that glyphosate and its typical formulations do not present significant genotoxic risk under normal conditions of human or environmental exposures.

Epidemiology

Available epidemiological studies of glyphosate and cancer endpoints were recently reviewed (Mink et al. 2012). Seven cohort studies and fourteen case-control studies examining a potential association between glyphosate and one or more cancer outcomes were subjected to a qualitative analysis. The review found no consistent pattern of positive associations between total cancer (in adults or children) or any site-specific cancer, and exposure to glyphosate. A recent review article (Alavanja et al. 2013) cites one epidemiology study associating glyphosate use with non-Hodgkin's lymphoma (NHL), and accepts the study findings *prima facie*. However, Alavanja et al. (2013) did not highlight six other published epidemiology studies which evaluated glyphosate use and NHL, noting that any association between NHL and glyphosate use was null or not statistically significant. All seven studies were scrutinized by Mink et al. (2012). NHL is not a specific disease, as mentioned in both the epidemiology review publications above, but is rather multiple presentations of lymphoma which are simplistically classified as not being Hodgkin's lymphoma (HL). This dichotomous classification of HL/NHL was rejected by the World Health Organization in 2001, whereby 43 different lymphomas of various etiologies were precisely characterized (Berry 2010). The Bradford Hill criteria are often applied in efforts to determine whether an association between a health effect and human exposure may be deemed causal. However, an important premise often overlooked from Sir Austin Bradford Hill's famous speech of 1965, is that before applying these criteria, the observations should "reveal an association between two variables, perfectly clear-cut and beyond what we care to attribute to the play of chance" (Bradford Hill 1965). This predicate of the association being "perfectly clear-cut"

Table 1. Summary of toxicological endpoints for glyphosate (Germany Rapporteur Member State 2015c).

Endpoint	Value	Remark
Oral absorption	*ca* 20%	Rat, *in vivo*
Dermal absorption	< 1%	Human, *in vitro*; 0.015 g glyphosate/L
Rat LD50 oral	> 2000 mg/kg bw	
Rat LD50 dermal	> 2000 mg/kg bw	
Rat LC50 inhalation	> 5 mg/L	4-h exposure
Skin irritation	Not irritating	
Eye irritation	Acid: moderately to severely irritating Salts: slight or non-irritating	
Skin sensitization	Not sensitizing (LLNA, Magnusson-Kligmant, and Buehler test)	
Genotoxicity	Not genotoxic (*in vitro* and *in vivo*)	
Chronic toxicity	BW gain, liver (organ weight ↑, clinical chemistry, histology); salivary glands (organ weight ↑, histology); stomach mucosa and bladder epithelium(histology); eye (cataracts), caecum (distention, organ weight ↑) NOAEL = 100 mg/kg bw/day (2-yr rat)	Critical study used for ADI setting
Reproductive toxicity	Reduced pup weight at parentally toxic doses. NOAEL = 300 mg/kg bw/day	
Developmental toxicity	Post-implantation loss, fetal BW & ossification ↓; effects confined to maternally toxic doses Rat NOAEL: 300 mg/kg bw/day Rabbit NOAEL: 50 mg/kg bw/day	
Delayed neurotoxicity	No relevant effects, NOAEL: 2000 mg/kg bw/day	
Acceptable Daily Intake (ADI)	0.5 mg/kg bw/day Based on developmental toxicity in rabbits	Safety factor 100
Acceptable Operator Exposure Level (AOEL)	0.1 mg/kg bw/day Based on maternal toxicity in rabbit teratogenicity study	Safety factor 100 Corrected for oral absorption of 20%

was recently highlighted as requiring statistical significance, wherein the confidence interval of a relative risk ratio is bracketed above 1.0, as well as concluding that the association may not be attributable to bias, confounding or sampling error (Woodside and Davis 2013). According to Bradford Hill, should an epidemiology study be considered to demonstrate a "perfectly clear-cut" association between glyphosate exposure and a human health outcome, only then should the Bradford Hill criteria be investigated to determine whether there is causality. To date, no such "perfectly clear-cut" association between glyphosate exposure and any cancer exists. However, investigative toxicology is an important discipline to evaluate chemicals before any human exposure occurs, and these data may inform subsequent considerations of whether associations are attributable to causality. One Bradford Hill criterion in establishing disease causality is plausibility, based on known disease etiologies. In the case of lymphoma, there are numerous etiologies for the numerous and different lymphoma diseases, and as such, each lymphoma type should be investigated for a plausible mechanism to determine whether causality may be attributed an appropriately qualified association. Another Bradford Hill criterion is identification of a biological gradient, or dose-response, which is a key consideration in the following data evaluation.

Chronic toxicity studies

Several one-year chronic studies have been undertaken in dogs and one in rats, in addition to the many chronic/carcinogenicity studies with one-year interim sacrifice groups. Current Test Guidelines (OECD, EPA, EU and JMAFF) for long-term studies clearly state that the highest dose tested should either be at the maximum tolerated dose (MTD), conventionally interpreted as a dose causing non-lethal toxicity, often noted as reduced body weight gain of 10% or more (IUPAC 1997). For test substances with low toxicity, a top dose not exceeding 1000 mg/kg bw/day may apply, except when human exposure indicates the need for a higher dose level to be used (OECD 2012a). All human exposure estimates are well below 1 mg/kg bw/day (see Discussion section), so that 1000 mg/kg bw/day is a practical limit dose for glyphosate in carcinogenicity studies. In the original pre-guideline chronic/carcinogenicity study, rats were dosed well below the MTD (Monsanto 1981), but in many subsequent studies, they were dosed well in excess of today's standard practice of not exceeding the dose limit.

Dog chronic studies

Five one-year oral toxicity studies have been conducted in Beagle dogs (Table 2). Studies in dogs are not designed to detect neoplastic effects; these studies are therefore not discussed in detail. Nonetheless, the histopathological investigations that are part of one-year dog studies according to OECD TG 452 did not identify (pre) neoplastic lesions related to the administration of glyphosate.

Treatment-related effects in dog studies with glyphosate were restricted to non-specific findings like small retardations in body weight gain and soft stools, which are common findings in this test species. The lowest relevant NOAEL (i.e. highest NOAEL below the lowest LOAEL) in dogs on a daily treatment regimen for one year was 500 mg/kg bw/day. These studies demonstrate that glyphosate is of very low toxicity following repeat exposures in dogs.

Rat chronic studies

The chronic toxicity potential of glyphosate acid was assessed in a 12-month feeding study (conducted in 1995 and 1996) in

DOI 10.3109/10408444.2014.1003423

Table 2. Summary of one-year toxicity studies with glyphosate.

Authors:	Monsanto (1985)
Reliability/Justification	2 Study performed according to GLP and OECD guideline requirements, with the following deviation: MTD not reached by highest dose
Substance:	Glyphosate (96.1% pure)
Species/Strain:	Dog/Beagle, groups of 6 ♂ and 6 ♀
Administration route:	Oral, capsule
Doses:	0, 20, 100, 500 mg/kg bw/day
Duration:	1 year
Findings:	≥ 500 mg/kg bw/day: NOAEL (♂ + ♀) no treatment-related effects
Authors:	**Cheminova (1990)**
Reliability/Justification	1 Study performed according to GLP and OECD guideline requirements, with no deviations
Substance:	Glyphosate (98.6–99.5% pure)
Species/Strain	Dog/Beagle, groups of 4 ♂ and 4 ♀
Administration route:	Oral, capsule
Doses:	0, 30, 300, 1000 mg/kg bw/day
Duration:	1 year
Findings:	300 mg/kg bw/day: NOAEL (♂ + ♀) 1000 mg/kg bw/day: soft, liquid stools (attributable to capsule administration); equivocal impact on body weight gain
Authors:	**Nufarm (2007)**
Reliability/Justification	2 Study performed according to GLP and OECD guideline requirements, with the following deviation: MTD not reached by highest dose
Substance:	Glyphosate (95.7% pure)
Species/Strain	Dog/Beagle, groups of 4 ♂ and 4 ♀
Administration route:	Oral, capsule
Doses:	0, 30, 125, 500 mg/kg bw/day
Duration:	1 year
Findings:	≥ 500 mg/kg bw/day: NOAEL (♂ + ♀) No treatment-related effects
Authors:	**Arysta Life Sciences (1997c)**
Reliability/Justification	2 Study performed according to GLP and OECD guideline requirements, with the following deviation: MTD not reached by highest dose
Substance:	Glyphosate (94.6% pure)
Species/Strain	Dog/Beagle, groups of 4 ♂ and 4 ♀
Administration route:	Oral, diet
Concentration:	0, 1600, 8000, 50 000 ppm diet (♂ about 34.1, 182, 1203 mg/kg bw/day; ♀ about 37.1, 184, 1259 mg/kg bw/day)
Duration:	1 year
Findings:	182/184 mg/kg bw/day: NOAEL (♂/♀) At high dose: loose stool, non-statistically significant retarded body weight gain, decreased urinary pH, slight and non-statistically significant focal pneumonia (♀), minor clinical chemistry changes of Cl ↑, albumin ↓, P ↓ (♀)
Authors:	**Syngenta (1996a)**
Reliability/Justification	1 Study performed according to GLP and OECD guideline requirements, with no deviations.
Substance:	Glyphosate (95.6% pure)
Species/Strain	Dog/Beagle, groups of 4 ♂ and 4 ♀
Administration route:	Oral, diet
Concentration:	0, 3000, 15 000, 30 000 ppm diet (♂ about 90.9, 440, 907 mg/kg bw/day; ♀ about 92.1, 448, 926 mg/kg bw/day)
Duration:	1 year
Findings:	15 000 ppm diet: NOAEL (♀) ≥ 30 000 ppm diet: NOAEL (♂): No treatment-related effects 30 000 ppm diet: slight body weight reduction (♀)
Authors:	**Syngenta (1996b)**
Reliability/Justification	1 Study performed according to GLP and OECD guideline requirements, with no deviations.
Substance:	Glyphosate (95.6% pure)
Species/Strain	Rat/Wistar Alpk: AP$_f$SD, groups of 24 ♂ and 24 ♀
Administration route:	Oral, diet
Concentration:	0, 2000, 8000, 20 000 ppm diet (♂ about 141, 560, 1409 mg/kg bw/day; ♀ about 167, 671, 1664 mg/kg bw/day)
Duration:	1 year
Findings:	8000 ppm diet: NOAEL (♂+♀) 20 000 ppm diet: parotid salivary glands (focal basophilia of the acinar cells considered non-adverse adaptive response, ♂: 13/24, ♀: 15/24), body weight reduction

24 male and female Wistar rats per group, dosed at 0, 2000, 8000 and 20 000 ppm (Syngenta 1996). The mean achieved dose levels were 0, 141, 560 and 1409 mg/kg bw/day for males, and 0, 167, 671 and 1664 mg/kg bw/day for females. Spastically significant reductions in bodyweight were evident in animals receiving 20 000 ppm glyphosate acid, together with a marginal reduction in bodyweight in rats receiving 8000 ppm, but food consumption relative to controls was lower for these dose groups, suggesting reduced palatability of the diets containing these doses of glyphosate. There were no toxicologically significant or treatment-related effects on hematology, blood and urine clinical chemistry, or organ weights (Table 2).

The treatment-related pathological finding, that is increased incidence of mild focal basophilia, and a hypertrophy of the acinar cells of the parotid salivary gland in both sexes which had received 20 000 ppm glyphosate acid, is considered an adaptive response due to oral irritation from the ingestion of glyphosate, an organic acid, in the diet. This was verified by

mode of action investigations and studies with dietary administration of citric acid, a non-toxic organic acid with irritation properties and pH dilution curve similar to those of glyphosate (Saltmiras et al. 2011), which elicited the same response in the acinar cells of the parotid salivary glands.

In conclusion, the 12-month NOAEL in rats for glyphosate acid, as determined from this study, is 8000 ppm (corresponding to 560 mg/kg bw/day in males and 671 mg/kg bw/day in females). This study does not cover neoplastic endpoints. These were addressed in a subsequent study by the same sponsor (Syngenta 2001). Consistent with the findings observed in dogs, this study demonstrates that glyphosate is of very low toxicological concern following long-term daily exposures.

Similarly, most of the following 2-year rat carcinogenicity studies included additional groups for 1-year interim sacrifice to evaluate chronic toxicity. These studies did not elucidate significant toxicological concerns for chronic dietary exposures to glyphosate in rats in multiple expert reviews by governmental agencies and several technical branches of the World Health Organization including the Joint Meeting on Pesticide Residues Toxicological Evaluations (WHO/FAO 2004a).

Carcinogenicity studies

Chronic/carcinogenicity tests are designed to simulate lifetime exposures to an individual chemical and represent the most robust *in vivo* assay to evaluate the effects of chronic exposure including carcinogenicity. These models are biological systems with natural background variability due to tumor formation as a natural consequence of aging. Glyphosate was found to have no carcinogenic potential, which is reflected in the data showing only background noise of spontaneous tumors across the wide range of doses. Normal biological variability should display various tumor types across all dose groups without an apparent dose-response. The study summaries discuss "select neoplasms", identified by the authors as having an elevated incidence above concurrent controls across one or more dose groups, most of which lacked statistical significance and/or dose-response within an individual study. These tumors are then evaluated in the context of the whole data set, to provide a robust weight of evidence overview for the doses spanning several orders of magnitude. While not all studies have select neoplasms identified in the individual study summary tables, select neoplasms for all studies are reported in Tables 20–23. Summary tables of the select neoplasms footnote the strain tested for each dose, to allow consideration of strain differences in spontaneous tumor susceptibility (Tables 20–23). In addition, complete tumor incidence summary tables have been extracted from the original eight rat (the published rat study, Study 9, is not included) and five mouse study reports or study files, and posted in their original format, as a comprehensive online data supplement to this manuscript.

Rat carcinogenicity

A total of nine chronic/carcinogenicity studies in the rat, including one peer-reviewed published study, were available for review. This duplication of large-scale studies in the same animal model using the same test substance is not consistent with today's broader appreciation for animal welfare and the reduction of unnecessary animal testing. However, these studies offer the opportunity for a critical discussion of findings in individual studies in the context of the larger body of data. Wistar and Sprague Dawley were the strains used for the bioassays in rats. Seven studies were conducted under conditions of GLP, and two studies were not under GLP (Study 1, conducted before the introduction of GLP; Study 9, non-GLP). Most studies in rats were designed as combined chronic toxicity/carcinogenicity studies, with interim sacrifices after 12 months of treatment for the assessment of non-neoplastic chronic toxicity. Statistical methods are noted in the manuscript tables where statistical significance was attained. Statistical differences in neoplasm incidence summary tables are reported in the online data supplements. Chronic endpoints and NOAEL values are captured in each study summary table; however, the following study reviews focus on carcinogenicity.

Study 1 (Monsanto 1981)

An early study into the long-term effects of orally administered glyphosate in the rat was conducted between 1978 and 1980 (Monsanto 1981), prior to the adoption of international test guidelines and GLP standards (Tables 3–6). Nonetheless, the test protocol was broadly compliant with OECD TG 453 (1981). However, an MTD was not reached and the high dose was well below an acceptable dose limit of 1000 mg/kg bw/day. Therefore, this study is rated Klimisch 3 for reliability, and is considered inadequate for carcinogenicity evaluation from a regulatory perspective.

Groups of 50 male and 50 female Sprague Dawley rats were administered glyphosate acid in the diet, at concentrations of 0, 30, 100 and 300 ppm, for up to least 26 months. The mean doses achieved were 0 (control), 3, 10, and 31 mg/kg bw/day for the males, and 0 (control), 3, 11, and 34 mg/kg bw/day for the females. Study results are summarized in Table 3.

In general, the incidences of all neoplasms observed in the treated and control animals were similar, or occurred at low incidence, such that a treatment-related association could not be made. The most common tumors found were common spontaneous neoplasms, as reported in the literature relating to rat (Johnson and Gad 2008), in the pituitary glands of both control and treated animals (Table 4). In the females, mammary gland tumors were the next most common neoplasm across control and dose groups (see data Supplementary Study 1 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423).

Table 3. Study 1–26-month feeding study of glyphosate in rats (Monsanto 1981).

Study owner:	Monsanto (1981)
Reliability/Justification:	3 Study not performed under GLP. High-dose well below MTD. Does not conform to modern testing standards.
Substance:	Glyphosate (98.7% pure)
Species/Strain:	Rat/Sprague-Dawley, groups of 50 ♂ and 50 ♀
Administration route:	Diet
Concentration:	0, 30, 100, 300 ppm diet (♂ about 0, 3, 10, 31 mg/kg bw/day; ♀ about 0, 3, 11, 34 mg/kg bw/day)
Duration:	26 months
Findings:	≥ 300 ppm diet: NOAEL (♂ + ♀) No treatment-related effects
Select neoplasms:	Pituitary adenoma, Testes interstitial cell

DOI: 10.3109/10408444.2014.1003423

Table 4. Study 1 – Pituitary tumor findings.

	Dose group (mg/kg bw/day)							
	Males				Females			
Tumors	0	3.05	10.3	31.49	0	3.37	11.22	34.02
Pituitary tumors	Number of animals/total number examined (% per group)							
Adenomas - B	16/48 (33)	19/49 (39)	20/48 (42)	18/47 (38)	34/48 (70)	29/48 (60)	31/50 (62)	26/49 (53)
Carcinomas - M	3/48 (6)	2/49 (4)	3/48 (6)	1/47 (2)	8/48 (17)	7/48 (14)	5/48 (19)	12/49 (24)
Combined	19/48 (40)	21/49 (43)	23/48 (48)	19/47 (40)	42/48 (88)	36/48 (75)	36/50 (72)	38/49 (78)

B benign, *M* malignant

The incidence of interstitial cell tumors of the testes in male rats in both the scheduled terminal sacrifice animals, as well as for all animals, suggested a possible treatment-related finding, and was presented along with contemporary historical control data for comparison (Tables 5 and 6). It was noted that at 12 months, the incidence of interstitial tumors was near zero; however, in animals aged 24–29 months at necropsy, the incidence increased to approximately 10%. The historical control data for chronic toxicity and carcinogenicity from 5 studies terminated at 24–29 months showed background levels of interstitial cell tumors comparable to those found at the highest dose in the study. Furthermore, the reported incidences in all dose groups reflect the normal range of interstitial cell tumors in rat testes, reported in the Registry of Industrial Toxicology Animal Data (Nolte et al. 2011). The incidence of interstitial cell hyperplasia did not provide evidence of a pre-neoplastic lesion. The investigators noted that at terminal sacrifice, the incidence of interstitial cell tumor was 15.4% (4/26), while the range in control animals from 5 contemporary studies (historical controls) was 6.2% (4/65) to 27.3% (3/11), with an overall mean value of 9.6% (16/166). When all animals on test are included, the incidence for the high-dose males was 12% (6/50), compared to a contemporary historical control range of 3.4% (4/116) to 6.7% (5/75), with a mean of 4.5% (24/535). The concurrent control incidence of interstitial cell tumors (0%) was not representative of the normal background incidence noted in contemporary historical control data. Therefore, the data suggest that the incidence in treated rats is within the normal biological variation observed for interstitial cell tumors at this site in this strain of rat. When evaluated in the context of the full data set for male rats (Table 20), a dose-response is clearly absent for the 25 doses evaluated in rats, ranging from 3 to 1290 mg/kg bw/day, which demonstrates that this tumor is clearly not a consequence of glyphosate exposure.

In conclusion, glyphosate was not considered carcinogenic in Sprague Dawley rats following continuous dietary exposure of upto 300 ppm, corresponding to 31 and 34 mg/kg bw/day in males and females, respectively, which is consistent with evaluations by the US EPA (US EPA 1993), the original Annex I listing in Europe (EC 2002), and WHO/FAO (WHO/FAO 2004a).

Based on the low doses tested in Study 1, Monsanto was obliged to conduct a second chronic/carcinogenicity study in rats (Study 2, discussed below) in accordance with OECD TG 453 (1981), which had been developed and instituted after this initial study was conducted.

Study 2 (Monsanto 1990)

In response to evolving regulatory requirements, this study was conducted in accordance with the contemporary version of OECD TG 453 (Monsanto 1990). The chronic toxicity and carcinogenic potential of glyphosate were assessed in a 24-month feeding study in 50 male and 50 female Sprague Dawley rats, dosed with 0, 2000, 8000 and 20 000 ppm (equivalent to mean achieved dose levels of 0, 89, 362 and 940 mg/kg bw/day for males and 0, 113, 457 and 1183 mg/kg bw/day for females (Table 7). In addition, 10 rats per sex per dose were included for interim sacrifice after 12 months. Observations covered clinical signs, ophthalmic examinations, body weight, food consumption, hematology, clinical chemistry and urinalysis, as well as organ weights, necropsy, and histopathological examination. This study was rated Klimisch 1 for reliability.

Treatment-related findings in this study were significantly reduced body weight in high-dose females, as well as increased liver weight in high-dose males and females, and a slight increase in incidence of cataract lens changes in high-dose males, which was not statistically significant for eye lesions confirmed by histopathology (Table 7). The body weight changes confirm that the MTD was achieved in the highest dose group. Benign thyroid C-cell adenomas were statistically higher than controls in the mid-dose terminally sacrificed males, but when pooled with unscheduled deaths, no statistically significant increase was noted. Benign pancreas islet cell adenomas were not statistically higher for the unscheduled or scheduled deaths, but when combined, were statistically higher than controls in the low and high dose males. In both cases, the benign tumors did not exhibit a dose-response, and did not progress to carcinomas, and thus the US EPA concluded that these tumors were not related to the administration

Table 5. Study 1 - Interstitial cell tumor findings in the testes.

	Dose (mg/kg bw/day)			
Tumors	0	3.05	10.3	31.49
Interstitial cell tumor – B	Number of animals/total number examined (% per group)			
Terminal sacrifice	0/15 (0)	2/26 (7.7)	1/16 (6.3)	4/26 (15.4)
All Animals	0/50 (0)	3/50 (6)	1/50 (2)	6/50 (12)
Interstitial cell hyperplasia	Number of animals (% per group)			
Terminal sacrifice	1/15 (6.7)	1/26 (3.8)	0/16 (0)	0/26 (0)
All Animals	1/50 (2)	1/50 (2)	1/50 (2)	0/50 (0)

B benign, *M* malignant

Table 6. Study 1 – Summary of the contemporary historical control data for interstitial cell tumors in the testes of rats in chronic toxicity studies.

	Study 1	Study 2	Study 3	Study 4	Study 5	
	Number of control animals/total number examined (% per study)					Range
Terminal sacrifice	4/65 (6.2)	3/11 (27.3)	3/26 (11.5)	3/24 (12.5)	3/40 (7.5)	6.2–27.3%
All animals	4/116 (3.4)	5/75 (6.7)	4/113 (3.5)	6/113 (5.3)	5/118 (4.2)	3.4–6.7%

of glyphosate (US EPA 1993). These neoplasms, in addition to skin keratoacanthoma in males, a common rat tumor, were selected for further weight of evidence evaluation (Tables 20 and 21). No evidence of a glyphosate-induced carcinogenic effect was noted in either sex (see data Supplementary Study 2 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423).

In conclusion, glyphosate was not carcinogenic in Sprague Dawley rats following continuous dietary exposure of up to 20 000 ppm for 24 months, corresponding to 940 and 1183 mg/kg bw/day in males and females, respectively, which is consistent with evaluations by the US EPA (US EPA 1993), European Authorities (EC 2002), and WHO/FAO (WHO/FAO 2004a).

Study 3 (Cheminova 1993a)

The chronic toxicity and carcinogenic potential of glyphosate technical acid were assessed in a 104-week feeding study in male and female Sprague Dawley rats (Cheminova 1993a). The study was conducted between 1990 and 1992. Groups of 50 rats per sex received daily dietary doses of 0, 10, 100, 300, or 1000 mg/kg bw/day of glyphosate technical acid for 24 months (Table 8). Five additional groups of 35 rats per sex, receiving daily dietary doses of, 0, 10, 100, 300 or 1000 mg/kg bw/day, were included for interim sacrifice at the 12th month for evaluation of chronic toxicity. The dietary glyphosate levels were adjusted weekly to ensure that animals were receiving the intended dose levels at all times. This study was rated Klimisch 1 for reliability.

At 1000 mg/kg bw/day, female mean liver weights were decreased, while males and females had statistically significant reductions in body weight throughout the study, confirming that the MTD was achieved (Table 8). Neoplasms were noted in control and treated groups, but dose-responses were not evident, and no statistically significant increases versus controls were noted for any tumor type ($p < 0.05$). No treatment-related neoplastic lesions were observed at termination,

Table 7. Study 2 – Two-year feeding study of glyphosate in rats (Monsanto 1990).

Study owner:	Monsanto (1990)			
Reliability/Justification:	1 Study performed according to GLP and OECD guideline requirements, with no deviations.			
Substance:	Glyphosate (96.5% pure)			
Species/Strain:	Rat/Sprague-Dawley, groups of 50 ♂ and 50 ♀ (10 rats per sex per dose were included for interim sacrifice after 12 months).			
Administration route:	Diet			
Concentration:	0, 2000, 8000, 20 000 ppm diet (♂ about 0, 89, 362, 940 mg/kg bw/day; ♀ about 0, 113, 457, 1183 mg/kg bw/day)			
Duration:	2 years			
Findings:	8000 ppm diet: NOAEL (♂+♀) 20 000 ppm diet: cataracts (♂), >20% reduced cumulative body weight gain through months 18–20 (♀), 13% increased liver weight (♂). Local effects: inflammation of gastric mucosa			
Select neoplasms:	Pancreatic islet cell adenoma, skin keratoacanthoma (males), thyroid C cell adenoma			
Tumor		Dose (mg/kg bw/day)		
Males	0	89	362	940
Findings for dead and moribund sacrificed animals				
Pancreas: Islet call adenoma – B	1/34 (3%)	4/28 (14%)	2/33 (6%)	4/32 (13%)
Skin: Keratoacanthoma – B	0/36	1/31 (3%)	2/33 (6%)	1/32 (3%)
Thyroid: C cell adenoma – B	0/36	2/29 (7%)	1/31 (3%)	1/33 (3%)
Thyroid: C cell carcinoma – M	0/36	1/29 (3%)	2/31 (6%)	1/33 (3%)
Findings for animals sacrificed at termination				
Pancreas: Islet call adenoma – B	0/14	4/19 (21%)	3/17 (6%)	3/17 (6%)
Skin: Keratoacanthoma – B	0/13	2/19 (11%)	2/17 (12%)	2/17 (12%)
Thyroid: C cell adenoma – B	0/14	2/19 (11%)	*7/17 (41%)	4/17 (24%)
Thyroid: C cell carcinoma – M	0/14	0/19	0/17	0/17
Females	0	113	457	1183
Findings for dead and moribund sacrificed animals				
Pancreas: Islet call adenoma – B	3/28 (11%)	0/28	3/33 (9%)	0/31
Thyroid: C cell adenoma – B	0/28	0/28	1/33 (3%)	2/32 (6%)
Thyroid: C cell carcinoma – M	0/28	0/28	1/33 (3%)	0/32
Findings for animals sacrificed at termination				
Pancreas: Islet call adenoma – B	2/22 (9%)	1/22 (5%)	1/17 (6%)	0/18
Thyroid: C cell adenoma – B	2/22 (9%)	2/22 (9%)	5/17 (29%)	4/18 (22%)
Thyroid: C cell carcinoma – M	0/22	0/22	0/17	0/18

B benign, *M* malignant
*Statistically higher than controls ($p < 0.05$, Fisher's Exact Test with the Bonferroni Inequality).

DOI 10.3109/10408444.2014.1003423

Table 8. Study 3 – Two-year feeding study of glyphosate in rats (Cheminova 1993a).

Study owner:	Cheminova (1993a)
Reliability/ Justification:	1 Study performed according to GLP and OECD guideline requirements, with no deviations.
Substance:	Glyphosate (98.7–98.9% pure)
Species/Strain:	Rat/Sprague-Dawley, groups of 50 ♂ and 50 ♀ (additional groups of 35 ♂ and 35 ♀per dose were included for 1-year interim sacrifice)
Administration route:	Diet
Achieved dose:	♂+♀: 0, 10, 100, 300, 1000 mg/kg bw/day (weekly adjustment of dietary concentration for the first 13 weeks and 4-weekly thereafter)
Duration:	2 years
Findings:	300 mg/kg bw/day: NOAEL (♂+♀) 1000 mg/kg bw/day: body weights ↓, urinary pH ↓, salivary glands (histopathology, organ weight ↑); evidence of weak liver toxicity (alkaline phosphatase ↑, ♀: organ weight ↓)
Select neoplasms:	No neoplasms from this study were identified for further consideration.

and no select neoplasms were identified in this study for further consideration (see data Supplementary Study 3 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423). Glyphosate was not considered carcinogenic in male and female Sprague Dawley rats following 104 weeks of continuous dietary exposure of up to 1000 mg/kg bw/day, the limit dose, which is consistent with evaluations by the European Authorities (EC 2002, Germany Rapporteur Member State 2015b) and WHO/FAO (WHO/FAO 2004a).

Study 4 (Feinchemie Schwebda 1996)

A 2-year bioassay in the Wistar rat used dietary glyphosate levels of 0, 100, 1000, and 10 000 ppm (Feinchemie Schwebda 1996). Groups of 50 rats per sex were fed for 24 months. The mean achieved dose levels were 0, 7.4, 73.9, and 740.6 mg/kg bw/day (Table 9). This study was rated Klimisch 1 for reliability.

In addition, one vehicle control with ten rats per sex and one high dose (10 000 ppm) group with 20 rats per sex were included for interim sacrifice after one year of treatment, to study non-neoplastic histopathological changes. The mean achieved dose level in the treated group was 764.8 mg/kg bw/day. Observations covered clinical signs, body weight, food consumption, hematology, clinical chemistry, and urinalysis, as well as organ weights, necropsy, and histopathological examination.

There were no treatment-related deaths or clinical signs in any of the dose-groups. Moreover, there were no treatment-related effects on body weight gain or food consumption noted. This suggests that the MTD may not have been reached by the applied dosing regimen.

There was some background variation in the incidences of benign tumors (e.g. reduced tumor incidence in low and mid-dose males, increased tumor incidence in middose females), which was considered incidental in absence of a dose-response relationship (see data Supplementary Study 4 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423).

The different liver tumors observed in the dead and moribund sacrificed and terminally sacrificed rats included hepatocellular adenoma, intrahepatic bile duct adenomas, cholangiocarcinoma, hepatocellular carcinoma, histiocytic sarcoma, fibrosarcoma, and lymphosarcoma. Among these, hepatocellular adenomas and carcinomas occurred more frequently, as often observed in aging rats (Thoolen et al. 2010). These tumors appeared to be incidental and not compound-related, as their frequency of occurrence was not dependent on dose. Hepatocellular adenomas and carcinomas were considered select neoplasms (Table 9), based on increased incidence above controls for total animals, albeit non-dose

Table 9. Study 4 – Two-year feeding study of glyphosate in rats (Feinchemie Schwebda 1996).

Study owner:	Feinchemie Schwebda (1996)			
Reliability/Justification:	1 Study performed according to GLP and OECD guideline requirements, with no deviations.			
Substance:	Glyphosate (96.0–96.8% pure)			
Species/Strain:	Rat/Wistar, groups of 50 ♂ and 50 ♀			
Administration route:	Diet			
Concentration:	0, 100, 1000, 10 000 ppm diet (♂ about 0, 6.3, 59.4, 595 mg/kg bw/day; ♀ about 0, 8.6, 88.5, 886 mg/kg bw/day)			
Duration:	2 years			
Findings:	10 000 ppm diet: ≥ NOAEL (♂+♀) Only mild effects on clinical chemistry (liver enzymes), without histopathological changes.			
Select neoplasms:	Hepatocellular adenoma, hepatocellular carcinoma			
Tumor		Dose (mg/kg bw/day)		
Males	0	7.4	73.9	741
Findings for dead and moribund sacrificed animals				
Hepatocellular adenoma – B	9/30 (30%)	9/30 (30%)	6/32 (19%)	6/21 (29%)
Hepatocellular carcinoma – M	12/30 (40%)	12/30 (40%)	9/32 (28%)	5/21 (24%)
Findings for animals sacrificed at termination				
Hepatocellular adenoma – B	15/20 (75%)	13/20 (65%)	4/16 (25%)	15/20 (75%)
Hepatocellular carcinoma – M	9/20 (45%)	16/20 (80%)	9/16 (56%)	19/29 (66%)
		Dose (mg/kg bw/day)		
Females	0	7.4	73.9	741
Findings for dead and moribund sacrificed animals				
Hepatocellular adenoma – B	2/26 (8%)	8/23 (3%)	3/17 (18%)	5/29 (17%)
Hepatocellular carcinoma – M	4/26 (15%)	4/23 (17%)	2/17 (12%)	5/29 (17%)
Findings for animals sacrificed at termination				
Hepatocellular adenoma – B	16/24 (67%)	10/25 (40%)	16/32 (50%)	8/21 (38%)
Hepatocellular carcinoma – M	6/24 (25%)	11/25 (44%)	12/32 (38%)	4/21 (19%)

B benign, *M* malignant

responsive, for adenoma in mid-dose females, carcinoma in low- and high-dose males, and carcinoma in low- and mid-dose females. These liver neoplasms are considered in the weight of evidence evaluation (Tables 20 and 21).

The study report concluded that glyphosate technical acid was not carcinogenic in Wistar rats following continuous dietary exposure of up to 595 and 886 mg/kg bw/day in males and females, respectively, for 24 months, which is consistent with evaluations by the European Authorities (EC 2002, Germany Rapporteur Member State 2015b).

Study 5 (Excel 1997)

A 2-year feeding study in the Sprague Dawley rats (Excel 1997) featured dietary concentrations of 0, 3000, 15 000, and 25 000 ppm glyphosate technical acid. Groups of 50 rats per sex were fed for 24 months, and mean dose levels of 0, 150, 780 and 1290 mg/kg bw/day (males) and 0, 210, 1060 and 1740 mg/kg bw/day (females) were achieved (Table 10).

In addition, 20 rats/sex/group were included for interim sacrifice at week-52, to study non-neoplastic histopathological changes with a different high-dose level of 30 000 ppm. The dietary doses correspond to 180, 920 and 1920 mg/kg bw/day (males) and 240, 1130 and 2540 mg/kg bw/day (females), for 3000, 15 000 and 30 000 ppm, respectively. Thus, a limit dose above 1000 mg/kg bw/day was achieved.

The study report notes that glyphosate technical acid was not carcinogenic in Sprague Dawley rats following continuous dietary exposure to up to 1290 mg/kg bw/day, and 1740 mg/kg bw/day for males and females, respectively, for 24 months. However, this study was rated Klimisch 3 for reliability (Germany Rapporteur Member State 2015b), and therefore, is considered unreliable for carcinogenicity evaluation based on lower than expected background tumor incidences (see data Supplementary Study 5 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423). In addition, the test substance was not adequately characterized, and several deviations from the OECD Test Guideline 453 were noted.

Study 6 (Arysta Life Sciences 1997b)

A combined chronic toxicity/carcinogenicity study in Sprague Dawley rats (Arysta Life Sciences 1997b) was conducted between December 1994 and December 1996. The rats were fed 0, 3000, 10 000, and 30 000 ppm glyphosate for two years (equivalent to 0, 104, 354 and 1127 mg/kg bw/day for males and 0, 115, 393 and 1247 mg/kg bw/day for females (Table 11). Thus, a limit dose was achieved, and the MTD was noted at the high dose in males and females with decreased body weight, increased cecum weight, distention of the cecum, loose stool and skin lesions. In addition, 30 rats/sex/group were included for interim sacrifice at 26, 52 and 78 weeks, to study non-neoplastic histopathological changes. Observations covered clinical signs, body weight, food consumption, hematology, clinical chemistry, and urinalysis, as well as organ weights, necropsy, and histopathological examination. This study was rated Klimisch 1 for reliability.

Non-statistically significant increases versus controls ($p < 0.05$) were noted for pituitary adenomas, skin keratoacanthoma in high-dose males, and mammary gland fibroadenoma in low and mid-dose females (Table 11). These neoplasms were considered for the weight of evidence evaluation (Tables 20 and 21), and the full tumor summary data are available online (see data Supplementary Study 6 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423). As mentioned under Study 1, pituitary and mammary tumors are common spontaneous neoplasms in aging rats (Johnson and Gad 2008), and skin keratoacanthoma is noted as one of the most common spontaneous benign neoplasms in male Sprague Dawley rats (Chandra et al. 1992). The study report concluded that glyphosate was not carcinogenic in Sprague Dawley rats following continuous dietary exposure to up to 30 000 ppm for 24 months, corresponding to 1127 mg/kg bw/day and 1247 mg/kg bw/day for males and females, respectively, which is consistent with the recent evaluation in Europe under the Annex I Renewal of glyphosate (Germany Rapporteur Member State 2015b).

Table 10. Study 5 – Two-year feeding study of glyphosate in rats (Excel 1997).

Study owner:	Excel (1997)			
Reliability/Justification:	3 Test substance not characterized and other deviations from OECD 453, lower than expected background tumor incidence			
Substance:	Glyphosate (no purity reported)			
Species/Strain:	Rat/Sprague-Dawley, groups of 50 ♂ and 50 ♀, additional groups of 20 rats per sex and group were included for interim sacrifice after 52 weeks			
Administration route:	Diet			
Concentration:	2-year group: 0, 3000, 15 000, 25 000 ppm diet (♂ about 0, 150, 780, 1290 mg/kg bw/day; ♀ about 0, 210, 1060, 1740 mg/kg bw/day) 1-year group: 0, 3000, 15 000, 30 000 ppm diet (♂ about 0, 180, 920, 1920 mg/kg bw/day; ♀ about 0, 240, 1130, 2540 mg/kg bw/day)			
Duration:	2 years			
Findings:	≥ 25 000 ppm diet: NOAEL (♂ + ♀) Only mild toxic effects, such as clinical chemistry of questionable relevance in aged rats, without correlating histopathological organ changes.			
Select neoplasms:	No neoplasms from this study were identified for further consideration. Low background tumor incidence indicates low study reliability with no relevant increases in the incidence of tumors.			
Males		Dose (mg/kg bw/day)		
	0	150	740.6	1290
Mortality	16/50 (32%)	17/50 (34%)	18/50 (36%)	23/50 (46%)
Females		Dose (mg/kg bw/day)		
	0	210	1060	1740
Mortality	19/50 (38%)	20/50 (40%)	20/50 (40%)	25/50 (50%)

DOI 10.3109/10408444.2014.1003423

Table 11. Study 6 – Two-year feeding study of glyphosate in rats (Arysta Life Sciences 1997b).

Study owner:	Arysta Life Sciences (1997b)			
Reliability/Justification:	1 Study performed according to GLP and OECD guideline requirements, with no deviations.			
Substance:	Glyphosate (94.6–97.6% pure)			
Species/Strain:	Rat/Sprague-Dawley, groups of 50 ♂ and 50 ♀; satellite groups of 30 ♂ and 30 ♀for interim investigations			
Administration route:	Diet			
Concentration:	0, 3000, 10 000, 30 000 ppm diet (♂ about 0, 104, 354, 1127 mg/kg bw/day; ♀ about 0, 115, 393, 1247 mg/kg bw/day)			
Duration:	2 years			
Findings:	3000 ppm diet: NOAEL (♂+♀) 10 000 ppm diet: cecum weight↑, distension of cecum, loose stool, follicular hyperkeratosis and/or folliculitis/follicular abscess of the skin, body weight ↓			
Select neoplasms:	Pituitary adenoma, skin keratoacanthoma (males), mammary gland fibroadenoma (females)			
Tumor		Dose (mg/kg bw/day)		
Males	0	104	354	1127
Findings for dead and moribund sacrificed animals (Table 25–10)				
Pituitary anterior adenoma – B	22/32 (69%)	21/30 (70%)	*14/32 (44%)	18/21 (86%)
Skin keratoacanthoma – B	2/32 (6%)	1/30 (3%)	0/32	1/21 (5%)
Findings for animals sacrificed at termination (after 104 weeks, Table 25–8)				
Lung adenoma – B	0/18	2/20 (10%)	1/18 (6%)	3/29 (10%)
Pituitary anterior adenoma – B	13/18 (72%)	14/20 (70%)	13/18 (72%)	21/29 (72%)
Pituitary adenoma in intermediate part – B	0/18	1/20 (5%)	0/18	0/29 (0%)
Skin keratoacanthoma – B	1/18 (6%)	2/20 (10%)	0/18	6/29 (21%)
Tumor		Dose (mg/kg bw/day)		
Females	0	115	393	1247
Findings for dead and moribund sacrificed animals				
Pituitary anterior adenoma – B	34/35 (97%)	29/31 (94%)	28/33 (82%)	31/36 (86%)
Thyroid follicular adenoma – B	0/35	2/31 (6%)	0/32	0/36
Mammary gland fibroadenoma – B	13/35 (37%)	14/31 (45%)	12/34 (35%)	20/36 (56%)
Findings for animals sacrificed at termination				
Pituitary anterior adenoma – B	12/15 (80%)	19/19 (100%)	12/16 (75%)	13/14 (93%)
Mammary gland fibroadenoma – B	10/15 (67%)	13/19 (68%)	12/16 (75%)	10/14 (71%)

B benign, *M* malignant
*Statistically lower than controls ($p < 0.05$).

Study 7 (Syngenta 2001)

The same rat model that was used in the previously discussed 12-month chronic rat study (Syngenta 1996b) was also employed in a 2-year feeding study (Syngenta 2001). A group of 52 male and 52 female Wistar rats received 0, 2000, 6000 or 20 000 ppm via feed (Table 12). The mean achieved dose levels were 0, 121, 361 and 1214 mg/kg bw/day for males, and 0, 145, 437 and 1498 mg/kg bw/day for females. Thus, a limit dose was achieved. In addition, three satellite groups with 12 rats per sex each were included for interim sacrifice after 12 months of treatment, to investigate potential non-neoplastic histopathological changes. Observations covered clinical signs, body weight, food consumption, hematology, clinical chemistry, and urinalysis, as well as organ weights, necropsy, and histopathological examination. This study was rated Klimisch 1 for reliability.

Treatment-related findings in this study were found in the liver and kidney, and were confined to animals (predominantly males) fed 20 000 ppm glyphosate acid. There were a number of changes in males and females fed 20 000 ppm glyphosate acid, notably renal papillary necrosis, prostatitis, periodontal inflammation, urinary acidosis, and hematuria, which may be attributed to the acidity of the test substance. Slight increases in proliferative cholangitis and hepatitis were noted in males at 20 000 ppm. Despite the findings at 20 000 ppm, survival was better in males fed 20 000 ppm than in the controls and lower dose groups. This improved survival was associated with a decreased severity of renal glomerular nephropathy and a 5% reduction in body weight (see data Supplementary Study 7 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423, for neoplastic and non-neoplastic findings).

A small increase in the incidence of hepatocellular adenoma was observed in males fed 20 000 ppm glyphosate acid. While not statistically significant using the Fisher's exact test, the difference was statistically significant for total male rats using the Peto Test for trend. However, there was no evidence of pre-neoplastic foci, no evidence of progression to adenocarcinomas, and no dose-response. In addition, the incidence was within the laboratory's historical control range for tumors of this type in the liver (Table 12). Therefore, the increased incidence was considered not to be related to treatment, yet these were considered select neoplasms (Table 12) and evaluated in context of the complete data set (Tables 20 and 21).

The study report concluded that glyphosate acid was not carcinogenic in the Wistar rats following continuous dietary exposure to up to 20 000 ppm for 24 months, at 1214 and 1498 mg/kg bw/day in males and females, respectively, which is consistent with the WHO/FAO review (WHO/FAO 2004a) and the recent evaluation in Europe under the Annex I Renewal of glyphosate (Germany Rapporteur Member State 2015b).

Study 8 (Nufarm 2009b)

The most recent study in this series of regulatory studies investigating the potential carcinogenicity of glyphosate in rats was conducted from September 2005 through March 2008 (Nufarm 2009b). The study was conducted by feeding dietary concentrations of 0, 1500, 5000 and 15 000 ppm glyphosate to groups of 51 Wistar rats per sex. To ensure that a received limit dose of 1000 mg/kg bw/day overall was achieved, the highest dose level was progressively increased to 24 000 ppm.

Table 12. Study 7 – Two-year feeding study of glyphosate in rats (Syngenta 2001).

Study owner:	Syngenta (2001)			
Reliability/Justification	1 Study performed according to GLP and OECD guideline requirements, with no deviations.			
Substance:	Glyphosate (97.6% pure)			
Species/Strain	Rat/Wistar Alpk: APfSD, groups of 52 ♂ and 52 ♀ (additional 12 animals per sex and dose for 1-year interim sacrifice)			
Administration route:	Diet			
Concentration:	0, 2000, 6000, 20 000 ppm diet (♂ about 0, 121, 361, 1214 mg/kg bw/day; ♀ about 0, 145, 437, 1498 mg/kg bw/day)			
Duration:	2 years			
Findings:	6000 ppm diet: NOAEL (♂+♀)			
	20 000 ppm diet: Kidney and liver findings. Increased survival due to reduction in CPN, prostatitis, periodontal inflammation			
Select neoplasms:	Hepatocellular adenoma (males), not a statistically significant increase for the high dose using the Fisher's exact test, but statistically significant using Peto trend analysis			
		Dose (mg/kg bw/day)		
Males	0	121	361	1214
Liver				
Hepatocyte fat vacuolation	6	7	11	11
Hepatitis	3	4	2	5
Kidney				
		Dose (mg/kg bw/day)		
Females	0	145	437	1498
Liver				
Hepatocyte fat vacuolation	7	5	6	6
Hepatitis	6	5	4	4
Tumors:		Dose (mg/kg bw/day)		
Males	0	121	361	1214
Findings for dead and moribund sacrificed animals				
*Hepatocellular adenoma – B	0/37	2/36 (6%)	0/35	3/26 (12%)
Hepatocellular carcinoma – M	0/37	0/36	0/35	0/26
Findings for animals sacrificed at termination				
*Hepatocellular adenoma – B	0/16	0/17	0/18	2/26 (8%)
Hepatocellular carcinoma – M	0/16	0/17	0/18	0/26

B benign, *M* malignant
*Historical Control Range: 0–11.5% total males with hepatocellular adenoma, 26 studies, 1984–2003

Mean dose levels of 86/105, 285/349, and 1077/1382 mg glyphosate/kg bw/day (males/females) were achieved (Table 13). This study was rated Klimisch 1 for reliability.

Non-neoplastic findings included transient liver enzyme activity for mid-dose males and high-dose males and females, and equivocal nephrocalcinosis depositions at the high-dose. Histopathology noted a statistically significant increase in adipose infiltration of the bone marrow in high-dose males compared to controls, suggestive of myeloid hypoplasia, which may be considered a stress response (Everds et al. 2013).

Skin keratoacanthoma in males and mammary gland adenocarcinoma in females (Table 13) were considered for evaluation in the context of the weight of evidence for rat tumor incidence (Tables 20 and 21), wherein dose-

Table 13. Study 8 – Two-year feeding study of glyphosate in rats (Nufarm 2009b).

Study owner:	Nufarm (2009a)			
Reliability/Justification:	1 Study performed according to GLP and OECD guideline requirements, with no deviations			
Substance:	Glyphosate (95.7% pure)			
Species/Strain:	Rat/Wistar, groups of 51 ♂ and 51 ♀			
Administration route:	Diet			
Concentration:	0, 3000, 10 000, 15 000 ppm diet, the top dose was progressively increased to reach 24 000 ppm diet by Week-40 (♂ about 0, 84, 285, 1077 mg/kg bw/day; ♀ about 0, 105, 349, 1382 mg/kg bw/day)			
Duration:	2 years			
Findings:	≥ 1077/1382 mg/kg bw/day: NOAEL (♂/♀)			
	Transient liver enzyme activity for mid-dose males and high-dose males and females; equivocal nephrocalcinosis depositions at the high-dose males and females; increased adipose infiltration of the bone marrow in high-dose males			
Select neoplasms:	Skin keratoacanthoma (males), mammary gland adenocarcinoma			
Tumor			Dose (mg/kg bw/day)	
Males	0	84	285	1077
Findings for all animals				
Skin keratoacanthoma – B	2/51 (4%)	3/51 (6%)	0/51	6/51 (12%)
			Dose (mg/kg bw/day)	
Females	0	105	349	1382
Findings for all animals				
Mammary gland adenocarcinoma – M	2/51 (4%)	3/51 (6%)	1/51 (2%)	6/51 (12%)

B benign, *M* malignant

DOI 10.3109/10408444.2014.1003423

responses were not evident. Tumor incidence summary data have been tabulated (see data Supplementary Study 8 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423). Microscopic evaluation of tissues did not reveal any indications of neoplastic lesions caused by glyphosate treatment. The study report concluded that glyphosate acid was not carcinogenic in Wistar rats following continuous dietary exposure to up to 24 000 ppm for 24 months, at 1077 and 1382 mg/kg bw/day in males and females, respectively, which is consistent with the recent evaluation in Europe under the Annex I Renewal of glyphosate (Germany Rapporteur Member State 2015b).

Study 9 Publication (Chruscielska et al. 2000a)

A two-year combined chronic toxicity and carcinogenicity study in Wistar rats was published by academic researchers from Warsaw, Poland. The study was conducted as a drinking-water study in Wistar-RIZ rats according to OECD TG 453. The test material was a 13.85% aqueous formulation of glyphosate as its ammonium salt (equivalent to 12.6% glyphosate acid). However, the ammonium salt of glyphosate tested is not commercially available, and the concentration of active ingredient suggests that a glyphosate-formulated product was tested; this is supported by a concurrent genotoxicity publication by the same lead author (Chruscielska et al. 2000b), previously reviewed by Kier and Kirkland (Kier and Kirkland 2013), in which a glyphosate formulation, Perzocyd, was tested. Deficiencies noted with respect to OECD TG 453 include insufficient dosing to elicit toxic effects, inadequate test material characterization, no reporting of water/feed consumption, body weights and diet composition, and no individual animal data. Although the manuscript reporting deficiencies may have been included in the study, they were not reported in the manuscript, and could warrant a Klimisch reliability score of 4 (not assignable), but the low doses employed in this study justify a Klimisch reliability score of 3.

The test material was administered in water at glyphosate salt concentrations of 0, 300, 900, and 2700 mg/L. Each dose group consisted of 85 animals per sex. Ten animals per sex and dose were sacrificed after 6, 12, and 18 months of exposure, for evaluation of general toxicity. The remaining 55 animals per sex and dose were scheduled for sacrifice after 2 years of exposure.

Water consumption was claimed to have been measured, but these data have not been reported. To estimate the glyphosate doses received via drinking water, the assumed default water consumptions were 50 and 57 mL/kg bw/day by male and female rats, respectively (Gold et al. 1984). Using these standard figures and the glyphosate content of the tested formulation (12.6%), daily doses are estimated at 0, 1.9, 5.7, and 17 mg of glyphosate/kg bw/day for males and 0, 2.2, 6.5, and 19 mg of glyphosate/kg bw/day for females. As this study appears to have tested a formulated product, data were not included in the weight of evidence review (Tables 20 and 21), but given the very low glyphosate doses and reported low tumor incidence, these were of no consequence to the overall data review.

Exposure to glyphosate ammonium salt had no effect on body weight, appearance and behavior, and hematological parameters, which is consistent with glyphosate chronic toxicity data regulatory reviews. Even though there seems to be a trend towards higher 2-year mortality in treated females (Table 14), this difference had no statistical significance according to the authors. There were sporadic alterations of clinical-chemical and urinalysis parameters, but not in a consistent fashion over time and without dose-dependence. These alterations were not interpreted as treatment-related. There was no effect of glyphosate on the incidence of neoplastic lesions (Table 14). Thus, the NOAEL for chronic toxicity and carcinogenicity in this study was greater than or equal to 17 and 19 mg glyphosate/kg bw/day, in males and females, respectively.

Due to the lack of systemic effects in the highest dose group, the MTD was not reached by this study. Judging from other rat studies reviewed here, the MTD is likely to be greater than 1000 mg/kg bw/day. Thus, the top glyphosate dose of an estimated 19 mg/kg bw/day in this study is too low to satisfy regulatory validity criteria for a carcinogenicity study.

Mouse carcinogenicity

There are a total of five carcinogenicity studies with glyphosate in mice, that have been submitted to support glyphosate Annex I renewal in the European Union. All but the oldest study (Study 10) were considered reliable without restriction, and were performed under conditions of GLP following OECD TGs. Most studies were conducted in the CD-1 strain. Each study was sponsored by a different manufacturer. In each case, technical grade glyphosate was administered via diet for at least 18 months. Select neoplasms, mostly lymphoreticular, liver and lung, are summarized for all mouse chronic studies in Tables 22 and 23. These neoplasms are widely recognized as occurring spontaneously in aging mice (Gad et al. 2008, Son and Gopinath 2004). Lymphomas have been recognized for many years as one of the most common, if not the most common category of spontaneous neoplastic lesions in aging mice (Brayton et al. 2012, Gad et al. 2008, Son and Gopinath 2004). The subclassification of malignant lymphomas is not a typical diagnostic feature in rodent studies, likely due to either expense and/or feasibility. It is, however, important to recognize that lymphomas are not a single type of neoplasm, rather they are a grouping of different neoplasms arising from different pathogeneses, and should be considered as different diseases (Bradley et al. 2012). As is the case for NHL in humans, these different immune system neoplasms are clustered together based on manifestation in lymphocytes, despite their very different etiologies; for example, the most common subset of NHL lymphomas clustered together as "diffuse large B cell lymphomas", have for many years been considered multiple clinical-pathologic entities (Armitage 1997), and therefore may be considered attributable to different modes of action. Chronic endpoints and NOAEL values are captured in each study summary table; however, the following study reviews focus on carcinogenicity.

Study 10 (Monsanto 1983)

The first chronic-carcinogenicity mouse study with glyphosate was conducted between March 1980 and March 1982 (Monsanto 1983), prior to the institution of GLP (Table 15). The study design was essentially in compliance with OECD TG 451 for carcinogenicity studies, adopted in 1981, when

Table 14. Publication, Study 9 – Two-year drinking water study in rats with 13.85% glyphosate ammonium salt (Chruscielska et al. 2000a).

Authors:	Chruscielska et al. (2000a)
Reliability/Justification:	3 Study not performed according to GLP, but according to OECD TG 453, with the following deficiencies: Reporting deficits (water and feed consumption, body weights, diet composition, individual animal data, substance composition, purity, and stability) Highest dose did not elicit toxicity.
Substance:	Ammonium salt of glyphosate, 13.85% solution
Species/Strain:	Rat/Wistar -RIZ outbred, 85 ♂ and 85 ♀ per dose group. 10 ♂ and 10 ♀each were sacrificed after 6, 12, and 18 months of exposure.
Administration route:	Drinking water
Concentration:	0, 300, 900, and 2700 mg/L Estimated glyphosate intake: ♂: 0, 1.9, 5.7, and 17 mg/kg bw/day, ♀: 0, 2.2, 6.5, and 19 mg/kg bw/day, based on assumed water consumptions of 50/57 mL/kg bw/day (♂/♀), (Gold, et al. 1984)
Duration:	2 years
Findings:	17/19 mg glyphosate/kg bw/day: NOAEL (♂/♀) No treatment-related effects
Tumors reported for 85 rats/sex/dose:	No increase in the incidence of tumors attributable to glyphosate administration

	Estimated dose (mg/kg bw/day)							
	0		1.9/2.2		5.7/6.5		17/19	
	♂	♀	♂	♀	♂	♀	♂	♀
Two-year mortality	42%	38%	42%	45%	54%	53%	44%	60%
Lungs								
Lymphoma	2	–	2	–	1	–	3	1
Histiocytoma	–	–	–	–	–	–	–	1
Adenocarcinoma	1	–	–	–	–	–	–	–
Histiocytoma, malignant	–	1	–	–	1	–	–	–
Spleen, leukemia	0	–	2	–	0	–	1	–
Kidneys, Fibrous histiocytoma	–	–	–	–	–	–	1	–
Pituitary gland								
Adenoma	4	10	4	6	2	8	0	3
Adenoma, malignant (assumed to be carcinoma)	0	1	0	3	1	2	1	5
Carcinoma	0	–	0	–	1	–	0	–
Thyroid								
Adenoma	1	1	1	2	0	0	3	3
Carcinoma	0	–	1	–	0	–	0	–
Uterus, cervix carcinoma	–	0	–	0	–	0	–	1
Uterus, body, histiocytoma	–	3	–	1	–	0	–	1
Mammary gland								
Fibroma	–	0	–	0	–	0	–	0
Fibroadenoma	–	3	–	2	–	3	–	3
Adrenal medulla, adenoma	1	2	2	2	1	2	0	2
Thymus, lymphoma	0		0		0		1	
Testis, Leydigoma	–		3		6		1	
Subcutaneous tissue								
Fibroma	0		1		1		3	
Lipoma	–	–	–	–	–	–	–	1
Cystadenoma	–	1	–	–	–	–	–	–
Lymph nodes								
Lymphoma	0		0		0		1	
Lymphoma, malignant	–	1	–	–	–	–	–	–
Skin, carcinoma	2	–	–	–	–	–	–	–
Prostate, adenoma	1	–	–	–	–	–	–	–

the study was already ongoing. Groups of 50 male and female CD-1 mice received glyphosate at dietary levels of 1000, 5000, and 30 000 ppm, over a period of nearly two years. The mean achieved doses were 157/190, 814/955, and 4841/5874 mg/kg bw/day in males and females, respectively, exceeding the limit dose. Based on this study predating both GLP and OECD TG 451, a reliability score of Klimisch 2 has been assigned.

In addition to post-mortem pathological examinations after terminal sacrifice, hematological investigations were performed on 10 mice per sex and dose at months 12 and 18, and on 12 male animals/group, as well as all surviving females at scheduled termination.

Two non-neoplastic histological changes affecting the liver and urinary bladder were assumed to be treatment-related. There was a higher incidence of centrilobular hepatocyte hypertrophy in high-dose males, and a more frequent occurrence of slight-to-mild bladder epithelial hyperplasia in the mid and high dose; however, a clear dose-response was lacking. Tumor incidences, which did not significantly increase with dose, were mostly bronchiolar-alveolar, hepatocellular, or lymphoreticular, all of which are commonly noted spontaneously occurring tumors in aging mice (Table 15). Lymphoreticular tumors combined for males and females totaled 7, 12, 10 and 12 for control, low, mid- and high-dose groups respectively, and were not considered as being related to test substance.

A more frequent occurrence of slight-to-mild bladder epithelial hyperplasia was observed in the mid and high-dose groups; however, clear dose-response was lacking (Table 15) and no urinary bladder neoplasms were noted at these doses (see data Supplementary Study 10 to be found online at http://

DOI 10.3109/10408444.2014.1003423

Table 15. Study 10 – Two-year feeding study with glyphosate in mice (Monsanto 1983).

Study owner:	Monsanto (1983)			
Reliability/Justification	2 Study was performed prior to institution of GLP and OECD guideline requirements			
Substance:	Glyphosate (99.7% pure)			
Species/Strain:	Mouse/CD-1, groups of 50 ♂ and 50 ♀			
Administration route:	Diet			
Concentration:	0, 1000, 5000, 10 000 ppm diet (♂ about 0, 157, 814, 4841 mg/kg bw/day; ♀ about 0, 190, 955, 5874 mg/kg bw/day)			
Duration:	24 months			
Findings:	1000 ppm diet: NOAEL (♂ + ♀) 5000 ppm diet: body weight ↓, histological changes in liver and urinary bladder (slight to mild epithelial hyperplasia in males at mid and high doses)			
Select neoplasms:	Lymphoreticular neoplasms, bronchiolar-alveolar adenocarcinoma			
		Dose (mg/kg bw/day)		
Males	0	157	814	4841
Lymphoreticular system				
Lymphoblastic lymphosarcoma with leukemia – M	1/48 (2%)	4/49 (8%)	3/50 (6%)	2/49 (4%)
Lymphoblastic lymphosarcoma without leukemia – M	0/48	1/49 (2%)	0/50 (0%)	0/49
Composite lymphosarcoma – M	1/48 (2%)	0/49	1/50 (2%)	0/49
Histiocytic sarcoma – M	0/48	1/49 (2%)	0/50	0/49
Total lymphoreticular neoplasms[#]	2/48 (4%)	6/49 (12%)	4/50 (8%)	2/49 (4%)
		Dose (mg/kg bw/day)		
Females	0	190	955	5873
Lymphoreticular system				
Lymphoblastic lymphosarcoma with leukemia – M	1/50 (2%)	4/48 (8%)	5/49 (10%)	1/49 (2%)
Lymphoblastic lymphosarcoma without leukemia – M	0/50 (0%)	1/48 (2%)	0/49 (0%)	3/49 (6%)
Composite lymphosarcoma – M	4/50 (8%)	1/48 (2%)	1/49 (2%)	6/49 (12%)
Histiocytic sarcoma – M	0/50 (0%)	0/48 (0%)	0/49 (0%)	0/49 (0%)
[#] Total lymphoreticular neoplasms	5/50 (10%)	6/48 (13%)	6/49 (12%)	10/49 (20%)

[#]Sum of lymphoblastic lymphosarcoma, composite lymphosarcoma, and histiocytic sarcoma.
M malignant

informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423). Benign renal tubule adenomas were noted in mid- and high-dose males at incidences of 1/50 and 3/50 respectively. These neoplasms were not observed in females, lacked statistical significance, and were considered spontaneous and unrelated to glyphosate administration by the study pathologists; this neoplasm, while not seen in the concurrent control group, had previously been noted in control male CD-1 mice of comparable age by the author of the study. As an additional measure of diligence, a Pathology Working Group was convened, and it concluded that the absence of any pre-neoplastic kidney lesion in all male animals provided sufficient evidence that this finding was spurious and not related to glyphosate administration. This is reflected in the US EPA review of glyphosate (US EPA 1993). This neoplasm was not observed in the other four mouse carcinogenicity studies discussed.

The author of the study also reported a trend towards a non-statistically significant increased occurrence of lymphoreticular neoplasia in treated female mice (Table 15). However, these consisted of three different categories of lymphoreticular neoplasms. Regulatory reviews confirmed that there is no apparent dose-dependence for these endpoints (EC 2002, US EPA 1993, WHO/FAO 2004a). Summary tables of incidence of neoplastic findings are available (see data Supplementary Study 10 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423).

Glyphosate was reported as not carcinogenic in CD-1 mice up to doses well in excess of the limit dose for carcinogenicity testing, which is consistent with evaluations by the US EPA (US EPA 1993), European Commission (EC 2002), recent EU Annex I Renewal evaluation by the Rapporteur (Germany Rapporteur Member State 2015b), and WHO/FAO (WHO/FAO 2004a).

Study 11 (Cheminova 1993b)

Another carcinogenicity bioassay in mice was conducted between December 1989 and December 1991 (Table 16) (Cheminova 1993b). In this assay, 50 male and 50 female CD-1 mice per dose group received glyphosate via their diet over a period of approximately two years. This treatment period is 6 months longer than the 18 months stipulated for mice by OECD TG 451 (1981 version). The dietary levels were adjusted regularly to achieve constant dose levels of 0, 100, 300 and 1000 mg/kg bw/day, achieving the limit dose. This study was rated Klimisch 1 for reliability.

Slight non-statistically significant increases in bronchiolar-alveolar adenomas were noted for all male dose groups above controls in a non-dose-responsive manner. Bronchiolar-alveolar neoplasms are evaluated in the context of the full data set (Tables 22 and 23), demonstrating a lack of dose-response across doses ranging from approximately 15 mg/kg bw/day to 5000 mg/kg bw/day. Although the number of pituitary adenomas were low and considered incidental, they were conservatively included in the select neoplasms, based on being slightly higher in high dose females than concurrent controls (Table 16). The data summary of all histological findings, including tumor incidence, is available (see data Supplementary Study 11 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423).

There were no statistically significant increases in the occurrence of any tumor type in this study. The observed variations did not show a dose relationship, and were within the range of historical control data. Glyphosate was determined to be not carcinogenic to CD-1 mice at up to 1000 mg/kg bw/day, which is consistent with evaluations by the European Commission (EC 2002) and WHO/FAO (WHO/FAO 2004a).

Table 16. Study 11 – Two-year feeding study with glyphosate in mice (Cheminova 1993b).

Study owner:	Cheminova (1993b)
Reliability/Justification:	1 Study performed according to GLP and OECD guideline requirements
Substance:	Glyphosate (98.6% pure)
Species/Strain:	Mouse/CD-1, groups of 50 ♂ and 50 ♀
Administration route:	Diet
Concentration:	♂+♀: 0, 100, 300, 1000 mg/kg bw/day (regular adjustment of dietary concentration)
Duration:	24 months
Findings:	≥1000 mg/kg bw/day: NOAEL (♂+♀) no treatment-related effects
Select neoplasms:	Bronchiolar-alveolar adenoma, bronchiolar-alveolar carcinoma, pituitary adenoma (females)

	Dose (mg/kg bw/day)			
Males	0	10	300	1000
Bronchiolar-alveolar adenoma – B	9/50 (18%)	15/50 (30%)	11/50 (22%)	13/50 (26%)
Bronchiolar-alveolar carcinoma – M	10/50 (20%)	7/50 (14%)	8/50 (16%)	9/50 (18%)
	Dose (mg/kg bw/day)			
Females	0	100	300	1000
Bronchiolar-alveolar adenoma – B	7/50 (14%)	3/50 (6%)	3/50 (6%)	6/50 (12%)
Bronchiolar-alveolar carcinoma – M	3/50 (6%)	2/50 (4%)	1/50 (2%)	5/50 (10%)
Pituitary adenoma – B	1/41 (2%)	0/32	0/23	3/43 (6%)

B benign, *M* malignant

Study 12 (Arysta Life Sciences 1997a)

An 18-month feeding study in ICR-CD-1 mice, conducted between February 1995 and September 1996, investigated higher doses by admixing 1600, 8000, or 40 000 ppm glyphosate into the diet fed to groups of 50 male and 50 female mice per dose (Arysta Life Sciences 1997a). The calculated test substance intake was 165/153, 838/787, and 4348/4116 mg/kg bw/day (males/females, Table 17), exceeding the limit dose. This study was rated Klimisch 1 for reliability.

Histopathological examinations did not show statistically significant increases for any type of neoplastic lesion in all treatment groups of both sexes (see data Supplementary Study 12 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423). Select neoplasms evaluated across the data set with some non-statistically significant increases above concurrent controls included lymphoma and lung tumors, all of which lacked a clear dose-response. Glyphosate was considered not carcinogenic in CD-1 mice up to doses well in excess of the limit dose for carcinogenicity testing, which is consistent with the recent evaluation in Europe under the Annex I Renewal of glyphosate (Germany Rapporteur Member State 2015b).

Study 13 (Feinchemie Schwebda 2001)

An 18-month feeding study in Swiss albino mice (Feinchemie Schwebda 2001), conducted between December 1997 and June 1999, featured treatment groups, each with 50 animals per sex, receiving 100, 1000, and 10 000 ppm technical grade glyphosate

Table 17. Study 12 – Two-year feeding study with glyphosate in mice (Arysta Life Sciences 1997a).

Study owner:	Arysta Life Sciences (1997b)
Reliability/ Justification:	1 Study performed according to GLP and OECD guideline requirements, with no deviations.
Substance:	Glyphosate (94.6–97.6% pure)
Species/Strain	Mouse/CD-1, groups of 50 ♂ and 50 ♀
Administration route:	Diet
Concentration:	0, 1600, 8000, or 40 000 ppm diet (♂ about 0, 165, 838, 4348 mg/kg bw/day; ♀ about 0, 153, 787, 4116 mg/kg bw/day)
Duration:	18 months
Findings:	8000/1600 ppm diet: NOAEL (♂/♀) 8000 ppm diet (♀): retarded growth 40 000 ppm diet: pale-colored skin ♂, loose stool, retarded growth, reduced food consumption and food efficiency, cecum distension and increased absolute and relative cecum weight, without histopathological findings of increased incidence of anal prolapse, consistent with histopathological erosion/ulcer of the anus
Select neoplasms:	Lung adenoma, lung adenocarcinoma, lymphoma

	Dose (mg/kg bw/day)			
Males	0	165	838	4348
Lung adenoma – B	8/50 (16%)	14/50 (28%)	13/50 (26%)	11/50 (11%)
Lung adenocarcinoma – M	1/50 (2%)	1/50 (2%)	6/50 (12%)	4/50 (8%)
Lymphoma – M	2/50 (4%)	2/50 (4%)	0/50	6/50 (12%)
	Dose (mg/kg bw/day)			
Females	0	153	787	4116
Lung adenoma – B	8/50 (16%)	5/50 (10%)	12/50 (24%)	5/50 (10%)
Lung adenocarcinoma – M	1/50 (2%)	2/50 (4%)	3/50 (6%)	1/50 (2%)
Lymphoma – M	6/50 (12%)	4/50 (8%)	8/50 (16%)	7/50 (14%)

B benign, *M* malignant

DOI 10.3109/10408444.2014.1003423

Table 18. Study 13–18-Month feeding study with glyphosate in mice (Feinchemie Schwebda 2001).

Study owner:	Feinchemie Schwebda (2001)
Reliability/Justification	2 Study performed according to GLP and OECD guideline requirements, with no deviations, but possible viral infection may have confounded interpretation of results
Substance:	Glyphosate (>95% pure)
Species/Strain	Mouse/Swiss albino, groups of 50 ♂ and 50 ♀
Administration route:	Diet
Concentration:	0, 100, 1000, 10 000 ppm diet (♂ about 0, 14.5, 150, 1454 mg/kg bw/day; ♀ about 0, 15.0, 151, 1467 mg/kg bw/day)
Duration:	18 months
Findings:	1000 ppm diet: NOAEL (♂+♀) 10 000 ppm diet (♂+♀): increased mortality
Select neoplasms:	Bronchiolar/alveolar adenoma, lymphoma

	Historical controls		Dose (mg/kg bw/day)			
			0	14.5	150	1454
Males						
Mortality	[§]11/50–27/50		[+]22/50 (6)	20/50 (6)	22/50 (8)	27/50 (8)
Findings for dead and moribund sacrificed animals						
Lymphoma – M	[#]20/75	26.7% [0–44]	9/22 (41.0%)	*12/20 (60.0%)	*13/22 (59.0%)	13/27 (48.0%)
Findings in animals sacrificed at termination						
Lymphoma – M	26/175	14.9% [8–24]	1/28 (3.6%)	3/30 (10.0%)	3/28 (10.7%)	*6/23 (26.1%)
Total animals						
Lymphoma – M	46/250	18.4% [6–30]	10/50 (20.0%)	15/50 (30.0%)	16/50 (32.0%)	*19/50 (38.0%)
	Historical controls		Dose (mg/kg bw/day)			
			0	15.0	151	1467
Females						
Mortality	12/50–20/50		16/50 (7)	16/50 (7)	20/50 (2)	20/50 (3)
Findings for dead and moribund sacrificed animals						
Bronchiolar/alveolar adenoma – B	–	–	0/16	0/16	1/20 (5%)	2/20 (10%)
Lymphoma – M	49/77	63.6% [0–100]	9/16 (56.0%)	10/16 (63.0%)	13/20 (65.0%)	12/20 (60.0%)
Findings in animals sacrificed at termination						
Bronchiolar/alveolar adenoma – B			1/34 (3%)	0/0	1/1 (100%)	1/30 (3%)
Lymphoma – M	50/175	28.9% [2043]	9/34 (26.5%)	10/30 (29.4%)	6/30 (20.0%)	*13/28 (43.3%)
Total animals						
Bronchiolar/alveolar adenoma – B			1/50 (2%)	0/16	2/21 (10%)	3/50 (6%)
Lymphoma – M	99/250	39.6% [1458]	18/50 (36.0%)	20/50 (40.0%)	19/50 (38.0%)	*25/50 (50.0%)

B benign, *M* malignant.
[§]Nine studies, performed by the same laboratory in the timeframe encompassing the study summarized here.
[+](Number of animals killed in extremis).
[#]Five studies, conducted in the same laboratory between 1996 and 1999.
*Statistically higher than concurrent controls ($p < 0.05$).

in the diet. Control mice received a plain diet. The calculated test substance intake was 14.5/15.0, 150/151, 1454/1467 mg/kg bw/day (males/females, Table 18), exceeding the limit dose, as reflected in elevated mortality in the high dose groups. This study was rated Klimisch 2 for reliability, based on speculation of a viral infection within the colony, discussed below.

Based on the slightly higher mortality and lower survival rates in the high dose groups, the NOAEL was considered 1000 ppm (151 mg/kg bw/day). There were no treatment-related effects on clinical signs, behavior, eyes, body weight, body weight gain, food consumption, and differential white blood cell counts in both sexes. Gross pathology, organ weight data, and histopathological examination demonstrated no treatment-related effects. An increase in the number of malignant lymphomas, the most common spontaneously occurring tumor category in the mouse, was statistically significant in the high-dose groups compared to controls (Table 18). The Germany Rapporteur Member State concluded that the malignant lymphoma increase in high-dose males was inconclusive but unrelated to treatment in the context of similar higher dosed studies (Germany Rapporteur Member State 2015b), and considered this endpoint irrelevant to carcinogenic risk in humans (Germany Rapporteur Member State 2015a). Whether or not a viral component (Taddesse-Heath et al. 2000) may have contributed to this endpoint, the finding was considered incidental background variation based on historical control data, and in agreement with the study director. As in Study 11, bronchiolar-alveolar adenoma was also considered a select neoplasm for evaluation in the broader data set (Tables 22 and 23), and as previously discussed, demonstrates a lack of dose-response across doses ranging from approximately 15 mg/kg bw/day to 5000 mg/kg bw/day. Summary tables of all histopathological neoplastic findings are available (see data Supplementary Study 13 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423).

Technical grade glyphosate was reported as not carcinogenic in Swiss albino mice, following continuous dietary exposure of up to 1460 mg/kg bw/day (average for both sexes) for 18 months. The NOAEL for general chronic toxicity was 151 mg/kg bw/day for both sexes combined.

Study 14 (Nufarm 2009a)

The most recent mouse carcinogenicity assay was conducted between October 2005 and November 2007 (Nufarm 2009a).

Table 19. Study 14–18-Month feeding study with glyphosate in mice (Nufarm 2009a).

Study owner:	Nufarm (2009b)			
Reliability/Justification:	1 Study performed according to GLP and OECD guideline requirements, with no deviations			
Substance:	Glyphosate (94.6–97.6% pure)			
Species/Strain:	mouse/CD-1, groups of 51 ♂ and 51 ♀			
Administration route:	Diet			
Concentration:	0, 500, 1500, and 5000 ppm diet (♂ about 0, 0, 71.4, 234, 810 mg/kg bw/day; ♀ about 0, 97.9, 300, 1081 mg/kg bw/day)			
Duration:	18 months			
Findings:	≥ 5000 ppm diet: NOAEL (♂/♀) No treatment-related effects			
Select neoplasms:	Bronchiolar-alveolar adenoma, Bronchiolar-alveolar adenocarcinoma, hepatocellular adenoma (males), hepatocellular carcinoma (males), lymphoma, pituitary adenoma (females)			
		Dose (mg/kg bw/day)		
Males	0	157	814	4841
Bronchiolar-alveolar adenoma – B	9/51 (18%)	7/51 (14%)	9/51 (18%)	4/51 (8%)
Bronchiolar-alveolar adenocarcinoma – M	5/51 (10%)	5/51 (10%)	7/51 (14%)	11/51 (22%)
Hepatocellular adenoma – B	1/51 (2%)	1/51 (2%)	4/51 (8%)	2/51 (4%)
Hepatocellular carcinoma – M	6/51 (12%)	11/51 (22%)	7/51 (14%)	4/51 (8%)
Lymphoma – M	0/51	1/50 (2%)	2/51 (4%)	5/51 (10%)
		Dose (mg/kg bw/day)		
Females	0	190	955	5873
Bronchiolar-alveolar adenoma – B	2/51 (4%)	4/51 (8%)	2/51 (4%)	2/51 (4%)
Bronchiolar-alveolar adenocarcinoma – M	5/51 (10%)	2/51 (4%)	2/51 (4%)	3/51 (6%)
Lymphoma – M	11/51 (22%)	8/51 (16)	10/51 (20%)	11/51 (22%)
Pituitary adenoma – B	0/51	1/50 (2%)	0/51	2/51 (4%)

B benign, *M* malignant

Groups of 51 CD-1 mice per sex received daily dietary doses of 0, 500, 1500, and 5000 ppm technical grade glyphosate (equivalent to an average intake of 85, 267 and 946 mg/kg bw/day, Table 19). The MTD was apparently not reached in the high-dose group, which is more indicative of low general toxicity of the test substance rather than a flaw in the study design. The NOAEL for chronic toxicity was 810 mg/kg bw/day for male mice and 1081 mg/kg bw/day for female mice, the highest dosage tested. Despite not quite achieving a limit dose in males, this study was arguably rated Klimisch 1 for reliability.

Several increases in common spontaneous mouse neoplasms in male mice were noted. Non-dose-response increases were noted for hepatocellular adenoma and carcinoma in males, and dose-responses were noted for bronchiolar-alveolar adenocarcinoma and malignant lymphoma in males, but not females. Pituitary adenoma incidences were low, and considered incidental in low and high-dose females, although they were slightly higher than controls (Table 19). These neoplasms were all evaluated in context of the broader data set (Tables 22 and 23). The summary of neoplastic findings is available (see data Supplementary Study 14 to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.1003423).

Glyphosate was considered not carcinogenic in the CD-1 mice, following continuous average dietary exposure for males and females, to quantities up to 945.6 mg/kg bw/day for 18 months, which is consistent with the recent evaluation in Europe under the Annex I Renewal of glyphosate (Germany Rapporteur Member State 2015b).

Discussion

An extraordinarily large volume of animal data has been compiled to evaluate the carcinogenic potential of glyphosate. The expected normal biological variability for spontaneous tumor formation is reflected across this extensive data set (Tables 20–23). However, no specific neoplasm stands out as a consequence of glyphosate exposures. While some individual studies may note an increase in a specific neoplasm at the high dose, the pooled data fail to identify any consistent pattern of neoplasm formation, demonstrating that the effect is not reproducible and not treatment-related. The lack of a dose-response across the several orders of magnitude suggests that no individual tumor of single etiology is attributable to glyphosate administration.

Glyphosate has undergone repeated and extensive review by the United States Environmental Protection Agency (US EPA 1993), the European Union (EC 2002, Germany Rapporteur Member State 2015b) and the World Health Organization/Food and Agriculture Organization of the United Nations (WHO/FAO 2004b, WHO/FAO 2004a). With regard to potential carcinogenic effects of glyphosate, the unanimous outcome of these reviews has been that the data provide sufficient evidence to conclude that glyphosate should not be considered a carcinogen. Genotoxicity studies with glyphosate, conducted under conditions stipulated by internationally accepted testing guidelines and GLP, as reviewed in 2000 (Williams et al. 2000) and recently updated (Kier and Kirkland 2013), indicate that glyphosate clearly does not exhibit the properties of a DNA-reactive genotoxic carcinogen. This lack of mutagenicity rules out an important concern for carcinogenicity.

Mink et al. published a review of the available epidemiological studies that investigated possible associations between glyphosate and cancer diagnosed in humans (Mink et al. 2012). No evidence was found for a statistically significant positive association between cancer and exposure to glyphosate. While one Agricultural Health Study (AHS) publication mentions a "suggested association" between glyphosate use and multiple myeloma (De Roos et al. 2005), a later summary of AHS

DOI 10.3109/10408444.2014.1003423

Table 20. Summary of select neoplasms in male rats (Studies 1–8).

	Tumor Incidence/number of animals examined, by dose (mg/kg bw/day)												
Select neoplasm	Controls – 0 [% range for studies]	[a]3	[d]7.4	[a]10	[c]10	[a]31	[d]73.9	[h]86	[b]89	[c]100	[f]104	[g]121
Pancreas islet cell adenoma	20/397 [0–14]	5/49	0/30	2/50	1/24	2/50	0/32	1/51	8/57	2/17	1/75	2/64
Pituitary adenoma	153/398 [6–57]	19/49	4/30	20/48	12/24	18/47	3/31	11/51	32/58	8/19	41/75	17/63
Pituitary carcinoma	4/98 [2–6]	2/49	NF	3/48	1/24	1/47	NF	NF	NF	0/19	NF	NF
Testes interstitial cell (Leydig)	14/447 [0–8]	3/50	0/37	1/50	1/25	6/50	2/32	3/51	0/60	0/19	2/75	2/63
Thyroid C cell adenoma	35/391 [4–18]	1/49	0/26	0/49	1/21	2/49	1/29	#1/51	5/58	1/17	10/74	#1/63
Hepatocellular adenoma	30/351 [0–48]	NF	22/50	NF	1/50	NF	10/48	2/51	2/60	1/49	0/75	2/64
Hepatocellular carcinoma	22/384 [0–42]	0/50	28/50	1/50	1/50	2/50	18/48	0/51	2/60	1/49	1/75	NF
Benign keratoacanthoma (skin)	8/250 [2–5]	NF	NF	NF	NF	NF	NF	3/51	3/60	NF	3/75	0/64

	Tumor Incidence/number of animals examined, by dose (mg/kg bw/day)												
Select neoplasm	[c]150	[b]285	[c]300	[f]354	[g]361	[b]362	[d]740.6	[e]780	[b]940	[c]1000	[h]1077	[f]1127	[g]1214	[e]1290
Pancreas islet cell adenoma	NF	2/51	2/21	1/80	0/64	5/60	1/49	NF	7/59	1/49	1/51	1/78	1/64	NF
Pituitary adenoma	NF	10/51	7/21	33/80	18/64	34/58	5/49	NF	32/59	17/50	20/51	42/78	19/63	NF
Pituitary carcinoma	NF	NF	1/21	NF	NF	NF	NF	NF	NF	0/50	NF	NF	NF	NF
Testes interstitial cell (Leydig)	1/49	1/51	0/21	0/80	2/63	3/60	3/50	2/49	2/60	2/50	1/51	2/78	2/64	0/47
Thyroid C cell adenoma	NF	#0/51	2/21	5/79	#1/63	8/58	1/50	NF	7/60	8/49	#3/51	6/78	#0/64	NF
Hepatocellular adenoma	NF	0/51	2/50	2/80	0/64	3/60	21/50	NF	8/60	2/50	1/51	1/78	5/64	NF
Hepatocellular carcinoma	1/49	0/51	0/50	2/80	NF	1/60	24/50	0/49	2/60	0/50	0/51	1/78	NF	0/47
Benign keratoacanthoma (skin)	NF	0/51	NF	0/80	1/64	4/60	NF	NF	5/59	NF	6/51	7/78	1/63	NF

[a]Study 1 (Monsanto) (CD) SD rats, rated unreliable for carcinogenicity evaluation.
[b]Study 2 (Monsanto) (CD) SD rats, including interim sacrifice groups.
[c]Study 3 (Cheminova) SD rats.
[d]Study 4 (Feinchemie Schwebda) Wistar rats.
[e]Study 5 (Excel) SD rats, rated unreliable for carcinogenicity evaluation.
[f]Study 6 (Arysta Life Sciences) Crj:CD SD rats, including interim sacrifice groups.
[g] Study 7 (Syngenta) Alpk:AP$_f$SD Wistar rats, including interim sacrifice groups.
[h] Study 8 (Nufarm) Wistar Han Crl:WI rats.
[#]Recorded as parafollicular adenoma.
NF not found/not reported

Table 21. Summary of select neoplasms in female rats (Studies 1–8).

	Tumor Incidence/number of animals examined, by dose (mg/kg bw/day)												
Select neoplasm	Controls – 0 [% range for studies]	[a]3	[d]7.4	[c]10	[a]11	[a]34	[d]73.9	[c]100	[b]105	[b]113	[f]115	[g]145
Pancreas islet cell adenoma	11/397 [0–9]	1/50	0/23	2/27	1/50	0/49	0/16	2/29	0/51	1/60	2/79	0/63
Pituitary adenoma	246/397 [14–78]	29/48	13/33	19/28	31/50	26/49	7/23	19/29	23/51	48/60	54/79	44/63
Pituitary carcinoma	16/155 [2–17]	7/48	NF	5/28	5/50	12/49	NF	5/28	NF	0/60	NF	NF
Thyroid C cell adenoma	25/302 [3% – 16%]	3/49	0/24	1/27	6/50	3/47	1/17	1/29	# 1/51	2/60	7/78	# 0/63
Hepatocellular adenoma	22/302 [0–36]	NF	18/48	1/50	NF	NF	19/49	3/50	0/51	2/60	1/79	0/64
Hepatocellular carcinoma	14/210 [0–20]	0/50	15/48	0/50	0/50	2/50	14/49	0/50	0/51	0/60	NF	NF
Mammary gland fibroadenoma	113/384 [6–58]	16/46	NF	12/28	20/48	16/44	NF	17/29	9/51	[$]24/54	30/79	4/63
Mammary gland adenocarcinoma	40/334 [2–22]	6/46	0/30	NF	5/48	8/44	0/33	NF	3/51	[~]10/54	8/79	0/63

	Tumor Incidence/number of animals examined, by dose (mg/kg bw/day)												
Select neoplasm	[c]210	[c]300	[b]349	[f]393	[g]437	[b]457	[d]740.6	[c]1000	[e]1060	[h]1183	[f]1247	[h]1382	[g]1498	[e]1740
Pancreas islet cell adenoma	NF	2/29	0/51	1/78	1/64	4/60	1/49	1/49	NF	0/59	1/78	0/51	0/64	NF
Pituitary adenoma	NF	25/30	16/51	47/77	46/63	46/60	6/50	34/49	NF	34/59	52/78	32/51	49/64	NF
Pituitary carcinoma	NF	2/30	NF	NF	NF	0/60	NF	7/49	NF	1/59	NF	NF	NF	NF
Thyroid C cell adenoma	NF	2/29	#1/50	8/76	#0/64	6/60	1/47	7/49	NF	6/60	4/78	# 0/51	# 2/64	NF
Hepatocellular adenoma	NF	1/50	1/51	0/78	1/64	6/60	13/50	2/50	NF	1/60	0/78	1/51	0/64	NF
Hepatocellular carcinoma	NF	0/50	1/51	NF	NF	1/60	9/50	0/50	NF	2/60	NF	0/51	NF	NF
Mammary gland fibroadenoma	1/22	19/30	7/51	27/77	6/64	[$]27/59	NF	29/50	5/22	[$]28/57	30/78	5/51	5/64	5/50
Mammary gland adenocarcinoma	0/22	NF	1/51	11/77	0/64	[~]14/59	0/48	NF	0/22	[~]9/57	8/78	6/51	2/64	0/50

[a]Study 1 (Monsanto) (CD) SD rats, rated unreliable for carcinogenicity evaluation.
[b]Study 2 (Monsanto) (CD) SD rats, including interim sacrifice groups.
[c]Study 3 (Cheminova) SD rats.
[d]Study 4 (Feinchemie Schwebda) Wistar rats.
[e]Study 5 (Excel) SD rats, rated unreliable for carcinogenicity evaluation.
[f]Study 6 (Arysta Life Sciences) Crj:CD SD rats, including interim sacrifice groups.
[g]Study 7 (Syngenta) Alpk:AP$_f$SD Wistar rats, including interim sacrifice groups.
[h]Study 8 (Nufarm) Wistar Han Crl:WI rats.
[$]Recorded as adenoma/adenofibroma/fibroma.
[~]Recorded as carcinoma/adenocarcinoma.
NF not found/not reported.

Table 22. Summary of select neoplasms in male mice (Studies 10–14).

Select neoplasm	Tumor Incidence/number of animals examined, by dose (mg/kg bw/day)							
	Controls – 0 [% range for studies]	[d]14.5	[c]85	[b]100	[d]150	[a]157	[c]165	[c]267
Bronchiolar-alveolar adenoma	31/249 [10–18]	2/22	[§]7/51	15/50	0/22	9/50	[§]14/50	[§]9/51
Bronchiolar-alveolar adenocarcinoma	10/149 [2–10]	NF	[§]5/51	NF	NF	3/50	[§]1/50	[§]7/51
Bronchiolar-alveolar carcinoma	10/100 [0–20]	0/22	NF	7/50	0/22	NF	NF	NF
Hepatocellular adenoma	27/250 [0–28]	5/25	1/51	12/50	3/28	0/50	15/50	4/51
Hepatocellular carcinoma	15/250 [0–16]	0/25	11/51	5/50	0/28	0/50	1/50	7/51
Malignant lymphoma	16/205 [0–100]	15/50	1/51	2/4	16/50	[#]5/50	2/50	2/51
Myeloid leukemia	3/101 [0–6]	1/50	1/51	NF	1/50	NF	NF	0/51
	Tumor Incidence/number of animals examined, by dose (mg/kg bw/day)							
Select neoplasm	[b]300	[a]814	[c]838	[c]946	[b]1000	[d]1454	[c]4348	[a]4841
Bronchiolar-alveolar adenoma	11/50	9/50	[§]13/50	[§]4/51	13/50	1/50	[§]11/50	9/50
Bronchiolar-alveolar adenocarcinoma	NF	2/50	[§]6/50	[§]11/51	NF	NF	[§]4/50	1/50
Bronchiolar-alveolar carcinoma	8/50	NF	NF	NF	9/50	1/50	NF	NF
Hepatocellular adenoma	11/50	1/50	15/50	2/51	9/50	3/50	7/50	0/50
Hepatocellular carcinoma	6/50	0/50	3/50	4/51	7/50	2/50	1/50	2/50
Malignant lymphoma	1/1	[#]4/50	0/50	5/51	6/8	19/50	6/50	[#]2/50
Myeloid leukemia	NF	NF	NF	0/51	NF	1/50	NF	NF

[a]Study 10 (Monsanto) CD-1 mice.
[b]Study 11 (Cheminova) CD-1 mice.
[c]Study 12 (Arysta Life Science) CD-1 mice.
[d]Study 13 (Feinchemie Schwebda) Swiss albino mice.
[e]Study 14 (Nufarm) CD-1 mice.
[§]Recorded as lung rather than bronchiolar-alveolar.
[#]Recorded as sum of malignant lymphoblastic lymphosarcoma with leukemia, lymphoblastic lymphosarcoma without leukemia and composite lymphosarcoma.
[§]Recorded as lymphoblastic lymphosarcoma with leukemia.
NF not found/not reported.

results note that there were no associations between glyphosate use and a number of cancers, including lymphohematopoietic cancers, leukemia, NHL, and multiple myeloma (Weichenthal et al. 2010). A subsequent reanalysis of AHS data obtained under the Freedom of Information Act notes no suggestion of an association between glyphosate use and multiple myeloma, with a relative risk of 1.1 and 95% and a confidence interval of 0.5–2.9 (Sorahan 2012). A recent review paper (Alavanja et al. 2013) cites another epidemiology study claiming an association between glyphosate use and NHL (Eriksson et al. 2008), but this research is strongly criticized in the recent Reevaluation Assessment Report for glyphosate Annex I Renewal in Europe (Germany Rapporteur Member State 2015b), highlighting potential referral bias, selection bias, uncontrolled confounding, limited data usage contrary to claims of including all new cases (living cases only, rather than living

Table 23. Summary of select neoplasms in female mice (Studies 10–14).

Select neoplasm	Tumor incidence/number of animals examined, by dose (mg/kg bw/day)							
	Controls – 0 [% range for studies]	[d]15.0	[c]85	[b]100	[d]151	[c]153	[a]190	[c]267
Bronchiolar-alveolar adenoma	28/250 [2–20]	0/16	[§]4/51	3/49	2/21	[§]5/50	9/50	[§]2/51
Bronchiolar-alveolar adenocarcinoma	2/99 [2]	NF	[§]2/51	NF	NF	[§]2/50	3/50	[§]2/51
Bronchiolar-alveolar carcinoma	9/151 [2–10]	0/16	NF	2/49	0/20	NF	NF	NF
Malignant lymphoma	54/215 [10–100]	20/50	8/51	12/15	19/50	4/50	[#]6/50	10/51
Myeloid leukemia	2/156 [0–4]	1/50	0/51	NF	2/50	0/50	NF	1/51
Pituitary adenoma	1/232 [0–2]	0/16	1/51	0/32	0/17	1/50	0/21	0/51
	Tumor incidence/number of animals examined, by dose (mg/kg bw/day)							
Select neoplasm	[b]300	[c]787	[d]946	[a]955	[b]1000	[d]1467	[c]4116	[a]5874
Bronchiolar-alveolar adenoma	3/50	[§]12/50	[§]2/51	10/49	6/50	3/50	[§]5/50	1/50
Bronchiolar-alveolar adenocarcinoma	NF	[§]3/50	[§]3/51	4/49	NF	NF	[§]1/50	4/50
Bronchiolar-alveolar carcinoma	1/50	NF	NF	NF	5/50	0/50	NF	NF
Malignant lymphoma	9/12	8/50	11/51	[#]6/50	13/14	25/50	7/50	[#]10/50
Myeloid leukemia	NF	0/50	0/51	NF	NF	1/50	1/50	NF
Pituitary adenoma	0/23	0/50	2/51	0/44	[~]3/50	1/48	0/50	0/37

[a]Study 10 (Monsanto) CD-1 mice.
[b]Study 11 (Cheminova) CD-1 mice.
[c]Study 12 (Arysta Life Science) CD-1 mice.
[d]Study 13 (Feinchemie Schwebda) Swiss albino mice.
[e]Study 14 (Nufarm) CD-1 mice.
[§]Recorded as lung rather than bronchiolar-alveolar.
[#]Recorded as sum of lymphoblastic lymphosarcoma with leukemia, lymphoblastic lymphosarcoma without leukemia and composite lymphosarcoma.
[~]2 animals in anterior lobe, 1 animal in intermediate lobe.
NF not found/not reported.

DOI 10.3109/10408444.2014.1003423

plus dead), and questionable definition/interpretation of dose-response. It is important to note that the Eriksson et al. study did detect statistically significant positive associations for small lymphocytic lymphoma/chronic lymphocytic leukemia and "unspecified NHL", while the following lymphomas were not statistically significantly associated with glyphosate use: B-cell lymphomas, grade I-III follicular lymphoma, diffuse large B-cell lymphoma, other specified B-cell lymphomas, unspecified B cell lymphomas, and T-cell lymphomas (Eriksson et al. 2008). As previously discussed, statistically significant associations need to be evaluated further for study bias, confounders and sampling error, before expending resources and energy on further evaluation of potential causality.

Epidemiological investigations face the difficulty of reliably determining the magnitude of exposure to the chemical in question, while ruling out confounders like co-exposure to other chemicals, and environmental and lifestyle factors. In contrast, carcinogenicity studies in experimental animals, when conducted according to appropriate testing guidelines, are designed in a fashion that allows a direct association between observed effects and substance exposure, yet the relevance of observed findings to humans is an important consideration. This manuscript collectively presents the scientific community with carcinogenicity results from a remarkably large body of data from fourteen long-term carcinogenicity studies on glyphosate.

Glyphosate is of very low acute toxicity with an oral LD_{50} in the rat in excess of 5000 mg/kg of body weight.. The sub-chronic NOAEL is 400 mg/kg bw/day, and is based on effects that do not impair long-term survival (WHO/FAO 2004b, WHO/FAO 2004a). This allows administration of very high glyphosate doses to rodents for a prolonged time. Dietary levels of up to 30 000 and 40 000 milligrams of glyphosate per kilogram of diet have been administered to rats and mice, respectively, in chronic feeding studies covering their expected lifespan without apparent effects on longevity.

One of the most critical aspects of designing a carcinogenicity study is the choice of dose levels, especially the top dose, at either the limit dose or MTD. The relevant OECD TGs 451 and 453 for carcinogenicity studies propose a body weight depression of approximately 10% as evidence for systemic toxicity. This is equivalent to the concept of the MTD, which is discussed in a supporting OECD guidance document (OECD 2012b). For chemicals which are well tolerated by the experimental animal, where no dose-limiting toxicity is observed, the respective OECD guidance suggests 1000 mg/kg bw/day as the highest dose level (OECD 2012a). Many of the carcinogenicity studies performed in rats and mice with glyphosate have been conducted with the high dose group receiving levels of glyphosate at, or in excess of the limit dose because of its very low toxicity following repeat exposure. Following this extensive testing, even at very high exposure levels, there was no evidence of a carcinogenic effect related to glyphosate treatment. The select neoplasms highlighted in Tables 20–23 show normal biological background levels of spontaneous neoplasms, with lack of dose-response across the data sets. The combined studies clearly indicate that glyphosate's carcinogenic potential is extremely low or non-existent in animal models up to very high doses.

By way of comparison, the worst-case calculated human dietary exposure to glyphosate, the Theoretical Maximum Daily Intake (TMDI) is 0.14 mg/kg bw/day (EFSA 2012). Systemic exposure of operators, as assessed for the EU reapproval of glyphosate, is predicted to be between 0.0034 (German BBA model, tractor-mounted ground-boom sprayer) and 0.226 mg/kg bw/day (UK POEM, hand-held-spraying to low targets, data not shown). The model estimates are supported by human biomonitoring data in farmers showing systemic exposures of 0.004 and 0.0001 mg/kg/day for worst-case and mean acute doses, respectively (Acquavella et al. 2004). The high doses in chronic rodent studies at which no evidence of carcinogenicity is demonstrated are at least hundreds of thousands fold greater than peak human systemic exposure levels. Clearly, there is no scientific basis for concern of carcinogenic risk to humans resulting from glyphosate exposure.

With over 40 years of scientific research on glyphosate, no compelling evidence exists for a mechanism for glyphosate to cause cancer. Mammalian metabolism does not activate glyphosate to a toxic metabolite (Anadon et al. 2009, WHO/FAO 2004a). The lack of glyphosate DNA reactivity supports the

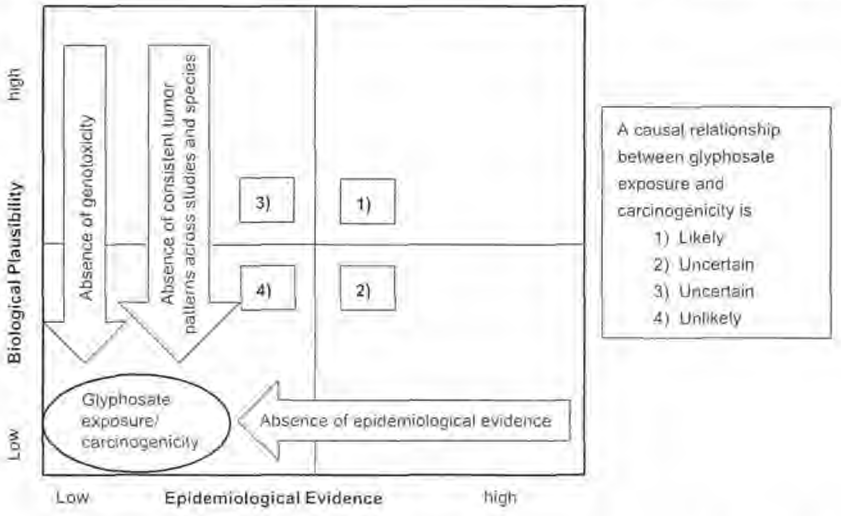


Figure 2. Likelihood of glyphosate carcinogenicity based on experimental and epidemiological data: a causal inference grid as proposed by Adami et al. (2011) to utilize both toxicological and epidemiological data.

lack of potential for an initiation event for carcinogenesis (Kier and Kirkland 2013). Clearly, there is a lack of potential for glyphosate to induce hormonal oncogenesis, based on both the tumor incidence data presented and the unequivocal evidence that glyphosate is not an endocrine disruptor (Bailey et al. 2013, Levine et al. 2012, Saltmiras and Tobia 2012, Webb et al. 2013, Williams et al. 2012).

The absence of test substance-related neoplastic findings in a total of 14 rodent cancer bioassays with glyphosate is in stark contrast to the recent dramatic media reports, internet postings, and YouTube videos of rat tumors, hypothesized to be caused by treatment with maize containing glyphosate residue or drinking water spiked with a glyphosate formulation (Seralini et al. 2014). Such reports, under the scrutiny of the global scientific community, demand greater data transparency and accountability within the peer review process.

The absence of a glyphosate-related mechanism for carcinogenesis, the huge volume of genotoxicity data studies indicating no likely mutagenic or DNA-reactive potential (Kier and Kirkland 2013), combined with the lack of epidemiological evidence for glyphosate-induced cancer (Mink et al. 2012), and the lack of carcinogenicity in multiple rodent carcinogenicity assays, are depicted in a causal inference grid in Figure 2, as put forth by Adami et al. (Adami et al. 2011). The overwhelming weight of the available evidence, demonstrating a lack of both biological plausibility and epidemiological effects, draws a compelling conclusion that glyphosate's carcinogenic potential is extremely low or non-existent.

Acknowledgements

Special thanks go to Elizabeth Webb, Monsanto Toxicologist, for her detailed attention to document and data table formatting and as the reference library curator. Quality control and review of data transcription were valued services provided by Carrie Leigh Logan and Apatna Desai Nemali, Monsanto Quality Assurance Specialists.

Declaration of interest

The employment affiliation of the authors is as shown on the cover page. Volker Mostert was an employee of the consulting group, Dr. Knoell Consult GmbH, involved in the preparation of the recent glyphosate Annex I Renewal dossier for the Glyphosate Task Force (GTF; a consortium of European glyphosate registrants http://www.glyphosatetaskforce.org/). Helmut Greim was funded as an independent consultant for his expert contributions to this manuscript. David Saltmiras and Christian Strupp are employed by member companies of the GTF, Monsanto and ADAMA Agriculture B.V. (formerly Feinchemie Schwebda GmbH) respectively. David Saltmiras is also Chair of the Toxicology Technical Working Group of the GTF. Christian Strupp is an expert member of the Toxicology Technical Working Group of the GTF. Monsanto Company was the original producer and marketer of glyphosate formulations. The authors had sole responsibility for the writing and content of the paper and the interpretations and opinions expressed in the paper are those of the authors and may not necessarily be those of the member companies of the Glyphosate Task Force.

References

Acquavella JF, Alexander BH, Mandel JS, Gustin C, Baker B, Chapman P, Bleeke M. (2004). Glyphosate biomonitoring for farmers and their families: Results from the farm family exposure study. Environ Health Perspect, 112, 321–6.

Adami HO, Berry SC, Breckenridge CB, Smith LL, Swenberg JA, Trichopoulos D, et al. (2011). Toxicology and epidemiology: improving the science with a framework for combining toxicological and epidemiological evidence to establish causal inference. Toxicol Sci, 122, 223–34.

Alavanja MC, Ross MK, Bonner MR (2013). Increased cancer burden among pesticide applicators and others due to pesticide exposure. CA Cancer J Clin, 63, 120–42.

Anadon A, Martinez-Larranaga MR, Martinez MA, Castellano VJ, Martinez M, Martin MT, et al. (2009). Toxicokinetics of glyphosate and its metabolite aminomethyl phosphonic acid in rats. Toxicol Lett, 190, 91–5.

Antoniou M, Habib M, Howard C, Jennings R, Leifert C, Nodari R, et al. (2012). Teratogenic effects of glyphosate-based herbicides: divergence of regulatory decisions from scientific evidence. J Environ Anal Toxicol, S4, 006.

Aris A, Leblanc S. (2011). Maternal and fetal exposure to pesticides associated to genetically modified foods in Eastern Townships of Quebec, Canada. Reprod Toxicol, 31, 528–33.

Aris A, Paris K. (2010). [Hypothetical link between endometriosis and xenobiotics-associated genetically modified food]. Gynecol Obstet Fertil, 38, 747–53.

Armitage JO. (1997). The changing classification of non-Hodgkin's lymphomas. CA Cancer J Clin, 47, 323–5.

Arysta Life Sciences. (1997a). HR-001: 18-Month Oral Oncogenicity Study in Mice, Tokyo, Japan: The Institute of Environmental Toxicology.

Arysta Life Sciences. (1997b). HR-001: 24-Month Oral Chronic Toxicity and Oncogenicity Study in Rats. Kodaira-shi, Tokyo, Japan: The Institute of Environmental Toxicology.

Arysta Life Sciences. (1997c). HR-001: 12-Month Oral Chronic Toxicity Study in Dogs. IET 94-0157

Bailey J, Hauswirth J, Stump D. (2013). No evidence of endocrine disruption by glyphosate in male and female pubertal assays (Conference Abstract PS 1937). The Toxicologist (supplement to Toxicological Sciences), 132, 412.

Barale-Thomas E. (2013). The SFPT feels compelled to point out weaknesses in the paper by Seralini et al. (2012). Food Chem Toxicol, 53, 473–4.

Benachour N, Seralini GE. (2009). Glyphosate formulations induce apoptosis and necrosis in human umbilical, embryonic, and placental cells. Chem Res Toxicol, 22, 97–105.

Berry C. (2013). Adverse effects in a feeding study of a GM derived corn in rats. Food Chem Toxicol, 53, 445–6.

Berry CL. (2010). Relativism, regulation and the dangers of indifferent science: the Sir Roy Cameron lecture of the Royal College of Pathologists. Toxicology, 267, 7–13.

Bradford Hill A. (1965). The environment and disease: association or causation? Proc R Soc Med, 58, 295–300.

Bradley A, Mukaratirwa S, Petersen-Jones M. (2012). Incidences and range of spontaneous findings in the lymphoid and haemopoietic system of control Charles River CD-1 mice (Crl: CD-1(ICR) BR) used in chronic toxicity studies. Toxicol Pathol, 40, 375–81.

Brayton CF, Treuting PM, Ward JM. (2012). Pathobiology of aging mice and GEM: background strains and experimental design. Vet Pathol, 49, 85–105.

Brock WJ. (2014). Strengthening the peer review process for the International Journal of Toxicology. Int J Toxicol, 33, 351–2.

Chandra M, Riley MG, Johnson DE. (1992). Spontaneous neoplasms in aged Sprague-Dawley rats. Arch Toxicol, 66, 496–502.

Cheminova. (1990). Glyphosate: 52 Week Oral Toxicity Study in Dogs. Inveresk Research International Report No.: 7502. Cheminova report No.: 52 GLY.

Cheminova. (1993a). Glyphosate – 104 Week Combined Chronic Feeding/Oncogenicity Study in Rats with 52 Week Interim Kill (Results After 104 Weeks). Tranent, UK: Inveresk Research International, Ltd.

Cheminova. (1993b). Glyphosate: 104 week dietary carcinogenicity study in mice. Tranent, UK: Inveresk Research International, Ltd.

Chruscielska K, Brzezinski J, Kita K, Kalhorn D, Kita I, Graffstein B, et al. (2000a). Glyphosate – evaluation of chronic activity and possible

far-reaching effects. Part 1. Studies on chronic toxicity. Pestycydy (Warsaw), 3–4, 11–20.
Chruscielska K, Graffstein B, Szarapinska-Kwaszewska J, Brzezinski J, Kalhorn D. (2000b). Glyphosate: Evaluation of chronic activity and possible far-reaching effects. Part 2. Studies on mutagenic activity. Pestycydy (Warsaw) 3–4, 21–5.
De Roos AJ, Blair A, Rusiecki JA, Hoppin JA, Svec M, Dosemeci M, et al. (2005). Cancer incidence among glyphosate-exposed pesticide applicators in the Agricultural Health Study. Environ Health Perspect, 113, 49–54. doi 10.1289/ehp.7340.
de Souza L, Oda LM. (2013). Letter to the editor. Food Chem Toxicol, 53, 440.
EC. (2002). Review report for the active substance glyphosate. European Commission., Directorate E – Food Safety: plant health, animal health and welfare, international questions. E1-Plant health.
EFSA. (2012). Reasoned Opinion – Modification of the existing MRL for glyphosate in lentils. EFSA J, 10, 2250.
Eriksson M, Hardell L, Carlberg M, Akerman M. (2008). Pesticide exposure as risk factor for non-Hodgkin lymphoma including histopathological subgroup analysis. Int J Cancer, 123, 1657–63.
EU. (2008). Council Regulation (EC) No 440/2008. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri = OJ:L:2008:142:0001:0739:en:PDF. Accessed on November 13, 2012.
Everds NE, Snyder PW, Bailey KL, Bolon B, Creasy DM, Foley GL, et al. (2013). Interpreting stress responses during routine toxicity studies: a review of the biology, impact, and assessment. Toxicol Pathol, 41, 560–614.
Excel. (1997). Combined Chronic Toxicity/Carcinogenicity Study of Glyphosate Technical in Sprague Dawley Rats. Pune, India: Indian Institute of Toxicology.
Feinchemie Schwebda. (1996). Combined Chronic Toxicity and Carcinogenicity Study with Glyphosate Technical in Wistar Rats. Bangalore, India: Rallis India, Ltd.
Feinchemie Schwebda. (2001). Carcinogenicity Study with Glyphosate Technical in Swiss Albino Mice. Bangalore, India: Rallis India, Ltd.
Gad SC, Frith CH, Goodman DG, Boysen BG. (2008). Chapter 2. The Mouse. In: Gad, SC, Ed. Animal Models in Toxicology, Second Edition. Informa Health Care USA, Inc. ISBN-13: 978-0-8247-5407-5, ISBN-10: 0-8247-5407-7.
Gasnier C, Benachour N, Clair E, Travert C, Langlois F, Laurant C, et al. (2010). Dig1 protects against cell death provoked by glyphosate-based herbicides in human liver cell lines. J Occup Med Toxicol, 5, 29.
Germany Rapporteur Member State. (2015a). Glyphosate Renewal Assessment Report, Volume 1. Report and Proposed Decision. Revised 29th, January 2015.
Germany Rapporteur Member State. (2015b). Glyphosate Renewal Assessment Report, Volume 3, Annex B.6.1 Toxicology and Metabolism. Revised 29th, January 2015.
Germany Rapporteur Member State. (2015c). List of Endpoints. Active Substance Glyphosate. Report and Proposed Decision. Revised 29th, January 2015.
Gold LS, Sawyer CB, Magaw R, Backman GM, de Veciana M, Levinson R, et al. (1984). A carcinogenic potency database of the standardized results of animal bioassays. Environ Health Perspect, 58, 9–319.
Grunewald W, Bury J. (2013). Comment on "Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize" by Seralini et al. Food Chem Toxicol, 53, 447–8.
Hammond B, Goldstein DA, Saltmiras D. (2013). Response to original research article, in press, corrected proof, "Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize". Food Chem Toxicol, 53, 459–64.
Hayes AW. (2014). Reply to letter to the editor: Editor in Chief of *Food and Chemical Toxicology* answers questions on retraction. Food Chem Toxicol, 65, 394–5.
IUPAC. (1997). Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: http://goldbook.iupac.org (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. doi 10.1351/goldbook. (http://goldbook.iupac.org/M03771.html).
JMAFF. (2000). Data Requirements for Supporting Registration of Pesticides. Appendix, Section 2-1-15. Available at: http://www.acis.famic.go.jp/eng/shinsei/8147annex.pdf. Accessed on March 8, 2013.
Johnson MD, Gad SC. (2008). Chapter 3. The Rat. In: Gad, SC, Ed. Animal Models in Toxicology, Second Edition. Informa Health Care USA, Inc. ISBN-13: 978-0-8247-5407-5, ISBN-10: 0-8247-5407-7.
Kier LD, Kirkland DJ. (2013). Review of genotoxicity studies of glyphosate and glyphosate-based formulations. Crit Rev Toxicol, 43, 283–315.
Klimisch HJ, Andreae M, Tillmann U. (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul Toxicol Pharmacol, 25, 1–5.
Langridge P. (2013). Problems lie at several levels and bring into serious question the quality and standard of the editorial processes in your journal. Food Chem Toxicol, 53, 441.
Le Tien D, Le Huy H. (2013). Comments on "Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize". Food Chem Toxicol, 53, 443–4.
Levine S, Saltmiras D, Webb E, Holmes C, Mortensen S, Honegger J, Bailey J. (2012). Tier 1 EDSP Assays and Regulatory Safety Studies Provide a Weight of Evidence that Glyphosate is Not an Endocrine Disruptor (Conference Abstract 529). Abstract Book for SETAC North America 33rd Annual Meeting, Long Beach, California, November 11–15, 2012.
Mink PJ, Mandel JS, Sceurman BK, Lundin JI. (2012). Epidemiologic studies of glyphosate and cancer: a review. Regul Toxicol Pharmacol, 63, 440–52.
Monsanto. (1981). A Lifetime Feeding Study of Glyphosate (ROUNDUP Technical) in Rats. East Millstone, New Jersey, USA: Bio/dynamics Inc.
Monsanto. (1983). A Chronic Feeding Study of Glyphosate (Roundup® Technical) in Mice. East Millstone, New Jersey, USA: Bio/dynamics, Inc.
Monsanto. (1985). Twelve-month Study of Glyphosate administered by Gelatin Capsule to Beagle Dogs. Monsanto report ML-83-137.
Monsanto. (1990). Chronic Study of Glyphosate Administered in Feed to Albino Rats. St. Louis, MO, USA: Monsanto Agricultural Company.
Nolte T, Rittinghausen S, Kellner R, Karbe E, Kittel B, Rinke M, Deschl U. (2011). RITA–Registry of Industrial Toxicology Animal data: the application of historical control data for Leydig cell tumors in rats. Exp Toxicol Pathol, 63, 645–56.
Nufarm. (2007). Glyphosate technical: 52-week Toxicity Study by Oral Route (Capsule) in Beagle Dogs. Study No. 29647 TCC.
Nufarm. (2009a). Glyphosate Technical: Dietary Carcinogenicity Study in the Mouse. Derbyshire, UK: Harlan Laboratories Ltd.
Nufarm. (2009b). Glyphosate Technical: Dietary Combined Chronic Toxicity/Carcinogenicity in the Rat. Shardlow, Derbyshire, UK: Harlan Laboratories Ltd.
OECD. (2012a). Guidance Document 116 on the Conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 and 453, 2nd Edition. Available at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote = ENV/JM/MONO(2011)47&doclanguage = en. Accessed on 20th, January 2015.
OECD. (2012b). OECD Guidelines for the Testing of Chemicals. Section 4 - Health Effects. Available at: http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788. Accessed on 20th, January 2015.
Ollivier L. (2013). A comment on "Seralini, G.-E., et al., Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. Food Chem. Toxicol. (2012)". Food Chem Toxicol, 53, 458.
Paganelli A, Gnazzo V, Acosta H, Lopez SL, Carrasco AE. (2010). Glyphosate-based herbicides produce teratogenic effects on vertebrates by impairing retinoic acid signaling. Chem Res Toxicol, 23, 1586–95.
Panchin AY. (2013). Toxicity of Roundup-tolerant genetically modified maize is not supported by statistical tests. Food Chem Toxicol, 53, 475.
Romano MA, Romano RM, Santos LD, Wisniewski P, Campos DA, de Souza PB, et al. (2012). Glyphosate impairs male offspring reproductive development by disrupting gonadotropin expression. Arch Toxicol, 86, 663–73.
Romano RM, Romano MA, Bernardi MM, Furtado PV, Oliveira CA. (2010). Prepubertal exposure to commercial formulation of the herbicide glyphosate alters testosterone levels and testicular morphology. Arch Toxicol, 84, 309–17.

Saltmiras D, Remick A, Haas M. (2011). Repeat dietary administration of an organic acid causes salivary gland alterations. Toxicol Lett, 205S, S233.

Saltmiras D, Tobia A. (2012). No evidence of endocrine disruption by glyphosate in Hershberger and Uterotrophic assays (Conference Abstract PS 2198). The Toxicologist (supplement to Toxicological Sciences), 126, 474.

Sanders D, Kamoun S, Williams B, Festing M. (2013). Re: Seralini, G.-E., et al. Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. Food Chem. Toxicol. (2012). Food Chem Toxicol, 53, 450–3.

Schmitz P, Harvert H. (2012). Die ökonomische Bedeutung des Wirkstoffes Glyphosat für den Ackerbau in Deutschland [Agro-Economic Analysis of the use of Glyphosate in Germany]. J Kulturpflanzen, 64, 150–62.

Schorsch F. (2013). Serious inadequacies regarding the pathology data presented in the paper by Seralini et al. (2012). Food Chem Toxicol, 53, 465–6.

Seralini GE, Clair E, Mesnage R, Gress S, Defarge N, Malatesta M, et al. (2014). Republished study: long-term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. Environ Sci Eur [Online], 26, 14.

Son WC, Gopinath C. (2004). Early occurrence of spontaneous tumors in CD-1 mice and Sprague-Dawley rats. Toxicol Pathol, 32, 371–4.

Sorahan T. (2012). Multiple myeloma and glyphosate use: a re-analysis of US Agricultural Health Study data. Abstract P27-02. Toxicol Lett, 211S, S127.

Syngenta. (1996a). Glyphosate Acid: 1 Year Dietary Toxicity Study in Dogs. Alderley Park, Macclesfield, Cheshire, UK: Central Toxicology Laboratory.

Syngenta. (1996b). Glyphosate Acid: One Year Dietary Toxicity Study in Rats. CTL/P/5143.

Syngenta. (2001). Glyphosate Acid: Two Year Dietary Toxicity and Oncogenicity Study in Rats. Central Toxicology Laboratory, Alderley Park Macclesfield, Cheshire, UK: Syngenta.

Taddesse-Heath L, Chattopadhyay SK, Dillehay DL, Lander MR, Nagashfar Z, Morse HC, III, Hartley JW. (2000). Lymphomas and high-level expression of murine leukemia viruses in CFW mice. J Virol, 74, 6832–7.

Tester M. (2013). It does not become the quality of a journal such as Food and Chemical Toxicology to publish such poor work. Food Chem Toxicol, 53, 457.

Thoolen B, Maronpot RR, Harada T, Nyska A, Rousseaux C, Nolte T, et al. (2010). Proliferative and nonproliferative lesions of the rat and mouse hepatobiliary system. Toxicol Pathol, 38, 5S–81S.

Trewavas A. (2013). Science requires the dispassionate presentation of information. Food Chem Toxicol, 53, 449.

Tribe D. (2013). My comments about the paper do not adequately describe the serious failures that have occurred in the peer review process at FCT. Food Chem Toxicol, 53, 467–72.

US EPA. (1993). Reregistration eligibility decision (RED): glyphosate. United States Environmental Protection Agency, Office of Prevention, Pesticides, and Toxic Substances. Washington, DC, USA.

US EPA. (2012). OCSPP Harmonized Test Guidelines - Series 870 - Health Effects Test Guidelines. Available at: http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series870.htm. Accessed on November 13, 2012.

Webb E, Saltmiras D, Levine S. (2013). Endocrine Disruptor Screening Program (EDSP) Tier 1 In Vitro Assays Indicate Glyphosate Does Not Interact with Estrogen and Androgen Receptor Nor Inhibit Steroidogenesis (Poster abstract P500. American College of Toxicology 33rd Annual Meeting, November 4–7, 2012). Int J Toxicol, 32, 77.

Weichenthal S, Moase C, Chan P. (2010). A review of pesticide exposure and cancer incidence in the Agricultural Health Study cohort. Environ Health Perspect, 118, 1117–25.

WHO/FAO. (2004a). Pesticides residues in food – 2004. Part II Toxicological Evaluations. Joint meeting of the FAO Panel of Experts on pesticide residues in food and the environment and the WHO Core Assessment Group (JMPR). World Health Organization/Food and Agriculture Organization of the United Nations, Rome, Italy. http://www.inchem.org/documents/jmpr/jmpmono/v2004pr01.pdf. Accessed on 20th, January 2015.

WHO/FAO. (2004b). Pesticides residues in food – 2004. Report of the joint meeting of the FAO Panel of Experts on pesticide residues in food and the environment and the WHO Core Assessment Group on Pesticide Residues (JMPR). World Health Organization/Food and Agriculture Organization of the United Nations, Rome, Italy. http://www.fao.org/fileadmin/templates/agphome/documents/Pests_Pesticides/JMPR/Reports_1991-2006/report2004jmpr.pdf. Accessed on 20th, January 2015.

Williams AL, Watson RE, DeSesso JM. (2012). Developmental and reproductive outcomes in humans and animals after glyphosate exposure: a critical analysis. J Toxicol Environ Health B Crit Rev, 15, 39–96.

Williams GM, Kroes R, Munro IC. (2000). Safety evaluation and risk assessment of the herbicide Roundup and its active ingredient, glyphosate, for humans. Regul Toxicol Pharmacol, 31, 117–65.

Woodside FC, Davis AG. (2013). The Bradford Hill Criteria: The Forgotten Predicate. Thomas Jefferson Law Rev, 35, 103–25.

Supplementary material available online

Data Supplementary Study 1–14.

Toxicologic Pathology, 33:283–291, 2005

ISSN: 0192-6233 print / 1533-1601 online
DOI: 10.1080/01926230908371

Carcinogenicity Evaluation: Comparison of Tumor Data from Dual Control Groups in the Sprague–Dawley Rat

PAUL BALDRICK

Scientific and Regulatory Consulting, Covance Laboratories Ltd., England, United Kingdom

ABSTRACT

Following recent clarification in Europe that a single control group is now acceptable for rodent carcinogenicity studies, the use of dual controls may be reduced or disappear. To date, virtually nothing has been published on whether this latter situation has improved the identification of tumorigenic risk potential in these studies. In this paper, the results of 13 rat carcinogenicity studies, performed between 1991 and 2002, with 2 control groups, are presented. Although no major differences in tumor incidences between these dual control groups were found, some interstudy variation occurred. In cases where a notable difference was seen, the use of 2 control groups, as well as robust, contemporary background data, allowed an easier interpretation of findings in drug-treated groups. Thus, the continued use of dual control groups has a vital role in the assessment of tumorigenic risk. The paper also presents an update on survival, on the range and extent of background spontaneous neoplasms, and comments on genetic drift in this commonly used rat strain.

Keywords. Sprague–Dawley rat; benign and malignant neoplasms; carcinogenicity; survival; dual control groups; genetic drift.

INTRODUCTION

Preclinical drug testing has benefited from a range of national, region-specific (e.g., Europe, United States, and Japan) or international directives, regulations, or guidelines. Furthermore, the emergence of the International Conference on Harmonization (ICH) has resulted in guidance documents that have helped to eliminate region-specific differences. However, until recently, a continued difference has been in the use of double control groups in Europe for carcinogenicity studies vs. the common use of only single control groups in other regions, such as the United States.

A review of guidance documents relating to carcinogenicity studies highlights why this situation occurred. Until recently, European guidelines indicated 2 control groups dosed with vehicle (Guidelines Volume 3B, 1998). However in the United States, 1 concurrent control group is mentioned (FDA Red Book, 1982) that agrees with OECD requirements (OECD, 1993); in Japan, a negative control group should be included (Japan, 1999). Relevant ICH documents do not mention the number of required control groups (ICH S1C, 1995; ICH S1A, 1996; ICH S1C(R), 1997; ICH S1B, 1998). As a consequence of these differences, some pharmaceutical companies have routinely included dual control groups to support global introduction of a new drug. It appears that single control groups may now become more popular with the recent clarification in Europe that 1 control group of the same number for each sex dosed with the vehicle by the same route is acceptable (CPMP/SWP/2879/00). Interestingly, a recent review of practices within carcinogenicity studies in the United States has indicated that 9/15 companies that responded to a questionnaire use 2 vehicle control groups (Perry and Menton, 2004).

Carcinogenicity studies obviously have an important role in identifying tumorigenic potential in animals, as part of the assessment of risk in humans for new drugs. However, little published data are available on how the use of 2 control groups has assisted this process. It has been pointed out that additional controls give a better assessment of how rare each tumor type is and allows a more accurate comparison to historical control data (Fairweather et al., 1998). Additionally, it is said that the results from 2 identical controls can be used as a mechanism for identifying the extent of control variability and to help evaluate the biological significance of increases in tumor incidence in the treated groups (CDER, 2001).

This paper presents data from 13 in-house carcinogenicity studies using Sprague–Dawley rats with 2 identical control groups. All the studies involved oral administration and were performed in the period 1991–2002. As well as examining potential tumor incidence differences between the dual controls, the paper also provides an update on the range and extent of background neoplasms in this commonly used rat strain and includes comment any on genetic drift.

MATERIALS AND METHODS

Selection of studies for evaluation from our in-house database was limited to carcinogenicity investigations in gang-housed Sprague–Dawley (Crl: CDBR strain) rats with 2 identical control groups using oral administration (gavage tube or dietary) in the period 1991–2002. A total of 13 studies were identified giving a total of 765 males and 765 females for each control group. Group sizes ranged from 50–65 animals/sex. Treatment comprised gavage dosing at 5 or 10 mg/kg of vehicle material or dietary intake of standard chow. Rats were treated for a 2-year period (actual period 98–105 weeks).

Rats of approximately 6 weeks of age at the start of dosing were obtained from Charles River UK and were housed 5 to a cage. They were fed SQC Rat and Mouse Maintenance Diet No. 1 (Special Diets Service Ltd., United Kingdom)

Address correspondence to: Paul Baldrick, Covance Laboratories Ltd., Otley Road, Harrogate, HG3 1PY, North Yorkshire, United Kingdom; e-mail: paul.baldrick@covance.com

ad libitium and had free access to water. Animals were maintained on a 12-hour light/dark cycle at 19–25°C and 40–70% relative humidity. At the end of the treatment period, complete necropsies were conducted on all the control animals and tissues were fixed and embedded in paraffin wax, sectioned (nominally 5 μm) and stained with hematoxylin and eosin. Tissues routinely examined in the 13 studies were adrenals, brain, cecum, colon, duodenum, eyes, femur with bone marrow and articular surface, gross lesions, Harderian glands, heart, ileum, jejunum, kidneys, liver, lungs with mainstem bronchi, mammary gland (female), mandibular and mesenteric lymph nodes, muscle, esophagus, optic nerves, ovaries, pancreas, pituitary, prostate, rectum, salivary glands, sciatic nerves, seminal vesicles, skin, spinal cord (at 3 levels), spleen, sternum with bone marrow, stomach, testes + epididymides, thymus, thyroids + parathyroids, tissue masses, tongue, trachea, urinary bladder, uterus, and vagina.

Survival after 2 years of treatment was measured. Rats were examined for nonneoplastic and neoplastic findings. Tumor identification was based on in-house nomenclature, which followed that commonly used elsewhere, e.g., International Agency for Research on Cancer (IARC) and Society of Toxicologic Pathologists (STP). To allow for in-house consistency, "older" synonyms for some tumors have been used for some neoplasms. Statistical evaluation was performed for differences between the 2 control groups (designated control I and control II). Survival and tumor incidence were compared using a 2-sided chi-squared test. If the incidences were shown to be not significantly different, then the 2 groups were combined for comparison with drug-treated groups in each of the presented carcinogenicity studies. All studies were performed to Good Laboratory Practice (GLP).

RESULTS

Survival at study end is given in Table 1. Overall, mean survival was similar between the 2 groups with 38–57% and 32–60% survival seen for males and 22–48% and 25–57% survival seen for females in control groups I and II, respectively. Generally, only small differences between the 2 control groups were seen for individual studies with no statistically significant differences in survival. When survival was broken down to 4-year periods of 1991–1994, 1995–1998, and 1999–2002, there was no obvious time effect on survival values. Across all studies, the main causes of death were consistent with those generally associated with aging laboratory rats of this strain and comprised pituitary and skin/appendage tumors in both sexes (with the latter predominantly in males) plus mammary tumors predominantly in females.

TABLE 1.—Survival at study end.*

Date	Group	Mean survival (%) Males	Females
1991–1994 (n = 3)	Control I	41	30
	Control II	44	33
1995–1998 (n = 5)	Control I	48	35
	Control II	40	38
1999–2002 (n = 5)	Control I	49	38
	Control II	48	32
		Survival range (%)	
1991–2002 (n = 13)	Control I	38–57	22–48
	Control II	32–60	25–57

* Measured at 104 weeks on all occasions except females in 3 studies (96, 97, and 100 weeks).

TABLE 2.—Common nonneoplastic lesions.

System	Lesion
Endocrine	Adrenal: basophilic focus, eosinophilic focus/nodule, hemangiectasis, medullary hyperplasia, normochromic focus, pigment, vacuolated focus
	Pancreas: islet fibrosis
	Pituitary: hyperplasia
	Thyroid: C-cell hyperplasia
Integumentary	Mammary gland: hyperplasia
Reproductive	Epididymis: inflammatory cell foci, oligospermia
	Ovary: acyclic-follicular, atrophy, cyst, stromal hyperplasia
	Prostate: prostatitis
	Testis: atrophy
	Uterus: cystic glands
	Vagina: mucification
Digestive	Liver: basophilic focus, bile duct hyperplasia, eosinophilic focus, hepatocyte vacuolation, inflammatory cell foci/infiltration, microcystic degeneration, telangiectasis
	Pancreas: acinar cell hypertrophy, basophilic focus, inflammatory cell foci, lobular atrophy
	Stomach: cystic glands
Hematopoietic/Lymphoreticular	Mandibular lymph node: agonal congestion/hemorrhage, hyperplasia
	Marrow (femur and sternum): fatty atrophy, hyperplasia
	Spleen: hemopoiesis, pigment
	Thymus: atrophy, cyst, epithelial hyperplasia, involution
Nervous	Brain: compression
	Sciatic nerve: neuropathy
	Spinal cord: radiculoneuropathy
Urinary	Kidney: glomerulonephropathy, hydronephrosis, mineralization–various sites, polypoid urothelial hyperplasia, pyelitis
Cardiovascular	Heart: cardiomyopathy
Musculoskeletal	Muscle: atrophy
Respiratory	Lung: agonal congestion/hemorrhage, foamy histiocytes, inflammatory cell foci, pneumonitis
Others	Foot/leg: pododermatitis
	Harderian gland: adenitis, hyperplasia, inflammatory cell foci
	Tail: dermatitis/folliculitis

Common nonneoplastic lesions in decedents and rats surviving to termination are shown in Table 2. Similar findings were seen for both control groups across all the studies with no obvious group differences seen. The findings were consistent with the spectrum associated with aging rats of this strain and included inflammatory (leukocyte) cell foci in several organs, chronic nephropathy in the kidney, vacuolation in the liver, C-cell hyperplasia in the thyroid, myocarditis/fibrosis in the heart, and foamy histiocytes in the lung.

A summary of tumor incidences is given in Table 3 with details of major organ or common and less common tumors shown in Tables 4 and 5, respectively. Notable interstudy differences in tumor incidences between dual control groups are presented in Table 6. Table 3 shows that the total number of neoplasms was similar for male and female control groups with a slightly higher tumor burden in females compared to males. Across all the groups, the malignant tumor incidence was 12–15%. When tumor incidences were broken down to

TABLE 3.—Summary of neoplastic incidence (mean number of neoplasms per animal).

		Male		Female	
	Date	Control I	Control II	Control I	Control II
Total neoplasms (neoplasm incidence)*	1991–2002	1,245 (1.63)	1,264 (1.65)	1,487 (1.94)	1490 (1.95)
Benign neoplasms		1,063	1071	1,264	1,316
Malignant neoplasms		182	193	223	174
Neoplasms incidence*	1991–1994	2.00	1.92	1.91	1.96
	1995–1998	1.67	1.77	1.93	1.99
	1999–2002	1.49	1.55	2.21	2.12

*Neoplasm incidence calculated from total number of neoplasms divided by number of animals on study.

4-year periods of 1991–1994, 1995–1998, and 1999–2002, results were generally similar although a slight trend for a decreased incidence in tumor burden was seen for males vs. a slight increase in burden for females.

As seen from Table 4, the most common tumors were those associated with the endocrine system and comprised pituitary adenomas (which were more common in females), benign adrenal pheochromocytomas (largely in males), and thyroid C-cell adenomas. Other common tumors were those of the integumentary system comprising mammary gland fibroadenomas, adenocarcinomas, and adenomas (in females) plus skin fibromas/dermal fibromas, keratoacanthomas, and hair follicle tumors (largely in males). Common reproductive system tumors were testicular interstitial cell (Leydig) tumors and uterine stromal polyps; lymphocytic leukemia tumors in males were the most common tumors of the hematopoietic/lymphoreticular system. Remaining tumors were present at <5%.

Overall, slight differences in the number of individual tumors were only seen between the control groups (control I vs. control II) on a few occasions, most notably as follows:

- benign adrenal pheochromocytomas in males (91 vs. 134)
- skin fibromas/dermal fibromas in males and females (185 vs. 149 and 42 vs. 18, respectively)
- skin benign hair follicle tumors in males (7 vs. 19)
- skin sarcomas in males (11 vs. 21)
- uterine schwannomas (6 vs. 1)
- ovarian tubulostromal adenomas (0 vs. 3)
- pancreatic acinar cell adenomas in males (6 vs. 0)
- hepatocellular adenomas in females (3 vs. 8)
- granulocytic leukemia in males (6 vs. 2)
- brain astrocytomas in females (8 vs. 3)
- endocardial schwannomas in males (0 vs. 6)
- bronchiolo-alveolar adenomas in males (3 vs. 0)
- bronchiolo-alveolar carcinomas in males (0 vs. 3)

As seen from Table 5, there was no specific pattern of incidence for less common tumors that, in the most extreme case, were only seen on 1 occasion among all the animals examined. In agreement with the overall variation between the 2 control groups, all 13 studies individually showed cases of "notable" differences as can be seen from Table 6, especially for benign adrenal pheochromocytomas among males and skin fibromas/dermal fibromas among both sexes or indeed, thyroid C-cell adenomas in males with an incidence of 13 vs. 5 in control I and control II, respectively. As noted in the Discussion, it is the magnitude of the difference with relevance to findings in drug-treated groups and in relation to background data that is important.

DISCUSSION

This paper shows that, in general, no major differences occur for any tumor incidence between dual control groups across the full range of mammalian body systems when assessed in 13 in-house carcinogenicity studies over a period of 12 years. Interstudy variation occurred for some tumors but did not follow any particular pattern, although the incidence of benign adrenal pheochromocytomas in males was quite variable in a number of the studies. These results are interesting, as little has been published in this area. An evaluation of the Food and Drug Administration (FDA) and National Toxicology Program/National Cancer Institute (NTP/NCI) database showed that most rodent carcinogenicity studies had 1 control group and only 18 cases of duplicate control groups (Contrera et al., 1997). These latter studies appear to be a series of experiments performed with 18 color additives in the Charles River CD rat (Haseman et al., 1986). Interestingly, Contrera et al. (1997) reported that in 12 of these studies statistically significant differences were found for tumor findings between these groups.

However, this finding is not fully supported in the referenced Haseman et al. (1986) paper. The latter authors did indeed report marked study-to-study variability for certain tumors and found statistically significant ($p < 0.05$) paired control differences in 12 of the studies (see Table 7). However, when they analyzed these results to see if these findings exceeded chance expectation, no evidence of extrabinomial within-study variability between the 2 concurrent control groups was found. Indeed, the total number of observed significant paired-control differences was virtually identical to what would be expected from usual binomial model assumptions. The study authors also postulate that the marked study-to-study variability may be related to the fact that the studies were performed at 4 different testing laboratories and suffer from some histopathological deficiencies and inconsistencies.

Thus, from these study results and from the present reported studies, some interstudy variation can be expected in carcinogenicity study control group tumor incidences. In most cases, such variation does not cause an issue in assessment for potential carcinogenic risk. However, in those occasions where a low or high incidence of a tumor finding appears in a control group (as can occur even between concurrent control groups as indicated by e.g., 13 vs. 5 tumors for thyroid C-cell adenomas in 1 study and 0 vs. 3 or more tumors on various occasions in other studies), then a value for dual control groups can be argued when, for example, interpreting an increased incidence of the tumor in the high drug-treated group. The fact that such differences can be

TABLE 4.—Summary of major organ or common spontaneous neoplasms.

Body system	Neoplasm	Control I males		Control II males		Control I females		Control II females	
		Number (%)	% Range	Number (%)	% Range	Number (%)	% Range	Number (%)	% Range
Endocrine									
Pituitary	Adenoma (B)	363(49.3)	33.3–65.0	350(47.8)	34.5–65.1	523(70.3)	61.7–81.7	557(75.4)	66.7–83.6
	Carcinoma (M)	1(0.15)	0–2.0	0(0)	0	22(3.2)	0–10.0	14(2.1)	0–5.6
Adrenal	Pheochromocytoma (B)	91(12.4)	4.0–21.7	134(18.2)	6.7–24.0	24(3.2)	0–8.0	33(4.5)	1.7–11.1
	Cortical adenoma (B)	7(0.95)	0–4.0	5(0.68)	0–2.0	10(1.3)	0–3.3	9(1.2)	0–3.3
	Pheochromocytoma (M)	17(2.3)	0–6.7	11(1.5)	0–6.7	3(0.40)	0–1.7	3(0.41)	0–1.7
	Cortical carcinoma (M)	2(0.34)	0–1.7	5(0.78)	0–3.2	2(0.31)	0–1.7	5(0.78)	0–3.4
Thyroid	C-cell adenoma (B)	104(14.9)	8.3–20.0	84(11.8)	7.7–16.7	77(10.5)	5.3–20.0	62(8.5)	2.0–20.0
	Follicular cell adenoma (B)	20(2.9)	0–9.1	19(2.7)	0–6.9	5(0.68)	0–3.1	3(0.41)	0–3.4
	C-cell carcinoma (M)	6(0.92)	0–3.4	7(1.1)	0–3.3	3(0.44)	0–1.8	3(0.44)	0–2.0
Pancreatic	Islet cell adenoma (B)	31(4.2)	0–8.3	36(4.9)	1.7–10.3	10(1.3)	0–5.3	7(0.95)	0–3.3
	Islet cell carcinoma (M)	12(1.8)	0–6.2	14(2.1)	0–6.0	2(0.29)	0–2.0	1(0.15)	0–1.7
Parathyroid	Adenoma (B)	9(1.5)	0–3.2	9(1.5)	0–6.8	4(0.66)	0–2.9	4(0.67)	0–3.5
Integumentary									
Mammary gland	Fibroadenoma (B)	*	*	*	*	413(55.6)	32.3–71.7	424(57.7)	46.7–72.2
	Adenoma (B)	*	*	*	*	38(5.6)	0–13.3	36(5.3)	0–9.5
	Adenocarcinoma (M)	*	*	*	*	101(13.6)	10.0–23.3	82(11.2)	6.7–22.0
Skin	Squamous cell papilloma (B)	17(2.3)	0–6.7	14(1.9)	0–4.8	0(0)	0	2(0.27)	0–4.0
	Keratoacanthoma (B)	46(9.1)	1.5–16.7	63(12.5)	3.4–23.3	3(0.59)	0–2.9	9(1.8)	0–6.0
	Hair follicle tumor (B)	7(2.7)	0–5.0	19(7.3)	3.3–11.7	2(0.74)	0–5.7	1(0.38)	0–1.6
	Basal cell tumor (B)	6(0.81)	0–3.3	13(1.8)	0–8.3	1(0.13)	0–1.7	1(0.14)	0–1.5
	Fibroma/Dermal fibroma (B)	125(25.1)	6.2–41.7	149(20.2)	3.1–41.7	42(5.6)	0–15.0	18(2.4)	0–8.3
	Fibrolipoma (B)	8(1.5)	0–3.3	4(0.75)	0–3.3	0(0)	0	1(0.19)	0–1.5
	Lipoma (B)	26(3.5)	0–8.3	25(3.4)	0–11.7	12(1.6)	0–3.3	17(2.3)	0–6.7
	Sebaceous cell adenoma (B)	4(0.71)	0–3.3	3(0.53)	0–2.0	0(0)	0	0(0)	0
	Fibrosarcoma (M)	14(2.1)	0–8.3	8(1.2)	0–4.0	2(0.29)	0–1.7	4(0.59)	0–1.7
	Histiocytic sarcoma (M)	7(1.0)	0–3.3	11(1.6)	0–4.0	11(1.6)	0–4.6	7(1.0)	0–3.1
	Sarcoma (M)	11(1.5)	0–5.0	21(2.9)	0–5.0	7(0.94)	0–3.3	6(0.81)	0–5.0
Reproductive									
Testicular	Interstitial cell tumor (B)	27(3.7)	0–11.7	39(5.3)	0–10.0	—	—	—	—
Uterine	Stromal polyp (B)	—	—	—	—	47(6.3)	1.5–11.7	53(7.2)	3.1–15.0
	Granular cell tumor (B)	—	—	—	—	5(0.87)	0–4.0	8(1.4)	0–3.3
	Schwannoma (M)	—	—	—	—	6(1.5)	0–4.6	1(0.25)	0–1.6
Ovarian	Cystadenoma (B)	—	—	—	—	1(0.33)	0–1.5	2(0.68)	0–2.0
	Tubulostromal adenoma (B)	—	—	—	—	0(0)	0	3(0.85)	0–2.0
	Granulosa cell tumor (B)	—	—	—	—	3(0.99)	0–3.3	1(0.33)	0–1.7
Vagina	Granular cell tumor (B)	—	—	—	—	4(1.1)	0–3.3	5(1.4)	0–3.3
Digestive									
Pancreatic	Acinar cell adenoma (B)	6(1.1)	0–3.3	0(0)	0	0(0)	0	1(0.16)	0–1.7
	Acinar cell carcinoma (M)	2(0.44)	0–1.7	1(0.23)	0–1.7	0(0)	0	1(0.22)	0–1.7
Hepatic	Hepatocellular adenoma (B)	12(1.7)	0–6.7	16(2.3)	0–3.4	3(0.42)	0–1.7	8(1.1)	0–2.0
	Hepatocellular carcinoma (M)	2(0.49)	0–1.7	4(0.98)	0–3.4	0(0)	0	0(0)	0

(Continued on next page)

Haematopoietic/Lymphoreticular									
Thymic	Thymona (B)	6(0.85)	0–3.4	8(1.1)	0–3.6	12(1.6)	0–4.7	9(1.3)	0–5.6
Splenic	Hemangioma (B)	4(0.84)	0–1.7	4(0.84)	0–1.7	0(0)	0	1(0.21)	0–1.7
Mesenteric lymph node	Lymphangioma (B)	8(1.2)	0–3.4	5(0.74)	0–3.4	2(0.29)	0–2.0	2(0.29)	0–1.7
	Hemangioma (B)	10(1.6)	0–5.0	14(2.3)	0–8.2	1(0.16)	0–1.6	3(0.48)	0–1.7
Systemic	Granulocytic leukemia (M)	6(1.4)	0–6.7**	2(0.48)	0–12.5**	0(0)	0	1(0.24)	0–100**
	Lymphocytic leukemia (M)	26(5.2)	0–100**	25(5.0)	0–83.3**	6(1.2)	0–100**	6(1.2)	0–100**
Nervous									
Brain	Meningioma (B)	8(1.7)	0–3.3	6(1.3)	0–3.3	3(0.63)	0–5.0	3(0.64)	0–3.1
	Granular cell meningioma (B)	4(1.0)	0–3.3	5(1.3)	0–5.1	1(0.25)	0–2.0	3(0.77)	0–3.1
	Astrocytoma (M)	10(1.5)	0–4.0	11(1.6)	0–5.0	8(1.2)	0–5.0	3(0.44)	0–2.0
Urinary									
Renal	Tubular cell adenoma (B)	2(0.40)	0–1.7	1(0.20)	0–1.7	1(0.20)	0–2.0	0(0)	0
	Lipoma (B)	2(0.40)	0–1.7	1(0.20)	0–1.7	1(0.20)	0–1.7	2(0.40)	0–1.7
Cardiovascular									
Cardiac	Endocardial schwannoma (B)	0(0)	0	6(1.7)	0–3.3	0(0)	0	2(0.56)	0–1.7
	Endocardial schwannoma (M)	2(0.57)	0–1.7	0(0)	0	0(0)	0	1(0.28)	0–1.6
Respiratory									
	Bronchiolo-alveolar adenoma(B)	3(0.72)	0–2.0	0(0)	0	0(0)	0	2(0.48)	0–2.0
	Bronchiolo-alveolar carcinoma (M)	0(0)	0	3(0.75)	0–2.0	0(0)	0	1(0.25)	0–1.7

B, Benign; M, Malignant.
*See Table 5.
**Only reported in a limited number of studies in a few animals (which gave an artificially high upper range value).

TABLE 5.—Summary of less common spontaneous neoplasms.

Body system	Neoplasm (number of tumors/number of animals examined)
Endocrine	
Adrenal	Ganglioneuroma (B): 1/355 CIM, 1/355 CIIM, 1/353 CIIF
Thyroid	Follicular cell carcinoma (M): 3/526 CIM, 1/530 CIIM, 3/549 CIF
Parotid	Adenoma (B): 1/48 CIIF
Integumentary	
Mammary gland	Fibroadenoma (B): 8/118 CIM, 7/98 CIIM; Adenoma (B): 1/91 CIM, 1/78 CIIM; Adenolipoma (B): 2/275 CIIF; Adenocarcinoma (M): 3/118 CIM, 1/98 CIIM; Carcinosarcoma (M): 2/244 CIF, 1/239 CIIF
Skin	Hemangioma (B): 1/125 CIM; Neurofibroma (B): 1/270 CIM, 1/275 CIF, 1/275 CIIF; Squamous cell carcinoma (M): 2/677 CIM, 1/676 CIIM, 2/685 CIF, 1/680 CIIF; Basal cell tumor (M): 4/357 CIM, 1/357 CIIM, 2/360 CIF; Hemangiosarcoma (M): 3/387 CIM, 2/386 CIIM, 1/395 CIF; Schwannoma (M): 2/212 CIM, 1/211 CIIM, 2/220 CIF, 2/216 CIIF
Zymbal gland	Sebaceous cell carcinoma (M): 5/69 CIM, 3/67 CIIM
Reproductive	
Testicular	Mesothelioma (B): 2/290 CIM; Hemangioma (B): 1/60 CIIM; Seminoma (M): 1/60 CIM
Epididymidial	Sarcoma (M): 1/50 CIIM
Prostatic	Adenoma (B): 4/232 CIM, 1/232 CIIM; Squamous cell papilloma (B): 1/60 CIM; Carcinoma (M): 4/475 CIM, 4/474 CIIM; Adenocarcinoma (M): 2/112 CIIM
Uterine	Adenoma (B): 2/374 CIF, 2/375 CIIF; Fibroma (B): 1/120 CIF, 2/120 CIIF; Hemangioma (B): 2/159 CIIF; Leiomyoma (B): 1/121 CIIF; Stromal sarcoma (M): 1/440 CIF, 1/437 CIIF; Squamous cell carcinoma (M):1/94 CIF; Granular cell tumor (M): 1/96 CIIF; Adenocarcinoma (M): 2/560 CIF, 1/561 CIIF; Stromal sarcoma (M): 1/180 CIF; Leiomyosarcoma (M): 1/60 CIF
Ovarian	Luteoma (B): 1/65 CIF; Thecoma (B): 1/169 CIIF; Mixed sex cord stromal tumor (B): 3/230 CIF, 1/230 CIIF; Mesothelioma (B): 1/59 CIIF; Granulosa-theca tumor (M): 1/180 CIF
Cervix	Schwannoma (M): 1/185 CIF
Vagina	Sarcoma (M): 1/171 CIIF; Squamous cell carcinoma (M): 1/171 CIIF; Schwannoma (M): 1/125 CIIF
Seminal vesicle	Adenocarcinoma (M): 1/240 CIM, 2/242 CIIM
Preputial/Clitoral gland	Squamous cell papilloma (B): 1/3 CIF; Squamous cell carcinoma (M): 1/7 CIF
Vas deferens	Carcinoma (M): 1/1 CIIM
Digestive	
Pancreatic	Acinar-islet cell adenoma (B): 4/290 CIM
Hepatic	Histiocytic sarcinoma (M): 2/290 CIF
Tongue	Squamous cell papilloma (B): 1/65 CIIF; Squamous cell carcinoma (M): 1/49 CIM
Esophageal	Squamous cell papilloma (B): 1/62 CIM
Stomach	Squamous cell papilloma (B): 1/386 CIM, 3/379 CIIM, 1/397 CIF; Squamous cell carcinoma (M): 1/230 CIIM
Duodenal	Adenocarcinoma (M): 1/95 CIIF
Jejunal	Leiomyoma (B): 1/111 CIM; Leiomyosarcoma (M): 1/117 CIIM; Adenocarcinoma (M): 1/108 CIIM
Ileal	Fibrosarcoma (M): 1/55 CIF
Cecal	Fibroma (B): 1/104 CIM; Leiomyoma (B): 1/169 CIM, 1/159 CIIM, 1/179 CIF, 2/166 CIIF
Colonic	Adenoma (B): 1/56 CIM; Leiomyosarcoma (M): 1/59 CIIM
Rectal	Adenoma (B): 1/59 CIF; Adenocarcinoma (M): 1/60 CIF; Leiomyosarcoma (M): 1/117 CIM, 1/120 CIF
Haematopoietic/Lymphoreticular	
Thymic	Thymona (M): 1/373 CIIM, 1/390 CIF, 1/383 CIIF; Sarcoma (M): 1/60 CIF
Splenic	Mesothelioma (B): 1/118 CIIF; Sarcoma (M): 1/352 CIM, 1/352 CIIM, 2/350 CIIF; Hemangiosarcoma (M): 1/352 CIM, 2/351 CIIM
Systemic	Lymphoma lymphocytic (M): 2/429 CIM, 4/428 CIIM, 3/429 CIF, 2/423 CIIF; Lymphoma pleomorphic (M): 1/314 CIM, 2/314 CIIM, 1/306 CIF, 5/305 CIIF; Lymphoma (M): 2/60 CIM, 3/64 CIIM, 4/55 CIF, 1/53 CIIF; Leukemia (M): 2/75 CIM; Histiocytic sarcoma (M): 3/189 CIM, 4/189 CIIM, 2/183 CIF, 2/182 CIIF; Mast cell tumor (M): 1/65 CIIM
Femur + marrow	Histiocytic sarcoma (M): 1/65 CIF
Nervous	
Brain	Oligodendroglioma (M): 1/169 CIM; Meningioma (M): 1/119 CIM; Mixed glioma (M): 1/60 CIIM
Spinal	Astrocytoma (M): 1/310 CIF; Schwannoma (M): 1/1 CIM
Nerve	Schwannoma (M): 1/2 CIIM
Urinary	
Renal	Tubular cell carcinoma (M): 1/210 CIM, 1/206 CIIM, 1/214 CIIF; Liposarcoma (M): 2/392 CIM, 1/390 CIIM, 1/395 CIIF; Mesenchymal tumor (M): 1/150 CIM, 1/154 CIF
Urinary bladder	Papilloma (B): 2/181 CIIF; Leiomyoma (B): 1/60 CIF; Sarcoma (M): 1/120 CIM, 1/117 CIIF
Cardiovascular	
Cardiac	Mesothelioma (B): 2/230 CIIM; Hemangioma (B): 1/151 CIIM
Musculoskeletal	
Connective tissue	Lipoma (B): 1/3 CIIM; Histiocytic sarcoma (M): 1/5 CIM, 6/7 CIIM, 4/4 CIF, 2/4 CIIF; Sarcoma (M): 1/1 CIIF
Bone	Osteoma (B): 1/5 CIM; Osteosarcoma (M): 2/8 CIM, 6/9 CIIM, 1/3 CIF, 2/3 CIIF
Muscle	Lipoma (B): 1/4 CIM; Histocytic sarcoma (M): 1/8 CIIM
Respiratory	
Lung	Squamous cell carcinoma (M): 1/65 CIM
Nasal cavity	Squamous cell carcinoma (M): 1/4 CIIM, 1/5 CIF
Other Tumours	
Thoracic cavity	Sarcoma (M): 1/10 CIM, 1/11 CIIM, 1/4 CIF; Mesothelioma (M): 1/6 CIM, 1/5 CIIM; Liposarcoma (M): 1/9 CIIM; Granular cell sarcoma (M): 1/1 CIIF; Histiocytic sarcoma (M): 1/5 CIM
Abdominal cavity	Lipoma (B): 2/24 CIM, 1/23 CIIM, 1/12 CIIF; Hemangioma (B): 1/2 CIIM; Mesothelioma (M): 1/13 CIM, 1/5 CIF, 1/3 CIIF; Schwannoma (M): 1/6 CIIF; Histiocytic sarcinoma (M): 2/7 CIM, 1/8 CIIM; Carcinoma (M): 2/16 CIM; Hemangiosarcoma (M): 1/3 CIIM; Paraganglioma (M): 1/3 CIIM
Cranial cavity	Osteoma (B): 1/1 CIM, 2/4 CIIM; Hemangiosarcoma (M): 1/1 CIIM; Schwannoma (M): 1/1 CIIM; Histiocytic sarcoma (M): 1/4 CIM
Oral cavity	Odontogenic tumor (B): 1/8 CIIF; Neurofibrosarcoma (M): 1/5 CIIM; Squamous cell carcinoma (M): 1/20 CIIM; Sarcoma (M): 1/9 CIF
Tail	Neurofibroma (B): 1/126 CIM, 2/123 CIIM, 1/81 CIF; Fibroma/Dermal fibroma (B): 1/63 CIM; Keratoacanthoma (B): 2/168 CIM, 1/168 CIIM, 2/101 CIF; Squamous cell papilloma (B): 5/270 CIM, 6/269 CIIM, 2/224 CIF, 1/190 CIIF; Basal cell tumor (B): 1/41 CIM; Squamous cell carcinoma (M): 1/28 CIM
Foot/leg	Fibroepithelial polyp (B): 1/8 CIIF; Hair follicle tumor (B): 1/7 CIM
Ear	Neurofibroma (B): 1/20 CIM, 1/20 CIF
Blood vessel	Hemangioma (B): 1/59 CIM, 2/60 CIF, 1/58 CIIF

B, Benign; M, Malignant.
CIM, Control group I male; CIIM, Control group II male; CIF, Control group I female; CIIF, Control group II female.

TABLE 6.—Summary of notable dual control group tumor differences in the present investigation.

Study	Neoplasm	Males		Females	
		Control I	Control II	Control I	Control II
1	Adrenal pheochromocytoma (B)	7/60	14/60	—	—
	Hepatic hepatocellular adenoma	3/60	0/60	—	—
2	Leukaemia lymphocytic	0/50	3/50	—	—
3	Adrenal pheochromocytoma (B)	8/60	4/60	—	—
	Thyroid C-cell adenoma	10/59	5/57	—	—
	Skin fibroma/dermal fibroma	5/60	2/60	—	—
	Hepatic hepatocellular adenoma	4/60	1/60	—	—
	Brain meningioma	—	—	3/60	0/59
4	Adrenal pheochromocytoma (B)	6/65	13/65	—	—
	Thyroid C-cell adenoma	13/64	5/65	—	—
	Pancreatic islet cell carcinoma	4/65	1/65	—	—
	Uterine schwannoma	3/65	0/65	—	—
5	Adrenal pheochromocytoma (M)	4/60	1/60	—	—
	Thyroid C-cell adenoma	11/59	6/59	—	—
	Pancreatic islet cell adenoma	—	—	3/57	0/58
	Skin basal cell tumor (B)	1/60	4/60	—	—
	Skin fibroma/dermal fibroma	—	—	9/60	0/59
	Skin fibrosarcoma	5/60	1/60	—	—
6	Thyroid follicular cell adenoma	0/55	4/58	—	—
	Skin fibroma/dermal fibroma	22/60	13/60	—	—
	Skin lipoma	2/60	7/60	—	—
	Uterine stromal polyp	—	—	1/60	4/60
7	Adrenal pheochromocytoma (B)	—	—	0/60	3/59
	Mammary gland adenoma	—	—	8/60	3/60
	Mammary gland carcinoma	—	—	14/60	5/60
	Skin fibroma/dermal fibroma	—	—	2/60	5/60
	Skin sarcoma	—	—	0/60	3/60
	Uterine stromal polyp	—	—	2/60	5/60
8	Adrenal pheochromocytoma (B)	8/59	14/60	—	—
	Pancreatic islet cell adenoma	1/58	4/60	—	—
	Skin basal cell tumor (B)	0/60	5/60	—	—
	Skin fibroma/dermal fibroma	—	—	6/60	2/60
	Skin lipoma	4/60	0/60	0/60	3/60
	Leukemia granulocytic	4/60	0/60	—	—
9*	Adrenal pheochromocytoma (B)	2/30	7/29	—	—
	Pituitary adenoma	17/30	10/29	—	—
	Skin fibroma/dermal fibroma	—	—	5/35	2/35
10	Adrenal pheochromocytoma (B)	5/60	10/60	—	—
	Pituitary carcinoma	—	—	6/60	1/60
	Skin hair follicle tumor	0/60	7/60	-	—
11	Adrenal pheochromocytoma (B)	—	—	6/60	1/60
	Skin fibroma/dermal fibroma	—	—	4/60	1/60
	Testicular interstitial cell tumor	2/60	5/60	—	—
12	Thyroid C-cell adenoma	9/50	5/48	8/50	1/49
	Skin hair follicle tumor	0/50	4/50	—	—
	Mesenteric lymph node hemangioma	1/50	4/49	—	—
13	Adrenal pheochromocytoma (B)	8/62	14/62	—	—
	Pituitary carcinoma	—	—	5/59	1/61
	Thyroid follicular cell adenoma	0/60	3/62	—	—
	Skin squamous cell papilloma	0/62	3/62	—	—

*Decedent data only presented.

statistically significant is in itself not necessarily key; what is important is the magnitude of the difference in conjunction with factors such as preneoplastic/hyperplastic changes, evidence of increasing tumor incidence with increasing dose and background data incidence.

Thus, in the case of adrenal pheochromocytomas, which has an incidence of 7/60 and 14/60 for the 2 control groups in 1 study, an incidence of 14/60 in the high drug-treated group could easily be discounted as nontreatment related, although it could also be argued that background data would have allowed a similar conclusion. In a recent study in our laboratories, malignant lymphomas were seen at an incidence of 3, 8, 3, 4, and 7 (males) and 7, 22, 4, 8, and 20 (females) in control I, control II, low, mid, and high dose groups, respectively. In the absence of control II, the high tumor incidence seen in high-dose females plus the statistically significant dose trend could have resulted in the conclusion of evidence of a drug-induction for this tumor. The use of dual control groups to support interpretation of rare tumors is also valid, although results from the presented 13 carcinogenicity studies indicate that such tumors were often seen in both control groups anyway. Overall, in all these cases, scientific assessment would involve a careful evaluation of tumor incidence across all study groups (control and drug-treated) with reference to contemporary mean and/or overall range of tumor incidences. Indeed, it has been pointed out elsewhere that 2 control groups within the same study can occasionally show notable differences, stressing the importance of knowing the range of incidences rather than just mean incidence (Gopinath, 1994). Results from the present investigation show that the continued use of dual control groups in carcinogenicity studies has a role in the assessment of tumorigenic risk.

The literature contains a number of papers reporting survival and/or spontaneous neoplasm incidence for Sprague–Dawley rats (e.g., Chandra et al., 1992; McMartin et al., 1992; Keenan et al., 1995; Attia, 1996; Knight et al., 1996; Kaspareit

TABLE 7.—Summary of significant dual control group tumor differences in a carcinogenicity assessment of 18 color additives[†].

Study	Neoplasm	Males		Females	
		Control I	Control II	Control I	Control II
1	Brain glioma	0/69	5/68	—	—
	Thyroid C-cell	2/62	10/61	—	—
	Prostate	0/67	11/67	—	—
2	Pituitary	3/70	13/69	—	—
3	Pituitary	27/68	41/69	—	—
	Liver	2/70	8/70	—	—
4	Thyroid C-cell	10/65	3/64	—	—
5	Pancreatic islet cell	13/69	3/70	—	—
6	Adrenal cortical	1/68	7/70	—	—
7	Mammary gland epithelial	5/70	0/70	—	—
	Keratoacanthoma	6/52	1/57	—	—
8	Hematopoietic system	10/70	1/70	—	—
9	Pituitary	—	—	39/69	26/69
10	Adrenal cortical	—	—	23/70	12/70
11	Uterus endometrial stromal	—	—	7/70	1/70
12	Integumentary system fibroma/fibrosarcoma	—	—	13/70	2/70

[†] Adapted from Haseman et al. (1986).

and Rittinghausen, 1999; Hooks and Harling, 2001; Britton et al., 2004; Son and Gopinath, 2004; Tennekes et al., 2004). Chandra et al. (1992) showed survival of 42–65% (males) and 41–53% (females) from 17 carcinogenicity studies in the late 1980s/early 1990s while McMartin et al. (1992) reported values of 33–43% (males) and 27–44% (females) from a survey of 9 in-house studies conducted between 1984–1991. Further authors report survival of 44% (males) and 32% (females) from a range of studies in the period 1993–2000 (Hooks and Harling, 2001) plus contemporaneous survival values of 39–40% (males) and 33–39% (females) (Britton et al., 2004). These values compare with those of 32–60% (males) and 22–57% (females) seen in the present reported studies.

The present investigation shows that the majority of tumors were benign in nature and the incidence of malignant forms was 12–15%; this value compares well with the 17–19% incidence reported elsewhere (Chandra et al., 1992). This paper found pituitary adenomas to be the most frequent neoplasm at 48–49% (males) and 70–75% (females). These values are higher than reported by Chandra et al. (1992), namely 27% (males) and 49% (females) and by Kaspareit and Rittinghausen (1999), namely 20% (males) and 39% (females) but are less than the frequency reported by McMartin et al. (1992), namely 62% (males) and 85% (females). Attia (1996) reports a pituitary adenoma incidence of 47% and 62% for males and females, respectively, and recently values of 40% and 71% (Son and Gopinath, 2004) and 51–52% and 72–73% (Tennekes et al., 2004), respectively have been cited.

The next most common tumors in male rats were skin fibromas/dermal fibromas (20–25%), benign adrenal pheochromocytomas (12–18%), thyroid C-cell adenomas (12–15%), skin keratoacanthomas (9–13%), skin benign hair follicle tumors (3–7%), lymphocytic leukemia tumors (5%), and testicular interstitial cell tumors (4–5%). In females, the next most common neoplasms were mammary fibroadenomas (56–58%), mammary adenocarcinomas (11–14%), thyroid C-cell adenomas (9–11%), uterine stromal polyps (6–7%), and mammary adenomas (5–6%). This pattern of tumor findings and incidences generally compares well with those seen in other studies in Sprague–Dawley rats (e.g., Chandra et al., 1992; McMartin et al., 1992; Attia, 1996; Kaspareit and Rittinghausen, 1999) although the incidence of skin tumors in males is somewhat higher than reported in other studies in this strain. Values for less common tumors also compare well with those given in these papers.

The general similarity of tumor incidence in the present study to those reported elsewhere for Sprague–Dawley rats is comforting bearing in mind differences in study design. Such differences include Sprague–Dawley strain/supplier, diet (composition and amount), and caging (single vs. gang housed). Discussion of how these factors can affect tumor incidence is outside the goal of this publication but is well covered in the literature (e.g., Hardisty, 1985; Keenan et al., 1995; Greim et al., 2003; Haseman et al., 2003).

It has been highlighted that historical control data has a role in the evaluation of rare neoplasms, high-incidence tumors, marginal increases in incidence and in the assessment of the quality of a study (Greim et al., 2003). However, it is known that variation of incidence of tumor types within the same strain and sex and, within the same laboratory, occur due to the progression of time (Gopinath, 1994). In order to avoid such "genetic drift" when using historical data, it has been suggested that only studies performed during the last 5 years prior to the new study, should be accessed (CPMP/SWP/2879/00). Others have recommended that historical comparisons be limited to control data from the previous 3–4 years of the new study (Haseman et al., 1984; Fairweather et al., 1998). The NTP uses a historical control database with a time window of 7 years (Haseman et al., 1998).

From an analysis of certain tumor types over a period of 9 years, Gopinath (1994) showed that although there were fluctuations, incidence rates remained within a small band and no clear evidence of a time-related drift in rates was found. Recently, a review of the historical control data for 20 carcinogenicity studies generated in 1990–2002 in Sprague–Dawley rats showed little variability with time (Son and Gopinath, 2004). In a further evaluation of time-related changes in the incidence of spontaneous neoplasms in the adrenals, mammary gland, liver, pituitary, and pancreas of Sprague–Dawley rats from studies performed in 1986–1996, no obvious change was seen in the majority of tumor types (Tennekes et al., 2004). A small positive (increase with time) trend was seen for mammary gland fibroadenomas, while a small negative (decrease with time) trend was seen for pituitary adenomas (both sexes) and benign adrenal pheochromocytomas (males). An evaluation of tumor incidence using 4-year time points over the 12-year period in the present investigation has also shown no obvious difference in the number or pattern of tumors recorded. Standardizing tumor burden by incidence in the total number of animals assessed has not shown any strong evidence of drift although a slight trend for a lower incidence for males and for a higher incidence for females was seen.

In conclusion, this paper shows that no major differences in tumor incidences occur when dual control groups are used. However, occasional interstudy variation was seen for some tumors, indicating that the use of 2 control groups has a role in the interpretation of findings in drug-treated groups. Interstudy variation assists such interpretation when one control group, but not the other, mirrors elevated tumor incidence in one or more of the drug-treated groups. Thus, dual control

groups give greater reassurance in assessing for biological significance of any increased tumor incidence. Such use also allows the buildup of contemporary background data at a greater rate.

ACKNOWLEDGEMENT

I would like to thank Chris Barton for his kind and constructive comments during the writing of this paper.

REFERENCES

Attia, M. A. (1996). Neoplastic and non-neoplastic lesions in the mammary gland, endocrine and genital organs in aging male and female Sprague–Dawley rats. *Arch Toxicol* **70**, 461–473.

Britton, L. D., Downs, N., Mullins, P., and Mitchell, D. (2004). The most appropriate strain for rat carcinogenicity bioassays? *Toxicol Sci* **78** (1–S), 560.

CDER (2001). Guidance for industry: statistical aspects of the design, analysis, and interpretation of chronic rodent carcinogenicity studies of pharmaceuticals. Issued as draft document in May 2001. Available at: (http:\\www.fda.gov\cder\guidance).

Chandra, M., Riley, M. G. I., and Johnson, D. E. (1992). Spontaneous neoplasms in aged Sprague–Dawley rats. *Arch Toxicol* **66**, 496–502.

Contrera, J. F., Jacobs, A. C., and DeGeorge, J. J. (1997). Carcinogenicity testing and the evaluation of regulatory requirements for pharmaceuticals. *Reg Toxicol Pharm* **25**(2), 130–45.

CPMP/SWP/2877/00 Note for guidance on carcinogenic potential. Operational January 2003. Available at: (http:\\www.eudra.org\humandocs)

Fairweather, W. R., Bhattacharyya, A., Ceuppens, P. P., Heimann, G., Hothorn, L. A., Kodell, R. L., Lin, K. K., Mager, H., Middleton, B. J., Slob, W., Soper, K. A., Stallard, N., Ventre, J., and Wright, J. (1998). Biostatistical methodology in carcinogenicity studies. *Drug Inform Jour* **32**, 401–421.

FDA Red Book (1982). Toxicological principles for the safety assessment of direct food additives and color additives used in food. U.S. Food and Drug Administration, Bureau of Foods.

Gopinath, C. (1994). Spontaneous tumor rates: their use to support rodent bioassays. *Toxicol Pathol* **22**(2), 160–4.

Greim, H., Gelbke, H. P., Reuter, U., Thielmann, H. W., and Edler, L. (2003). Evaluation of historical controls in carcinogenicity studies. *Hum Exp Toxicol* **22**, 541–9.

Guidelines Volume 3B (1998). Medicinal products for human use. Safety, environment and information. In The Rules Governing Medicinal Products in the European Union. Office for Official Publications of the European Communities, Luxembourg.

Hardisty, J. F. (1985). Factors influencing laboratory animal spontaneous tumor profiles. *Toxicol Pathol* **13** (2), 95–104.

Haseman, J. K., Hailey, J. R., and Morris, R. W. (1998). Spontaneous neoplasm incidences in Fischer 344 rats and $B6C3F_1$ mice in two-year carcinogenicity studies: a National Toxicology Program update. *Toxicol Pathol* **26**(3), 428–41.

Haseman, J. K., Huff, J., and Boorman, G. A. (1984). Use of historical control data in carcinogenicity studies in rodents. *Toxicol Pathol* **12**(2), 126–35.

Haseman, J. K., Ney, E., Nyska, A., and Rao, G. N. (2003). Effect of diet and animal care/housing protocols on body weight, survival, tumor incidences, and nephropathy severity of F344 rats in chronic studies. *Toxicol Pathol* **31**, 674–81.

Haseman, J. K., Winbush, J. S., and O'Donnell, M. W. Jr. (1986). Use of dual control groups to estimate false positive rates in laboratory animal carcinogenicity studies. *Fund App Toxicol* **7**(4), 573–84.

Hooks, W. N., and Harling, R. J. (2001). A review of the choice of rat strains in tumorigenicity studies. *Toxicol Sci* **60**(1), 260.

ICH S1C. (1995). Dose selection for carcinogenicity studies of pharmaceuticals. Operational March 1995. Available at: (http:\\www.eudra.org\humandocs)

ICH S1A. (1996). Guideline on the need for carcinogenicity studies of pharmaceuticals. Operational July 1996. Available at: (http:\\www.eudra.org\humandocs)

ICH S1C(R). (1997). Guidance on dose selection for carcinogenicity studies of pharmaceuticals: addendum on a limit dose and related notes. Operational December 1997. Available at: (http:\\www.eudra.org\humandocs)

ICH S1B. (1998). Testing for carcinogenicity of pharmaceuticals. Operational February 1998. Available at: (http:\\www.eudra.org\humandocs)

Japan's and ICH Guidelines for New Drug Registration. (1999). Yakuji Nippon Ltd.

Kaspareit, J., and Rittinghausen, S. (1999). Spontaneous neoplastic lesions in Harlan Sprague–Dawley rats. *Exp Toxicol Pathol* **51**, 105–7.

Keenan, K. P., Soper, K. A., Smith, P. F., Ballam, G. D., and Clark, R. L. (1995). Diet, overfeeding and moderate dietary restriction in control Sprague–Dawley rats: 1. Effects on spontaneous neoplasms. *Toxicol Pathol* **23**(3), 269–86.

Knight, E. V., Bryant, S., Kimball, B., Eitzen, B. H., Legrand, E.K., and Barrett, D. S. (1996). Two-year evaluation of clinical and gross pathological parameters in Sprague–Dawley rats from different suppliers. *Toxicol Pathol* **24**, 788–9.

McMartin, D. N., Sahota, P. S., Gunson, D. E., Hsu, H. H., and Spaet, R. H. (1992). Neoplasms and related proliferative lesions in control Sprague–Dawley rats from carcinogenicity studies. Historical data and diagnostic considerations. *Toxicol Pathol* **20**(2), 212–25.

OECD 451. (1993). Carcinogenicity Studies, adopted May 1981. OECD guidelines for the testing of chemicals, vol 2. Organisation for economic cooperation and development, Paris.

Perry, R., and Menton, R. (2004). Statistical methods in 2-year rodent carcinogenicity studies. *Toxicol Sci* **32**(Suppl. 1), 157.

Son, W. C., and Gopinath, C. (2004). Early occurrence of spontaneous tumors in CD-1 mice and Sprague–Dawley rats. *Toxicol Pathol* **32**(4), 371–4.

Tennekes, H., Gembardt, C., Dammann, M., and van Ravenwaay, B. (2004). The stability of historical control data for common neoplasms in laboratory rats: adrenal gland (medulla), mammary gland, liver, endocrine pancreas and pituitary gland. *Reg Toxicol Pharm* **40**, 18–27.

Toxicologic Pathology, 35:562–569, 2007

ISSN: 0192-6233 print / 1533-1601 online
DOI: 10.1080/01926230701347330

Carcinogenicity Evaluation: Comparison of Tumor Data from Dual Control Groups in the CD–1 Mouse

PAUL BALDRICK AND LESLEY REEVE

Scientific and Regulatory Consulting, Covance Laboratories Ltd., England, HG3 1PY, United Kingdom

ABSTRACT

Current regulatory thinking allows for the use of single control groups for rodent carcinogenicity testing although there has been a trend until recently to use dual control groups. To date, virtually nothing has been published on whether a shift from dual to single control groups will affect the identification of tumorigenic risk potential in these studies. A recent evaluation of dual control carcinogenicity data in the rat (Baldrick, *Toxicol Pathol* 2005, 33: 283–291) showed that although no major differences in tumor incidences between the control groups were found, some interstudy variation occurred and in cases were a notable difference was seen, the use of 2 control groups, as well as robust, contemporary background data, allowed an easier interpretation of findings in drug-treated groups. In this paper, the results of 10 mouse carcinogenicity studies, performed between 1991 and 2004, with 2 control groups, are presented. As in the rat, interstudy variation was seen and in some cases, the use of dual control groups assisted in the tumor risk assessment. Thus, the continued use of 2 control groups can have a vital role in mouse carcinogenicity studies. The paper also presents an update on survival, on the range and extent of background spontaneous neoplasms and comments on genetic drift in this commonly used mouse strain.

Keywords. CD–1 mouse; benign and malignant neoplasms; carcinogenicity; survival; dual control groups; genetic drift

INTRODUCTION

As part of the drug development process, rodent carcinogenicity studies are required by regulatory agencies for products that will have chronic clinical use or for which a particular class or mechanistic safety alert has been identified. Over the years, major region-specific (e.g., Europe, United States and Japan) differences in the design of these studies have been eliminated by the emergence of guidance documents arising from the International Conference on Harmonization (ICH), namely ICH S1A, ICH S1B, ICH S1C and ICH S1C(R). However, such guidance does not address whether one or 2 control groups need to be included as part of the study design.

Until recently, the regional situation gave a mixed picture: in Europe 2 control groups dosed with vehicle was indicated (Guidelines Volume 3B, 1998), United States guidance state one concurrent control group (FDA Redbook, 2000), which agrees with OECD requirements (OECD, 1993) and in Japan, a negative control group is mentioned (Japan, 1999). However, due to recent clarification in Europe that one control group of the same number for each sex dosed with the vehicle by the same route, is acceptable (CPMP/SWP/2879/00), it seems that "globally," regulators will now accept the single control group model. It remains to be seen whether drug companies that have routinely used double control groups, whether for regulatory compliance, scientific or traditional reasons, continue with this trend.

Although carcinogenicity studies are performed to identify tumoriogenic potential in animals as part of the assessment of risk of use in humans, little published data are available on how the actual use of 2 control groups has assisted this process. A recent review of 13 carcinogenicity studies in the rat with dual control groups showed that on occasion, the presence of both groups assisted in the interpretation of tumor findings (Baldrick, 2005). Additional controls are reported as giving a better assessment of how rare each tumor type is and allows for a more accurate comparison to historical control data (Fairweather et al., 1998). It is also said that the results from 2 identical controls can be used as a mechanism for identifying the extent of control variability and to help evaluate the biological significance of increases in tumor incidence in the treated groups (CDER, 2001).

This paper presents data from 10 in-house carcinogenicity studies using CD–1 mice with 2 identical control groups. All the studies involved oral administration and were performed in the period 1991–2004. As well as examining potential tumor incidence differences between the dual controls, the paper also provides an update on the range and extent of background neoplasms in this commonly used mouse strain and will comment on genetic drift.

MATERIALS AND METHODS

Studies for evaluation from our in-house database were limited to carcinogenicity investigations in gang-housed CD–1 (Crl:CD–1(ICR)BR) mice with 2 identical control groups using oral administration (gavage tube or dietary) in the period 1991–2004. A total of 10 studies were identified giving a total of 564 males and 564 females for each control group. Group sizes ranged from 51–60 animals/sex. Treatment comprised gavage dosing at 10 mL/kg of vehicle material or dietary intake of standard chow. Mice were treated for an approximately 2-year period (up to 105 weeks).

Mice of approximately 6–8 weeks of age at the start of dosing were obtained from Charles River UK and were generally housed 3 to a cage (one study had males housed singly

Address correspondence to: Paul Baldrick, Covance Laboratories Ltd., Otley Road, Harrogate, HG3 1PY, North Yorkshire, United Kingdom; e-mail: paul.baldrick@covance.com

and one study had both sexes housed singly). They were fed SQC Rat and Mouse Maintenance Diet No. 1 (Special Diets Service Ltd., United Kingdom) ad libitium and had free access to water. Animals were maintained on a 12h light/dark cycle at 19–25°C and 40–70% relative humidity. At the end of the treatment period, complete necropsies were conducted on the control animals and tissues were fixed and embedded in paraffin wax, sectioned (nominally 5 μm) and stained with hematoxylin and eosin. Tissues routinely examined were adrenals, brain, cecum, colon, duodenum, eyes, femur with bone marrow and articular surface, gall bladder, gross lesions, Harderian glands, heart, ileum, jejunum, kidneys, lacrimal glands, liver, lungs with mainstem bronchi, mammary gland (female), mandibular and mesenteric lymph nodes, muscle, esophagus, optic nerve, ovaries, pancreas, pituitary, prostate, rectum, salivary glands, sciatic nerves, seminal vesicles, skin, spinal cord (at 3 levels), spleen, sternum with bone marrow, stomach, testes + epididymides, thymus, thyroids + parathyroids, tissue masses, tongue, trachea, urinary bladder, uterus and vagina.

Survival after 2 years of treatment was measured. Mice were examined for non–neoplastic and neoplastic findings. Tumor identification was based on in–house nomenclature which followed that commonly used elsewhere, e.g., International Agency for Research on Cancer (IARC) and Society of Toxicologic Pathologists (STP). To allow for in–house consistency, "older" tumor synonyms have been used for some neoplasms. Statistical evaluation was performed for differences between the 2 control groups (designated control I and control II). Survival and tumor incidence was compared using a 2-sided chi-squared test. If the incidences were shown to be not significantly different, then the 2 groups were combined for comparison with drug-treated groups in each of the presented carcinogenicity studies. All studies were performed to Good Laboratory Practice (GLP) and animal care conformed to the United Kingdom Animals (Scientific Procedures) Act of 1986.

RESULTS

Control group survival at study end is given in Table 1. Overall, mean survival was similar between the 2 groups with 30–72% and 27–73% survival seen for males and 25–78% and 33–65% survival seen for females in control groups I and II, respectively. Generally, only small differences between the 2 control groups were seen for individual studies with no statistically significant differences in survival. When survival was broken down to 7-year periods of 1991–1997 and 1998–2004 (the most appropriate split based on the dates when the studies were performed), there was no obvious time effect on survival values. Across all studies, the main causes of death were consistent with those generally associated with aging laboratory mice of this strain and comprised amyloidosis (both sexes) and urogenital tract lesions (predominantly in males) as well as tumors of the hemolymphoreticular system, lung and skin/subcutis (both sexes) and liver (predominantly males).

Common nonneoplastic lesions in decedents and mice surviving to termination are listed in Table 2. There were no obvious group differences seen across the studies. The findings were consistent with the spectrum associated with aging mice of this strain and included amyloid deposition, arteritis

TABLE 1.—Survival at study end.*

Date	Group	Mean survival (%) Males	Females
1991–1997 (n = 4)	Control I	49	47
	Control II	42	48
1998–2004 (n = 6)	Control I	39	35
	Control II	42	40
		Survival range (%)	
1991–2004 (n = 10)	Control I	30–72	25–78
	Control II	27–73	33–65

* Measured at 104/105 weeks on all occasions except in 3 studies (84, 85 and 93/94 weeks).

TABLE 2.—Common non–neoplastic lesions.

System	Lesion
Endocrine	Adrenal: amyloid deposit, eosinophilic focus/nodule, corticomedullary pigmentation, subcapsular cell hyperplasia
	Pancreas: inflammatory cell foci, edema
Integumentary	Mammary gland: ductal ectasia
	Preputial / clitoral gland: abscess, duct ectasia
Reproductive	Epididymis: oligospermia
	Prostate: prostatitis, inflammatory cell foci, coagulating gland distension
	Testis: interstitial cell hyperplasia, tubular atrophy
	Seminal vesicle: distension, vesiculitis
	Penis: balanoposthitis
	Ovary: arteritis, cyst
	Uterus: adenomyosis, arteritis, cystic glands, endometrial hyperplasia, squamous metaplasia
	Vagina: squamous cell hyperplasia
Digestive	Liver: amyloid deposit, basophilic focus, cytomegaly/polyploidy/karyomegaly, focal necrosis, inflammatory cell foci/infiltration, haemopoiesis, hepatocyte vacuolation, pigmented histiocytes
	Stomach: amyloid deposit, chronic gastritis, cystic glands, hyperkeratosis
	Duodenum: amyloid deposit, avillous hyperplasia
	Jejunum: amyloid deposit
	Ileum: amyloid deposit
	Gall bladder: distension
	Salivary gland: amyloid deposit, atrophy, inflammatory cell foci
Hematopoietic/ Lymphoreticular	Mandibular lymph node: lymphoid hyperplasia, pigment Mesenteric lymph node: amyloid deposit, hemangiectasis, lymphoid hyperplasia Marrow (femur and sternum): arthropathy, chondropathy, hyperplasia
	Spleen: atrophy, hemopoiesis, lymphoid atrophy, lymphoid hyperplasia
	Thymus: arteritis, atrophy, cyst, lymphoid hyperplasia
	Thyroid: amyloid deposit, cystic follicles
Nervous	Brain: mineralisation
	Sciatic nerve: neuropathy
Urinary	Kidney: amyloid deposit, arteritis, cyst, focal nephropathy, glomerulonephropathy, hyaline droplets, hydronephrosis, inflammatory cell foci, mineralisation - various sites, segmental cortical atrophy, tubular dilation
	Urinary bladder: inflammatory cell foci
Cardiovascular	Heart: arteritis, degeneration / fibrosis, thrombosis
Respiratory	Lung: eosinophilic histiocytes, inflammatory cell foci
Others	Eye: lenticular degeneration, retinal atrophy
	Skin / subcutis: dermatitis / folliculitis
	Tail: dislocation / fracture

TABLE 3.—Summary of neoplastic incidence (mean number of neoplasms per animal).

		Male		Female	
	Date	Control I	Control II	Control I	Control II
Total neoplasms (neoplasm incidence)	1991–2004	453 (0.80)	467 (0.83)	487 (0.86)	445 (0.79)
Benign neoplasms		316	309	213	211
Malignant neoplasms		137	158	274	234
Neoplasms incidence[a]	1991–1997	0.79	0.82	0.70	0.68
	1998–2004	0.83	0.88	0.97	0.87

[a]Neoplasm incidence calculated from total number of neoplasms divided by number of animals on study.

and inflammatory cell foci across various tissues and nephropathic findings.

A summary of tumor incidences is given in Table 3 with details of major organ or common and less common tumors shown in Tables 4 and 5, respectively. Notable interstudy differences in tumor incidences between dual control groups are presented in Table 6. Table 3 shows that the total number of neoplasms was similar to each other for male and female control groups; benign tumors occurred with approximately twice the frequency of malignant tumors in males but with a fairly similar incidence to malignant tumors among females. Across all the groups, the malignant tumor incidence was 30–34% (males) and 53–56% (females). When tumor incidences were broken down to 7-year periods of 1991–1997 and 1998–2004, results were generally similar although the incidence for females was slightly increased (through malignant tumor burden) in the most recent 7-year period compared to the earlier time period.

As seen from Table 4, the most common tumors were those associated with the digestive system (comprising hepatocellular adenomas and/or carcinomas in males), the respiratory system (comprising bronchiolo-alveolar adenomas or carcinomas with a higher incidence of adenomas in males compared to females) and the hematopoietic/lymphoreticular system (comprising lymphomas [lymphocytic or pleomorphic] with a higher incidence in females compared to males). Other more frequent tumors were those of the reproductive system (comprising uterine stromal polyps, leiomyomas or histiocystic sarcomas) and the integumentary system (comprising mammary gland adenocarcinomas in females). Remaining tumors (when assessed from a full complement of mice) were present at <5%.

Overall, "notable" differences in the number of individual tumors were only seen between the control groups (control I vs. control II) on a few occasions among the combined data for the 10 studies examined, most notably as follows:

- adrenal pheochromocytomas in females (1 vs. 4);
- pancreatic islet cell adenomas in males (5 vs. 0) and females (4 vs. 1);
- skin fibrous histiocytomas in females (7 vs. 2);
- uterine leiomyomas (19 vs. 29);
- uterine histiocytic sarcomas (25 vs. 15);
- hepatocellular adenomas in females (5 vs. 9);
- hepatic histiocytic sarcomas in males (0 vs. 3);
- lymphoma lymphocytic in males (10 vs. 18);
- granulocytic leukemias in males (6 vs. 19) and females (12 vs. 6);
- hematopoietic/lymphoreticular histiocytic sarcomas in males (2 vs. 6) and females (9 vs. 3);
- Harderian gland adenomas in males (13 vs. 20).

As seen from Table 5, there was no specific pattern of incidence for less common tumors which, in the most extreme case, were only seen on one occasion among all the animals examined. It should be noted, however, that there were various occasions when a rare tumor was seen in 2 animals in one control group and not at all in another group. In agreement with the overall variation between the 2 control groups, all 10 studies individually showed cases of "notable" differences as can be seen from Table 6; these differences tended to be among the more commonly seen tumor types. As will be pointed out in the Discussion, it is the magnitude of the difference with relevance to findings in drug-treated groups and in relation to background data that are important.

DISCUSSION

This paper shows that, in general, no major differences occur for any tumor incidence between dual control groups across the full range of mammalian body systems when assessed in 10 in-house mouse carcinogenicity studies over a period of 14 years. Interstudy variation occurred for some tumors but did not follow any particular pattern. A similar result was found in a recent evaluation of dual control carcinogenicity data in the rat (Baldrick, 2005). Unfortunately, due to the general lack of publications in this area it is difficult to assess whether other authors have considered such interstudy variation important.

An evaluation of the Food and Drug Administration (FDA) and National Toxicology Program/National Cancer Institute (NTP/NCI) database from 1979 to 1992 showed that most rodent carcinogenicity studies had one control group and only 18 cases of duplicate control groups were mentioned (Contrera et al., 1997). These latter studies appear to be a series of experiments performed with 18 color additives in the CD–1 mouse (Haseman et al., 1986). Review of the findings by these latter authors showed marked study-to-study variability for certain tumors with statistically significant ($p < 0.05$) paired control differences in 7 of the studies (see Table 8).

However, when these results were analyzed to see if these findings exceeded chance expectation, no evidence of extrabinomial within-study variability between the 2 concurrent control groups was found. Indeed, the total number of observed significant paired-control differences was virtually identical to what would be expected from usual binomial model assumptions. Examination of dual control data in CD–1 mice using an unidentified chemical by Selwyn (1989) showed 3 cases of statistically significant differences ($p < 0.05$) among tumors (see Table 8). Upon analysis, it was concluded that statistically significant differences between identically treated groups will occur with regular frequency but such data do not provide strong evidence of extrabinomial variation in tumor rates. Overall, these findings support the view that caution is needed when interpreting statistically significant findings.

It would appear that although some interstudy variation can be expected in carcinogenicity study control group tumor incidences, such variation does not cause an issue in assessment for potential carcinogenic risk in most cases. This situation

TABLE 4.—Summary of major organ or common spontaneous neoplasms.

Body system	Neoplasm	Control I males		Control II males		Control I females		Control II females	
		Number (%)	% Range	Number (%)	% Range	Number (%)	% Range	Number (%)	% Range
ENDOCRINE									
Pituitary	Adenoma (B)	3 (0.5)	0–2.0	3 (0.6)	0–2.0	14 (2.5)	0–13.7	20 (3.8)	0–9.0
Adrenal	Pheochromocytoma (B)	2 (0.6)	0–3.3	2 (0.6)	0–1.7	1 (0.3)	0–1.7	4 (1.1)	0–3.3
	Subcapsular cell adenoma (B)	17 (3.3)	0–8.6	15 (2.9)	0–8.0	1 (0.2)	0–1.6	3 (0.6)	0–2.0
Pancreatic	Islet cell adenoma (B)	5 (1.1)	0–7.8	0 (0)	0	4 (0.9)	0–2.0	1 (0.2)	0–1.6
INTEGUMENTARY									
Mammary gland	Adenocarcinoma (M)	0 (0)	0	0 (0)	0	20 (4.7)	0–9.3	20 (4.6)	0–7.5
Skin	Fibrous histiocytoma (M)	0 (0)	0	1 (1.5)*	1.5*	7 (10.9)*	10.9*	2 (3.2)*	3.2*
	Fibrosarcoma (M)	6 (1.7)	0–3.9	6 (1.7)	0–8.3	0 (0)	0	2 (0.6)	0–1.7
	Sarcoma (M)	10 (1.9)	0–5.3	11 (2.2)	0–5.0	11 (2.1)	0–6.7	15 (3.1)	0–6.7
REPRODUCTIVE									
Testicular	Interstitial cell tumour (B)	11 (1.9)	0–5.0	8 (1.5)	0–4.8	—	–	—	—
Uterine	Stromal polyp (B)	—	—	—	—	35 (6.2)	1.7–17.6	28 (5.0)	0–11.8
	Hemangioma (B)	—	—	—	—	7 (2.1)	0–3.4	5 (1.3)	0–3.9
	Leiomyoma (B)	—	—	—	—	19 (3.4)	0–7.8	29 (5.2)	0–9.8
	Stromal sarcoma (M)	—	—	—	—	5 (1.0)	0–1.7	8 (1.6)	0–3.9
	Histiocytic sarcoma (M)	—	—	—	—	25 (6.2)	1.7–12.5	15 (3.7)	0–9.8
	Leiomyosarcoma (M)	—	—	—	—	4 (0.9)	0–3.3	2 (0.5)	0–1.7
Ovarian	Cystadenoma (B)	—	—	—	—	5 (0.9)	0–2.0	4 (0.7)	0–3.3
	Luteoma (B)	—	—	—	—	6 (1.4)	0–2.0	3 (0.7)	0–3.9
	Granulosa cell tumor (B)	—	—	—	—	5 (1.1)	0–2.0	4 (0.9)	0–2.0
DIGESTIVE									
Hepatic	Hepatocellular adenoma (B)	120 (26)	1.7–32.8	98 (19.5)	2.0–26.7	5 (1.0)	0–3.1	9 (0.8)	0–4.0
	Hemangioma (B)	15 (3.3)	0–7.4	11 (2.4)	0–7.4	2 (0.4)	0–2.0	0 (0)	0
	Hepatocellular carcinoma (M)	16 (2.8)	0–5.9	18 (3.2)	0–5.9	2 (0.4)	0–1.7	0 (0)	0
	Histiocytic sarcoma (M)	3 (0.7)	0–2.0	0 (0)	0	3 (0.7)	0–2.0	1 (0.2)	0–4.5
HAEMATOPOIETIC/LYMPHORETICULAR									
Splenic	Haemangioma (B)	3 (0.7)	0–3.3	5 (1.1)	0–2.0	2 (0.4)	0–2.0	2 (0.4)	0–1.7
Systemic	Lymphoma lymphocytic (M)	10 (3.3)	0–20	18 (6.4)	0–27.8	30 (9.2)	0–29.2	28 (9.6)	0–40
	Lymphoma pleomorphic (M)	24 (7.9)*	0–100*	27 (9.6)*	0–71.4*	54 (16.5)*	2–62.5*	54 (18.6)*	7.8–52.4*
	Lymphoma (M)	13 (3.6)	0–13.3	15 (4.5)	0–18.6	32 (8.4)	0–23.3	27 (7.8)	0–21.7
	Lymphoma lymphoblastic (M)	4 (3.2)	0–6.8	3 (2.3)	1.7–5.6	2 (1.4)	0–3.3	3 (2.3)	0–3.9
	Granulocytic leukemia (M)	6 (1.4)*	0–20*	19 (4.8)*	0–27.8*	12 (2.7)*	0–8.3*	6 (1.5)*	0–19.1*
	Histiocytic sarcoma (M)	2 (0.6)	0–20	6 (1.9)	0–5.6	9 (2.7)	0–12.5	3 (0.9)	0–4.8
URINARY									
Renal	Tubular cell adenoma (B)	3 (0.5)	0–1.7	2 (0.4)	0–1.7	0 (0)	0	0 (0)	0
RESPIRATORY									
Lung	Adenoma (B)	5 (9.8)*	9.8*	8 (15.7)*	15.7*	9 (17.6)*	17.6*	9 (17.6)*	17.6*
	Bronchiolo–alveolar adenoma (B)	89 (17.1)	11.7–25	111 (22.2)	18.3–35.5	52 (10.0)	4.7–13.7	47 (9.7)	3.3–15.7
	Bronchiolo–alveolar carcinoma (M)	36 (6.9)	3.9–13.3	29 (5.8)	1.7–19.6	29 (5.6)	0–11.8	20 (4.1)	0–8.3
	Carcinoma (M)	5 (9.8)*	9.8*	2 (3.9)*	3.9*	2 (3.9)*	3.9	3 (5.9)*	5.9*
OTHER TUMORS									
Harderian gland	Adenoma (B)	13 (9.8)*	0–100*	20 (15.2)*	0–100*	12 (9.4)*	0–100	15 (11.8)*	0–100*

B, Benign; M, Malignant.
*Only examined in a limited number of studies and/or a reduced number of animals assessed (which gave an artificially high percentage incidence and upper range value)

TABLE 5.—Summary of less common spontaneous neoplasms.

Body system	Neoplasm (number of tumors/number of animals examined)
ENDOCRINE	
Pituitary	Carcinoma (M): 1/64 CIM
Adrenal	Adenoma (B): 1/51 CIIM; Cortical adenoma (B): 2/237 CIM, 1/204 CIIM, 1/234 CIF
Thyroid	Follicular cell adenoma (B): 1/294 CIM, 1/289 CIIM, 2/293 CIF
Parathyroid	Adenoma (B): 1/163 CIF, 1/163 CIIF
INTEGUMENTARY	
Mammary gland	Adenoma (B): 2/168 CIIF; Carcinosarcoma (M): 1/25 CIM, 3/280 CIF, 2/253 CIIF
Skin	Squamous cell papilloma (B): 1/568 CIM, 2/568 CIF, 1/542 CIIF; Keratoacanthoma (B): 1/51 CIF; Hair follicle tumour (B): 1/108 CIM; Histiocytoma (B): 2/188 CIM; Rhabdomyosarcoma (B): 1/68 CIIM; Sebaceous cell adenoma (B): 2/111 CIM, 1/111 CIIF; Hemangioma (B): 1/238 CIM; Trichoepithelioma (B): 1/51 CIF; Melanoma (B): 1/60 CIM; Squamous cell carcinoma (M): 1/291 CIIM, 2/291 CIF, 1/291 CIIF; Basal cell tumor (M): 1/162 CIF; Histiocytic sarcoma (M): 1/273 CIM, 1/275 CIIM, 1/275 CIF; Hemangiosarcoma (M): 1/349 CIM, 2/346 CIF, 2/345 CIIF; Osteosarcoma (M): 1/406 CIM, 1/381 CIIM, 1/406 CIF, 2/380 CIIF
Zymbal gland	Sebaceous cell carcinoma (M): 1/60 CIF; Carcinoma (M): 1/60 CIF
REPRODUCTIVE	
Testicular	Hemangioma (B): 1/180 CIM, 1/179 CIIM; Rete testis adenoma (B): 1/60 CIM, 1/60 CIIM
Epididymidial	Histiocytoma (B): 2/348 CIIM; Histiocytic sarcoma (M): 1/51 CIIM
Uterine	Adenoma (B): 2/335 CIIF; Granular cell tumor (B): 2/243 CIIF; Blood vessel hemangioma (B): 2/51 CIF; Hemangiosarcoma (M): 1/243 CIIF; Osteosarcoma (M): 1/60 CIIF; Sarcoma (M): 1/60 CIIF
Ovarian	Tubulostromal adenoma (B): 2/162 CIF, 1/162 CIIF; Sex cord stromal tumor (B): 2/352 CIF, 2/354 CIIF; Leiomyoma (B): 1/180 CIIF; Sertoli cell tumour (B): 1/64 CIF
Cervix	Stromal polyp (B): 1/59 CIF; Leiomyoma (B): 2/119 CIIF; Histiocytic sarcoma (M): 1/59 CIF
Vagina	Histiocytoma (B): 1/63 CIIF; Sarcoma (M): 1/60 CIF; Histiocytic sarcoma (M): 1/58 CIIF
Preputial/Clitoral gland	Histiocytoma (B): 1/32 CIM, 2/34 CIIM, 1/59 CIF
DIGESTIVE	
Hepatic	Blood vessel hemangioma (B): 1/20 CIIM, 1/22 CIIF; Histiocytic carcinoma (M): 1/51 CIM, Hemangiosarcoma (M): 2/288 CIIM, 1/285 CIIF
Gall bladder	Papilloma (B): 1/201 CIM, 2/177 CIIM, 2/184 CIIF
Tongue	Squamous cell papilloma (B): 2/131 CIM, 1/129 CIIM
Salivary gland	Adenocarcinoma (M): 1/63 CIF; Carcinoma (M): 1/60 CIIF
Esophageal	Leiomyosarcoma (M): 1/51 CIM
Stomach	Squamous cell papilloma (B): 2/324 CIM, 2/333 CIF, 1/334 CIIF; Adenoma (B): 2/210 CIIM; Hemangioma (B): 1/66 CIIM; Leiomyoma (B): 1/49 CIF; Osteosarcoma (M): 1/113 CIM, 1/111 CIIF
Duodenal	Adenoma (B): 1/223 CIIF
Jejunal	Adenoma (B): 1/48 CIIF
Cecal	Carcinoma (M): 1/99 CIIM
Colonic	Hemangiosarcoma (M): 1/50 CIIM
HAEMATOPOIETIC/LYMPHORETICULAR	
Thymic	Thymona (B): 1/155 CIF, 1/137 CIIF; Hemangioma (B): 2/107 CIF
Splenic	Blood vessel hemangioma (B): 1/51 CIM, 1/51 CIF; Lymphocytic leukemia (M): 1/32 CIIM, 3/72 CIF, 2/39 CIIF; Hemangiosarcoma (M): 2/350 CIM, 2/350 CIIM
Mesenteric lymph node	Hemangioma (B): 2/175 CIM; Histiocytic sarcoma (M): 1/51 CIIF
Systemic	Lymphoma plasmacytic (M): 1/135 CIF; Leukemia (M): 1/290 CIM, 2/259 CIIM, 1/291 CIF; Undifferentiated leukaemia (M): 1/5 CIM; Mast cell tumor (M): 1/126 CIM, 1/99 CIIM
Femur + marrow	Hemangioma (B): 1/187 CIM; Osteoma (B): 2/296 CIM, 1/298 CIIM, 1/294 CIF
Sternum + marrow	Hemangioma (B): 1/119 CIM, 1/118 CIF
URINARY	
Renal	Tubular cell carcinoma (M): 3/51 CIIM; Sarcoma (M): 1/111 CIF
MUSCULOSKELETAL	
Connective tissue	Histiocytic sarcoma (M): 2/10 CIF, 1/5 CIIF
Bone	Osteoma (B): 1/3 CIM, 1/9 CIIM, 1/6 CIF, 1/6 CIIF; Hemangiosarcoma (M): 1/20 CIF; Osteosarcoma (M): 2/9 CIM, 1/10 CIIM, 2/3 CIF, 1/5 CIIF
Muscle	Hemangiosarcoma (M): 1/111 CIF
OTHER TUMORS	
Thoracic cavity	Sarcoma (M): 1/24 CIF, 1/13 CIIF; Hemangiosarcoma (M): 1/8 CIF
Abdominal cavity	Hemangioma (B): 1/5 CIF; Histiocytic sarcinoma (M): 1/8 CIM, 1/10 CIIM; Sarcoma (M): 1/6 CIIM, 2/19 CIIF
Oral cavity	Odontogenic tumor (B): 2/2 CIIM; Squamous cell papilloma (B): 1/1 CIIM, 1/1 CIIF
Tail	Hemangioma (B): 1/76 CIM, 1/77 CIF, 1/79 CIIF
Foot/leg	Squamous cell papilloma (B): 1/2 CIIF; Hemangioma (B): 1/7 CIM
Ear	Histiocytoma (B): 1/8 CIIM; Sarcoma (M): 1/4 CIIM
Blood vessel (unspecified)	Hemangioma (B): 3/60 CIIM, 4/60 CIF, 2/60 CIIF; Hemangiosarcoma (M): 2/62 CIM, 1/61 CIF
Harderian gland	Adenocarcinoma (M): 1/129 CIIM, 1/125 CIF, 3/125 CIIF

B, Benign; M, Malignant; CIM, Control group I male; CIIM, Control group II male; CIF, Control group I female; CIIF, Control group II female.

is supported by the fact that dual control tumor data are usually added together to allow comparison with drug-treated groups. Where dual control data come into their own is when variation (as seen in the present study between concurrent control groups of for example, 3 vs. 11 systemic lymphomas for males in one study and 0 vs. 4 pituitary adenomas for females in another study) can assist interpretation of an increased incidence of the tumor in the high drug-treated group that would be "diluted" if the 2 control group data are summed.

As indicated earlier, it is not necessarily important if the difference in tumor incidence between dual controls is statistically significant in itself; what is important is the magnitude of the difference in conjunction with factors such as preneoplastic/hyperplastic changes, evidence of increasing tumor incidence with increasing dose and background data

TABLE 6.—Summary of notable dual control group tumor differences in the present investigation.

Study	Neoplasm	Males		Females	
		Control I	Control II	Control I	Control II
1	Pancreatic islet cell adenoma (B)	4/51	0/51	—	—
	Uterine leiomyoma (B)	—	—	2/51	5/51
	Lung carcinoma (M)	5/51	2/51	—	—
2	Uterine stromal polyp (B)	—	—	3/51	0/50
	Bronchiolo–alveolar adenoma (B)	7/51	12/51	—	—
3	Bronchiolo–alveolar adenoma (B)	—	—	5/51	1/22
4	Pituitary adenoma (B)	—	—	0/59	4/59
	Adrenal subcapsular cell adenoma (B)	5/58	2/58	—	—
	Skin fibrosarcoma (M)	2/59	5/60	—	—
	Uterine histiocytic sarcoma (M)	—	—	6/59	0/60
	Systemic lymphoma (M)	3/60	11/59	—	—
	Bronchiolo–alveolar adenoma (B)	—	—	3/60	8/60
	Bronchiolo–alveolar carcinoma (M)	4/60	1/60	—	—
5	Skin sarcoma (M)	—	—	1/60	4/60
	Hepatocellular adenoma (B)	12/58	7/59	—	—
	Systemic lymphoma (M)	8/60	3/60	—	—
	Bronchiolo–alveolar adenoma (B)	—	—	7/60	4/60
	Blood vessel hemangiosarcoma (B)	0/60	3/60	—	—
6	Bronchiolo–alveolar adenoma (B)	7/60	12/60	6/60	2/60
7	Skin sarcoma (M)	—	—	0/60	3/60
	Hepatocellular adenoma (B)	8/60	17/60	0/60	3/60
8	Skin malignant fibrous histiocytoma (M)	—	—	7/64	2/63
	Uterine histiocytic sarcoma (M)			8/64	3/63
	Bronchiolo–alveolar adenoma (B)	8/68	16/68	—	—
	Bronchiolo–alveolar carcinoma (M)	9/68	5/68	5/64	2/63
	Harderian gland adenoma (B)	5/68	13/67	—	—
9	Pituitary adenoma (B)	—	—	0/59	3/60
	Skin sarcoma (M)	—	—	4/60	1/60
	Uterine leiomyoma (B)			2/60	5/60
	Hepatocellular adenoma (B)	11/60	7/60	—	—
	Systemic lymphocytic lymphoma (M)	—	—	5/60	1/60
	Systemic pleomorphic lymphoma (M)	—	—	9/60	14/60
	Systemic lymphoblastic lymphoma (M)	4/59	1/59	—	—
	Systemic histiocytic sarcoma (M)	0/59	3/59	4/60	0/60
10	Uterine stromal polyp (B)			9/51	6/51
	Uterine histiocytic sarcoma (M)			2/51	5/51
	Hepatocellular adenoma (B)	8/51	2/51	—	—
	Systemic lymphocytic lymphoma (M)	—	—	2/51	5/51
	Renal tubular cell carcinoma (M)	0/51	3/51	—	—
	Bronchiolo–alveolar carcinoma (M)	—	—	6/51	1/51

incidence. However, such an argument can work both ways. In the case of an incidence of 8/51 vs. 2/51 hepatocellular adenomas in males for the 2 control groups in one study, an incidence of 8/51 in the high drug-treated group could be easily discounted as non–treatment-related and that one control group would have supported such a conclusion as the incidence was within the background data range.

In the case of an incidence of 4/51 vs. 0/51 pancreatic islet cell adenomas in males in the concurrent controls in another study, it would be more difficult to explain an incidence of for example, 4 such tumors in the high-dose group in the absence of both control groups as background data may not have been useful (Table 4 shows that only 5/452 males in Control I and 0/421 males in Control II had this tumor). The use of dual control groups to support interpretation of rare tumors is also valid, although results from the presented 10 carcinogenicity studies indicate that such tumors were often seen in both control groups anyway.

Overall, in all these cases, scientific assessment would involve a careful evaluation of tumor incidence across all study groups (control and drug-treated) with reference to contemporary mean and overall range tumor values. Indeed, it has been pointed out elsewhere that 2 control groups within the same study can occasionally show notable differences, stressing the importance of knowing the range of incidences rather than just mean incidence (Gopinath, 1994). Results from the present investigation show that the continued use of dual control groups in carcinogenicity studies has a role in the assessment of tumorigenic risk.

Various authors have reported survival, cause of mortality and/or spontaneous neoplasm data for CD–1 mice (e.g., Maita et al., 1988; Chandra and Frith, 1992; Criver, 1995; Gopinath, 1994; Son, 2003; Son and Gopinath, 2004; Criver, 2005; Hooks et al, 2005). Maita et al. (1988) showed survival of 20–44% (males) and 27–43% (females) from 11 carcinogenicity studies in the late 1970s/early 1980s while Chandra and Frith (1992) reported values of 41% (males) and 47% (females) from a survey of 11 in-house studies largely conducted in the late 1980s; further data from 9 studies of 2-year duration again largely from the late 1980s showed survival ranges of 28–72% (males) and 27–65% (females) (Criver, 1995). Further authors report survival of 48% (males) and 44% (females) from a range of studies in the period 1995–2001 (Hooks et al, 2005) plus survival values of 18–66% (males) and 19–55% (females) from 14 studies of 2–year duration from 1990–2000 (Criver, 2005). These latter values compare well with the survival range of 27–73% (males) and 25–78% (females) or overall survival means of 43% (males) and 42% (females) seen in the present reported studies.

TABLE 7.—Comparison of common spontaneous neoplasms found in the present investigation with that found by other authors.

	Incidence (%)											
	Present study		Maita et al (1988)		Chandra and Frith (1992)		Gopinath (2004)		Criver (1995)		Criver (2005)	
Tumor	M	F	M	F	M	F	M	F	M	F	M	F
Hepatocellular adenoma	20–26	++	26	5	11	2	16	3	19	3	11	++
Bronchiolo–alveolar adenoma	17–22	10	15	15	19	12	12	7	17	10	14	10
Lymphatic lymphoma	3–6	9–10	NM	NM	4	9	NM	NM	++	2	++	++
Lymphoma–total	5–6	10–11	8	22	4	11	NM	NM	7	16	5	11
Bronchiolo–alveolar carcinoma	6–7	4–6	19	12	3	2	11	8	11	7	7	5
Uterine stromal polyp	—	5–6	—	NM	—	4	—	NM	—	6	—	5
Uterine histiocytic sarcoma	—	4–6	—	NM	—	NM	—	NM	—	NM	—	NM
Uterine leiomyoma	—	3–5	—	NM	—	4	—	NM	—	2	—	2
Mammary gland adenocarcinoma	—	5	—	6	—	3	—	5	—	5	—	2
Pituitary adenoma	++	3–4	++	5	++	++	NM	NM	++	4	++	2
Hepatocellular carcinoma	3	++	9	++	6	++	17	++	13	2	6	++
Harderian gland adenoma	10–15+	9–12+	10	9	3	2	NM	NM	13+	5	5	2

* Tumor type only examined in a limited number of studies and/or a reduced number of animals assessed (which gave an artificially high percentage incidence).
** Incidence of 1% or less.
NM, Not mentioned; M, Male; F, Female.

The major factors contributory to death in the present study were amyloidosis and urogenital tract lesions as well as tumors of the hemolymphoreticular system, lung and skin/subcutis and liver. These factors compare favorably with those reported elsewhere in CD–1 mice (e.g., Maita et al., 1988; Chandra and Frith, 1992; Son, 2003; Son and Gopinath, 2004; Hooks et al., 2005).

The present investigation shows that benign tumors were just over twice as common as malignant tumors in males (68 vs. 32%) while a fairly similar incidence occurred in females (46 vs. 54%); these values are similar to the 70 vs. 30% (males) and 55 vs. 45% (females) incidence, respectively reported elsewhere (Chandra and Frith, 1992). Surprisingly few publications cite actual frequency of neoplasms but as can be seen from Table 7, a fairly consistent pattern of common tumors has been reported for the CD–1 mouse with the most frequent being hepatocellular adenomas and/or carcinomas in males, bronchiolo-alveolar adenomas and carcinomas in both sexes and systemic lymphomas which occurred at a higher incidence in females. A similar finding has also been reported by Hooks et al. (2005).

Uterine tumors (notably stromal polyp and leiomyomas), mammary gland adenocarcinomas and pituitary adenomas were also at a frequency of 2–6% across the literature and the present investigation. Harderian gland adenomas were also seen often in control animals but calculation of an overall background incidence is made difficult as, on occasion, not all the mice available in the data sets were examined (e.g., only about 25% of available mice were actually examined for this tumor type in the present study). As pointed out elsewhere for Harderian gland neoplasm evaluation, some tissues are not routinely included in pathology protocols but are only examined when grossly abnormal (Hardisty, 1985). Values for less common tumors were also found to compare well with those given in the published literature. Interestingly, no specific tumors of the nervous or cardiovascular systems were detected which confirmed the rare nature of such neoplasms reported elsewhere (e.g., Maita et al., 1988; Chandra and Frith, 1992; Criver, 1995, 2005).

Overall, the general similarity of tumor incidence in the present study to those reported elsewhere for CD–1 mice is comforting bearing in mind probable differences in study design. Such differences include mouse strain/supplier, diet (composition and amount), caging (single vs. gang housed), environmental conditions and pathology criteria. The lack of any major shifts in tumors incidences across different investigations raises the question of what role genetic drift may have. Gopinath (1994) has reported that variation of incidence of tumor types within the same strain and sex and, within the same laboratory, occur due to the progression of time. In order to avoid such "genetic drift" when using historical data, it has been suggested that only studies performed during the previous 3–4 years (Haseman et al., 1984; Fairweather et al., 1998), 3–5 years (Greim et al., 2003 referring to the

TABLE 8.—Summary of notable dual control group tumor differences in a carcinogenicity assessment of color additives* and one unidentified chemical**

		Males		Females	
Study	Neoplasm	Control I	Control II	Control I	Control II
		Color additives			
1	Lung	12/59	4/59	—	—
2	Lung	5/60	15/59	—	—
3	Liver	3/60	10/60	—	—
4	Hematopoietic system	2/60	9/60	—	—
5	Hematopoietic system	7/60	1/60	—	—
6	Uterus leiomyoma/leiomyosarcoma	—	—	0/60	5/60
7	Uterus endometrial stromal	—	—	6/46	1/49
		Chemical			
—	Total blood vessel tumors	3/100	13/100	—	—
	Reticulum cell sarcoma	7/100	1/100	—	—
	Lymphosarcoma	—	—	17/100	29/100

* Adapted from Haseman et al. (1986); ** Data from Selwyn (1989).

European Registry of Industrial Toxicology Animal database), 5 years (CPMP/SWP/2879/00) or 7 years (Haseman et al., 1998) prior to the new study, should be accessed.

From an analysis of certain tumor types over a period of 9 years (under conditions were the potential sources of variability were recognised and strictly controlled), Gopinath (1994) showed that although there were fluctuations, incidence rates remained within a small band and no clear evidence of a time-related drift in rates was found. Furthermore, a review of the historical control data for 20 carcinogenicity studies generated in 1990–2002 in CD–1 mice showed little variability with time (Son and Gopinath, 2004). An evaluation of tumor incidence using 7-year time points over the 14-year period in the present investigation has also shown no obvious difference in the number or pattern of tumors recorded. Standardizing tumor burden by incidence in the total number of animals assessed has not shown any strong evidence of drift although a slightly higher incidence of tumors for females was seen in the most recent 7–year period.

In conclusion, this paper shows that no major differences in tumor incidences occur when dual control groups are used for mouse carcinogenicity studies. However, occasional interstudy variation occurred for some tumors indicating that the use of 2 control groups has a role in the interpretation of findings in drug-treated groups. Such interstudy variation assists interpretation when one control group, but not the other, mirrors elevated tumor incidence in one or more of the drug-treated groups. Thus, dual control groups give greater reassurance in assessing for biological significance of any increased tumor incidence. Such use also allows the buildup of contemporary background data at a greater rate.

ACKNOWLEDGMENTS

We would like to thank Chris Barton for his kind and constructive comments during the writing of this paper.

REFERENCES

Baldrick, P. (2005). Carcinogenicity evaluation: comparison of tumor data from dual control groups in the Sprague–Dawley rat. *Toxicol Pathol* **33**, 283–91.

CDER (2001). Guidance for industry: Statistical aspects of the design, analysis, and interpretation of chronic rodent carcinogenicity studies of pharmaceuticals. Issued as draft document in May 2001. Available at: (http://www.fda.gov/cder/guidance).

Chandra, M., and Frith, C. H. (1992). Spontaneous neoplasms in aged CD–1 mice. *Toxicol Lett* **61**, 67–74.

Contrera, J. F., Jacobs, A. C., and DeGeorge, J. J. (1997). Carcinogenicity testing and the evaluation of regulatory requirements for pharmaceuticals. *Rev Toxicol Pharm* **25**(2), 130–45.

CPMP/SWP/2877/00: Note for guidance on carcinogenic potential. Operational January 2003. Available at: (http://www.eudra.org/humandocs).

Criver (1995). Spontaneous neoplastic lesions in Crl:CD–1BR mouse. Charles River Laboratories, March 1995. (www.criver.com/flex_content_area/documents/rm_rm_r_lesions_crl_cd1_br_mouse.pdf).

Criver (2005). Spontaneous neoplastic lesions in Crl:CD–1(ICR) mouse in control groups from 18 month to 2 year studies. Charles River Laboratories, March 2005. (www.criver.com/flex_content_area/documents/rm_rm_r_lesions_crlcd_1_icr_mouse.pdf).

Fairweather, W. R., Bhattacharyya, A., Ceuppens, P. P., Hermann, G., Hothorn, L. A., Kodell, R. L., Lin, K. K., Mager, H., Middleton, B. J., Slob, W., Soper, K. A., Stallard, N., Ventre, J., and Wright, J. (1998). Biostatistical methodology in carcinogenicity studies. *Drug Inform Jour* **32**, 401–21.

FDA Redbook 2000. Toxicological principles for the safety assessment of food ingredients. January 2006. IV.C.6. Carcinogenicity studies with rodents. US Food and Drug Administration, Office of Food Additive Safety.

Gopinath, C. (1994). Spontaneous tumor rates: Their use to support rodent bioassays. *Toxicol Pathol* **22**(2), 160–4.

Greim, H., Gelbke, H. P., Reuter, U., Thielmann, H. W., and Edler, L. (2003). Evaluation of historical controls in carcinogenicity studies. *Hum Exp Tox* **22**, 541–9.

Guidelines Volume 3B–medicinal products for human use. Safety, environment and information. In "The rules governing medicinal products in the European Union." Office for Official Publications of the European Communities, Luxembourg, 1998.

Hardisty, J. F. (1985). Factors influencing laboratory animal spontaneous tumor profiles. *Toxicol Pathol* **13**(2), 95–104.

Hooks, W., Groom, C. J., and Stewart, J. K. (2005). In–life parameters and tumour data from Charles River Crl:CD–1(ICR)BR mouse dietary and oral gavage tumorigenicity studies terminated after two years. *Toxicol Sci* **84**(S–1), 1639.

Haseman, J. K., Huff, J., and Boorman, G. A. (1984). Use of historical control data in carcinogenicity studies in rodents. *Toxicol Pathol* **12**(2), 126–35.

Haseman, J. K., Winbush, J. S., and O'Donnell, M. W. Jr. (1986). Use of dual control groups to estimate false positive rates in laboratory animal carcinogenicity studies. *Fund Appl Toxicol* **7**(4), 573–84.

Haseman, J. K., Hailey, J. R., and Morris, R. W. (1998). Spontaneous neoplasm incidences in Fischer 344 rats and $B6C3F_1$ mice in two–year carcinogenicity studies: a National Toxicology Program update. *Toxicol Pathol* **26**(3), 428–41.

ICH S1A: Guideline on the need for carcinogenicity studies of pharmaceuticals. Operational July 1996. Available at: (http://www.eudra.org /humandocs).

ICH S1B: Testing for carcinogenicity of pharmaceuticals. Operational February 1998. Available at: (http://www.eudra.org/humandocs).

ICH S1C: Dose selection for carcinogenicity studies of pharmaceuticals. Operational March 1995. Available at: (http://www.eudra.org/humandocs).

ICH S1C(R): Guidance on dose selection for carcinogenicity studies of pharmaceuticals: Addendum on a limit dose and related notes. Operational December 1997. Available at: (http://www.eudra.org/humandocs).

Japan's and ICH Guidelines for New Drug Registration, Yakuji Nippon Ltd. 1999.

Maita, K., Hirano, M., Harada, T., Mitsumori, K., Yoshida, A., Takahashi, K., Nakashima, N., Kitazawa, T., Enomoto, A., Inui, K., and Shirasu, Y. (1988). Mortality, major cause of moribundity and spontaneous tumors in CD–1 mice. *Toxicol Pathol* **16**(2), 340–9.

OECD 451 "Carcinogenicity Studies," adopted May 1981. OECD guidelines for the testing of chemicals, vol 2. Organisation for economic co–operation and development, Paris, 1993.

Selwyn, M. R. (1989). Dual controls, p–value plots and the multiple testing issue in carcinogenicity studies. *Environ Health Perspect* **82**, 337–44.

Son, W. C. (2003). Factors contributory to early death of young CD–1 mice in carcinogenicity studies. *Toxicol Lett* **145**, 88–98.

Son, W. C., and Gopinath, C. (2004). Early occurrence of spontaneous tumors in CD–1 mice and Sprague–Dawley rats. *Toxicol Path* **32**(4), 371–4.

Unclassified ENV/JM/MONO(2011)47

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

13-Apr-2012

English - Or. English

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

Cancels & replaces the same document of 03 April 2012

GUIDANCE DOCUMENT 116 ON THE CONDUCT AND DESIGN OF CHRONIC TOXICITY AND CARCINOGENICITY STUDIES, SUPPORTING TEST GUIDELINES 451, 452 AND 453 2ND EDITION

Series on Testing and Assessment

No. 116

JT03319769

ENV/JM/MONO(2011)47
Unclassified
English - Or. English

ENV/JM/MONO(2011)47

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 116

GUIDANCE DOCUMENT 116 ON THE CONDUCT AND DESIGN OF CHRONIC TOXICITY AND CARCINOGENICITY STUDIES, SUPPORTING TEST GUIDELINES 451, 452 AND 453 – 2ND EDITION

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris 2012

Also published in the Series on Testing and Assessment:

No. 1, *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals (1993; reformatted 1995, revised 2006)*

No. 2, *Detailed Review Paper on Biodegradability Testing (1995)*

No. 3, *Guidance Document for Aquatic Effects Assessment (1995)*

No. 4, *Report of the OECD Workshop on Environmental Hazard/Risk Assessment (1995)*

No. 5, *Report of the SETAC/OECD Workshop on Avian Toxicity Testing (1996)*

No. 6, *Report of the Final Ring-test of the Daphnia magna Reproduction Test (1997)*

No. 7, *Guidance Document on Direct Phototransformation of Chemicals in Water (1997)*

No. 8, *Report of the OECD Workshop on Sharing Information about New Industrial Chemicals Assessment (1997)*

No. 9, *Guidance Document for the Conduct of Studies of Occupational Exposure to Pesticides during Agricultural Application (1997)*

No. 10, *Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data (1998)*

No. 11, *Detailed Review Paper on Aquatic Testing Methods for Pesticides and industrial Chemicals (1998)*

No. 12, *Detailed Review Document on Classification Systems for Germ Cell Mutagenicity in OECD Member Countries (1998)*

No. 13, *Detailed Review Document on Classification Systems for Sensitising Substances in OECD Member Countries 1998)*

No. 14, *Detailed Review Document on Classification Systems for Eye Irritation/Corrosion in OECD Member Countries (1998)*

No. 15, *Detailed Review Document on Classification Systems for Reproductive Toxicity in OECD Member Countries (1998)*

No. 16, *Detailed Review Document on Classification Systems for Skin Irritation/Corrosion in OECD Member Countries (1998)*

No. 17, *Environmental Exposure Assessment Strategies for Existing Industrial Chemicals in OECD Member Countries (1999)*

No. 18, *Report of the OECD Workshop on Improving the Use of Monitoring Data in the Exposure Assessment of Industrial Chemicals (2000)*

No. 19, *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (1999)*

No. 20, *Revised Draft Guidance Document for Neurotoxicity Testing (2004)*

No. 21, *Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals (2000)*

No. 22, *Guidance Document for the Performance of Out-door Monolith Lysimeter Studies (2000)*

No. 23, *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures (2000)*

No. 24, *Guidance Document on Acute Oral Toxicity Testing (2001)*

No. 25, *Detailed Review Document on Hazard Classification Systems for Specifics Target Organ Systemic Toxicity Repeated Exposure in OECD Member Countries (2001)*

No. 26, *Revised Analysis of Responses Received from Member Countries to the Questionnaire on Regulatory Acute Toxicity Data Needs (2001)*

No 27, *Guidance Document on the Use of the Harmonised System for the Classification of Chemicals which are Hazardous for the Aquatic Environment (2001)*

No 28, *Guidance Document for the Conduct of Skin Absorption Studies (2004)*

No 29, *Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (2001)*

No 30, *Detailed Review Document on Hazard Classification Systems for Mixtures (2001)*

No 31, *Detailed Review Paper on Non-Genotoxic Carcinogens Detection: The Performance of In-Vitro Cell Transformation Assays (2007)*

No. 32, *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (2000)*

No. 33, *Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures (2001)*

No. 34, *Guidance Document on the Development, Validation and Regulatory Acceptance of New and Updated Internationally Acceptable Test Methods in Hazard Assessment (2005)*

No. 35, *Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies (2002)*

No. 36, *Report of the OECD/UNEP Workshop on the use of Multimedia Models for estimating overall Environmental Persistence and long range Transport in the context of PBTS/POPS Assessment (2002)*

No. 37, *Detailed Review Document on Classification Systems for Substances Which Pose an Aspiration Hazard (2002)*

No. 38, *Detailed Background Review of the Uterotrophic Assay Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Assay (2003)*

No. 39, *Guidance Document on Acute Inhalation Toxicity Testing (in preparation)*

No. 40, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures Which Cause Respiratory Tract Irritation and Corrosion (2003)*

No. 41, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which in Contact with Water Release Toxic Gases (2003)*

No. 42, *Guidance Document on Reporting Summary Information on Environmental, Occupational and Consumer Exposure (2003)*

No. 43, *Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment (2008)*

No. 44, *Description of Selected Key Generic Terms Used in Chemical Hazard/Risk Assessment (2003)*

No. 45, *Guidance Document on the Use of Multimedia Models for Estimating Overall Environmental Persistence and Long-range Transport (2004)*

No. 46, *Detailed Review Paper on Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances (2004)*

No. 47, *Detailed Review Paper on Fish Screening Assays for the Detection of Endocrine Active Substances (2004)*

No. 48, *New Chemical Assessment Comparisons and Implications for Work Sharing (2004)*

No. 49, *Report from the Expert Group on (Quantitative) Structure-Activity Relationships [(Q)SARs] on the Principles for the Validation of (Q)SARs (2004)*

No. 50, *Report of the OECD/IPCS Workshop on Toxicogenomics (2005)*

No. 51, *Approaches to Exposure Assessment in OECD Member Countries: Report from the Policy Dialogue on Exposure Assessment in June 2005 (2006)*

No. 52, *Comparison of emission estimation methods used in Pollutant Release and Transfer Registers (PRTRs) and Emission Scenario Documents (ESDs): Case study of pulp and paper and textile sectors (2006)*

No. 53, *Guidance Document on Simulated Freshwater Lentic Field Tests (Outdoor Microcosms and Mesocosms) (2006)*

No. 54, *Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application (2006)*

No. 55, *Detailed Review Paper on Aquatic Arthropods in Life Cycle Toxicity Tests with an Emphasis on Developmental, Reproductive and Endocrine Disruptive Effects (2006)*

No. 56, *Guidance Document on the Breakdown of Organic Matter in Litter Bags (2006)*

No. 57, *Detailed Review Paper on Thyroid Hormone Disruption Assays (2006)*

No. 58, *Report on the Regulatory Uses and Applications in OECD Member Countries of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models in the Assessment of New and Existing Chemicals (2006)*

No. 59, *Report of the Validation of the Updated Test Guideline 407: Repeat Dose 28-Day Oral Toxicity Study in Laboratory Rats (2006)*

No. 60, *Report of the Initial Work Towards the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1A) (2006)*

No. 61, *Report of the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1B) (2006)*

No. 62, *Final OECD Report of the Initial Work Towards the Validation of the Rat Hershberger Assay: Phase-1, Androgenic Response to Testosterone Propionate, and Anti-Androgenic Effects of Flutamide (2006)*

No. 63, *Guidance Document on the Definition of Residue (2006)*

No. 64, *Guidance Document on Overview of Residue Chemistry Studies (2006)*

No. 65, *OECD Report of the Initial Work Towards the Validation of the Rodent Uterotrophic Assay - Phase 1 (2006)*

No. 66, *OECD Report of the Validation of the Rodent Uterotrophic Bioassay: Phase 2. Testing of Potent and Weak Oestrogen Agonists by Multiple Laboratories (2006)*

No. 67, *Additional data supporting the Test Guideline on the Uterotrophic Bioassay in rodents (2007)*

No. 68, *Summary Report of the Uterotrophic Bioassay Peer Review Panel, including Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the follow up of this report (2006)*

No. 69, *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models (2007)*

No. 70, *Report on the Preparation of GHS Implementation by the OECD Countries (2007)*

No. 71, *Guidance Document on the Uterotrophic Bioassay - Procedure to Test for Antioestrogenicity (2007)*

No. 72, *Guidance Document on Pesticide Residue Analytical Methods (2007)*

No. 73, *Report of the Validation of the Rat Hershberger Assay: Phase 3: Coded Testing of Androgen Agonists, Androgen*

Antagonists and Negative Reference Chemicals by Multiple Laboratories. Surgical Castrate Model Protocol (2007)

No. 74, *Detailed Review Paper for Avian Two-generation Toxicity Testing (2007)*

No. 75, *Guidance Document on the Honey Bee (Apis Mellifera L.) Brood test Under Semi-field Conditions (2007)*

No. 76, *Final Report of the Validation of the Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances: Phase 1 - Optimisation of the Test Protocol (2007)*

No. 77, *Final Report of the Validation of the Amphibian Metamorphosis Assay: Phase 2 - Multi-chemical Interlaboratory Study (2007)*

No. 78, *Final Report of the Validation of the 21-day Fish Screening Assay for the Detection of Endocrine Active Substances. Phase 2: Testing Negative Substances (2007)*

No. 79, *Validation Report of the Full Life-cycle Test with the Harpacticoid Copepods Nitocra Spinipes and Amphiascus Tenuiremis and the Calanoid Copepod Acartia Tonsa - Phase 1 (2007)*

No. 80, *Guidance on Grouping of Chemicals (2007)*

No. 81, *Summary Report of the Validation Peer Review for the Updated Test Guideline 407, and Agreement of the Working Group of National Coordinators of the Test Guidelines Programme on the follow-up of this report (2007)*

No. 82, *Guidance Document on Amphibian Thyroid Histology (2007)*

No. 83, *Summary Report of the Peer Review Panel on the Stably Transfected Transcriptional Activation Assay for Detecting Estrogenic Activity of Chemicals, and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2007)*

No. 84, *Report on the Workshop on the Application of the GHS Classification Criteria to HPV Chemicals, 5-6 July Bern Switzerland (2007)*

No. 85, *Report of the Validation Peer Review for the Hershberger Bioassay, and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2007)*

No. 86, *Report of the OECD Validation of the Rodent Hershberger Bioassay: Phase 2: Testing of Androgen Agonists, Androgen Antagonists and a 5 α-Reductase Inhibitor in Dose Response Studies by Multiple Laboratories (2008)*

No. 87, *Report of the Ring Test and Statistical Analysis of Performance of the Guidance on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (Transformation/ Dissolution Protocol) (2008)*

No. 88, *Workshop on Integrated Approaches to Testing and Assessment (2008)*

No. 89, *Retrospective Performance Assessment of the Test Guideline 426 on Developmental Neurotoxicity (2008)*

No.90, *Background Review Document on the Rodent Hershberger Bioassay (2008)*

No. 91, *Report of the Validation of the Amphibian Metamorphosis Assay (Phase 3) (2008)*

No. 92, *Report of the Validation Peer Review for the Amphibian Metamorphosis Assay and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-Up of this Report (2008)*

No. 93, *Report of the Validation of an Enhancement of OECD TG 211: Daphnia Magna Reproduction Test (2008)*

No. 94, *Report of the Validation Peer Review for the 21-Day Fish Endocrine Screening Assay and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2008)*

No. 95, *Detailed Review Paper on Fish Life-Cycle Tests (2008)*

No.96, *Guidance Document on Magnitude of Pesticide Residues in Processed Commodities (2008)*

No.97, *Detailed Review Paper on the use of Metabolising Systems for In Vitro Testing of Endocrine Disruptors (2008)*

No. 98, *Considerations Regarding Applicability of the Guidance on Transformation/Dissolution of Metals Compounds in Aqueous Media (Transformation/Dissolution Protocol) (2008)*

No. 99, *Comparison between OECD Test Guidelines and ISO Standards in the Areas of Ecotoxicology and Health Effects (2008)*

No. 100, *Report of the Second Survey on Available Omics Tools (2009)*

No. 101, *Report of the Workshop on Structural Alerts for the OECD (Q)SAR Application Toolbox, 15-16 May 2008, Utrecht, the Netherlands (2009)*

No. 102, *Guidance Document for using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping of Chemicals (2009)*

No. 103, *Detailed Review Paper on Transgenic Rodent Mutation Assays (2009)*

No. 104, *Performance Assessment: Comparsion of 403 and CxT Protocols via Simulation and for Selected Real Data Sets (2009)*

No. 105, *Report on Biostatistical Performance Assessment of the draft TG 436 Acute Toxic Class Testing Method for Acute Inhalation Toxicity (2009)*

No. 106, *Guidance Document for Histologic Evaluation of Endocrine and Reproductive Test in Rodents (2009)*

No. 107, *Preservative treated wood to the environment for wood held in storage after treatment and for wooden commodities that are not cover and are not in contact with ground. (2009)*

No. 108, *Report of the validation of the Hershberger Bioassay (weanling model) (2009)*

No. 109, *Literature review on the 21-Day Fish Assay and the Fish Short-Term Reproduction Assay (2009)*

No. 110, *Report of the validation peer review for the weanling Hershberger Bioassay and agreement of the working of national coordinators of the test guidelines programme on the follow-up of this report (2009)*

No. 111, *Report of the Expert Consultation to Evaluate an Estrogen Receptor Binding Affinity Model for Hazard Identification (2009)*

No. 112, *The 2007 OECD List of High Production Volume Chemicals (2009)*

No. 113, *Report of The Focus Session on Current and Forthcoming Approaches for Chemical Safety and Animal Welfare (2010)*

No. 114, *Performance Assessment of Different Cytotoxic and Cytostatic Measures for the In Vitro Micronucleus Test (MNVIT): Summary of results in the collaborative trial (2010)*

No. 115, *Guidance Document on the Weanling Hershberger Bioassay in Rats: A Short-term Screening Assay for (Anti) Androgenic Properties (2009)*

No. 116, *Guidance Document 116 on the Conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 and 453 – 2nd Edition (2011)*

No. 117, *Guidance Document 117 on the Current Implementation of Internal Triggers in Test Guideline 443 for an Extended One Generation Reproductive Toxicity Study, in the United States and Canada (2011)*

No. 118, *Workshop Report on OECD Countries Activities Regarding Testing, Assessment and Management of Endocrine Disrupters Part I and Part II (2010)*

No. 119, *Classification and Labelling of chemicals according to the UN Globally Harmonized System: Outcome of the Analysis of Classification of Selected Chemicals listed in Annex III of the Rotterdam Convention (2010)*

No. 120, *Part 1: Report of the Expert Consultation on Scientific and Regulatory Evaluation of Organic Chemistry Mechanism-based Structural Alerts for the Identification of DNA Binding Chemicals*

No. 120, *Part 2: Report of the Expert Consultation on Scientific and Regulatory Evaluation of Organic Chemistry Mechanism-based Structural Alerts for the Identification of DNA Binding Chemicals*

No. 121, *Detailed review paper (DRP) on Molluscs life-cycle Toxicity Testing (2010)*

No. 122, *Guidance Document on the determination of the Toxicity of a Test Chemical to the Dung Beetle Aphodius Constans (2010)*

No. 123, *Guidance Document on the Diagnosis of Endocrine-related Histopathology in Fish Gonads (2010)*

No. 124, *Guidance for the Derivation of an Acute Reference Dose (2010)*

No. 125, *Guidance Document on Histopathology for Inhalation Toxicity Studies, Supporting TG 412 (Subacute Inhalation Toxicity: 28-Day) and TG 413 (Subchronic Inhalation Toxicity: 90-Day) (2010)*

No. 126, *Short Guidance on the Threshold approach for Acute Fish Toxicity (2010)*

No. 127, *Peer review report of the validation of the 21-day androgenised female stickleback screening assay (2010)*

No. 128, *Validation Report of the 21-day Androgenised Female Stickleback Screening Assay (2010)*

No. 129, *Guidance Document on using Cytotoxicity Tests to Estimate Starting Doses for Acute Oral Systemic Toxicity Tests(2010)*

No. 131, *Report of the Test Method Validation of Avian Acute Oral Toxicity Test (OECD test guideline 223) (2010)*

No. 132, *Report of the Multi-Laboratory Validation of the H295R Steroidogenesis Assay to Identify Modulators (2010)*

No.133, *Peer Review Report for the H295R Cell-Based Assay for Steroidogenesis (2010)*

No.134, *Report of the Validation of a Soil Bioaccumulation Test with Terrestrial Oligochaetes by an International ring test (2010)*

No.135, *Detailed Review Paper on Environmental Endocrine Disruptor Screening: The use of Estrogen and Androgen Receptor Binding and Transactivation Assays in Fish (2010)*

No. 136, *Validation Report of t he Chironomid Full Life-Cycle Toxicity Test (2010)*

No. 137, *Explanatory Background Document to the OECD Test Guideline On In Vitro Skin Irritation Testing (2010)*

No. 138, *Report of the Workshop on Using Mechanistic Information in Forming Chemical Categories (2011)*

No. 139, *Report of the Expert Consultation on Scientific and Regulatory Evaluation of Organic Chemistry Mechanism Based Structural Alerts for the Identification of Protein-binding Chemicals (2011)*

No. 140, *Report of the WHO/OECD/ILSI (Hesi) Workshop on Risk Assessment of Combined Exposures to Multiple Chemicals (2011)*

No. 141, *Report of the Phase 1 of the Validation of the Fish Sexual Development Test for the Detection of Endocrine Active Substances (2011)*

No. 142, *Report of the Phase 2 of the Validation of the Fish Sexual Development Test for the Detection of Endocrine Active Substances (2011)*

No. 143, *Peer Review Report for the Validation of the Fish Sexual Development Test and Agreement of the Working Group of National Co-ordinators of the Test Guideline Programme on the Follow-up of the Peer Review (2011)*

No. 144, *Validation Report for the Acute Chironomid Assay (2011)*

No. 145, *Transgenic Rodent Somatic and Germ Cell Gene Mutation Assay: Retrospective Performance Assessment (2011)*

No. 146, *Syrian Hamster Embryonic (SHE) Cell PH 6.7 Cell Transformation Assay Prevalidation Study Report (2011)*

No. 147, *Syrian Hamster Embryonic (SHE) Cell PH 7.0 Cell Transformation Assay Prevalidation Study Report (2011)*

No. 148, *Guidance Document on the Androngenised Female Stickleback Screen (2011)*

No. 149, *Validation Report of the Balb/c 3T3 Cell Transformation Assay (2011)*

No. 150, *Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption (2011)*

No. 152, *Case Study: Assessment of an Extended Chemical Category, the Short-chain Methacrylates, Targeted on Bioaccumulation (2011)*

No. 153, *Guidance Document for the Derivation of an Acute Reference Concentration (Arfc) (2011)*

No. 154, *Draft Validation Report: Part 1 – Validation of Efficacy Methods for Antimicrobials used on Hard Surfaces (2011)*

No. 154, *Draft Validation Report: Part 2 – Validation of Efficacy Methods for Antimicrobials used on Hard Surfaces (2011)*

No. 155, *Peer Review for the Validation of the Modified Skin Irritation Test Method using LabyCyte EPI-MODEL24; additional studies; and agreement of the Working Group of National Coordinators on the follow-up to the Peer Review*

No. 156, *Guidance Notes on Dermal Absorption* (2011)

No. 157, *Validation Report Phase 1 for the Zebrafish Embryo Toxicity Test (2011)*

No. 158, *Report of Progress on the Interlaboratory Validation of the OECD Harpacticoid Copepod Development and Reproduction Test (2011)*

No. 159, *Validation Report for the Skin Irritation Test Method using Labcyte Epi-Model24 (2011)*

No. 160, *Guidance Document on the Bovine Corneal Opacity and Permeability (Bcop) and Isolated Chicken Eye (Ice) Test Methods: Collection of Tissues For Histological Evaluation And Collection Of Data On Non-Severe Irritants (2011)*

No. 161, *Peer Review Report for the Validation of the Stably Transfected Transcriptional Activation Assay for the Detection Androgenic and Anti-Androgenic Activity of Chemicals (2011)*

ABOUT THE OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 34 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in ten different series: **Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides and Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/ehs/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

FOREWORD

This document is the second edition of the Guidance Document (GD) 116 on the Design and Conduct of chronic toxicity and carcinogenicity studies, supporting Test Guidelines 451, 452 and 453 (carcinogenicity, chronic toxicity and combined chronic toxicity/carcinogenicity studies).

The proposal for developing this GD was approved by the WNT in 1997. At that time the project only included the development of a GD on dose selection. In 2008, the objective of the project was revised. The WNT agreed that the GD should be developed in parallel with the update of the Test Guidelines 451, 452 and 453 as a supporting GD for these Test Guidelines. Thus, although the WNT agreed that the section on dose selection should be developed as a priority, work also started to develop other areas of guidance.

The proposal for this Guidance Document was first discussed, with the draft updated Test Guidelines 451, 452 and 453, at a workshop held in Washington D.C. in 2008. The 1st edition including the general introduction (Chapter 1) and the section on dose selection (Section 3.1) was finalized at an expert meeting held in Paris on 7-8 October 2009 and published in June 2010. This 2nd edition, including the entire GD, was finalised at an expert meeting held in Paris in November 2010. WNT comments on two successive drafts were requested before the expert meeting and the entire draft GD was circulated for a third commenting round to the WNT in December 2010.

Comments from the WNT have been addressed and the second edition of the GD was approved by the WNT at its meeting held in April 2011. The Joint Meeting of the Chemicals Committee and the working Party on Chemicals, Pesticides and Biotechnology agreed to the declassification of the second edition of this document in Ocotber 2011.

This document is published under the responsibility of the Joint Meeting of the Chemicals Committee and the working Party on Chemicals, Pesticides and Biotechnology.

TABLE OF CONTENTS

1. GENERAL INTRODUCTION

1.1 Guiding principles and considerations

1. Chronic toxicity and carcinogenicity studies are intended to identify toxic effects and potential health hazards following prolonged, repeated exposure. This type of study is usually required if humans are likely to be exposed to a substance over a significant portion of their life span. In the 1960s, long-term animal bioassays (chronic toxicity and carcinogenicity studies) began to be routinely used for hazard identification, to assess the qualitative potential of a chemical to cause chronic toxicity and cancer.

2. The objectives of the long-term bioassays have however expanded beyond hazard identification and are now focused primarily on hazard characterization for use in the assessment of risk for humans. In addition there has been increasing pressure for the long-term bioassay designs to consider financial constraints and societal desires to minimize the number of animals needed for scientific interpretation of results. There is a growing desire for long-term studies to provide data that cover a number of objectives including characterization of the nature of specific toxic responses, description of dose–response relationship, establishment of inflection points, and provision of insight into the roles of toxicokinetics and mechanisms of toxic action. In practice, it is likely that the bioassay design will be a compromise among a set of different purposes; to the extent that the ability to address one question is enhanced, the ability to address others may be diminished. For example, it may be necessary to achieve a balance between the power to detect toxicity and the ability to estimate the dose–response relationship of any observed effects. If information on carcinogenicity hazard identification is not available, this should be the main objective of the study.

3. The use of formal risk assessment procedures by government regulatory bodies began to emerge in the late 1970s and early 1980s bringing with it a strong interest in using data for quantitative as well as qualitative purposes. The need to gather data that allowed an understanding of the shape and slope of the dose-response curve focused attention on the number of doses in a bioassay and their spacing. Advances in knowledge of how chemicals perturbed or otherwise modulated biological processes in the development of tumours or other forms of toxicity provided bases for further improving the risk assessment process. Through meetings held primarily under the auspices of the International Programme on Chemical Safety (IPCS), a Mode of Action (MOA) framework was developed and refined (Sonnich-Mullin *et al.*, 2001; Cohen *et al.*, 2003; Meek *et al.*, 2003; Holsapple *et al.*, 2006; Boobis *et al.*, 2006; EPA, 2005), as will be further developed in Chapter 2 of this guidance. The key purpose of this work was to introduce greater transparency into the process of assessing human relevance, and the goal was to use a broad array of relevant data to determine the predictive value of a bioassay tumour response to risk in humans.

4. The broadened range and complexity of scientific data used to evaluate chemical toxicity and carcinogenicity potential for humans highlighted the need to revise and update the following OECD Test Guidelines (TGs): TG 451 (Carcinogenicity Studies), TG 452 (Chronic Toxicity Studies), and TG 453–(Combined Chronic Toxicity/Carcinogenicity Studies), originally adopted in 1981. These TGs have therefore recently been revised in the light of scientific progress and the updating of related OECD Guidelines such as TG 408 (90-day oral toxicity study in rodents) and TG 407 (28-day oral toxicity study in rodents).

5. During the revision of the Test Guidelines, an emphasis was placed on providing guidance on factors that influence the selection of test doses, particularly for carcinogenicity studies. It was recognized that while general principles of dose selection should be contained in the Test Guidelines themselves, there was a need for additional guidance on these principles. The revision took into

account two publications by the International Life Sciences Institute (ILSI), "*Principles for the Selection of Doses in Chronic Rodent Bioassays*" (ILSI, 1997), and "*Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection*" (Rhomberg *et al.*, 2007). These reports provided theoretical and practical guidance on factors that influence dose selection in long-term bioassays.

6. A summary of the principles contained in these two publications, to underpin the texts on dose selection contained in the Test Guidelines, is provided in this guidance (Section 3.1, Appendix 1). During the development of this material, suggestions were made for additional guidance on specific aspects of study design in relation to core objectives of these studies, and how they might impact on other aspects of the study (e.g., designing for optimal collection of carcinogenicity data versus chronic toxicity data, design of studies for risk estimation rather than hazard assessment). It was generally agreed that the scope of the guidance should be wider than principles of dose selection, and should cover a number of key issues related to carcinogenicity and chronic toxicity testing.

7. This guidance therefore provides additional information on the conduct of studies performed using TG 451, 452 and TG 453. Its objective is to assist users of the TGs to select the most appropriate methodology to assess the chronic toxicity and carcinogenicity of a test chemical so that particular data requirements can be met while reducing animal usage if possible/appropriate.

8. The guidance is intended to foster a common approach among those carrying out chronic toxicity and carcinogenicity studies, and thereby contributes to the harmonisation activities undertaken by the OECD and other agencies, such as the WHO[1]. It should be consulted in addition to other guidance and requirements. It provides broad guidance on approaches to the execution of chronic toxicity and carcinogenicity studies. The text reflects current scientific understanding and standards. In time, the scientific community will gain a better understanding of the mechanisms of toxicity, and this may lead to changes in both methodology and interpretation of results, which should reflect scientific consensus at the time data are reviewed.

9. It should be noted that the basic principles for the conduct of chronic toxicity and carcinogenicity studies will differ, given that the endpoints are different. While the guidance provided in this document can be taken as generally applicable to the conduct of either a chronic toxicity or a carcinogenicity bioassay, or a combined chronic and carcinogenicity study, users of the guidance should be mindful of the primary objectives of the study. It should also be noted that much of the information provided in the guidance reflects general principles for conducting animal studies, e.g., information on route of exposure and dosing considerations, choice of species and strain, the importance of toxicokinetic studies, housing and feeding and animal welfare issues, and is not unique to chronic toxicity or carcinogenicity studies. The objective in restating this general information in this guidance is to provide as far as possible a stand-alone document, avoiding the need for the reader to make cross reference between several guidance documents covering different aspects of such studies.

10. Two other OECD documents also provide guidance on aspects of the conduct and interpretation of repeat-dose toxicity studies, chronic toxicity and carcinogenicity studies. Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (OECD, 2002a) and Guidance

---

[1] An example, already referred to, is the IPCS project *Harmonisation of Approaches to the Assessment of Risk from Exposure to Chemicals*, which has developed a *Conceptual Framework for Cancer Risk Assessment*. The framework is an analytical tool for judging whether the available data support a postulated mode of carcinogenic action.

Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies (OECD, 2002b). The Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies (GD 35), provides valuable information on assessment of the quality, integrity, and completeness of the experimental data from chronic toxicity and carcinogenicity studies and determining the acceptability of the study reports, including aspects such as reliability, relevance and adequacy (OECD, 2002b). GD 35 also provides guidance on the reporting of the study and on the use of historical control data as an adjunct to the internal controls of the study in question. These aspects are therefore considered to lie outside the scope of the current Guidance Document, and GD 35 should be consulted for further information.

1.2 Scope of application of the guidance

11. The Test Guidelines 451, 452 and 453 and this guidance document are designed to be used in the testing of a wide range of chemicals, whatever their field of application, including pesticides, industrial chemicals and pharmaceuticals. However, as noted in the Test Guideline 451, some testing requirements may differ for pharmaceuticals. The International Conference on Harmonisation of Testing for Pharmaceuticals (ICH) has produced a series of safety guidelines for the testing of pharmaceuticals, including guidelines on toxicokinetics (ICH, S3A), genotoxicity testing (ICH, S2), duration of chronic toxicity testing in animals (rodent and non-rodent toxicity testing) (ICH, M3(R2)), testing for carcinogenicity of pharmaceuticals (ICH, S1B) and dose selection for carcinogenicity studies of pharmaceuticals (ICH, S1C(R2)). These guidelines should always be consulted for specific guidance when testing pharmaceuticals using the approaches outlined in TG 451, 452 and 453. This guidance document provides, in various sections, examples of where the testing requirements may be different for pharmaceuticals.

1.3 Objectives of a chronic toxicity study

12. The objective of a chronic toxicity study, such as described by TG 452, is to characterize the toxicological response of a substance in a mammalian species following prolonged and repeated exposure. The chronic toxicity study provides information on the possible health hazards likely to arise from repeated exposure over a prolonged period of time. Key objectives of the study are to provide information useful for classification and labelling and an estimate of a point of departure (e.g., lower confidence limit of the benchmark dose (BMDL) or the no-observed-adverse-effect level (NOAEL)) for any adverse effects, which can be used for establishing safety criteria for human exposure. The study will provide information on the major toxic effects, and indicate target organs, progressive toxic effects and the possibility of delayed toxicity. The need for careful clinical observations of the animals, so as to obtain as much information as possible, is also stressed. Previous repeated dose 28-day and/or 90-day toxicity tests on a chemical may, among others, have indicated the potential to cause neurotoxic/neurobehavioural effects, or effects on the endocrine system, warranting further in-depth investigation as part of a chronic toxicity study.

1.4 Objectives of a carcinogenicity study

13. The objective of a long-term carcinogenicity study, such as described by TG 451, is to observe test animals for a major portion of their life span for the development of neoplastic lesions during or after exposure to various doses of a test substance by an appropriate route of administration. The carcinogenicity study may also provide information on the possible health hazards likely to arise from repeated exposure for a period lasting up to the entire lifespan of the species used. The study will provide information on potential carcinogenicity, and may indicate toxic effects, target organs, progressive toxic effects and the possibility of delayed toxicity. It can provide information useful for classification and labelling, and an estimate of a point of departure for toxic effects and, in the case of non-genotoxic carcinogens, for tumour responses, which can be used for establishing safety levels for human exposure. The need for careful clinical observations of the animals, so as to obtain as much

information as possible, is also stressed. Such an assay requires careful planning and documentation of the experimental design, a high standard of pathology, and unbiased statistical analysis. These requirements are well known and have not undergone any significant changes in recent years.

1.5 Objectives of a combined chronic toxicity/carcinogenicity study

14. The objective of a combined chronic toxicity/carcinogenicity study, such as described by TG 453, is to determine the toxicological effects of a substance (including carcinogenic potential) in a mammalian species following prolonged and repeated exposure. The combined chronic toxicity/carcinogenicity study provides information on the possible health hazards likely to arise from repeated exposure over the majority of the entire lifespan (in rodents). The study will provide information on the major toxic effects of the substance including potential carcinogenicity, and indicate target organs, progressive toxic effects and the possibility of delayed toxicity. The need for careful clinical observations of the animals, so as to obtain as much information as possible, is also stressed. The application of the TG 453 should generate data on which to identify the majority of chronic and carcinogenic effects and to determine dose-response relationships. Ideally, the design and conduct should allow for the detection of neoplastic effects and a determination of carcinogenic potential as well as general toxicity, including neurological, physiological, biochemical, and haematological effects and exposure-related morphological (pathology) effects.

15. The design of the updated TG 453 recommends, for the chronic phase of the study, at least three dose groups and a control group, each group containing at least 10 males and 10 females per group. The reduction of the number of animals per sex in the updated TG 453 compared to the initial version (1981) is justified on the basis of further information being available from animals on the carcinogenicity phase of the study and a more careful use of animals in laboratories than in 1981. The design of the carcinogenicity phase of the revised TG 453 is identical to the revised TG 451. The study will thus provide similar information on chronic toxicity and carcinogenicity as TG 452 and TG 451. It will allow derivation of a point of departure (e.g., BMDL or NOAEL), and will offer greater efficiency in terms of time and cost compared to conducting two separate studies, without compromising the quality of the data in either the chronic phase or the carcinogenicity phase. The data from both phases (chronic toxicity and carcinogenicity) will reinforce each other, as the animals used in the studies are drawn from the same stock and have similar characteristics at the start of the study. Measurements carried out on the animals in one phase will be relevant for the animals in the other phase, e.g., clinical signs, body weights, haematology and biochemistry (if carried out), pathology. The terminal kill of the chronic phase can act as an interim kill for the carcinogenicity phase.

1.6 Consideration of testing strategies

16. This section refers to testing strategies that may be applicable for certain regulatory authorities but have not been formally adopted by all. It does not recommend any particular testing strategy or approach, but suggests consideration of such approaches as part of an ongoing strategy to assess the toxic potential of a substance in an intelligent and iterative manner. As new validated approaches become scientifically appropriate for use in chronic toxicity or carcinogenicity assessment, and accepted by the relevant regulatory authorities, the study sponsor is encouraged to implement them where possible.

17. A reasoned scientific approach to the assessment of substances for chronic toxicity or carcinogenicity must first include an assessment of all available information that has the potential to influence the study design. This can include the identity, molecular structure, class, and physico-chemical properties of the test substance; any information regarding mode of action; results of relevant *in vitro* or *in vivo* toxicity tests such as genotoxicity, subchronic toxicity and toxicokinetics studies; anticipated use(s) and potential for human exposure; available (Q)SAR data; and relevant

toxicological data on structurally-related substances. This analysis can focus the study parameters, but may also lead to the conclusion that a study can be refined in some way, or not conducted at all based on a weight of evidence (Carmichael *et al.*, 2006; Doe *et al.*, 2006; Barton *et al.*, 2006; Cooper *et al.*, 2006).

18. Integrating a wide range of information to determine the potential toxicity of a substance is becoming more common as the gap between assessments that need to be conducted and the resources with which to conduct such assessments widens. Efforts are underway in many OECD countries to determine ways in which assessments of substances can be satisfactorily completed, and protection of public health and the environment achieved, while minimising costs in terms of time, money and animal use. However, the acceptability and use of testing strategies and weight-of-evidence approaches differ among OECD countries and regulatory sectors; thus, application of these approaches should always occur in consultation with appropriate regulatory authorities.

19. Shorter-term *in vitro* or *in vivo* tests may provide information regarding potency, mode of action, metabolism, and/or target organ that can help refine the chronic toxicity study protocol parameters or priorities for observation. Tiered approaches using a combination of *in silico*, *in vitro*, and *in vivo* tests have been proposed but are not yet widely implemented (Worth and Balls, 2002; Becker *et al.*, 2007).

20. A phased or tiered approach to the assessment of the carcinogenic potential of a substance should also be considered (Ashby, 1996). A number of shorter-term tests can be conducted which will provide useful information for determining whether and how a substance may be carcinogenic, including genetic toxicity assays, cell transformation or other cell-based assays, short-term cancer initiation-promotion tests which may or may not include toxicogenomic analyses (Ellinger-Ziegelbauer *et al.*, 2005; 2008), and *in vivo* repeated dose 28- or 90-day toxicity tests (for a review on *in vitro* and *in vivo* short term test see Maurici *et al.*, 2005). (Q)SAR prediction models have been used in a regulatory context to predict the carcinogenic potential of chemicals for several decades. There are also commercial (Q)SAR models available for predicting rodent carcinogenicity (Benigni *et al.*, 2007). (Q)SAR models should be validated according to OECD principles (OECD 2007, GD No. 69).

21. A number of different strategies for assessing carcinogenicity have been proposed (Langley 2001; Worth and Balls 2002; Knight *et al.*, 2006; Combes *et al.*, 2008). All feature a stepwise process or decision tree that prescribes information analysis and stopping points where classification and labelling and/or risk assessment could be possible. However, specific approaches have not yet been optimised or validated.

22. Consideration of particular tests or approaches should always be made within the context of whether the results will contribute mechanistic information that will be useful in the weight-of-evidence assessment of carcinogenic potential (OECD, 2002b).

23. The US National Toxicology Program, along with institutes in other OECD countries, has had a longstanding interest in the use of transgenic or knockout mouse models for the assessment of carcinogenicity (Bucher and Portier, 2004), as they consider that these models offer potential refinements, in terms of study duration and animal numbers, over the traditional long-term bioassay. At the time this guidance was prepared, some regulatory authorities in the pharmaceutical sector may accept studies with these models in combination with a full long-term rat bioassay in lieu of a second full bioassay in mice (ICH, 1997). In the past, the predictive ability of the models, and any refinements or animal reductions, has been questioned (Goodman, 2001; van Zeller and Combes, 1999; RIVM, 2004). A detailed review paper (DRP) on transgenic rodent mutation assays prepared by Canada was published by OECD (OECD, 2009).

1.7 Animal welfare considerations

24. The principles of the "3Rs" (Replacement, Reduction, and Refinement), first articulated by Russell and Burch in 1959 (Russell and Burch, 1959), should be considered as integral to the assessment of carcinogenicity or chronic toxicity in mammals, in order to ensure sound science, maximize animal welfare, and minimize animal use. Animals in a condition of stress or distress have a documented effect on the outcome of the study (Olsson and Dahlborn, 2001; Reinhardt and Reinhardt, 2002; Hurst and West, 2010). For these reasons the following principles should be implemented as much as practicably possible.

25. First and foremost, as discussed above, consideration of documented existing information from any reliable source that could provide a refinement in the testing protocol or procedure is recommended. Existing information could be used to inform dose spacing or selection, exposure route, observation priorities, potential modes of action or target organs of the test substance, and/or study design. Use of this information to focus the study before it begins ensures that the study will meet the expectations of the study sponsor and/or regulatory authorities, decreasing the likelihood of repeat studies.

26. The use of the combined chronic toxicity/carcinogenicity study (TG 453) is also recommended, which can in most cases accomplish the objectives of both studies, and offers savings in the numbers of animals used. This is due to the use of 10 animals per sex per dose group for the chronic toxicity phase of the study instead of 20 animals per sex per dose group when the carcinogenicity and the chronic toxicity studies are performed separately.

27. Any studies involving animals should abide by the principles of humane euthanasia as detailed in the OECD Guidance Document 19 on the recognition, assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation, and in particular paragraph 62 thereof (OECD, 2000). This paragraph states that "In studies involving repeated dosing, when an animal shows clinical signs that are progressive, leading to further deterioration in condition, an informed decision as to whether or not to humanely kill the animal should be made. The decision should include consideration as to the value of the information to be gained from the continued maintenance of that animal on study relative to its overall condition. If a decision is made to leave the animal on test, the frequency of observations should be increased, as needed. It may also be possible, without adversely affecting the purpose of the test, to temporarily stop dosing if it will relieve the pain or distress, or reduce the test dose." Close and frequent observations are recommended in order to determine the status of the animals, and any animals exhibiting clear signs of severe pain or distress should be humanely killed.

28. Animals may be housed individually, or be caged in small groups of the same sex; individual housing should be considered only if scientifically justified. Further detailed information on housing, feeding, and handling will be provided in Section 3.5.

29. As will be further discussed in Section 3.2, while the route of administration will depend on the physical and chemical characteristics of the test substance and expected route of human exposure, mixing the test substance into the diet or water is normally recommended for rodent studies. Administration of the test substance by oral gavage in carcinogenicity and chronic toxicity testing is normally not recommended for the reasons outlined in Section 3.2. If the oral gavage route is employed then its use should be justified. The testing of substances at potentially irritating or corrosive concentrations/doses should be avoided, as administering such substances could result in severe pain and tissue damage at point-of-entry, which would compromise both animal welfare and the integrity of the study.

30. Testing the chronic toxicity or carcinogenicity of inhaled substances can be achieved using either of two exposure conditions: whole-body or nose-only/snout-only. For studies of liquid or solid aerosols and for vapor that may condense to form aerosols, the nose-only exposure method allows the avoidance of oral exposure due to grooming of particles deposited on the fur. However, the welfare implications of a 1- or 2-year nose-only exposure study, and the potential for physiological effects of stress experienced by the animals to affect the results of the study, can lead to a preference for the use of the whole-body mode of exposure (Thomson *et al.,* 2009). Reasons for choice of exposure system should be justified in the study report. If rodent species other than rats are exposed nose-only, maximum exposure durations may be adjusted to minimise species-specific distress (GD 39).

31. Guidelines providing practical advice on animal welfare specifically for today's cancer researchers were updated in 2010. They define and encourage sharing of best practice in laboratory work in the field of cancer research (Workman *et al.*, 2010). Relevant information can be found in these guidelines, that also applies to toxicological studies.

REFERENCES

Ashby, J. (1996), "Alternatives to the 2-Species Bioassay for the Identification of Potential Human Carcinogens", *Hum Exp Toxicol,* Vol. 15: 183-202.

Barton, H.A., T.P. Pastoor, T. Baetcke, J.E. Chambers, J. Diliberto, N.G. Doerrer, J.H. Driver, C.E. Hastings, S. Iyengar, R. Krieger, B. Stahl and C. Timchalk (2006), "The Acquisition and Application of Absorption, Distribution, Metabolism, and Excretion (ADME) Data in Agricultural Chemical Safety Assessments", *Crit Rev Toxicol,* Vol. 36: 9-35.

Becker, R.A., L.M. Plunkett, J.F. Borzelleca and A.M. Kaplan (2007), "Tiered Toxicity Testing: Evaluation of Toxicity-Based Decision Triggers for Human Health Hazard Characterization", *Food and Chemical Toxicology,* Vol. 45: 2454–2469.

Benigni, R., T.I. Netzeva, E. Benfenati, C. Bossa, R. Franke, C. Helma, E. Hulzebos, C. Merchant, A. Richard, Y.T. Woo and C. Yang (2007), "The Expanding Role of Predictive Toxicology: an Update on the (Q)SAR Models for Mutagens and Carcinogens", *J Env Sci Hlth* Part C, Vol.25: 53-97.

Boobis, A.R., S.M. Cohen, V. Dellarco, D. McGregor, M.E. Meek, C. Vickers, D. Willcocks and W. Farland (2006), "IPCS Framework for Analyzing the Relevance of a Cancer Mode of Action for Humans", *Crit. Rev. in Toxicol.*, 36:793-801.

Bucher, J. and C. Portier (2004), "Human Carcinogenic Risk Evaluation, Part V: The NationalToxicology Program Vision for Assessing the Human Carcinogenic Hazard of Chemicals", *Tox Sci,* Vol. 82: 363–366.

Cohen, S.M., M.E. Meek, J.E. Klaunig, D.E. Patton and P.A. Fenner-Crisp (2003), "The Human Relevance of Information on Carcinogenic Modes of Action: An Overview", *Crit. Rev. Toxicol.,* Vol. 33:581-589.

Cooper, R.L., J.S. Lamb, S.M. Barlow, K. Bentley, A.M. Brady, N.G. Doerrer, D.L Eisenbrandt, P.A. Fenner-Crisp, R.N. Hines, L.F. Irvine, C.A. Kimmel, H. Koeter, A.A. Li, S.L. Makris, L.P. Sheets, G. Speijers and K.E. Whitby (2006), "A Tiered Approach to Life Stages Testing for Agricultural Chemical Safety Assessment", *Crit Rev Toxicol,* Vol. 36: 69-98.

Carmichael, N.G., H.A. Barton, A.R. Boobis, R.L. Cooper, V. Dellarco, P.A. Fenner-Crisp, J.E. Doe, J.C. Lamb and T.P. Pastoor (2006), "Agricultural Chemical Safety Assessment: A Multisector Approach to the Modernization of Human Safety Requirements", *Crit Rev Toxicol,* 36: 1-7.

Combes, R., C. Grindon, M.T.D. Cronin, D.W. Roberts and J.F. Garrod (2008), "Integrated Decision-tree Testing Strategies for Mutagenicity and Carcinogenicity with Respect to the Requirements of the EU REACH Legislation", *Alt Lab Anim,* Vol. 36: 43-63.

Doe, J.E., A.R. Boobis, A. Blacker V. Dellarco, N.G. Doerrer, C. Franklin, J.I. Goodman, J.M. Kronenberg, R. Lewis, E.E. Mcconnell, T. Mercier, A. Moretto, C. Nolan, S. Padilla, W. Phang, R. Solecki, L. Tilbury, B. van Ravenzwaay and D.C. Wolf (2006), "A Tiered Approach to Systemic Toxicity Testing for Agricultural Chemical Safety Assessment", *Crit Rev Toxicol,* Vol. 36: 37-68.

Ellinger-Ziegelbauer, H., H. Gmuende, A. Bandenburg and H.J. Ahr (2008), "Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term *in vivo* studies", *Mutat Res,* Vol. 637: 23-39.

Ellinger-Ziegelbauer, H., B. Stuart, B. Wahle, W. Bomann and H.J. Ahr (2005), "Comparison of the expression profiles induced by genotoxic and non-genotoxic carcinogens in rat liver", *Mutat Res / Fundamental and Molecular Mechanisms of Mutagenesis*, Vol. 575: 61-84.

EPA (2005), *Guidelines for Carcinogen Risk Assessment*, Risk Assessment Forum, U.S. Environmental Protection Agency, Washington DC.

Goodman, J.I. (2001), "A Perspective on Current and Future Uses of Alternative Models for Carcinogenicity Testing", *Toxicol Pathol*, Vol. 29 (Suppl): 173-176.

Holsapple, M.P., H.C. Pitot, S.N. Cohen, A.R. Boobis, J.E. Klaunig, T. Pastoor, V.L. Dellarco and Y.P. Dragan (2006), "Mode of Action in Relevance of Rodent Liver Tumors to Human Cancer Risk". *Toxicol. Sci.*, Vol. 89:51-56.

Hurst, J.L. and R.S. West (2010), "Taming Anxiety in Laboratory Mice", *Nature Methods*, 7(10), 825-826.

ICH (International Conference on Harmonisation) (2009), *ICH Harmonised Tripartite Guideline - Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals – M3(R2)*. Available at: http://www.ich.org/cache/html/501-272-1.html.

ICH (International Conference on Harmonisation) (1994), *ICH Harmonised Tripartite Guideline - Note for Guidance on Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies – S3A*. Available at: http://www.ich.org/LOB/media/MEDIA495.pdf.

ICH (International Conference on Harmonisation) (1997), *ICH Harmonised Tripartite Guideline - Testing for Carcinogenicity of Pharmaceuticals - S1B*.
Available at: http://www.ich.org/LOB/media/MEDIA490.pdf.

ICH (International Conference on Harmonisation) (1998), *ICH Harmonised Tripartite Guideline - Duration of Chronic Toxicity Testing in Animals (Rodent and Non-Rodent Toxicity Testing) – S4*.
Available at: http://www.ich.org/LOB/media/MEDIA497.pdf.

ICH (International Conference on Harmonisation) (2008), *ICH Harmonised Tripartite Guideline - Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use – S2*. Available at: http://www.ich.org/LOB/media/MEDIA4474.pdf.

ICH (International Conference on Harmonisation) (2008), *ICH Harmonised Tripartite Guideline - Dose Selection for Carcinogenicity Studies of Pharmaceuticals – S1C*.
Available at: http://www.ich.org/LOB/media/MEDIA491.pdf.

ILSI (International Life Sciences Institute) (1997), *Principles for the Selection of Doses in Chronic Rodent Bioassays*, Foran JA (Ed.), ILSI Press, Washington, DC.

Knight, A., J. Bailey and J. Balcombe (2006), "Animal Carcinogenicity Studies: 3. Alternatives to the Bioassay", *Alt Lab Anim*, 34: 30-48.

Langley, G. (2001), *The way forward - Action to end animal toxicity testing*, London, UK: BUAV.

Maurici, D., M. Aardema, R. Corvi, M. Kleber, C. Krul, C. Laurent, N. Loprieno, M. Pasanen, S. Pfuhler, B. Phillips, D. Prentice, E. Sabbioni, T. Sanner and P. Vanparys (2005), "Carcinogenicity", *Alt Lab Anim*, Vol. 33 (Suppl 1): 177-182.

Meek, E.M., J.R. Bucher, S.M. Cohen, V. Dellarco, R.N. Hill, L.D. Lehman-McKemmon, D.G. Longfellow, T. Pastoor, J. Seed, and D.E. Patton (2003), "A Framework for Human Relevance analysis of Information on Carcinogenic Modes of Action", *Crit. Rev. Toxicol.,* Vol. 33:591-653.

OECD (2000), *Guidance Document on the Recognition, Assessment, and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation,* Series on Testing and Assessment N° 19, OECD, Paris.

OECD (2002a), *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies,* Series on Testing and Assessment N° 32, OECD, Paris.

OECD (2002b) *Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies,* Series on Testing and Assessment N° 35, OECD, Paris.

OECD (2007), *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(QSAR)] Models*, Series on testing and assessment N° 69, OECD Paris.

OECD (2009), *Detailed Review Paper on Transgenic Rodent Mutation Assays*, Series on testing and assessment N° 103, OECD Paris.

OECD (2009), *Carcinogenicity Studies*, OECD Guideline for the Testing of Chemicals, No. 451, OECD, Paris.

OECD (2009), *Chronic Toxicity Studies,* OECD Guideline for the Testing of Chemicals, No. 452, OECD, Paris.

OECD (2009), *Combined Chronic Toxicity/Carcinogenicity Studies,* OECD Guideline for the Testing of Chemicals, No. 453, OECD, Paris.

Olsson, A. and K. Dahlborn (2001), "Improving housing conditions for laboratory mice: a review of environmental enrichment", *Lab Animals,* Vol. 36: 243-270.

Reinhardt, V. and A. Reinhardt (2002), *Comfortable Quarters for Laboratory Animals*, Animal Welfare Institute, Washington, D.C.

Rhomberg, L., K. Baetcke, J. Blancato, J. Bus, S. Cohen, R. Conolly, R. Dixit, J. Doe, K. Ekelman, P. Fenner-Crisp, P. Harvey, D. Hattis, A. Jacobs, D. Jacobson-Kram, T. Lewandowski, R. Liteplo, O. Pelkonen, J. Rice, D. Somers, A. Turturro, W. West and S. Olin (2007), "Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection" *Crit. Rev. Toxicol.,*Vol. 37:729-837.

RIVM (2004), *Transgenic Mice as Alternative in Carcinogenicity Testing: Current Status*, RIVM report 340700001/2004, Netherlands.
Available at: http://rivm.openrepository.com/rivm/bitstream/10029/8911/1/340700001.pdf

Russell, W.M.S. and R.L. Burch (1959), *The Principles of Humane Experimental Technique*, Methuen Publishing Ltd., London.

Sonnich-Mullin, C., R. Fielder, J. Wiltse, K. Baetcke, J. Dempsey, P. Fenner-Crisp, D Grant, M. Hartley, A. Knaap, D. Kroese, I. Mangelsdorf, E. Meek, J. Rice, and M. Younes (2001), "IPCS Conceptual Framework for Evaluating a Mode of Action for Chemical Carcinogenesis" *Reg. Toxicol. & Pharm.,* Vol. 34:146-152.

Thomson, E.M., A. Williams, C.L. Yauk and R. Vincent (2009), "Impact of Nose-Only Exposure System on Pulmonary Gene Expression", *Inhal Toxicol,* 21(S1): 74-82.

Van Leeuwen, C.J., G.Y. Patlewicz and A.P. Worth (2007), *Intelligent Testing Strategies in Risk Assessment of Chemicals, an Introduction*, 2nd Edition, C.J. van Leeuwen and T.G.Vermeire, eds. Springer Netherlands, Rotterdam, pp 467-509.

Van Zeller, A.M. and R.D. Combes (1999), "Transgenic Mouse Bioassays for Carcinogenicity Testing: A Step in the Right Direction?", *Alt Lab Anim,* Vol. 27: 39–846.

Workman, P., E.O. Aboagye, F. Balkwill, A. Balmain, G. Bruder, D.J. Chaplin, J.A. Double, J. Everitt, D.A.H. Farningham, M.J. Glennie, L.F. Kelland, V. Robinson, I.J. Stratford, G.M. Tozer, S. Watson, S.R. Wedge and S.A. Eccles (2010), "Guidelines for the Welfare and Use of Animals in Cancer Research", *British Journal of Cancer,* 102, 1555–1577.

Worth, A. and M. Balls (2002), "Alternative (Non-animal) Methods for Chemicals Testing: Current Status and Future Prospects. A Report prepared by ECVAM and the ECVAM Working Group on Chemicals", *Alt Lab Anim,* Vol. 30 (Suppl 1): 1-125.

2. MODE OF TOXICOLOGIC ACTION

2.1 Introduction

32. This chapter provides an overview of the mode of action framework that has been developed to test hypothesized pathways of carcinogenicity and toxicity in recent years. It is based largely on the U.S. EPA Guidelines for Carcinogen Risk Assessment, (USEPA, 2005) and the International Programme on Chemical Safety (IPCS) Workshop Report (IPCS, 2005). It is provided primarily for the information of those carrying out long-term bioassays, rather than as specific guidance on determination of mode of action. It is important to note that while this chapter focuses primarily on the cancer endpoint, the mode of toxicologic action may be more generically applied to both cancer and noncancer endpoints.

33. The integrative approach to understanding of a chemical's toxicologic pathway, by weighing data from short-term toxicity, subchronic toxicity, genotoxicity and toxicokinetic data, becomes critical when designing a long-term bioassay. If a mode of action can be proposed, this information could greatly enhance the design of the long-term bioassay, in order to offer better insight into how a chemical elicits its toxicity and to help elucidate the shape of the dose-response curve. To fully utilize information on the hypothesized mode of action, the design of the study needs to account for doses that need to be placed carefully so as to yield observations of subtle precursor effects or other biomarkers of toxicity without inducing confounding effects related to frank toxicity (see also Section 3.1, Dose Selection).

2.2 Background

34. Animal long-term bioassays have been used for more than a half century to determine whether pesticides, pharmaceuticals, industrial chemicals, and other types of substances might cause cancer or other adverse chronic health problems in humans. As such, cancer bioassays, recognized by national and international regulatory groups, have become the default for testing the carcinogenic potential of products when human use or exposure is anticipated. Inherent in these animal-based assessments is the assumption that the observation of tumours in animals is directly relevant to the risk of cancer in humans, and that the responses observed at high doses in animals could be meaningfully extrapolated to doses with relevance for humans (IPCS, 2005). In order to predict responses in people more accurately, information is preferred from animals that are as similar to people as possible. More specifically, the use of other mammals, such as dogs, rats, and mice, as models for responses in humans is based on the assumption that there are important similarities among mammals in the way they respond to chemicals. In fact, a qualitative similarity has been established in the response of laboratory animals and humans to carcinogenic substances. Most known human carcinogens have been shown to be positive for tumourigenicity in well-conducted animal studies (Parekh and Dearfield, 2007).

35. Based on this assumption, such animal to human extrapolations, (dose and species) while necessary and practical, have been surrounded by intense discussion and debate. Through the evaluation of key events on the molecular and cellular level, a clearer understanding of how chemicals induce neoplasia has been realized. Such mechanistic data have called into question the appropriateness of extrapolating certain rodent tumour responses to humans. Thus, cancer risk assessments are rarely without controversy and are often heatedly debated within the scientific community (Holsapple *et al.*, 2005).

36. More recently, the understanding of the pathogenesis of neoplasia has evolved significantly. It is now recognized that cancers originating from at least some cell types may arise by a variety of independent pathways. It is also established that different carcinogens may have different modes of

action and that some carcinogens act through more than one mode of action in different tissues. While some modes of action lead to cancers in both rodents and humans, others that are carcinogenic in rodents are not in humans, at least under realistic circumstances of human exposure. To refine and improve the process of carcinogenic hazard identification, and to avoid misidentification of non-tumourigenic compounds as possible human carcinogens, it has become crucial that mode of action analysis be undertaken and that data to support such analysis be collected in a thorough and scientifically rigorous manner (Rice, 2004).

37. Risk assessments have benefited from our understanding of the mode of action of carcinogenesis in both animals and in humans and from the use of pharmacokinetic and pharmacodynamic data to determine the appropriateness of assumptions and to characterize the biological basis underlying the use of such assumptions. There has been increased recognition of advancements in scientific thinking on cancer that is reflected in some regulatory agency's adoption of the mode of action paradigm (USEPA, 2005). In accordance with this line of thinking, some regulatory agencies have also provided a systematic framework to test hypothesized toxicity pathways. Because of the many benefits of using this mode of action framework, implementation has been widespread and is now commonly used by additional regulatory agencies and international organizations (Meek *et al.*, 2003; Boobis *et al.*, 2006). In the United Kingdom, the mode of action framework is being applied to the assessments of pesticides and industrial chemicals. The UK Committee on Carcinogenicity (COC) has noted its value with regard to both harmonization between agencies and internal consistency in its latest Guidelines (COC, 2004). It has also been adopted and is being implemented by Canadian agencies, for example in the evaluation of Existing Chemicals under the Canadian Environmental Protection Act. The European Union has incorporated the framework into the technical guidance of assessments that are being updated for toxic effects, including carcinogenicity, of new and existing industrial chemicals (ECHA, 2009). The framework has been featured by the WHO/FAO Joint Meeting on Pesticide Residues in its evaluation of pyrethrin extract and its incorporation into the resulting monograph. Finally, IPCS in cooperation with international partners has taken steps to move the framework forward by melding it with the human relevance concept (IPCS Workshop, 2005).

2.3 "Mode of Action" Framework

38. The term "mode of action" is defined as a sequence of key events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in e.g., cancer formation. A "key event" is an empirically observable precursor step that is a necessary element of the mode of action or is a biologically based marker for such an element. The term "mechanism of action" implies a more detailed understanding and description of events, often at the molecular level, than is meant by mode of action. The toxicokinetic processes that lead to formation or distribution of the active agent to the target tissue are considered in estimating dose but are not part of the mode of action as the term is used here. This chapter focuses on the carcinogenic mode of action. There are many examples of possible modes of carcinogenic action, such as mutagenicity, mitogenesis, inhibition of cell death, cytotoxicity with reparative cell proliferation, and immune suppression.

39. Elucidation of a mode of action for a particular carcinogenic response in animals or humans is a data rich determination. Significant information should be developed to ensure that a scientifically justifiable mode of action underlies the process leading to cancer at a given target site. In the absence of sufficiently scientifically justifiable mode of action information, regulatory scientists generally take a public health protective default position regarding interpretation of toxicologic and epidemiologic data; animal tumour findings are judged to be relevant to humans and cancer risks are assumed to conform with low dose linearity.

Mode of Action Framework: Animal Tumours

40. The framework is intended to be an analytic tool for systematically judging whether available data support a mode of carcinogenic action hypothesized for an agent in a transparent manner. It is not designed to give an absolute answer on sufficiency of the information as this will vary depending on the circumstance (IPCS, 2005). Amongst the strengths of the framework are its flexibility, general applicability to carcinogens acting by any mechanism and the ability to explore the impact of each key event on the carcinogenic response (IPCS, 2005). It is primarily based upon considerations for causality in epidemiologic investigations originally articulated by Hill (1965) but later modified by others and extended to experimental studies. The modified Hill criteria are useful in organizing thinking about aspects of causation, and they are consistent with scientific method of developing hypotheses and testing those hypotheses experimentally. A key question is whether the data to support a mode of action meet the standards generally applied in experimental biology regarding inference of causation.

Components of a Mode of Action Analysis

41. To perform a mode of action analysis, the key biochemical, cellular and molecular events need to be established, and the temporal and dose-dependent concordance of each of the key events in the mode of action can then be determined. The key events can be used to bridge species and dose for a given mode of action. The next step in the mode of action analysis is the assessment of biological plausibility for determining the relevance of the specified mode of action in an animal model for human cancer risk based on kinetic and dynamic parameters. (Holsapple *et al.*, 2005)

Postulated Mode of Action: Key Events

42. The postulated mode of action is a biologically plausible hypothesis/basis for the sequence of key events leading to an observed effect supported by robust experimental observations and mechanistic data (IPCS, 2005). Key events are critical to the induction of tumours as hypothesized in the postulated mode of action. To support an association, a body of experiments needs to define and measure an event consistently.

43. To evaluate whether a hypothesized or postulated mode of action is operative, an analysis starts with an outline of the scientific findings regarding the hypothesized key events leading to cancer, and then weighing information to determine whether there is a causal relationship between these events and cancer formation. Again, it is not generally expected that the complete sequence will be known at the molecular level. Instead, observations made at different levels of biological organization (e.g., biochemical, cellular, physiological, tissue, organ, and system) are analyzed.

44. For each tumour site being evaluated, the mode of action analysis should begin with a description of the relevant data and key events that may be associated with a hypothesized mode of action and its sequence of key events. This can be followed by discussion of various aspects of the experimental support for hypothesized modes of action in animals and humans. (See Appendix I. for examples of mode(s) of action.)

Experimental Support for the Postulated Mode of Action

45. Experimental support addressing the strength, consistency and specificity of association, dose response concordance, temporal relationship and if the mode of action is biologically plausible all add to establishing a clear mode of action.

1. Strength, Consistency and Specificity of Association

46. A statistically significant association between key events and tumour response, observed in well conducted studies is generally supportive of causation. Consistent observations in a number of such studies with differing experimental designs increase that support, because different designs may reduce unknown biases. Studies showing absence/reduction of carcinogenicity when the rate limiting event is reversed, blocked or diminished, are particularly useful tests of association. Conversely, if enhancement of rate limiting key events increases the tumour response, this evidence would also provide strong support for the postulated mode of action. Pertinent observations include tumour response and key events in the same cell type, sites of action biologically related to key event(s), and results from multistage studies and from stop/recovery studies (Boobis *et al.*, 2006). Specificity of the association without evidence of other modes of action also strengthens a causal conclusion. And while these factors provide additional confidence that the primary mode of action has been identified, conversely, a lack of strength, consistency and specificity of an association tends to weaken the overall causal conclusions for a particular mode of action.

2. Dose Response Concordance

47. If a key event and tumour endpoints increase with dose such that the key events forecast the appearance of tumours at a later time or higher dose, the shape of the dose/response curve could be revealed and a causal association can be strengthened. Dose-response associations of the key event with other precursor events can add further strength.

3. Temporal Concordance

48. If a key event is shown to be causally linked to tumourigenesis, it should precede tumour appearance. An event may also be observed contemporaneously or after tumour appearance; these observations may add to the strength of association but not to the temporal association. Pertinent observations include studies of varying durations observing the temporal sequence of events and development of tumours (see paragraphs 67 and 68 on precursor/key events for application to long-term studies).

4. Biological Plausibility and Coherence

49. The biological plausibility of any postulated mode of action in humans depends on a consideration of dose-effect and dose-response relationships (IPCS, 2005). The postulated mode of action and key events should be based on contemporaneous understanding of the biology of cancer. If the body of information under scrutiny is consistent with other chemical agents for which the hypothesized mode of action is accepted, the case is strengthened. Note: Because some modes of action can be anticipated to evoke effects other than cancer, the available toxicity database on noncancer effects can contribute to this evaluation.

Alternative Mode(s) of Action

50. The possibility of other modes of action should also be considered and discussed. If there is evidence for more than one mode of action, each mode should receive a separate analysis. Furthermore, different modes of action can operate in different dose ranges; for example, an agent can act predominately at lower doses where cytotoxicity may not occur. Ultimately, however, information on all modes of action should be integrated to better understand how and when each mode acts, and which modes may be of primary interest for exposure levels relevant to human exposure of interest.

Uncertainties, Inconsistencies and Data Gaps

51. Uncertainties should be stated clearly, fully and explicitly. They should include those related to the biology of the toxicological response and those for the database on the specific chemical being evaluated. Any inconsistencies should be clearly noted and characterized with respect to the impact on the weight of evidence in support of the postulated mode action. Data gaps should also be identified and characterized. It should be clearly stated whether the identified data gaps are critical in supporting the postulated mode of action and what recommendations can be provided to address those data deficiencies in the future (Boobis *et al.*, 2008).

Conclusion of Postulated Mode of Action Analysis

52. Conclusions about each postulated mode of action should address (1) whether the mode of action is supported in animals, (2) whether it is relevant to humans and (3) which populations or life stages can be particularly susceptible. Special attention should be paid to whether tumours can arise from childhood exposure, considering various aspects of development during these life stages. Because the cancer studies are usually performed with young adult or juvenile animals, conclusions about relevance during early childhood generally rely on inference.

Relevance of rodent mode of action for humans

53. "Relevance" of a potential mode of action is considered in the context of characterization of hazard and not at the level of risk. Anticipated levels of human exposure are not used to determine whether the postulated mode of action is operative in a particular population or life stage, for example, in those with pre-existing disease. (USEPA, 2005) Human relevance is discussed in the following section (section 2.4), in the context of the Human Relevance Framework.

54. Other populations or life stages may not be analogous to the test animals, in which case the question of relevance would be decided by inference. Although agent specific data would be preferable, this review may also rely on general knowledge about the precursor events and characteristics of individuals susceptible to these key precursor events. Any information suggesting quantitative differences between populations or life stages should be flagged for consideration in the dose-response assessment, and a separate risk estimate should be quantified for susceptible populations or life stage if data suggests a quantitative difference.

2.4 Human Relevance Framework

55. Considerable effort has been expended during the past several decades to evaluate the mode of action for specific chemicals causing cancer in rodents. However, the key question is the relevance of this postulated mode of action to human risk assessment. A framework was developed by an ILSI/RSI working group sponsored by the U.S. EPA and Health Canada to address this issue and to provide direction in determining the relevance of rodent tumours to human health (Cohen *et al.*, 2003; Meek *et al.*, 2003; Cohen *et al.*, 2004). This human relevance framework is not prescriptive and does not provide a check list of criteria; it is an analytical tool that describes methods and a decision tree logic to establish a relationship between early cellular events and the development of cancer and its relevance to humans. Knowledge of key events and the identification of a mode of action provide a more rational basis for human hazard and risk assessment.

56. The human relevance framework is based on three questions: (1) is the weight of evidence sufficient to establish the mode of action in animals? (2) are key events in the animal mode of action plausible in humans? And (3) taking into account kinetic and dynamic factors, are key events in the

animal mode of action plausible in humans? This is a more quantitative analysis which addresses the relevance of tumourigenicity to a level of exposure, and again relies on a concordance analysis between animal model and humans. This approach focuses not only on dose response but also on quantitative differences between species in fundamental biologic processes that can affect exposure.

57. Presentation in tabular form referred to as a concordance table can be particularly useful. The information in these tables should be relatively brief, as a narrative explanation. There should be one column for the effect on humans for each key events evaluated and another column for the results in a different strain, species, or sex or for a different route of administration that does not result in toxicity. Factors may be identified that are not key events but can modulate key events and contribute to differences between species or individuals. Examples include genetic differences in pathways of metabolism, competing pathways of metabolism, and effects induced by concurrent pathology. While information for evaluating key events in humans may come from *in vitro* and *in vivo* studies on the chemical, basic information on anatomy, physiology, endocrinology, genetic disorders, human epidemiology, and other information that is known regarding the key events should be considered in this framework (Boobis *et al.*, 2008) .

58. This human relevance framework is focused on hazard identification and evaluation. If the second and third questions are answered in the negative, then there is not a cancer hazard for humans and therefore no cancer risk (Holsapple *et al.*, 2005). It is clearly acknowledged that departure from the default assumption of human relevance is a data rich determination. If a conclusion is strongly supported by empirical data, exposure to chemicals producing the toxicity only via that mode of action would not pose a risk to humans. Therefore, no additional risk characterization for this endpoint of carcinogenicity is further warranted (Boobis *et al.*, 2008).

59. Appendix II includes an example where Pastoor *et al.* (2005) describe a rodent mode of action they judge not relevant for humans. This determination for thiamethoxam-related mouse liver tumours was based on the quantitation of key metabolites *in vivo* and *in vitro* that showed mice, but not rats or humans to be capable of generating sufficient amounts of these metabolites to initiate the hepatic toxicity necessary for tumour formation.

Hazard Characterization

60. The hazard characterization provides the overall weight of evidence summary of the assessment. It summarizes the conclusions about the agent's potential effects, whether they can be expected to depend qualitatively on the circumstances of exposure, and if anyone can be expected to be especially susceptible. It discusses the extent to which these conclusions are supported by data or are the result of default options invoked because the data are inconclusive. It explains how complex cases with differing results in different studies were resolved. The hazard characterization highlights the major issues addressed in the hazard assessment and discusses alternative interpretations of the data and the degree to which they are supported scientifically.

61. When the conclusion is supported by mode of action information, the hazard characterization also provides a clear summary of the mode of action conclusions, including the completeness of the data, the strengths and limitations of the inferences made, the potential for other modes of action, and the implications of the mode of action for selecting viable approaches to the dose response assessment. The hazard characterization also discusses the extent to which mode of action information is available to address the potential for disproportionate risks in specific populations or lifestages or the potential for enhanced risks on the basis of interactions with other agents or stressors.

2.5 Consideration of "Mode of Action" information to Optimise the Design of Long-term Studies

An "Integrative" Approach

62. Before embarking on the design of a long-term chronic rodent study, the objective or goal of the study must be clearly defined. If the objective is to determine a chemical's carcinogenic potential, and understand how it elicits its carcinogenicity (mode of toxicological action), a weight of evidence, integrative approach needs to be considered. It is critical to explore potential challenges that might arise from having multiple objectives for the bioassay. For practical reasons, it is often the case that there are multiple objectives in this one study design. To this end, multiple objectives can be accomplished by considering how the design of the study could be optimized for each individual objective. This may necessitate the use of additional animals in the study, compared with the standard design of 50 animals of each sex per group laid down in the Test Guideline TG 451. However such use of additional animals in the study design may be justified if it can avoid the need for a separate study, involving the use of more animals overall. The use of additional animals or the introduction of new experimental groups must however be scientifically justified.

63. Using an integrative approach, the following information on the chemical would be weighed:

- What are the basic physicochemical properties of the chemical? Can the compound be administered orally for a long-term duration?
- Is the material a direct acting DNA mutagen; is it genotoxic?
- Does the chemical induce liver enzymes, liver weight increase, hypertrophy?
- How does the chemical causes its toxicity in shorter term studies and is this consistent across multiple species?
- Does the chemical cause hyperplasia or toxicity in particular target organs?
- Does the chemical cause cell proliferation, inflammation, cellular necrosis, apoptosis?
- Does the chemical cause hormonal perturbation in shorter term or subchronic animal studies? or
- Are there similar analogs, QSAR analyses or structural alerts for organ toxicity? Cancer?

All these data become relevant to the consideration of a chemical's primary mode of action and the determination of carcinogenic potential (Jacobs, 2005).

Hypothesized Mode of Action

64. If the objective of the study is to provide data to test hypotheses regarding a mode of toxicologic action based on information from a structurally similar analog, SAR prediction, or if a chemical "fits" structurally in a particular class of compounds that is known to have a specific mode of action and similar target organ, the study design should include the consideration of dose spacing, temporal sequence of key events specific to the proposed or suggested mode of action, and subsequent precursor key events. Some key questions include: what properties of the dose-response curve are most important, the steepness of the slope? How would one determine the placement of animals into the various dose categories and what are the necessary precursor events in the mode of action for carcinogenicity? The issues of a chemical's structural activity, dose selection and inclusion of precursor key events will be discussed in the ensuing paragraphs.

Consideration of Structure Activity Relationship (SAR)

65. If information on potential modes of action on the chemical of interest or similarly structured compound analog (e.g., SAR) is available, this could assist in identifying the possible, proposed mode of action. In general, the results of shorter term, subchronic studies can also provide information needed to identify necessary precursor events, and inform the selection of adequate doses for a subsequent carcinogenicity study. However, the dose levels administered in a subchronic study in the parent or structurally similar compound that induce a particular targeted effect may need to be adjusted (increased or decreased) for incorporation in a carcinogenicity study.

Dose Selection/Placement

66. Possible insights as to the postulated mode of action can be provided by obtaining information on how the toxicity of the chemical changes with increasing dose, which in turn can be provided by appropriate dose selection and placement (see also Section 3.1). This approach requires some previously generated information (e.g., shorter term studies or toxicokinetic data). If prior evidence allows, it may be possible to optimize the study design in terms of the location of the doses and the allocation of animals to the doses. Depending on the state of the scientific knowledge regarding possible mode(s) of action, sensitivity may be a less important constraint on dose group number than the need for an experimental design, that can compare the various alternative modes of actions. This is a method driven by hypotheses regarding the mode of action.

Precursor/Key Events

67. For mode of action determinations, there may be value in conducting additional, specific studies as a preliminary step to conducting a chronic/carcinogenicity study. With a hypothesized mode of action, the study design of the long-term bioassay can be modified to include additional precursor key events (enzyme induction, cellular proliferation, hormonal perturbation, necrosis and/or apoptosis, hyperplasia, foci development, clonal expansion, etc.) that would support a mode of action. This will normally involve the use of additional satellite animals in the study design, which can be justified if it can avoid the need for a separate study, involving the use of more animals overall. It may however be more appropriate to examine some of these key events in separate, shorter term, studies. Additional analyses of cellular proliferation at various time points (1, 7, 14, 28 days), or cellular necrosis and/or apoptosis and additional clinical chemistry parameters (e.g., gamma glutamyl transpeptidase (GGT), alanine transaminase (ALT), etc.) may be investigated, again in satellite groups or in a separate study, to provide a more robust mode of action analysis. In including these additional parameters in the study design of the long-term bioassay, it is important to consider dose selection and key events at doses lower than the top dose and observations occurring earlier than the time of appearance of the first tumours in the study.

68. Results from investigations applying non-standard test methods such as DNA-microarrays and other 'omics' tools, e.g., methods to globally analyze the expression of genes, proteins or metabolites may provide supplementary information to support the identification of primary targets and mechanisms on the molecular level. Characterization of key events and dose-related responses in a presumed toxic/carcinogenic mode of action may be used to predict the relevance of toxic effects observed in conventional tests used for human health-related end points. In a case-by case approach such data could contribute to the optimisation of the design of long-term studies. This way, the weight of evidence on the potential of the test substance to induce cancer can be evaluated. In order to ensure the validity of such non-standard test methods, any microscopic or measurable substance-related adverse effects of the conventional methods should correlate with the observed changes in expression of targeted genes, proteins or metabolites, taking into account the high sensitivity and appropriateness of the technique applied.

69. The use of information on mode of action and human relevance frameworks for interpretation of carcinogenicity studies in rodents is relatively new. Therefore Appendix I and Appendix II to this Chapter provide examples in which these approaches have been useful, in order to indicate the type of additional information that is valuable. These approaches should be taken into consideration in the design of a chronic toxicity or carcinogenicity study to ensure that the maximum information is obtained from the minimum number of animals.

REFERENCES

Boobis, A.R., S.M. Cohen, V. Dellarco, M.E. Meek, C. Vickers, D. Willcocks and W. Farland (2006), "IPCS Framework for Analyzing the Relevance of a Cancer Mode of Action for Humans", *Critical Reviews in Toxicology*, 36:781-792.

Boobis, A.R., J.E. Doe, B. Heinrich-Hirsch, M.E. Meek, S. Munn, M. Ruchirawat, J. Schlatter, J. Seed and C. Vickers (2008), "IPCS Framework for Analyzing the Relevance of a Noncancer Mode of Action for Humans", *Critical Reviews in Toxicology*, 38:87-96.

Cattley, R., J. DeLuca, C. Elcombe, P. Fenner-Crisp, B. Lake, D. Marsman, T. Oastoor, J.A. Popp, D.E. Robinson, B. Schwatz, J. Tugwood and W. Wahli (1998), "Do Peroxisome Proliferating Compounds Pose a Hepatocarcinogenic Hazard to Humans?" *Regul. Toxicol. Pharmacol.*, 27:47-60.

Cohen, S.M., M.E. Meek, J.E. Klaunig, D.E. Patton and P.A. Fenner-Crisp, (2003), "The Human Relevance of Information on Carcinogenic Modes of Action: Overview", *Crit. Rev. Toxicol.* **33,** 581–589.

Cohen, S.M., J. Klaunig, M.E. Meek, R.N. Hill, T. Pastoor, L. Lehman-McKeeman, J. Bucher, D.G. Longfellow, J. Seed, V. Dellarco, P. Fenner-Crisp and D. Patton (2004), "Evaluating the Human Relevance of Chemically Induced Animal Tumors", *Toxicol. Sci.*, 78: 181-186.

Committee on carcinogenicity of chemicals (COC) in food, consumer products and the environment (2004), Guidance on a strategy for the risk assessment of chemical carcinogens. http://www.iacoc.org.uk/publications/documents/guideline04.pdf

Guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures. http://guidance.echa.europa.eu/docs/guidance_document/clp_en.pdf

Ellinger-Ziegelbauer, H., B. Stuart, B. Wahle, W. Bormann and H.J. Ahr (2004), "Characteristic Eexpression Profiles Induced by Genotoxic Carcinogens in Rat Liver", *Toxicol. Sci.*, 77: 19-34.

Ellinger-Ziegelbauer, H., B. Stuart, B. Wahle, W. Bormann and H.J. Ahr (2005), "Comparison of the Expression Profiles Induced by Genotoxic and Nongenotoxic Carcinogens in Rat Liver", *Mutat. Res.* 575:61-84.

Hill, A.B. (1965), "The Environment and Disease: Association or Causation?" *Proceedings of the Royal Society of Medicine*, 58:295–300.

Holsapple, M.P., H.C. Pitot, S.M. Cohen, A.R. Boobis, J.E. Klaunig, T. Pastoor, V.L. Dellarco and Y.P. Dragan (2006), "Mode of Action in Relevance of Rodent Liver Tumors to Human Cancer Risk", *Toxicol. Sci.*, 89:51-56.

IPCS (2005), *Record of the Cancer Framework Workshop, Bradford, United Kingdom, 21–23 April 2005.* Geneva, World Health Organization, International Programme on Chemical Safety.

Jacobs, A. (2005), "Prediction of 2-year Carcinogenicity Study Results for Pharmaceutical Products: How Are We Doing?" *Toxicological sciences*, 88(1), 18-23.

Meek, E.M., J.R. Bucher, S.M. Cohen, V. Dellarco, R.N. Hill, L.D. Lehman-McKemmon, D.G. Longfellow, T. Pastoor, J. Seed and D.E. Patton (2003), "A Framework for Human Relevance Analysis of Information on Carcinogenic Modes of Action", *Crit. Rev. Toxicol.*, 33:591-653.

Parekh, A.A. and K.L. Dearfield (2007), "USEPA's Risk Assessment Practice: Extrapolation Between Species", *Human and Ecological Risk Assessment*, 13:96-100.

Pastoor T., P. Rose, S. Lloyd, R. Peffer and T. Green (2005), "Case Study: Weight of Evidence Evaluation of the Human Health Relevance of Thiamethoxam-Related Mouse Liver Tumors", *Toxicological Sciences*, 86(1), 56-60.

Rhomberg L.R., K. Baetke, J. Blancato, J. Bus, S. Cohen, R. Conolly, R. Dixit, J. Doe, K. Ekelman, P. Fenner Crisp, P. Harvey, D. Hattis, A. Jacobs and D. Jacobson-Kram (2007), "Issues in the design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection", *Critical Reviews in Toxicology*, 37:729-837.

Rice, J.M. (2004), "On the Application of Data on Mode of Action to Carcinogenic Risk Assessment", *Toxicol. Sci.*, 78, 175–177.

USEPA (2005), *Guidelines for Carcinogen Risk Assessmen,* Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC, USA.

EAP/630/P-03/001F http://cfpub.epa.gov/ncea/raf/recorddisplay.cfm?deid=116283

Ward, W., D. Delker, S. Hester, S.-F. Thai, D. Wolf, J. Allen and S. Nesnow (2006), "Transcriptional Profiles in Liver from Mice Treated with Hepatotumorigenic and Nonhepatotumorigenic Triazole Conazole Fungicides: Propiconazole, Triadimefon, and Myclobutanil", *Toxicol. Path.*, 34:863-878.

Appendix I. Examples of Animal Mode(s) of Action (MOA) Framework

1. Example(s) of Liver Cytotoxic Mode of Action: Chloroform and Carbon Tetrachloride

This paper summarizes recent developments in the continuing evolution of Human Relevance Frameworks to systematically consider the weight of evidence of hypothesized modes of action in animals and their potential human relevance for both cancer and non-cancer effects. These frameworks have been developed in initiatives of the International Life Sciences Institute Risk Sciences Institute and the International Programme on Chemical Safety engaging large numbers of scientists internationally. They are analytical tools designed to organize information in hazard characterization as a basis to clarify the extent of the weight of evidence for mode of action in animals and human relevance and subsequent implications for dose-response. They are also extremely helpful in identifying critical data gaps. These frameworks which are illustrated by an increasing number of case studies, have been widely adopted into international and national guidance and assessments and continue to evolve, as experience increases in their application (Meek, 2008).

Under the 2005 U.S. EPA Guidelines for Carcinogen Risk Assessment, evaluations of carcinogens rely on mode of action data to better inform dose response assessments. A reassessment of carbon tetrachloride, a model hepatotoxicant and carcinogen, provides an opportunity to incorporate into the assessment biologically relevant mode of action data on its carcinogenesis. Mechanistic studies provide evidence that metabolism of carbon tetrachloride via CYP2E1 to highly reactive free radical metabolites plays a critical role in the postulated mode of action. The primary metabolites, trichloromethyl and trichloromethyl peroxy free radicals, are highly reactive and are capable of covalently binding locally to cellular macromolecules, with preference for fatty acids from membrane phospholipids. The free radicals initiate lipid peroxidation by attacking polyunsaturated fatty acids in membranes, setting off a free radical chain reaction sequence. Lipid peroxidation is known to cause membrane disruption, resulting in the loss of membrane integrity and leakage of microsomal enzymes. By-products of lipid peroxidation include reactive aldehydes that can form protein and DNA adducts and may contribute to hepatotoxicity and carcinogenicity, respectively. Natural antioxidants, including glutathione, are capable of quenching the lipid peroxidation reaction. When glutathione and other antioxidants are depleted, however, opportunities for lipid peroxidation are enhanced. Weakened cellular membranes allow sufficient leakage of calcium into the cytosol to disrupt intracellular calcium homeostasis. High calcium levels in the cytosol activate calcium-dependent proteases and phospholipases that further increase the breakdown of the membranes. Similarly, the increase in intracellular calcium can activate endonucleases that can cause chromosomal damage and also contribute to cell death. Sustained cell regeneration and proliferation following cell death may increase the likelihood of unrepaired spontaneous, lipid peroxidation- or endonuclease-derived mutations that can lead to cancer. Based on this body of scientific evidence, doses that do not cause sustained cytotoxicity and regenerative cell proliferation would subsequently be protective of liver tumours if this is the primary mode of action. To fulfill the mode of action framework, additional research may be necessary to determine alternative mode(s) of action for liver tumours formed via carbon tetrachloride exposure (Manibusan *et al.*, 2007).

References

Manibusan, MK, M. Odin and D.A. Eastmond (2007), "Postulated Carbon Tetrachloride Mode of Action: a Review", *J Environ Sci Health C. Environ Carcinog Ecotoxicol Rev* Jul-Sep; 25(3), 185-209.

Meek, B. (2008), "Recent Developments in Frameworks to Consider Human Relevance of Hypothesized Modes of Action for Tumors in Animals", *Environmental and Molecular Mutagenesis*, Volume 49, Issue 2, 110-116.

SAB (1999), *Minutes of the Chloroform Risk Assessment Review Subcommittee Meeting*, October 27-28, 1999, U.S. Environmental Protection Agency, Science Advisory Board. See website: http://www.epa.gov/sab/chloro00.htm

2. Example of Urothelial Cytotoxicity and Increased Cellular Proliferation: Dimethylarsinic Acid

The dose-response relationship for Dimethyl Arsenic (DMA) tumourigenesis based on mode of action considerations will be nonlinear as it is dependent on genetic, biochemical and histopathological events for which dose-response relationships are nonlinear. There must be a sufficient concentration of DMAIII in the bladder to produce cell death and regenerative proliferation. The dose-response assessment would ideally be based on use of DMAIII dosimetry at the target tissue because it represents the rate-limiting event of reductive metabolism to DMAIII to provide a level of exposure that will be protective against the key event of regenerative proliferation. Therefore, the mode of action analysis shows that sufficient DMAIII must be present to result in sufficient urothelial cytotoxicity and cell killing to result in increase cell poliferation and associated chromosomal aberrations. All of these events must occur to result in a neoplastic response. Any one event alone is not sufficient to lead to tumours.

References

USEPA (2005), Science Issue Paper: Mode of Carcinogenic Action for Cacodylic Acid (Dimethylarsinic Acid, DMAV) and Recommendations for Dose Response Extrapolation July 26, 2005, Health Effects Division, Office of Pesticide Programs, US Environmental Protection Program. See website;
http://www.epa.gov/oppsrrd1/reregistration/cacodylic_acid/dma_moa.pdf

SAB (2007), Advisory on EPA's Assessments of Carcinogenic Effects of Organic and Inorganic Arsenic: A Report of the US EPA Science Advisory Board http://www.epa.gov/sab/pdf/sab-07-008.pdf

3. Example of a Neuroendocrine Mode of Action: Atrazine

In 2000, EPA presented a proposed MOA for atrazine to the FIFRA SAP which supported the Agency's approach. EPA described this MOA and the relevant cancer and reproductive toxicity data in the "Atrazine: Hazard and Dose-Response Assessment and Characterization" (FIFRA SAP, 2000a). In brief, upon high levels of exposure to atrazine, the release of gonadotropin releasing hormone (GnRH) from the hypothalamus is reduced, thereby lessening the afternoon pituitary luteinizing hormone surge in female Sprague Dawley rats. As a result, the estrus cycle lengthens. This, in turn, leads to increased estrogen levels and an increased incidence of mammary tumours in female Sprague Dawley rats. (SAP, 2000)

References

SAP (2000), SAP Report No. 2000-05, FIFRA Scientific Advisory Panel Meeting, June 27-29, 2000, Held at the Sheraton Crystal City Hotel, Arlington, Virginia, Atrazine Hazard and Dose Response Assessment and Characterization. See website:
http://www.epa.gov/oscpmont/sap/meetings/2000/june27/finalatrazine.pdf

4. Example of a Mutagenic Mode of Action: Chromium (VI)

Based on the findings from the National Toxicology Program (NTP) 2-year drinking water studies in rodents that Cr (VI) significantly increased the incidence of carcinomas in the small intestine of male and female mice, a weight-of-the-evidence (WOE) approach was applied to judge the mutagenic data of Cr (VI) relative to the induction of gene mutations and/or chromosome aberrations. The next step was to determine whether the induction of gene mutations and chromosome aberrations, which was seen across *in vitro* studies, in animals (mice and rats) and in humans, was an early key event in the carcinogenic process. The synthesis and critical analysis of these data were instrumental in establishing whether the genetic damage occurs early and at doses that are within the tumourigenic range. These two criteria are the hallmarks of the *Cancer Guidelines* MOA framework and were clearly supported by the Cr (VI) data. While this is largely a data-rich undertaking, most of the studies evaluated in this document were primarily designed for hazard identification. Consequently, studies that specifically address various aspects of the MOA framework analysis are not always available for review. Despite this limitation, the WOE approach taken with Cr (VI) demonstrates the utility of this strategy in identifying missing data and establishes the influence missing data can have on the final conclusion. Based, on these considerations, it was concluded that there is plausible evidence that Cr (VI), administer via drinking water, acts via a mutagenic MOA for carcinogenicity (Akerman *et al.*, 2009).

References

Akerman, G., N. McCarroll, J. Chen, N. Keshava, A. Kligerman and E. Rinde (2009), *An Evaluation of the Mode of Action Framework for Mutagenic Carcinogens: Chromium (VI)*, Society of Toxicology, Baltimore, MD.

McCarroll, N., N. Keshava, J. Chen, G. Akerman, A. Kligerman and E. Rinde (2009), "An Evaluation of the Mode of Action Framework for Mutagenic Carcinogens Case Study II: Chromium (VI)", *Environ Mol Mutag*, 51(2):89-111.

McCarroll, N., N. Keshava, M. Cimino, M. Chu, K. Dearfiled, C. Keshava, A. Kligerman, R. Owen, A. Protzel, R. Putzrath, R. Schoney, (2008), "An Evaluation of the Mode of Action Framework for Mutagenic Carcinogens Case Study: Cyclophosphamide", *Environ Mol Mutag*, 49:117-131.

5. Example of liver mitogenic mode of action: conazole group of antifungal agents

Many of the conazole agents induce liver tumours in male mice. The key events leading to tumour induction are known: activation of constitutive receptors (CAR), modulation of cytochrome P450 CYP isoforms, increased liver weights, increased liver hypertrophy, induction of cell proliferation, and liver tumours. By expanding the sampling times and adding additional measurements such as liver enzymes (e.g., PROD, EROD, BROD and possible other liver function assays) and cell proliferation tests, the mode of action analysis can proceed and generate meaningful results long before the data from chronic studies are generated. Similarly, the organ weight data and histopathologic data on hypertrophy and hyperplasia can be examined to determine if a time-related increase in liver weight occurs and if the development of hyperplasia is sustained. It is also important to note that data pertinent to the MOA analysis may also be found in other toxicology tests such as other subchronic, chronic, carcinogenicity, developmental and/or reproductive toxicity studies. This is particularly true for target organ support, organ weights and histopathology and consistency across animal species.

References

Allen, J.W., D.C. Wolf, M.H. George, S.D. Hester, G. Sun, S.F. Thai, D.A. Delker, T. Moore, C. Jones, G. Nelson, *et al.* (2006), "Toxicity Profiles in Mice Treated with Hepatotumorigenic and Non-

hepatotumorigenic Triazole Conazole Fungicides: Propiconazole, Triadimefon, and Myclobutanil", *Toxicol. Pathol.*, 34, 853–862.

Wolf, D.C., J.W. Allen, M.H. George, S.D. Hester, G. Sun, T. Moore, S.F. Thai, D.A. Delker, E. Winkfield, S. Leavitt, *et al.* (2006),. "Toxicity Profiles in Rats Treated with Tumorigenic and Non-tumorigenic Triazole Conazole Fungicides: Propiconazole, Triadimefon, and Myclobutanil", *Toxicol. Pathol.*, 34, 895–902.

Nesnow, S., W. Ward, T. Moore, H. Ren and S.D. Hester (2009), "Discrimination of Tumorigenic Triazole Conazoles from Phenobarbital by Transcriptional Analyses of Mouse Liver Gene Expression", *Toxicol. Sci.*, 110, 68-83.

Appendix II. Example of Human Relevance Framework

Evidence Evaluation of the Human Health Relevance of Thiamethoxam-Related Mouse Liver Tumours (Pastoor *et al.*, 2005)

Thiamethoxam was shown to increase the incidence of mouse liver tumours in an 18-month study; however, thiamethoxam was not hepatocarcinogenic in rats. Thiamethoxam is not genotoxic, and given the late life generation of mouse liver tumours, suggests a time related progression of key hepatic events that leads to the tumours. These key events were identified in a series of studies of up to 50 weeks that showed the time dependent evolution of relatively mild liver dysfunction within 10 weeks of dosing, followed by frank signs of hepatotoxicity after 20 weeks leading to cellular attrition and regenerative hyperplasia. A metabolite CGA330050 was identified as generating the mild hepatic toxicity, and another metabolite, CGA265307, exacerbated the initial toxicity by inhibiting inducible nitric oxide synthase. This combination of metabolite generating hepatotoxicity and increase in cell replication rates is postulated as the mode of action for thiamethoxam. The relevance of these mouse specific tumours to human health was assessed by using the framework and decision logic developed by ILSI/RSI. The postulated mode of action was tested against the Hill criteria and found to fulfil the comprehensive requirements of strength, consistency, specificity, temporality, dose response, and the collective criteria of being a plausible mode of action that fits with known and similar modes of action. Whereas the postulated mode of action could theoretically operate in human liver, quantitation of key metabolites *in vivo* and *in vitro* showed that mice, but not rats or humans, generate sufficient amounts of these metabolites to initiate the hepatic toxicity and consequent tumours. Indeed rats fed 3000 ppm thiamethoxam for a lifetime did not develop hepatotoxicity or tumours. In conclusion, the coherence and extent of the database clearly demonstrates the mode of action for mouse liver tumourigenesis and also allows for the conclusion that thiamethoxam does not pose a carcinogenic risk to humans because of toxicokinetic differences between mice and humans.

3. STUDY DESIGN

3.1 DOSE SELECTION

3.1.1 Introduction

70. The purpose of a long-term bioassay (chronic toxicity and/or carcinogenicity studies) is the detection of biological evidence of any toxic and/or carcinogenic potential of the substance being investigated. Protocols should therefore maximise the sensitivity of the test without significantly altering the accuracy and interpretability of the biological data obtained. The dose regimen has an extremely important bearing on these two critical elements. Since one of the objectives is determination of the dose–response relationship in respect to any endpoints, the OECD TGs 451 (Carcinogenicity Studies), 452 (Chronic Toxicity Studies) and 453 (Combined Chronic Toxicity/Carcinogenicity Studies) normally require at least three dose levels, as well as controls.

71. OECD TGs 451, 452 and 453 outline general principles for dose selection in their respective bioassays. Provision of in depth guidance and a strategy for dose selection is however beyond the scope of the Test Guideline texts. This section of the Guidance Document is designed to underpin and expand the principles of dose selection for chronic toxicity and carcinogenicity studies outlined in the Test Guidelines.

72. These principles of dose selection are generally applicable to a wide range of chemicals, whatever their field of application e.g., pesticides, industrial chemicals and pharmaceuticals. However, although this document provides a number of references to specific requirements for dose selection for pharmaceuticals, the principles applied in studies on pharmaceuticals may differ from that for other agents (Rhomberg *et al.*, 2007; ICH, 2008). More information is generally available on the pharmacodynamic effects of pharmaceuticals, including the results of controlled clinical studies, than for other types of chemicals. The intended systemic human exposure is known and detailed pharmacokinetic studies enable valid comparisons to be made between the systemic exposures in rodents at the chosen dose levels and those in humans under therapeutic administration of the drug, as measured by the comparative areas under the curve (AUC) of blood concentrations over time (Rhomberg *et al.*, 2007). Users of the Guidance should therefore consult the Guideline S1C (R2) on dose selection for carcinogenicity studies of pharmaceuticals for specific information on testing of such chemicals (ICH, 2008).

73. General principles and guidance on dose selection for chronic toxicity and carcinogenicity studies in rodents are provided in two publications of the International Life Sciences Institute (ILSI). An initial 1997 report, entitled *Principles for the Selection of Doses in Chronic Rodent Bioassays* (ILSI, 1997), presented common views on the selection of doses for carcinogenicity and chronic toxicity studies while a second ILSI working group publication in 2007, entitled *Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection* (Rhomberg *et al.*, 2007) provides additional discussion of the factors that influence dose selection in long-term bioassays (Rhomberg *et al.*, 2007). The latter publication incorporates concepts included in other documents prepared by national and international organisations (OECD, ECETOC, NTP and USEPA), and places emphasis on the influence of the objectives of a long-term bioassay on dose selection, as summarised in section 3.1.3 of this guidance. Users of this Guidance Document are recommended to consult these publications for more information on the factors influencing dose

selection. The following sections provide guidance on (a) the principles for dose selection in the Test Guidelines 451, 452, 453, and (b) the influence of the objectives of a long-term bioassay on dose selection.

74. The basic principles for the conduct of chronic toxicity and carcinogenicity studies will differ, given that the endpoints are different. However, given the drive to reduce the number of animals for welfare reasons and the cost of carcinogenicity bioassays, there is a need to maximise the results to assess non-cancer effects that may arise during the study, as these may be critical to the interpretation of any carcinogenic effects. The possibilities for considering non cancer effects in the interpretation of carcinogenic effects are maximised in the TG 453, Combined Chronic Toxicity/Carcinogenicity Study.

75. While the guidance provided in this chapter can be taken as generally applicable to dose selection for either a chronic toxicity or a carcinogenicity bioassay, or a combined chronic and carcinogenicity study, users of the guidance should be mindful of the primary objectives of the study in establishing the optimum study design in terms of dose selection.

76. In selecting appropriate dose levels for long-term bioassays (e.g., TG 451, TG 452, TG 453), a balance has to be achieved between hazard identification/characterization on the one hand and characterization of low-dose responses and their relevance on the other. This is particularly relevant in the situation where a combined chronic toxicity and carcinogenicity study (TG 453) is to be carried out.

3.1.2 Principles for Dose Selection

77. The general principles for dose selection laid down in the TGs are summarised as follows:

- Dose levels should generally be based on the results of shorter-term repeated dose or range finding studies and should take into account any existing toxicological and toxicokinetic data available for the test substance or related materials (Barton at al. 2006).

- The highest dose level should be chosen to identify toxic effects including the principal target organs while avoiding severe toxicity, morbidity, or death (OECD 2000, GD No.19). It should be noted that the severity of toxicity and survival in a two year study may be underestimated from the short-term study; for this reason, Test Guidelines indicate that a top dose lower than the dose providing evidence of toxicity in a short-term study may be chosen. When there is no toxicity in shorter-term studies it is recommended to consult with the relevant regulatory authorities.

- Dose levels should be selected to reflect the purpose of the study. In most cases, dose levels and dose level spacing may be selected to establish a dose-response and to derive a point of departure (e.g., BMDL or NOAEL).

78. These principles for dose selection are broadly similar to the key principles for dose selection outlined in the ILSI publications (ILSI, 1997; Rhomberg et al., 2007), as listed in full in Appendix I to this section. They are further discussed in the following sections.

3.1.2.1 Key information for the selection of doses in chronic toxicity and carcinogenicity studies

79. Identifying/characterizing carcinogenic effect is the primary objective of the OECD TG 451 on Carcinogenicity Studies while identifying/characterizing other toxic effects is the primary objective of the OECD TG 452 on Chronic Toxicity Studies. The OECD TG 453 Combined Chronic Toxicity/Carcinogenicity Studies combines the objective of OECD TG 451 and OECD TG 452. For all three studies the core minimum study design comprises at least one control group and three dose groups, each of which is exposed to different concentrations of the test substance.

80. The robustness of a carcinogenicity or chronic toxicity study, in particular the former, is dependent on a demonstration that the dose levels selected in the study are adequate to show an effect or effects of the test substance, without producing either false negative results (because the doses selected were too low) or false positive results (because metabolic/homeostatic mechanisms are overwhelmed, etc), which may be problematic in assessing risk in humans.

81. The data provided by shorter-term repeated dose or range finding studies, including 28-day or 90-day studies, are important in selecting the dose levels for a longer-term chronic toxicity or carcinogenicity study. The dose levels used in such studies and the NOAELs established can be used as a starting point for dose selection, both in relation to the highest dose level to be chosen in the study and possibly (but not necessarily) to the lower dose levels. Considerations that should be taken into account in determining whether similar, lower or higher dose levels than those used in a short-term study should be selected for a chronic toxicity or carcinogenicity study include (Rhomberg *et al.*, 2007):

- whether the effect is an adaptive response (e.g., liver hypertrophy in the absence of any other evidence of hepatotoxicity);

- potential of the toxic effect(s) observed in repeat dose toxicity studies of shorter duration to progress to neoplasia. A dose that induces a marked effect in such study should not be excluded from a carcinogenicity study if the effect or effects can reasonably be anticipated to be a precursor event in the development of neoplasia (e.g., a key event for the mode of action of the test substance). However, care should be taken that selection of a dose level that induces such effects will not result in excessive toxicity in the carcinogenicity study;

- the potential that an effect may limit the sensitivity of the chronic/carcinogenicity study to detect tumours (e.g., haemolytic anemia may limit the duration of the study due to an increase in mortality or to severe toxicity that may compromise the health of the animals; for more examples see Appendix 1, *Principle 5*, ILSI Principles for Dose Selection in Chronic Rodent Bioassays);

- the duration of the short-term study (e.g., repeated dose 90-day study, repeated dose 28-day study, two-generation reproduction study) and the potential for a toxic effect to progress in severity (e.g., progression from focal to multifocal necrosis);

- evidence of transitory effects that may be life-threatening: if prechronic studies revealed transitory effects that may last during some days or weeks until metabolic capacity (e.g., by liver enzyme induction) is adapted, testing of high dose is limited by transitory effects of life-threatening nature (e.g., sedation);

- use of gavage for administration of the test substance in studies of shorter duration. A dose that induces overt toxicity in a gavage study may be tolerated if a dietary route of

administration is selected for a carcinogenicity study because of the differences in toxicokinetics and toxicodynamics resulting from the two methods of administration.

82. Additional evidence on the extent to which dose levels should be increased or decreased in a long-term study relative to a short-term study or studies may be provided by dose–response data from the latter studies. For example, a marked reduction in dose levels would be warranted if results from short-term studies show that a minor increase in dose is associated with a pronounced increase in severity or incidence of a lesion (i.e., a steep dose–response). It is recommended that all the information from such short-term studies, (rather than the use of an arbitrary factor e.g., one-tenth the highest dose tested in a short-term study that induced a severe toxic effect) should be used when selecting the high dose (or mid and low dose levels) for a proposed carcinogenicity study.

83. Available toxicokinetic data (ADME) should always be taken into account when selecting dose levels for a chronic toxicity or carcinogenicity study, although such data may not be readily available for all chemicals, as they are not required under all regulatory schemes. Many toxicokinetic processes influencing absorption, distribution, elimination and metabolic activation or detoxication may become saturated at higher doses, resulting in systemic exposures to parent compound or metabolites that would not be expected in the real life human exposures for which risk assessments are needed. The effect of repeated exposures on the pattern of absorption, metabolism, detoxification, and clearance of a compound will provide information on the internal dose achieved during chronic exposure under conditions of the bioassay. The importance of having data on toxicokinetics in reaching a decision on the design most suitable for a chronic toxicity or carcinogenicity study is stressed in this guidance and the use of such data are discussed in more detail in Chapter 3, Section 3.4 of this Guidance Document.

84. Physiologically-based toxicokinetic (PBTK) modelling is also a valuable tool for defining doses where non-linear toxicokinetics may occur, thus allowing this to be considered in selecting the highest and other dose levels in the study. The use of PBTK modelling is explored in more detail in Section 3.4. Finally, specific mechanistic studies (where available) may provide useful information regarding target tissues affected by the test substance and the doses associated with effects on key events, and should be taken into account when selecting doses for a chronic toxicity or carcinogenicity study.

85. Additional considerations in selecting dose levels for chronic toxicity or carcinogenicity studies arise as a result of practical constraints such as the physicochemical characteristics of the substance to be tested (e.g., solubility, vapour pressure), palatability of the compound in food or drinking water, and other factors such as the potential for the test substance to cause adverse effects such as irritancy at the site of administration (Rhomberg *et al.*, 2007). Further guidance on these aspects is provided in the ILSI publications (ILSI, 1997; Rhomberg *et al.*, 2007) and in Chapter 3, Section 3.5 of this Guidance Document.

86. Information on, and consideration of, the mode of action (MOA) of a suspected carcinogen is particularly important, since the dose selection may differ depending on the known or suspected mode of action (Sonnich-Mullin *et al.*, 2002; Cohen *et al.*, 2003; Meek *et al.*, 2003; Holsapple *et al.*, 2006; Boobis *et al.*, 2006; EPA 2005). In selecting dose levels for such a study, doses will need to be placed carefully, to yield observations of subtle precursor effects or other biomarkers of toxicity without inducing confounding effects related to frank toxicity. This approach requires some previously gathered information on potential modes of action, e.g., from genotoxicity studies. Further guidance on these aspects is discussed in Chapter 2 of this Guidance Document.

3.1.2.2 Selection of the top dose

87. Dose selection should be based on the findings of subchronic or range-finding studies. The highest dose level to be used in a chronic toxicity or carcinogenicity study needs to be carefully considered and the reasons for the final choice clearly defined. Ideally, the dose levels selected will maximise the detection of dose–response relationships and facilitate the extrapolation of these to potential hazards for other species, including humans.

88. The selection of the highest dose level to be used in a chronic toxicity or carcinogenicity study has long been a matter of controversy. At the time when long-term animal bioassays began to be routinely used to assess the qualitative potential of a test substance to cause chronic toxicity and cancer, the emphasis was on testing at high levels in order to maximise the potential of such studies to detect effects. The concept of the Maximum Tolerated Dose (MTD), conventionally defined as the highest dose to produce toxic effects without causing death and to decrease body weight gain by no more than 10% relative to controls (OECD 2002, GD No. 35) became well established. The MTD is often used in the assessment of a chronic toxicity or a carcinogenicity study to decide whether the top dose tested was adequate to give confidence in a negative result. This Guidance Document focuses on the selection of the top dose, rather than attempting to define an MTD.

89. While some regulatory bodies or organisations interpret an adequate high dose to be a minimally toxic dose, others emphasize the need to select a dose level that is a maximally tolerated dose (i.e., more severe toxicity should be demonstrated). Thus, because of differences in views regarding the severity of toxic effects that are interpreted as providing evidence that an adequate high dose has been attained or exceeded, a completed carcinogenicity bioassay may be considered to be acceptable by one organisation but not by another. Many carcinogenicity studies can be challenged on the basis of selection of a top dose that is too high, particularly if there is a large interval to the next highest dose. This results in data that are difficult to interpret and cannot be used for regulatory purposes. Appendix 2 of Rhomberg (Rhomberg *et al.*, 2007) provides detailed guidance on criteria that can be applied in order to assess the acceptability of the high dose level or MTD.

90. If the main objective of the study is to identify a cancer hazard, there is broad acceptance that the top dose should ideally provide some signs of toxicity such as slight depression of body weight gain (not more than 10%), without causing e.g., tissue necrosis or metabolic saturation and without substantially altering normal life span due to effects other than tumours. Excessive toxicity at the top dose level (or any other dose level) may compromise the usefulness of the study and/or quality of data generated. Criteria that have evolved for the selection of an adequate top dose level include: (in particular) toxicokinetics; saturation of absorption; results of previous repeated dose toxicity studies; the MOA and the MTD.

91. Toxicokinetic non-linearity should also be considered in the selection of the top dose to be used. Although top dose selection based on identification of inflection points in toxicokinetic non-linearity may result in study designs that fail to identify traditional target organ or body weight effects, it must be appreciated that metabolic saturation in fact represents an equivalent indicator of biological stress. In this case, the stress is evidenced by appearance of non-linear toxicokinetics rather than appearance of histological damage, adverse changes in clinical chemistry, haematology parameters or decrease in body weight gain (Toxicokinetics is discussed in Section 3.4).

92. For compounds that are (or potentially are) genotoxic, conventional considerations of top dose given above would apply. For compounds that are not genotoxic, the top dose should be informed by considerations of MOA (see Chapter 2). For a given compound, for which the mode of action is known or suspected, establishing a point of departure based on precursor key events, may also be protective against any carcinogenic effect. This is because non-genotoxic carcinogens produce

cancer by perturbing normal physiology or biochemistry. The long-term assay should be designed to also identify and characterize these key events.

93. Nutritional effects, physiological factors, physical-chemical factors and compound bioavailability can influence selection of the top dose level to be used in a long-term bioassay. For nutritional and possibly other physiological reasons a maximum level is imposed, commonly 5% concentration in the diet (Sontag *et al.*, 1976; Chhabra *et al.*, 1990).

94. Palatability of a compound in either feed or water can also lead to perturbation of physiological homeostasis or nutritional status. A compound's solubility limit or vapour pressure may constrain selection of the top dose level. Irritation at the site of compound deposition may constrain dose or otherwise confound cross species extrapolation. Inhalation of doses that overwhelm pulmonary clearance may lead to tissue responses that are specific to the species being tested; however, this does not apply to asbestos-like substances. These limitations may influence selection of the top dose.

95. The top dose used in the study may be based on a defined level of the target population's exposure of interest and multiples of that exposure (e.g., 100 times or 150 times higher based on dose ratios expressed in terms of body weight). If toxicokinetic data are available, dose levels based on internalised doses (e.g., AUC) can be used. It has been shown that the relative systemic exposure corresponds (AUC ratios) better with dose ratios expressed in terms of body surface area rather than ratios expressed in terms of body weight. It should be noted that the use of systemic exposure comparison between rodents and humans to derive the top dose may be useful in the case of pharmaceuticals testing (ICH, 3(R2), S1C(R2)) but is not likely to be useful for testing plant protection products or commodity chemicals. Given the uncertainties regarding exposure levels in scenarios where these are used, and given the need for these chemicals that the top dose be sufficiently high for the purpose of classification for carcinogenic effect, an inherent potential of the substance in question to cause cancer irrespective of the dose should be demonstrated.

96. The relevance of the top dose level recommended to be used in the study to potential human exposures can also be debated. Mechanistic information gleaned from this type of study may be irrelevant. Positive results in the high dose group may be difficult to interpret as they may reflect a high-dose-only phenomenon, not relevant for human exposure. On the other hand, if the top dose level is set lower, to ensure relevance, the power of the study to detect effects may be compromised.

3.1.2.3 Dose level spacing

97. Selection of dose intervals is influenced by the study objectives (see section 3.1.3) and the available information. Dose levels and dose level spacing may be selected to establish a dose-response and to derive a point of departure (e.g., BMDL or NOAEL). The dose level spacing does not need to be regular. It may be reduced in regions of the dose-response curve where particularly robust estimation is needed, e.g., in the range of the anticipated BMD or a suspected threshold. The increasing emphasis on consideration where the lower dose levels used in the study are placed, and the number of such dose levels, reflects the changing purposes of lifetime bioassays.

98. If the primary purpose is identification of hazard, whether this is chronic toxicity or carcinogenicity, the focus of dose selection should be on maximizing the power of the study and on the top doses tested. As the risk assessment process becomes increasingly concerned with characterization of human risk, there has been a corresponding need to characterize whether and how high- dose effects extend to responses at lower exposure levels as well, with a consequent interest in how the lower dose levels are placed in bioassays (Rhomberg *et al.*, 2007).

99. Dose selection and dose level spacing need to be based, where possible, on the following considerations:

- known or suspected non-linearities or inflection points in the dose–response;
- toxicokinetics, and dose ranges where metabolic induction, saturation, or non-linearity between external and internal doses does or does not occur;
- precursor lesions, markers of effect, or indicators of the operation of key underlying biological processes;
- key (or suspected) aspects of mode of action, such as doses at which cytotoxicity begins to arise, hormone levels are perturbed, homeostatic mechanisms are overwhelmed, etc.;
- regions of the dose–response curve where particularly robust estimation is required, e.g., in the neighbourhood of the anticipated point of departure;
- consideration of anticipated human exposure level;
- a suspected threshold.

100. Dose levels should be selected to reflect the purposes of the study, and they should use available knowledge on how dose-dependent biological and impacted physiological factors may affect study outcomes. The Test Guidelines (TG 451, paragraph 24; TG 452, paragraph 24; and TG 453, paragraph 26) indicate that "The dose level spacing selected will depend on the characteristics of the test substance, and cannot be prescribed in this Guideline, but two to four fold intervals frequently provide good test performance when used for setting the descending dose levels and addition of a fourth test group is often preferable to using very large intervals (e.g., more than a factor of about 6-10) between dosages. In general, the use of factors greater than 10 should be avoided, and must be justified if used".

101. If prior evidence allows, it may be possible to optimise the design in terms of the location of the dose levels. A design often applied uses a mid dose that is half of the top dose, or the geometric mean of the low and high dose. This will ensure that the power and sensitivity of the assay is maximised and that at least one dose is unlikely to have a carcinogenic or other effect. This approach minimises the chance of a false negative (failing to detect an effect that actually exists) at some increased risk of a false positive (finding a high-dose effect that is an artefact of excessively high doses and is not relevant to the dose range of interest).

102. Limited information may be obtained regarding the shape of the dose–response curve, particularly if non-linearity is seen in the middle of the dose range. The power of the assay at lower dose levels will also be limited if the incidences of the responses of interest are low (e.g., rare tumours) and not markedly different from the controls. Information on the dose-response relationship would depend on how well the dose range of interest is anticipated.

103. The issue of where to place the lowest dose should receive comparable attention as to the placement of the top dose. If the lowest dose is too low, it may be insufficiently powerful and therefore uninformative; if too high, it may lose opportunities to characterize effects as near as possible to environmental exposure levels. For example, for pharmaceuticals, a dose sufficient to produce a pharmacodynamic effect or result in systemic exposure comparable with that expected at the intended clinical use is normally selected for the low dose level (see also ICH guideline S1C(R2)). For agrochemical products, food additives and similar products, selection of the lowest dose/dietary concentration usually takes into account a "desired" NOAEL reflecting the likely use of this NOAEL for risk assessment purposes by application of an Uncertainty Factor (UF) to obtain an Acceptable Daily Intake (ADI).

104. It may be possible to place adjacent dose levels somewhat above and below the levels at which a key transition in underlying biological actions, including considerations of the mode of action, is believed to lie, thereby revealing its influence on response. Transitions need not be sharp; typically, there are ranges of doses over which an underlying biological factor, such as metabolic saturation or cytotoxicity, comes increasingly into play.

105. When evaluating threshold effects in a chronic toxicity bioassay, the doses selected will include at least one dose high enough to show toxicity, at least one dose low enough to show lack of toxicity, and usually one but occasionally more than one in between to help characterize the shape of the curve near the point where the threshold appears to lie (Rhomberg *et al.*, 2007). These dose placement concerns differ from those in the carcinogenicity bioassay for substances where genotoxicity is known or suspected; however, this difference disappears if the BMD approach is used. The same dose range is preferred for both phases of a combined chronic toxicity/carcinogenic study, particularly if the MOA is under investigation.

3.1.3 Integration of the objectives of a long-term bioassay

106. The ILSI publications (ILSI, 1997; Rhomberg *et al.*, 2007) provide practical guidance on factors that influence dose selection in long-term bioassays, with particular emphasis on how the varying objectives of a chronic toxicity/carcinogenicity bioassay influence dose level selection.

107. Test Guideline 453 (combined chronic toxicity/carcinogenicity studies) identifies nine possible objectives:

- The identification of the carcinogenic properties of a chemical, resulting in an increased incidence of neoplasms, increased proportion of malignant neoplasms or a reduction in the time to appearance of neoplasms, compared with concurrent control groups;
- The identification of the time to appearance of neoplasms;
- The identification of the chronic toxicity of a chemical;
- The identification of target organ(s) of chronic toxicity and carcinogenicity;
- Characterization of the dose:response relationship;
- Identification of a point of departure (e.g., BMDL or NOAEL);
- Extrapolation of carcinogenic effects to low dose human exposure levels;
- Prediction of chronic toxicity effects at human exposure levels;
- Provision of data to test hypotheses regarding mode of action (EPA 2005; OECD 2009, GD No.116; Boobis *et al.*, 2006; Cohen *et al.*, 2003; Holsapple *et al.*, 2006; Meek *et al.*, 2003).

108. Various study designs have been proposed to address these objectives, as described by Rhomberg *et al.*, 2007. The core study design for a long-term bioassay as laid out in TGs 451, 452, 453 primarily addresses the objective of identification/characterization of carcinogenic substances or those causing other toxic effects, while seeking to integrate the other objectives as far as possible. Modifications of the core study design in TGs 451, 452, 453 in order to optimise data for the other objectives may compromise the Mutual Acceptance of Data and should be discussed with the relevant regulatory authorities before the commencement of the study. More generally, it is recommended to

ensure regulatory acceptance of the study design before performing any long-term bioassay to be submitted to a competent authority.

109. The different objectives outlined above seek to maximise the statistical power of the study at different points on the dose-response curve. The focus may be on a level of response or on the shape and slope of the overall curve. The situation is also complicated by the fact that, below a certain dose, attempts to increase statistical power by increasing animal numbers in particular dose groups become futile.

110. For the majority of bioassays, there will be one primary objective (typically the identification of carcinogenic potential and/or chronic toxicity) and several subsidiary objectives such as characterizing the dose-response curve, extrapolating to low doses, or identifying a point of departure. The nature of the subsidiary objectives will be contingent on the intended outcome. If a valid negative result is obtained in a carcinogenicity study (e.g., OECD TG 451), and this was the only objective of the study, there may be no further questions to be answered. If a positive result is obtained, however, a number of issues arise regarding the nature of the carcinogenic responses and their relevance to the levels of exposure of target populations, requiring further investigation into the nature and interpretation of the effects seen. Consideration of the mode of action framework before embarking on a long-term bioassay will provide guidance on optimising the design to collect the information necessary to the interpretation (see Chapter 2).

111. The choice of dose levels for the identification of a point of departure will depend on the type of point of departure sought. For a NOAEL, a dose without effects is required but ideally at the highest dose at which this can be observed. For a BMDL, the data from all dose levels is used but it is important that the responses differ at the different doses.

112. The study design selected at the outset should include dose levels that combine several objectives. One approach to achieve this is to include additional dose groups in such a way that the optimal doses for a number of different objectives are all included (Rhomberg *et al.*, 2007). Some doses would be optimised for some objectives and others for other objectives, essentially running several bioassays in tandem.

113. However, this is not feasible, given animal welfare, economic and time constraints. When attempting to combine these various objectives into a single study, selection of dose levels must be done in a way that does not compromise the primary objective while still allowing a secondary objective to be pursued in an acceptable albeit suboptimal manner. There may be embellishments to the core design based on study objectives but it would be a rare event when an erosion of the core minimum would be acceptable.

114. Based on Rhomberg *et al.*, 2007, four core selection schemes are presented below:

1. **Hazard Identification/Characterization Plus Dose–Response:** The top dose is chosen to increase the study's statistical power to detect effects that may be rare. A second dose combines two functions: (1) hedging against the top dose being found to have been too high in retrospect, and (2) providing the opportunity for dose–response characterization of any effects found. The lowest dose level or, if considered necessary, other lower dose levels can be placed so as to inform dose–response, no-effect levels, or other purposes. Key challenges will be balancing statistical power and toxicological relevance of the top dose level and compromising among subsidiary objectives while accounting for relevant dose-related physiological changes when setting lower dose levels.

2. **NOAEL/BMDL-Seeking for Threshold Effects:** The main aim is to identify no-effect (or low-effect) levels for the more sensitive adverse threshold effects. The top dose should aim at

engendering an adverse effect, the lowest dose should aim at constituting a NOAEL/BMDL, and the intermediate dose or doses should be set so as to identify the dose levels at which the top dose responses begin to manifest.

3. **Assessment of Safety of Human Exposure Levels:** This is modeled on safety assessment studies for nutrients and pharmaceuticals. For agents that are not genotoxic, show low toxicity, and evince no known difference in metabolic profile between rodents and humans, one can test multiples of anticipated human exposure. Lack of adverse effects at doses sufficiently above human exposure (and the perceived implausibility of non-threshold effects) gives evidence supporting the safety of the anticipated exposures. The bioassay exposures should be selected on an appropriate basis for animal:human comparison; for instance, the application to pharmaceuticals is typically based on area under the blood concentration-time curve that results from anticipated human exposures.

4. **Special-Purpose Bioassays**: this case would be beyond the OECD Test Guidelines and therefore is not developed in this document.

REFERENCES

Barton, H.A., T.P Pastoor, K. Baetcke, J.E. Chambers, J. Diliberto, N.G. Doerrer, J.H. Driver, C.E. Hastings, S. Iyengar, R. Krieger, B. Stahl and C. Timchalk (2006), "The Acquisition and Application of Absorption, Distribution, Metabolism, and Excretion (ADME) Data in Agricultural Chemical Safety Assessments", *Crit Rev Toxicol,* Vol. 36: 9-35.

Boobis, A.R., S.M. Cohen, V. Dellarco, D. McGregor, M.E. Meek, C. Vickers, D. Willcocks and W. Farland (2006), "IPCS Framework for Analyzing the Relevance of a Cancer Mode of Action for Humans", *Crit Rev in Toxicol*, 36:793-801.

Chhabra, R.S., J.E. Huff, B.S. Schwetz and J. Selkirk (1990), "An Overview of Prechronic and Chronic Toxicity/Carcinogenicity Experimental Study Designs and Criteria Used by the National Toxicology Program", *Environmental Health Perspectives,* 86: 313-321.

Cohen, S.M., M.E. Meek, J.E. Klaunig, D.E. Patton and P.A. Fenner-Crisp (2003), "The Human Relevance of Information on Carcinogenic Modes of Action: An Overview", *Crit Rev Toxicol,* 33:581-589.

ECETOC (1996), *Practical concepts for dose selection in chronic toxicity and carcinogenicity studies in rodents*, ECETOC Monograph No. 25. European Centre for Ecotoxicology and Toxicology of Chemicals. Brussels, Belgium.

EPA (2005), *Guidelines for Carcinogen Risk Assessment,* Risk Assessment Forum U.S. Environmental Protection Agency, Washington, DC
http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=116283&CFID=1267360&CFTOKEN=65052793&jsessionid=9830b2c4116e3d8fbbf017414e1a782e7f79TR

Holsapple, M.P., H.C. Pitot, S.N. Cohen, A.R. Boobis, J.E. Klaunig, T. Pastoor, V.L. Dellarco and Y.P. Dragan (2006), "Mode of Action in Relevance of Rodent Liver Tumors to Human Cancer Risk", *Toxicol Sci,* 89:51-56.

ICH (International Conference on Harmonisation) (2008), *Guidance for Industry. S1C(R2) Dose Selection for Carcinogenicity Studies,* U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER), September 2008 ICH. Revision 1

ICH (International Conference on Harmonisation) (2009), *ICH Harmonised Tripartite Guideline - Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorisation for Pharmaceuticals – M3(R2)*
Available at: http://www.ich.org/cache/html/501-272-1.html

ILSI (International Life Sciences Institute) (1997), *Principles for the Selection of Doses in Chronic Rodent Bioassays*, Foran JA (Ed.). ILSI Press, Washington, DC.

Meek, E.M., J.R. Bucher, S.M. Cohen, V. Dellarco, R.N. Hill, L.D. Lehman-McKemmon, D.G. Longfellow, T. Pastoor, J. Seed and D.E. Patton (2003), "A Framework for Human Relevance analysis of Information on Carcinogenic Modes of Action", *Crit Rev Toxicol,* 33:591-653.

OECD (2000), *Guidance Document on the Recognition, Assessment, and use of clinical signs as humane endpoints for experimental animals used in safety evaluation,* OECD Series on Testing and Assessment No. 19, OECD, Paris.

OECD (2002), *Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies*, Series on Testing and Assessment No. 35, OECD, Paris. Available at:

http://www.olis.oecd.org/olis/2002doc.nsf/LinkTo/NT00002BE2/$FILE/JT00130828.PDF.

OECD (2009), *Carcinogenicity Studies*, OECD Guideline for the Testing of Chemicals, TG No. 451, OECD, Paris.

OECD (2009), *Chronic Toxicity Studies,* OECD Guideline for the Testing of Chemicals, TG No. 452, OECD, Paris.

OECD (2009), *Combined Chronic Toxicity\Carcinogenicity Studies*. OECD Guideline for the Testing of Chemicals, TG No. 453, OECD, Paris.

Rhomberg, L., K. Baetcke, J. Blancato, J. Bus, S. Cohen, R. Conolly, R. Dixit, J. Doe, K. Ekelman, P. Fenner-Crisp, P. Harvey, D. Hattis, A. Jacobs, D. Jacobson-Kram, T. Lewandowski, R. Liteplo, O. Pelkonen, J. Rice, D. Somers, A. Turturro, W. West and S. Olin (2007), "Issues in the Design and Interpretation of Chronic Toxicity and Carcinogenicity Studies in Rodents: Approaches to Dose Selection", *Crit Rev Toxicol*, 37:729-837.

Smith, G. (2002), "Scientific Review of Chronic Rodent Bioassays. Criteria for Maximum Tolerated Dose (MTD) Attainment: Clinical Pathology- and Pathology-related Endpoints", Unpublished draft, submitted to Pest Management Regulatory Agency, Canada, September 18, 2002.

Sonnich-Mullin, C., R. Fielder, J. Wiltse, K. Baetcke, J. Dempsey, P. Fenner-Crisp, D. Grant, M. Hartley, A. Knaap, D. Kroese, I. Mangelsdorf, E. Meek, J. Rice and M. Younes (2001), "IPCS Conceptual Framework for Evaluating a Mode of Action for Chemical Carcinogenesis", *Reg Toxicol & Pharm,* 34:146-152.

Sontag, J.M., N.P. Page and U. Saffiotti (1976), *Guidelines for Carcinogen Bioassay in Small Rodents*, U.S. Dept. of Health and human services publication (NIH 76-801), National Cancer Institute, Bethesda.

Appendix 1: ILSI Principles for Dose Selection in Chronic Rodent Bioassays[2] (Rhomberg et al., 2007)

Principle 1

Dose selection for chronic studies must be based on sound toxicologic principles. Within a reasonable dose range, increasing the dose can increase the ability to detect an effect; therefore, doses for chronic rodent bioassays should be selected within this range to maximize the sensitivity of a chronic bioassay. However, trying to increase study sensitivity by increasing doses into ranges that do not reflect application of sound toxicologic principles could lead to results that are inappropriate for human risk assessment.

Increasing the highest dose in a chronic bioassay may increase sensitivity within some defined dose range, but the potential exists that different mechanisms of toxicity or chemical mode of action are active at higher doses, which may not be relevant to humans exposed to lower doses. In this case, selection of the highest dose may be influenced by consideration of the mechanism/ mode of action and other factors discussed in Principle 4. However, when the highest dose in a carcinogenicity assay is limited by effects (e.g., a mode of action in one organ system) that are thought not to occur in humans, one must be aware that it still is possible that a higher dose of the chemical may be carcinogenic in other animal/organ systems.

To address these issues, the ILSI working group encourages an approach to dose selection that incorporates all relevant information from prechronic studies and other sources, uses toxicologic tools associated with an understanding of the mechanisms or mode of action by which a chemical produces an effect (e.g., genotoxicity, cell proliferation, etc.), and uses good scientific principles to enhance the accuracy of judgments of potential human risks. In the case of negative studies (particularly where the highest dose is chosen based on a full characterization of the chemical's toxicity in prechronic studies), use of sound scientific principles as well as all available chemical, physical, and toxicologic data will lessen concern that the result may be a false negative. Similarly, in positive carcinogenicity studies, this approach will lessen concern that the result may be a false positive. In both cases, the predictiveness of the bioassay for human health effects will be improved.

Principle 2

A chronic bioassay requires a major investment in resources and time, and the objective of such a study should be broader than hazard identification. Scientists who conduct chronic bioassays and those who use data from bioassays, including regulatory agencies, should encourage innovative approaches to dose selection by considering appropriate study designs, mechanistic data, and other information in the design and interpretation of studies. Use of additional endpoints and other information must be based on sound scientific rationale, and such designs should be evaluated on the basis of their individual merits.

A goal of high-dose selection in carcinogenicity bioassays is, in the context of hazard identification, to reduce the likelihood of a false-negative result. However, it is recognized that the qualitative nature of the hazard (e.g., carcinogenic response) may itself be dose dependent. This principle encourages approaches to dose selection that incorporate consideration of mechanistic and other toxicologic information. Such approaches should improve the scientific basis for dose selection and aid in interpretation of data generated from chronic bioassays.

---

[2] The terminology used in the Guidance Document regarding "long-term bioassay" designing both chronic toxicity and carcinogenicity studies doesn't apply to this annex.

Principle 3

Human exposure should be considered in dose selection, particularly for selection of the middle and lowest doses. Further, the middle and lowest doses should be selected to characterize the shape of the dose response curve as much as possible. Selection of the middle and lower doses should take into account factors such as the mechanism or mode of action, toxicokinetics, and others listed in Principles 4 and 5 and should not be based solely on a fraction of the highest dose.

Issues that should be considered when incorporating potential human exposure in dose selection include the human exposure route and mode, the dose range in the chronic bioassay in relation to human exposure, and the duration and frequency of human exposure, if known. Subpopulations that may be more highly exposed than the general population, or that are genetically more susceptible, also should be considered. The relationship between external and delivered (internal) dose (e.g., ingested dose versus dose delivered to the target organ, toxicokinetics) in both humans and test organisms may influence dose selection. Further, for substances expected to exhibit a toxicity threshold, or if the evaluation of carcinogenic potential is being combined with an evaluation of chronic toxicity, the study should be designed to include one dose that does not elicit adverse effects; that is, one dose should be a NOAEL. Of course, caution must be exercised to ensure that the NOAEL is not simply an artifact of small sample size or poor study design.

Principle 4

The [ILSI] working group has recommended the use of innovative approaches, additional endpoints, and other information in the selection of doses for chronic rodent bioassays. The following endpoints, generally determined in prechronic studies, should be considered in dose selection for chronic rodent bioassays. Further, it is recognized that endpoints other than those listed below may provide important information for dose selection, and use of those endpoints, where they are based on sound toxicologic principles, is encouraged. Such endpoints may be available presently, or they may be developed as the science of toxicology advances.

- **Histopathology**

  The site, morphology, and severity of the treatment-related effects observed in the pre-chronic study should be taken into account in setting dose levels for the chronic study. Histopathological examination of tissues, especially the liver, kidneys, gastrointestinal tract, urinary tract, respiratory tract, skin, spleen/bone marrow/blood, and endocrine tissues derived from properly designed pre-chronic studies, often provides information that is crucial for dose selection in chronic bioassays.

- **Toxicokinetics**

  Studies to determine the effect of dose (or exposure concentration) on absorption, tissue distribution, metabolism, and clearance of a compound are helpful in selecting appropriate doses for the chronic bioassay. The kinetics of absorption will determine the internal exposure dose achieved. The absorption and clearance of the compound and its metabolites will determine the systemic and target organ exposure resulting from a single dose and can be used to design the treatment regimen required to achieve a desired internal dose. The effect of repeated exposures on the pattern of absorption, metabolism, biotransformation, and clearance of a compound will provide information on the internal dose achieved during chronic exposure under conditions of the bioassay. The nutritional status of the chronically exposed animals may be affected during the experimental

period which is why adequate information on interactions between the exposure chemical and nutrionally important compounds may be of great value in the interpretation of the final results of the chronic study.

- **Cell Proliferation**

In the process of chemical carcinogenesis, events related to induced cell proliferation might be critical in fixing mutations and in providing a selective growth advantage to pre-cancerous cells. Considerations may be different for direct mitogenic stimulation of organ growth versus regenerative cell proliferation, and these modes of action should be distinguished for the test agent. Further, apoptosis can be a strong determinant of normal and pre-cancerous cell turnover kinetics and should be considered. Information on the dose dependence of regenerative cell proliferation is a useful adjunct to histological observations in determining the shapes of organ-specific toxic response curves. This information, when available, can be of value in selecting high, middle, and low doses and in interpreting the results of the study.

- **Physiological Functions**

Disturbances of physiology or homeostasis that would compromise the validity of the study should be considered in the dose-selection process. Examples include hypotension, inhibition of blood clotting, overwhelming normal pulmonary clearance mechanisms, immune system effects, and in some cases hormonal imbalance. Such disturbances, and their effects on the validity of a study, may be difficult to determine and may apply differentially to different categories of chemicals (e.g., pharmaceuticals in which the desired pharmacological action is a physiological effect).

- **Body Weight**

It is suggested that body-weight changes are the primary factor in the selection of the highest-dose group (that is, when no other toxic effects are observed), a decrement in body-weight gain of no more than 5–10% in pre-chronic studies should be used in the selection of the highest dose for chronic assays of carcinogenicity.

Historically, scientists have adopted a 10% decreased body-weight gain at the end of pre-chronic studies (typically 90 days duration) as the target that should not be exceeded in chronic (carcinogenicity) studies. It is now recognised that there is a positive correlation between body weight and the occurrence of certain tumours in rodent species and strains used in safety assessment or for hazard identification; i.e., the higher the body weight between 6 and 18 months on test, the higher the probability that the animal will develop some tumours. Moreover, the lower the body weight, the less sensitive the animal may be to agent-induced toxicity, including cancer. A significant decrease in body-weight gain therefore could reduce the animal's ability to respond to compound-induced toxicities.

- **Clinical Chemistry, Haematology, Urinalysis**

Clinical chemistry and urinalysis results are best used to support dose-selection decisions based on other criteria/parameters. Changes in serum clinical chemistry in the absence of histopathological observations may not affect high-dose selection but may complement dose-selection decisions based on toxicokinetics, cell proliferation, and other parameters. Haematology results also are more often affected secondarily to other processes more relevant for dose selection (e.g., inflammation). However, when haematological tissues are determined to be a target organ in pre-chronic studies, haematology results may be an appropriate basis for dose selection.

- **Organ Weights**

Organ weights are not often the critical factor in selection of doses for chronic rodent bioassays. Chemically induced changes in organ weights, however, should be considered in conjunction with other data in the dose-selection process.

Ideally, data from the factors and endpoints listed above would be collected from pre-chronic studies and used to select doses for chronic studies; however, not all parameters may be useful or necessary for every compound. Even when based on information concerning the points described above, dose selection for a chronic bioassay will remain an inexact process; thus, reconsideration of these same points must be made

when interpreting and assessing the significance of the effects obtained in the bioassay.

Principle 5

Physicochemical factors (e.g., solubility, vapour pressure), the bioavailability of the compound, the palatability of the compound in food or drinking water, and other factors such as the potential for the substance to cause adverse effects at the site of administration (e.g., irritation, erosion, and ulceration) will influence the selection of the highest dose for chronic rodent bioassays. It is recommended that doses for chronic rodent bioassays be selected to minimise or avoid adverse nutritional, physical, organoleptic, and irritant effects.

3.2 ROUTES OF EXPOSURE AND DOSE ADMINISTRATION CONSIDERATIONS

Introduction

115. The choice of the route of administration depends on the physical and chemical characteristics of the test substance, its intended field of application, the availability of information on shorter-term repeat dose studies and the predominant route of human exposure. Comparison across the set of studies available for a given chemical will be more robust if the parameters of the study e.g., route, dose volume, vehicle, remain constant between all the repeat dose toxicity studies. The three main routes of administration used in chronic toxicity and carcinogenicity studies are oral, dermal and inhalation. For example, if human exposure to the test substance is likely to be through food or is a pharmaceutical intended to be taken by mouth, the relevant route of administration will be the oral route, while for a workplace gas, inhalable dust or volatile liquid, inhalation should be the route of choice. The dermal route may be chosen, e.g., for substances used in the workplace, where skin contact is likely, or for pharmaceuticals applied to the skin. Other routes such as subcutaneous or intraperitoneal injection have been used when they are considered to be more appropriate for the anticipated route of exposure of humans (see paragraphs 136 and 137). In choice of route of administration of the test chemical due consideration should be given to animal welfare including choice of routes different from the predominant route of human exposure, if relevant.

116. Given the potential for oral exposure to a wide range of chemicals and also the practical experimental considerations associated with the long duration of chronic toxicity and carcinogenicity studies, the oral route is the route most commonly used in chronic toxicity and carcinogenicity studies. Route-to-route extrapolation may be considered for systemic effects when reliable data on ADME are available, rather than carrying out an additional study by a second route. For example, it may be possible to carry out an assessment of systemic effects via inhalation exposure based on the results of an oral chronic toxicity or carcinogenicity study (Gerrity and Henry, 1990). The use of route-to-route extrapolation should be decided on a case-by-case basis (Nielsen *et al.*, 2008) and is not, however, relevant for the assessment of local toxicity.

3.2.1 The oral route of exposure

117. Test substances may be administered via the diet or drinking water, by oral administration in capsules (in non-rodents) or by gavage, normally in a vehicle, depending on the physical and chemical characteristics of the test substance, its intended field of application and the predominant oral route of exposure of humans. Each method has advantages and disadvantages, and it should in particular be kept in mind that the toxicokinetics of the test substance may be affected by the method of oral administration. Data from previous short-term toxicological studies, including data on toxicokinetics, can provide information on potential local gastrointestinal effects and the extent of bioavailability of the test substance, in order to select the most appropriate route of oral administration and to demonstrate that systemic exposure is adequate (see also Section 3.4). As indicated in the Test Guidelines, a top dose not exceeding 1000 mg/kg body weight/day may apply except when human exposure indicates the need for a higher dose level to be used.

118. The animals are dosed with the test substance daily (seven days per week), normally for the entire duration of the study. Any other dosing regime, e.g., five days per week, needs to be justified. In the case of rodents, dosing of the animals should begin as soon as possible after weaning and acclimatisation and preferably before the animals are 8 weeks old.

3.2.1.1 Administration via the diet

119. Oral administration via the diet is the preferred route of administration if human exposure to the test substance is also likely to be via the diet. This route of administration may be appropriate if the objective is to establish an Acceptable Daily Intake (ADI) or Tolerable Daily Intake (TDI), for example for substances deliberately added to food, substances released from food contact materials or for environmental contaminants entering the food chain, and the pattern of exposure is continuous ingestion of small doses. However oral gavage studies may also be used to derive an ADI or a TDI.

120. When using oral administration via the diet, the test substance is administered in the diet either as a constant dietary concentration (mg/kg diet), or as a constant dose level in terms of the animal's body weight. In the latter case the dietary concentration must be adjusted regularly based on anticipated food consumption and body weight of the animals. While doses are expressed in terms of mg/kg diet, food consumption must be monitored on a cage basis at least weekly in order to be able to derive the intake of the test substance on mg/kg body weight per day or mg/m^2 per day. The food intake e.g., in the rat decreases from above 100 g per kg bw per day in early life (6-8 weeks, at the commencement of the study) to about little above 50 g per kg bw per day for older females (e.g., 6 months or more) and below 50 g per kg bw per day for older males. This will lead to a gradual decrease in intake of dietary administered test substance over age when keeping the dietary concentration of the test chemical constant. The concentration of the chemical in the feed should not normally exceed an upper limit of 5% of the total diet (FDA, 1982; Borzelleca, 1992), although higher levels are feasible (e.g., when testing carbohydrates or proteins) as long as the diet is adapted nutritionally adequately, e.g., the test substance is incorporated, at the expense of other components in a purified diet (Howlett *et al*., 2003).

121. Oral administration via the diet has the advantage that no handling of the animals is required. However, the palatability of the diet may be reduced at high dietary levels due to the taste or odour of the test substance, resulting in reduced food intake and thus reduced exposure to the test substance. This is likely to have been identified in previous shorter-tem studies and may require the introduction into the study design of an additional control group, pair fed (i.e., having matched food intake) in parallel with the high dietary level test group (see section 3.5.2 for further details). The substance should be stable during the preparation, storage and period of administration of the diet, for example it should not react chemically with dietary constituents, and analytical data must be provided to demonstrate this. It is also essential to ensure that the substance is mixed homogeneously in test diet at the desired level and, again, analytical data must be provided to demonstrate this, as required under Good Laboratory Practice (OECD, 1998).

3.2.1.2 Administration via drinking water

122. Oral administration in drinking water is the method of choice if human exposure to the test substance is likely to be via drinking water (e.g., drinking water contaminants) or in liquids (e.g., for substances that are volatile, or reactive with feed components, or any case where drinking water has an advantage over diet administration such as for soft drinks or beverages). The test substance is normally incorporated at a fixed concentration in the drinking water, at the approximate levels (in mg/ml water) required to provide the dose levels selected for the study (in mg/kg body weight per day), based on anticipated water consumption of the animals. Suspensions of test chemicals have been used where a solution in water has not provided a sufficiently high dose. Care should be taken to prevent precipitation of the test substance, with consequent accumulation in the drinking water valve, resulting in a disproportionately high dose being administered to the animal. Realistically, the test material should be soluble in water at all concentrations tested. While doses are expressed in terms of mg/ml water, water consumption must be monitored on a cage basis at least weekly in order to be able

to derive the intake of the test substance on mg/kg body weight per day. Concerning possible adjustment of the concentration of the test substance in the drinking water e.g., when this route of administration causes changes in water consumption due to palatibility of the drinking water, similar measures as described for dietary dosing are appropriate (see paragraphs 120 and 121) (Sharp and Regina, 1998; Wolfensohn and Lloyd, 1998; Pool, 1999; Nielsen *et al.*, 2008). The test substance should not markedly affect the palatability of the drinking water or cause marked changes in the pH, and its content and stability must be demonstrated analytically, as required under GLP (OECD, 1998).

3.2.1.3 Administration via gavage or encapsulation

123. Oral intubation (gavage) may be used if administration in the diet or drinking water is not possible, e.g., because of stability or palatability considerations. However, in the interests of animal welfare in particular, administration of the test substance by oral gavage should preferentially be restricted to those agents for which a bolus dose administration reasonably represents potential human exposure (e.g., administration of pharmaceuticals or food supplements orally at one or more doses per day) (Craig and Elliott, 1999; Brown *et al.*, 2000). Gavage dosing is experimentally more difficult than dietary administration, inducing stress in the animals which has toxicological implications (Brown *et al.*, 2000), and also requires daily handling of the animals, which may interfere with experimental parameters e.g., if neurobehavioural assessments are carried out during the study. A second control group to address this potential confounding factor is, however, rarely required and should be justified.

124. If the test substance is administered by gavage, this should be done using a stomach tube or a suitable intubation cannula, at similar times each day. The test substance may be administered in capsules, dissolved or suspended in a suitable vehicle. Administration by encapsulation rather than gavage dosing is a common route for dogs, but is not a preferred route for rodents, due to the associated technical difficulties. Vehicles of choice include aqueous solutions of thickeners such as methyl cellulose or carboxymethylcellulose, although other vehicles may be used (Gad *et al.*, 2006). Methylcellulose (0.5-1.0%) may have advantages over carboxymethylcellulose due to its superior wetting properties and polyethylene glycol (e.g., PEG 400) may also be used. Preferred vehicles are those that do not induce effects in their own right in long-term studies. Oil (e.g., corn oil) has been used when it is not possible to prepare aqueous solutions or homogeneous suspensions. When oil is used as a vehicle for gavage administration of the test substance dietary adjustment, for example a low fat diet should be considered, to compensate for the additional caloric intake. Induction of pancreatic adenomas in F-344 rats has been reported after administration of corn oil alone by gavage, with male F-344 rats being more sensitive than females (Boorman *et al.*, 1987; Haseman & Rao 1992). The maximum volume of solution that can be given by gavage in one dose depends on the size of the test animal (Diehl *et al.*, 2001; Gad *et al.*, 2006). For rodents, the volume ordinarily should not exceed 0.5-1 ml/100 g body weight, except in the case of aqueous solutions where 2 ml/100g body weight may be used (Diehl *et al.*, 2001). It should be noted, however that dosing volumes above 1 ml/100 g body weight may result in reflux of the dose.

125. Normally a single dose will be administered once daily, but where, for example, a substance is a local irritant or the pattern of human dosing is multiple doses per day, the daily dose may be administered as a split dose e.g., twice a day, within a 6 hour period. Variability in dose volume should be minimised by adjusting the concentration to ensure a constant volume at all dose levels. Potentially corrosive or irritant substances may however need to be diluted to avoid severe local effects, and testing at concentrations that are likely to be corrosive or irritant to the gastrointestinal tract should be avoided. The frequency and length of time for which the animals in a chronic toxicity or carcinogenicity study are dosed can lead to irritation in the oesophageal tissue and distress of the animals, potentially compromising the integrity of the study. If oral gavage is used, careful

observation should be conducted after dosing to watch for signs of distress such as laboured breathing, sudden lethargy, or poor mucous membrane colour.

3.2.2 The dermal route of exposure

126. The dermal route of exposure has been used in long-term carcinogenicity studies, primarily in the assessment of carcinogens such as polycyclic aromatic hydrocarbons in skin painting studies generally carried out in the mouse. Assessment of systemic toxicity or carcinogenicity using the dermal route is only appropriate if it has been demonstrated that the test substance is bioavailable via the skin, i.e., it crosses the skin barrier and the tested concentration of the substance shows either no or minimal irritation potential. Although the dermal route may be used in assessing the chronic toxicity and carcinogenicity of substances such as workplace chemicals, where skin contact is likely, or for pharmaceuticals applied to the skin, for which continuous dermal contact is anticipated, it is in practice a difficult route for long-term administration of a test substance, and the animal welfare implications involved in carrying out a 1- or 2-year dermal exposure study should be considered.

127. Bioavailability by the dermal route may be assessed initially via a dermal penetration study to determine extent of absorption through the skin *in vivo* or *in vitro* in accordance with OECD TG 427 or OECD TG 428 (OECD 2004a, 2004b) and following the guidance laid down in GD 28, the OECD Guidance Document for the conduct of skin absorption studies (OECD 2004c). This should be followed by a short term 21/28 day dermal toxicity study (TG 410, OECD, 1981a) or a subchronic dermal toxicity study (TG 411, OECD, 1981b) before conducting the longer term study. If the chemical does not penetrate the dermal layer, it is not appropriate to use this route for examination of chronic toxicity.

128. The method is based on the repeated application of the test substance, generally at a defined concentration in mg/ml in a suitable vehicle, to a clipped or shaved area of skin of approximately 10% of the total body surface area, to provide the desired dose in mg/kg body weight per day. Application is for at least 6 hours per day, 7 days per week, for a period of 24 months. Animals are normally housed separately in dermal studies to prevent grooming behaviors and oral ingestion of the test substance. TG 410 on Repeat Dose Dermal Toxicity: 21/28 day study (OECD, 1981a) or TG 411: Subchronic Dermal Toxicity: 90 day study (OECD, 1981b) should be consulted in the case of testing carried out by the dermal route,

129. The site may be occluded with polyethylene sheeting and gauze patches or semi-occluded, in order to prevent dislodgement of material and oral ingestion, which could affect the validity or usefulness of the study. With volatile or semi-volatile materials, application and covering procedures should minimise the possibility of evaporation. However, long-term dermal studies without occlusion are acceptable when justified scientifically, considering the potential difficulties of occlusion in long-term studies, including stress in the animals and the resource demands of repeated occlusion. In particular long-term dermal studies in the mouse without occlusion are recommended, considering the technical difficulties of occlusion when using this species.

3.2.3 The inhalation route of exposure

130. If it is likely that humans may be exposed by inhalation to a test substance, either as a gas, a vapour, or a liquid or solid aerosol (or a mixture thereof), the inhalation route should be used to evaluate the chronic toxicity or carcinogenicity of the substance in animals. The results of acute, subacute (28 days), subchronic (90 days) and range finding inhalation studies should be considered

when designing these studies and selecting concentration levels that will yield robust data regarding local and systemic toxicity. When testing substances which are irritants and/or corrosive, existing information should be used in selecting an appropriate dilution ratio for testing (see paragraph 133). This guidance does not extend to the testing of nanoparticles which can pose challenging physiological and methodological problems.

131. A chronic inhalation toxicity or carcinogenicity study should follow the principles described in Test Guideline 413 (subchronic inhalation toxicity: 90 day study, OECD, 2009a) in all respects except for the number of animals per group and study duration. Exposure by the inhalation route is normally carried out for 6 hours per day, 7 days per week, but exposure for 5 days per week may also be used, if justified. A rationale must be provided when using an exposure duration less than 6 hours per day. If rodent species smaller than rats, e.g., mice, are exposed nose-only, maximum exposure durations may be adjusted to minimise species-specific distress. The exposure restraining tubes must be adapted to the size of this species. Historical data, and/or published information in scientific literature may be used to demonstrate that the exposure technology chosen does not impose undue stress to the animals exposed. Endpoints suggestive of undue immobilization stress-related effects include marked changes body temperature, changes in ventilation, or decrease in body weight gains. These effects have been seen in animals exposed to clean air in nose-only exposure tubes for varying lengths (van Eijl *et al.*, 2006; Narciso *et al.*, 2003). There is the potential for stress experienced by animals in nose-only exposure studies to affect the results of the study by affecting immune competence (Bernet *et al.*, 1998; Jakab and Hemenway 1989) or xenobiotic metabolism (Fechter *et al.*, 2008), or by modulating gene expression in key organs (Thomson *et al.*, 2009; Ha *et al.*, 2003; Laconi *et al.*, 2000; Panuganti *et al.*, 2006; Sato *et al.*, 2006; Yin *et al.*, 2006), although some did not find confounding effects (Rothenberg *et al.*, 2000). The duration of a chronic toxicity study by the inhalation route will normally be 12 months and that for a carcinogenicity study will be 18-24 months, dependent on the species used (see also Section 3.3.3 of this guidance). Further general guidance on the performance of an inhalation toxicity study can be found in the Guidance Document 39 on Acute Inhalation Toxicity Testing (OECD, 2009b) and the OECD Guidance Document 125 on Histopathology for Inhalation Toxicity Studies (OECD, 2010). Although GD 39 is intended to provide guidance for acute inhalation studies, the technical aspects of exposing animals and generating and characterizing test atmospheres are similar for repeated exposures and single exposures and are therefore also applicable for chronic and carcinogenicity inhalation studies.

132. The nature of the test substance and the object of the test should be considered when selecting an inhalation chamber. For studies of liquid or solid aerosols and for vapor that may condense to form aerosols, the nose-only exposure method allows the avoidance of oral exposure due to grooming of particles deposited on the fur. However, the welfare implications of a 1- or 2-year nose-only exposure study, and the potential for physiological effects of stress experienced by the animals to affect the results of the study, can lead to preference for the use of the whole-body mode of exposure (Thomson *et al.*, 2009). Reasons for choice of exposure system should be justified in the study report. Particular attention should be paid to the technical problems that may arise from the large numbers of animals in whole body inhalation chambers (e.g., time required to attain inhalation chamber steady-state, heat and CO_2 production, and adsorption of test article on inhalation chamber walls and other surfaces). To ensure atmosphere stability when using a whole-body chamber, the total volume of the test animals should not exceed 5% of the chamber volume, and there should be a sustained dynamic airflow of at least 10 air changes per hour. Principles of the nose-only and whole body exposure techniques and their particular advantages and disadvantages are addressed in GD 39.

133. Test substances that are irritating or corrosive should always be tested using methodology laid out in Test Guideline TG 413 because it provides the study director or principal investigator with control over the selection of target concentrations. Corrosive or irritating test substances should be tested at concentrations that will yield the desired degree of toxicity without affecting longevity or undue stress to respiratory tract irritation (GD 39: OECD, 2009b). Any information available on the

corrosive or irritancy potential of the substance, including existing *in vivo* and *in vitro* data, pH values, and data from similar substances, should be considered in the design of a long-term study. Results from existing inhalation studies can be used to determine appropriate dose levels and should be reviewed carefully before any new range finding studies are undertaken. When exposing animals to corrosive or irritating substances, the targeted concentrations should be low enough to not cause marked pain and distress, yet sufficient to extend the concentration-response curve to levels that reach the regulatory and scientific objective of the test. These concentrations should be selected on a case-by-case basis, preferably based upon adequately designed range-finding studies that provide information regarding the critical location of irritation within the respiratory tract and endpoint for probing it. Adequately designed range-finding studies should demonstrate whether respiratory tract irritation depends on any irritation threshold (concentration-dependent) or on the total daily exposure intensity (concentration x time – dependent), and whether carry-over effects from one exposure day to another may lead to time-dependent exacerbations. Some irritant effects are instant in onset and others require time to accumulate. These factors need to be identified and may serve as justification for dose selection.

134. Species selection should be carefully considered for test substances causing upper respiratory tract irritation because numerous secondary species-specific physiological responses make the extrapolation from small rodents to humans more difficult (GD 39, OECD, 2009b). In depth justification for species-selection is necessary when using species other than rats for inhalation studies of irritant test substances.

135. For substances likely to accumulate in the lung over time due to poor solubility or other properties, the degree of lung-overload and delay in clearance needs to be estimated based on adequately designed pre-studies; ideally a 90-day study with postexposure periods long enough to encompass at least one elimination half-time. The use of concentrations exceeding an elimination half-time of approximately 1 year due to lung-overload at the end of study is discouraged.

3.2.4 Other routes of exposure

136. Other routes of exposure e.g., subcutaneous or intraperitoneal injection are generally only used in chronic toxicity or carcinogenicity studies when they mirror the anticipated route of administration in humans, such as in the case of pharmaceuticals. For example subcutaneous or intramuscular injections may be used for pharmaceuticals and for materials designed to be used as implants or prostheses. The subcutaneous and intraperitoneal routes have also been used in carcinogenicity bioassays for some solid-state, insoluble materials (e.g., fibres and plastics).

137. For substances administered parenterally, the dose volume used, stability of the formulation before and after administration, pH, viscosity, osmolality, buffering capacity, sterility and biocompatibility of the formulation are factors to consider (Diehl *et al.*, 2001). The smallest needle size should be used for administration, taking into account the dose volume, viscosity of injection material, speed of injection and species. (Diehl *et al.*, 2001). The use of parenteral injections in chronic toxicity and carcinogenicity studies is likely to result in local inflammation, and has significant animal welfare implications. Study investigators should document compelling reasons for using this method of administration.

REFERENCES

Bernet, F., J.F. Dedieu, C. Laborie, V. Montel and J.P. Dupouy (1998), "Circulating Neuropeptide Y (NPY) and Catecholamines in Rat under Resting and Stress Conditions. Arguments for Extra-adrenal Origin of NPY, Adrenal and Extra-adrenal Sources of Catecholamines", *Neurosci Lett,* 250(1):45-8.

Boorman, G.A., D.A. Banas, S.L. Eutis and J.K. Haseman (1987), "Proliferative Exocrine Pancreatic Lesions in Rats. The Effect of Sample Size on Incidence of Lesions", *Toxicol Pathol,* 15: 451-456.

Borzelleca, J.F. (1992), "Macronutrient Substitutes: Safety Evaluation", *Regulatory Toxicology and Pharmacology*, 16, 253-264.

Brown, A.P., N. Dinger and B.S. Levine (2000), "Stress Produced by Gavage Administration in the Rat", *Contemporary Topics in Laboratory Animal Science,* 39(1): 17-21.

Craig, M.A. and J.F. Elliott (1999), "Mice Fed Radiolabeled Protein by Gavage Show Sporadic Passage of Large Quantities of Intact Material into the Blood, an Artifact not Associated with Voluntary Feeding", *Contemporary Topics in Laboratory Animal Science,* 38(5): 18-23. ISSN: 1060-0558.

FDA (1982), *Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food*, US Food and Drug Administration, Washington, DC.

Derelanko, M.J. and M.A. Hollinger (Eds) (1995), *CRC Handbook of Toxicology*, CRC Press, Boca Raton, Florida.

Diehl, K.-H., R. Hull, D. Morton, R. Pfister, Y. Rabemampianina, D. Smith, J.-M. Vidal and C. van de Vorstenbosch (2001), "A Good Practice Guide to the Administration of Substances and Removal of Blood, Including Routes and Volumes", *Journal of Applied Toxicology*, 21:15-23. Available at: http://www.ff.up.pt/farmacologia/pdf/good_practice_lab_animals.pdf

Fechter, L.D., A. Nelson-Miller and C. Gearhart (2008), "Depletion of Liver Gluthathione Levels in Rats: a Potential Confound of Nose-only Inhalation", *Inhal Toxicol*, 20:885-890.

Gad, S.C., C.D. Cassidy, N. Aubert, B. Spainhour and H. Robbe (2006), "Nonclinical Vehicle Use in Studies by Multiple Routes in Multiple Species", *International Journal of Toxicology*, 25: 499-521.

Gerrity, T.R. and C.J. Henry (1990), *Principles of Route-to-route Extrapolation for Risk Assessment: Proceedings of the Workshops on Principles of Route-to-route Extrapolation for Risk Assessment*, Hilton Head, SC and Durham, NC. New York: Elsevier.

Ha, H.Y., K.S. Kim, Y.I. Yeom, J.K. Lee and P.L Han (2003), "Chronic Restraint Stress Massively Alters the Expression of Genes Important for Lipid Metabolism and Detoxification in Liver", *Toxicol Lett,* 146(1):49-63.

Haseman, J.K. and G.N. Rao (1992), "Effects of Corn Oil, Time-related Changes and Inter-laboratory Variability on Tumour Occurrence in Control Fischer 344 (F344/N) Rats", *Toxicol Pathol*, 20 (1): 52-60.

Hayes, A.W. (Ed) (1994), *Principles and Methods of Toxicology*, Third Edition, Raven Press, New York.

Howlett, J., D.G. Edwards, A. Cockburn, P. Hepburn, J. Kleiner, D. Knorr, G. Kozianowski, D. Müller, A. Peijnenburg, I. Perrin, M. Poulsen and R. Walker (2003), "The Safety Assessment of Novel Foods and Concepts to Determine their Safety in Use", *International Journal of Food Science and Nutrition*, 54 (Suppl.1), 1-32.

Jakab G.J. and D.R. Hemenway (1989), "Restraint of Animals Required for Nose-Only Inhalation Toxicologic Studies Suppresses Pulmonary Antibacterial Defenses", *Inhal Toxicol.* 1(4):289-300.

Laconi, E., C. Tomasi, F. Curreli, S. Diana, S. Laconi, G. Serra, M. Collu and P. Pani (2000), "Early Exposure to Restraint Stress Enhances Chemical Carcinogenesis in Rat Liver", *Cancer Lett,* 161(2):215-20.

Larson, J.L., D.C. Wolf and B.E. Butterworth (1994), "Induced Cytotoxicity and Cell Proliferation in the Hepatocarcinogenicity of Chloroform in Female B6C3F1 Mice: Comparison of Administration by Gavage in Corn Oil vs *ad libitum* in Drinking Water", *Fundam Appl Toxicol,* 22:90–102.

McClellan, R.O. and R.F. Henderson (Eds) (1989), *Concepts in Inhalation Toxicology*, Hemisphere Publishing Corp, New York.

Narciso, S.P., E. Nadziejko, L.C. Chen, T. Gordon and C. Nadziejko (2003), "Adaptation to Stress Induced by Restraining Rats and Mice in Nose-only Inhalation Holders", *Inhal Toxicol,* 15:1133-1143.

Nielsen, E., G. Østergaard and J.C. Larsen (2008), *Toxicological Risk Assessment of Chemicals*, Informa Healthcare, US, ISBN-13: 978-0-8493-7265-0.

OECD (1981a), *Repeated Dose Dermal Toxicity: 21/28-day Study*, OECD Guideline for the Testing of Chemicals, No. 410.

OECD (1981b), *Subchronic Dermal Toxicity: 90-day Study*, OECD Guideline for the Testing of Chemicals, No 411.

OECD (1998), *OECD Principles on Good Laboratory Practice,* Series on Principles of Good Laboratory Practice and Compliance Monitoring No 1, OECD, Paris.

OECD (2004a), *Skin absorption: in vivo method*, OECD Guideline for the Testing of Chemicals No 427, OECD, Paris.

OECD (2004b), *Skin absorption: in vitro method*, OECD Guideline for the Testing of Chemicals No 428, OECD, Paris.

OECD (2004c), *Guidance Document for the conduct of skin absorption studies,* Series on testing and assessment N°28, OECD, Paris.

OECD (2009a), *Subchronic Inhalation Toxicity: 90-day Study,* OECD Guideline for the Testing of Chemicals No 413, OECD, Paris.

OECD (2009b), *Guidance Document on Acute Inhalation Toxicity Testing,* Series on Testing and Assessment No 39, OECD, Paris.

OECD (2010), *Guidance Document on Histopathology for Inhalation Toxicity Studies, Supporting TG 41 (Subacute Inhalation Toxicity: 28-day Study) and TG 413 (Subchronic Inhalation Toxicity: 90-day Study)*, Series on testing and assessment No 125, OECD, Paris.

Panuganti, S.D., F.D. Khan and C.K. Svensson (2006), "Enhanced Xenobiotic-Induced Hepatotoxicity and Kupffer Cell Activation by Restraint-Induced Stress" *JPET*, 318:26–34.

Pool, T. (1999), *The UFAW Handbook of the Care and Management of Laboratory Animals,* Blackwell Science, Oxford UK.

Rothenberg, S., R.M. Parker, R.G. York, G.E. Dearlove, M.M. Martin, K.H. Denny, S.D. Lief, A.M. Hoverman and M.S Christian (2000), "Lack of Effects of Nose-only Inhalation Exposure on Testicular Toxicity in Male Rats", *Tox Sci*, 53:127-134.

Sato, T., H. Yamamoto, N. Sawada, K. Nashiki, M. Tsuji, K. Muto, H. Kume, H. Sasaki, H. Arai, T. Nikawa, Y. Taketani and E. Takeda (2006), "Restraint Stress Alters the Duodenal Expression of Genes Important for Lipid Metabolism in Rat", *Toxicology*, 227(3):248-61.

Sharp, P.E. and M.C.L. Regina (1998), *The Laboratory Rat,* Boca Raton FL USA: CRC Press.

Thomson, E.M., A. Williams, C.L. Yauk and R. Vincent (2009), "Impact of Nose-only Exposure System on Pulmonary Gene Expression", *Inhal Toxicol*, 21(S1): 74-82.

van Eijl, S., R. van Oorschot, B. Oliver, F.P. Nijkamp and N. Bloksma (2006), "Stress and Hypothermia in Mice in a Nose-only Cigarette Smoke Exposure System", *Inhal Toxicol,* 18:911-918.

Wolfensohn, S. and M. Lloyd (1998), *Handbook of Laboratory Animal Management and Welfare,* Blackwell Science, Oxford UK.

Yin, D., Y. Zhang, C. Stuart, J. Miao, Y. Zhang, C. Li, X. Zeng, G. Hanley, J. Moorman, Z. Yao and M. Woodruff (2006), "Chronic Restraint Stress Modulates Expression of Genes in Murine Spleen", *J Neuroimmunol,* 177(1-2):11-7.

3.3 CHOICE OF SPECIES AND STRAIN, NUMBERS AND SEX OF ANIMALS, STUDY DURATION, ALTERNATIVE *IN VIVO* MODELS

General issues

138. The choice of species to be used in a chronic toxicity or a carcinogenicity study is dictated by a number of factors, including the following:

- physiological and metabolic similarity to humans, in order to provide a valid model for extrapolation of the findings,
- familiarity with the species,
- availability of existing data on the species chosen,
- lifespan of the animals,
- ease of handling under experimental conditions,
- other issues such as cost of maintenance

139. Rodents (e.g., rats, mice or hamsters) have been used extensively, also dogs and primates. The choice of species must be justified and may be dictated by the purpose of the study (e.g., chronic toxicity or carcinogenicity) and by regulatory requirements. It should also be noted that mechanistic studies should be performed on the same species and strain as the cancer/chronic toxicity studies unless otherwise scientifically justified. With few exceptions, carcinogenicity and combined chronic toxicity/carcinogenicity studies are normally carried out in rodent species. Similarly, chronic toxicity studies are normally carried out in rodents. Chronic toxicity testing in non-rodents may however be required under certain regulatory regimes.

3.3.1 Testing in rodents

140. As indicated, rodent species have been used in the majority of chronic toxicity studies and in almost all carcinogenicity testing. The standard approach to carcinogenicity testing has been to use two species, the rat and the mouse, although in recent years a number of alternative approaches have been initiated to refine this approach in hazard identification and for risk assessment purposes, as discussed further in paragraphs 148 onwards. Advantages deriving from the use of rats and mice include the (relatively) low cost of maintenance, their short lifespan, meaning that a lifetime study can be completed in 2 – 3 years, and the availability of a large amount of historical data on age-related biochemical, haematological and pathological changes including on spontaneous tumours at specific organ sites.

141. Rodents have a number of metabolic pathways, physiological and pathological responses in common with humans. However, there are a number of instances where the chronic toxicity and carcinogenicity findings in rodents have been demonstrated not to be relevant to humans, because of toxicokinetic and toxicodynamic differences including species-specific pathways of metabolism, genetic differences, enzyme differences, differences in toxicologic pathways etc. If differences in toxicokinetics or toxicodynamics and/or other relevant parameters are suspected between the test species and humans that may have an impact on the relevance of the outcome of the study, these should be explored to determine if another test species may be more appropriate. Syrian golden hamsters have been used in some studies, for testing substances for which there is evidence that the toxicokinetics or toxicodynamics in humans are more similar to that in hamster than rats or mice. For example, hamsters have been used when considering a specific mode of action (PPAR-alpha, etc.).

Hamsters have also been used for studies of carcinogenesis in the respiratory and urinary tract, particularly for administration by parenteral routes such as intraperitoneal and intratracheal installation. It should be noted that a number of rodent tumour types are considered to be species-specific or strain-specific tumours with no relevance for humans (ECHA, 2009) (See Section 2.2).

142. It is important to consider the general sensitivity of the test animals, their background pathology and hence the responsiveness of particular organs and tissues to the chemicals under test when selecting rodent species, strains or stocks for toxicity studies. In general the selected rodent strain or stock should be well-characterized preferably including data on e.g., body and organ weight, haematological and biochemical parameters and background pathology. Additionally, it is important that test animals come from healthy colonies. Normally Specific Pathogen Free (SPF) animals are used, being SPF derived at birth and maintained under barrier conditions.

Rodent species and strain specificity

143. Assessment of chronic toxicity in rodents using the Test Guideline 452 is generally carried out in the rat, although other rodent species, e.g., the mouse, may be used. Since the normal lifespan of the rat is longer than that of the mouse, the potential for development of age-related background pathologies that may be influenced by the choice of a particular strain is less for rats than mice. In practice the strain commonly in use in the testing laboratory will be selected, since the laboratory will have historical data that will aid in the interpretation of any test substance-related change.

144. The Fischer 344 rat is a particularly well-characterized rat strain in carcinogenicity studies, since it has been the selected rat strain for the National Toxicology Programme (NTP) studies for over 20 years. However it has recently been reported (King-Herbert *et al.*, 2010) that the NTP is currently evaluating the Harlan Sprague Dawley (Hsd: Sprague Dawley SD) as the primary rat model for NTP studies, due to a number of health issues with the Fischer 344 rat and and decreased fecundity inherent in Wistar rats.

145. Importantly, in selecting a suitable rat strain for carcinogenicity testing, test animals should be selected that are likely to survive for the recommended duration of the study (see Section 3.3.2). Britton *et al.* (2004) reported that of the three rat strains studied (Harlan Hsd:Sprague-Dawley SD, Harlan Han Wistar Hsd:BrlHan:WIST, Charles River Crl:CD), Harlan Wistar strain survived in much greater numbers in 104-week carcinogenicity studies. The improved survival rate, according to the authors, appeared to be independent of body weight and food consumption and was reflected in the spontaneous pathology profile. Other authors believe this phenomenon to be attributable to a combination of obesity and genetic susceptibility and advocated dietary restriction as a method of extending survival in long-term carcinogenicity bioassays (Keenan, 1996). However, there is no scientific consensus on applying dietary restriction in such studies.

146. As discussed further in Section 3.5, and as reported by many investigators, dietary restriction results in a delay in age-related degenerative diseases such as nephropathy, which is commonly seen in all rat strains and has been shown to be diet-related. Dietary restriction may however result in a lower susceptibility of the animals to the development of tumours in carcinogenicity studies and to development of chemically-induced toxicity, thus presenting problems in extrapolation of the results of such studies to humans and, as indicated above, there is no scientific consensus on the application of dietary restriction.

147. Mouse strains used in carcinogenicity testing include the B6C3F1 mouse, as used by NTP, the ICR Swiss (CD-1), BALB/c, etc. The use of multiple strains of mice is being explored by NTP (King-Herbert *et al.*, 2010). The CD-1 mouse has been used by the US EPA OPP for chronic long-term toxicity studies. Notably, different mouse inbred strains show a variation in susceptibility to

tumourigenesis in different organs. The commonly used strains, in particular the B6C3F1 mouse used by NTP, carry hepatocellular tumour susceptibility loci that result in a high susceptibility to chemically induced hepatocarcinogenesis (Gariboldi *et al.*, 1993; Manenti *et al.*, 1994), which has limited their usefulness in carcinogenicity testing, while CD-1, an outbred mouse line derived from the Swiss strain has a relatively high incidence of spontaneous lung tumours and a high susceptibility to chemically induced lung tumourigenesis (Manenti *et al.*, 2003).

148. In recent years there has been considerable debate about the value of the two rodent species approach to carcinogenicity and about the continued use of the mouse as a second species, within the ICH (ICH, Proceedings of the Third International Conference, 1995) and in other fora (e.g., Huff and Haseman, 1991; Gold and Stone, 1993; Ennever *et al.*, 2003; Cohen, 2004, Billington *et al.*, 2010; Storer *et al.*, 2010). A number of studies have assessed the relative individual contribution of rat and mouse carcinogenicity studies and whether the use of rats or mice alone, or alternatively the reduced protocol using male rats and female mice would result in a significant loss of information on carcinogenicity relevant to human risk assessment. This debate has led to the suggestion that there may be no need for routine conduct of two long-term rodent carcinogenicity studies, since the use of the mouse in carcinogenicity testing may have limited utility (References above, also Griffiths *et al.*, 1994, Usui *et al.*, 1996, Carmichael *et al.*, 1997; Meyer, 2003, Doe *et al.*, 2006). However, testing in a second species is still acceptable, and is required under some current regulatory programmes.

149. Other experimental approaches to the evaluation of carcinogenic potential have been recommended, that may obviate the requirement to test in a second species (see also paragraphs 164 onwards). These approaches include short or medium-term *in vivo* rodent test systems providing insight into carcinogenic endpoints, such as models of initiation-promotion in rodents, or models of carcinogenesis using transgenic or neonatal rodents.

Numbers and sex of animals to be used in rodent studies

150. The Test Guidelines TG 451, TG 452 and TG 453 specify the core number of animals to be used in chronic toxicity and carcinogenicity studies. In a stand-alone chronic toxicity study, the core number indicated is normally at least 20 animals of each sex per group. Additional animals may also be included in the study design for interim kills during the study, also satellite animals for investigation of reversibility of any toxicological changes and sentinel animals for investigation of disease status. Smaller numbers of animals per sex and dose group are acceptable for these supplementary groups, as indicated in TG 452. It is unlikely that a regulatory authority would find a study using a lower core number of animals per sex and per group acceptable for regulatory purposes, unless a robust scientific justification is provided, since a sufficient number of animals should be used so that a thorough biological and statistical evaluation can be carried out. Furthermore, the Mutual Acceptance of Data (Council Decision C(81)30/FINAL – 12 May 1981, amended on 26 November 1997 - C(97)186/FINAL) would not apply to such a study, which may result in test duplication and many more animals being used. The key issues of importance in carrying out a statistical evaluation of the results of a chronic toxicity study are discussed further in Chapter 4 of this Guidance Document. In the case of a combined chronic toxicity and carcinogenicity study, there is provision for a smaller number of animals (at least 10 per sex and per group) to be used in the chronic toxicity phase, since interpretation of the data from the reduced number of animals per group in the chronic toxicity phase of this combined study will be supported by the data from the larger number of animals in the carcinogenicity phase of the study.

151. Similarly, TG 451 on the conduct of a carcinogenicity study specifies that at least 50 animals of each sex per dose group should be used, plus a concurrent control. Again, it is unlikely that a regulatory authority would find a study using a lower core number of animals per sex and per group acceptable for regulatory purposes, since a sufficient number of animals should be used so that a thorough biological and statistical evaluation can be carried out. It is however possible to increase

numbers of animals in all groups, in particular the lower dose groups, in order to increase the sensitivity of the study. In general use of additional numbers of animals above the 50 males and 50 females per group indicated in the TG for carcinogenicity testing (OECD TGs 451 and 453) has to be justified, considering e.g., animal strain, survival rate and statistical power. This is discussed further in Chapter 4. However, a number of publications have indicated that survivability problems exist for certain strains, notably the Sprague-Dawley rat (Nohynek *et al.*, 1993; Keenan, 1996). For strains with poor survival such as Sprague Dawley rats, higher numbers of animals per group may be needed in order to maximise the duration of treatment (typically at least 65/sex/group).

3.3.2 Testing in non-rodents, including considerations of numbers of animals to be used

152. The use of non-rodent species may be considered when available data suggest that they are more relevant for the prediction of health effects in humans. Non-rodents are used generally only for chronic toxicity testing and not for carcinogenicity testing. The principles and procedures outlined in this Guidance Document, together with those outlined in OECD TG 409, Repeated Dose 90-day Oral Toxicity Study in Non-Rodents (OECD, 1998) should be applied, with appropriate modifications. The use of non-rodent species should be, in the main, restricted to special purpose studies, rather than for basic screening of chronic toxicity. As indicated in the Test Guideline, a second, non-rodent species should only be used:

- where effects observed in other studies indicate a need for clarification/characterization in a second species, or
- where toxicokinetic/toxicodynamic studies indicate that the use of a specific non-rodent species is the most relevant choice of laboratory animal, or
- where other specific reasons justify the use of a non-rodent species.

153. The dog has been a commonly used non-rodent species in chronic toxicity studies in the past. There has been extensive debate about the need for, and added-value of, chronic toxicity studies in the dog compared with a 90 days subchronic toxicity dog study (Box and Spielmann, 2005, Doe *et al.*, 2006, ESAC, 2006, EFSA 2007; Kobel *et al.*, 2010). As a result of analyses carried out by these authors and also by the US EPA (Baetcke *et al.*, 2005), it has been suggested that tests of longer duration than 3 months using typical non-rodent species such as the dog do not have a substantial added value for making regulatory decisions. Some regulatory regimes have therefore now discontinued this requirement. For other regulatory sectors, such as small molecule pharmaceuticals, a large retrospective analysis confirmed the need for a 9 month study in the dog and provided examples for which even 6 months was not sufficient. For these reasons ICH (ICHM3R2, 2009) generally recommends 9-month studies in non rodents for chronic use pharmaceuticals.

154. Dogs used for chronic toxicity testing should be of a defined breed. Beagles are the most commonly used dog breed. The study design should minimise the numbers of animals used, and for a chronic toxicity study normally 4-6 animals per dose level are used. Dosing should begin preferably at four to six months and not later than nine months of age. Of particular importance when using dogs for toxicity testing, are considerations of appropriate housing, exercise, the need for environmental enrichment and for social contact. These aspects are discussed further in Section 3.5.

155. Other non-rodent species used include minipigs, as their basic physiology is considered to be very similar to humans, and they may therefore provide a better model than e.g., dogs or rodents. Rabbits, although used in the area of skin and eye irritation testing and reproductive toxicity testing, are rarely if ever used as a second species for chronic toxicity and carcinogenicity testing, and their use is therefore not discussed further in this Guidance Document.

156. Minipigs used for chronic toxicity testing should be of a defined breed. Gøttingen Minipigs are the most commonly used minipig strain. The study design should minimise the number of animals used; for a chronic toxicity study, normally 4-6 animals per dose level are used. Dosing should begin preferably at three to four months of age. Where the study is conducted as a preliminary to a long-term chronic toxicity study, the same species/breed should be used in both studies. Animal welfare considerations are of the utmost importance when using minipigs for toxicity testing, including housing, exercise, the need for environmental enrichment and for social contact. These aspects are discussed further in Section 3.5.

3.3.3 Study duration

157. The duration of the chronic toxicity study and of the chronic toxicity phase in the combined chronic toxicity/carcinogenicity study is normally 12 months, although longer or shorter periods may be used if scientifically justified, and for pharmaceuticals, chronic studies of 6 months duration in rats are required.

158. In the carcinogenicity study, mice are generally exposed to the test chemical for 18–24 months and rats for 24–30 months with exposure being longer for strains of greater longevity or with a lower spontaneous tumour rate. However, exposure for longer than 24 months is unusual and should be justified. TG 451 specifies that the duration of the study will normally be 24 months for rodents, representing the majority of the normal life span of the animals to be used. Shorter or longer study durations may be used, dependent on the lifespan of the strain of the animal species in the study, but should be justified. For specific strains of mice, e.g., AKR/J, C3H/J, CD-1 or C57BL/6J strains, for which documentation exists showing that a duration of 18 months may be more appropriate (e.g., Giknis and Clifford, 2010), a reference to this information is sufficient for the justification of using a duration shorter than 24 months. Many carcinogenicity studies in mice are conducted for 18 months; therefore, there is rather limited historical control data available at 24 months. The study may also make provision for interim kills, e.g., at 12 months, to provide information on progression of neoplastic changes and mechanistic information, if scientifically justified. Where such information is already available from previous repeat dose toxicity studies on the substance, interim kills may not be scientifically justified.

159. Termination of the study should be considered when the number of survivors in the lower dose groups or the control group falls below 25 per cent, considering the survival of each sex separately. The US EPA Health Effects Test Guidelines 870.4200 (US EPA, 1998b) specify that survival in any group should not fall below 50% at 15 months in the case of mice and 18 months in the case of rats, or below 25% at 18 and 24 months respectively. In addition, the WHO (1990) has recognised a further type of carcinogenicity study that continues until mortality in the most susceptible group reaches a fixed level, usually 80%. In addition the OECD GD 19 gives guidance on when to kill the animals in a study, based upon the recognition, assessment and use of clinical signs as humane endpoints (OECD 2000).

160. The study should not normally be extended beyond the point when the data available from the study are no longer sufficient to enable a statistically valid evaluation to be made. However, in the case where only the high dose group dies prematurely for obvious reasons of toxicity, this should not trigger termination. While the validity of the study may be prejudiced by early mortality, e.g., in the high dose group, valuable information will still be obtained from it, and a decision to terminate the study in its entirety must be carefully weighed against the animal welfare implications of having to repeat the study. The lower dose groups, in particular, the next highest dose level, when continued to the scheduled end of the study may still be useful for the evaluation.

161. If the current dosing regime results in severe animal toxicity and the study must be terminated before the full duration of exposure, the study sponsor needs to contact the regulatory authority immediately. All data should be compiled and all available tissues preserved for further evaluation. While this study may not meet all test guideline requirements for chronic/carcinogenicity testing, the results may be useful and considered in the overall risk assessment. The determination of a retest will be made on a case-by-case basis by the regulatory authorities.

Consideration of the acceptability of a negative carcinogenicity result relative to survival in the study.

162. For a negative result to be acceptable in a rat carcinogenicity bioassay, survival in the study should ideally be no less than 50% in all groups at 24 months, while for "life span studies", studies continued to end of life/ death of the animals survival at study termination should not be less than 25%. In a mouse study, survival in all groups in the study should be no less than 50% at 18 months. It is the responsibility of the study director to use rodent strains that would ensure adequate survival at 18/24 months. Additionally, no more than 10% of any group should be lost due to autolysis, cannibalism, or management problems.

163. Acceptability of a negative result in a carcinogenicity study based on survival rates will vary depending on the study design. For example, a survival rate of 50% may not be appropriate for a BMD design with appreciably more dose groups than the more conventional design. A flexible approach is necessary to accommodate different situations, most importantly the distinction between confidence in a negative result for risk characterization (most often required for risk assessments) and obtaining a quantitative dose-response value e.g., a BMD (most often required for risk/benefit analysis). Survival of less than 50% of animals in the top dose group need not disqualify the evaluation of a negative study outcome, provided that the higher mortality in this group can be clearly attributed to another toxic effect, such as chronic undernutrition or malabsorption resulting from gastrointestinal irritation by too high a dietary concentration of the test substance and no trend/drift/change in tumour incidence is observed. Evaluation of a negative study outcome may be based on calculation of the power of the test for groups with lower mortality.

3.3.4 Alternative in vivo *models for carcinogenicity testing, including testing in transgenic animals*

164. Some of the medium-term tests for carcinogenicity involve the development of proliferative lesions in a single tissue, e.g., foci of alteration in the liver (Williams *et al.*, 1982; Goldsworthy *et al.*, 1986; Ito *et al.*, 1989). Others use tumour end-points, such as induction of lung adenomas in the A-strain mouse (Maronpot *et al.*, 1986) or induction of tumours in initiation–promotion studies using various organs, including the skin, bladder, intestine, liver, lung, mammary gland and thyroid (see reviews by Enzmann *et al.*, 1998a & 1998b; IARC, 1992 & 1999). A further category of study is the "start/stop" protocol. Here, an agent is administered for a limited period to induce particular effects or lesions; the progression or reversibility of these is then observed in the absence of further treatment (Todd, 1986; Marsman & Popp, 1994).

165. Transgenic assays in genetically engineered rodents have also been developed following the identification of genes, such as proto-oncogenes and tumour-suppressor genes that are highly conserved across species and associated with a wide variety of human and animal cancers. The genetically engineered rodent designs involve activated oncogenes that are introduced (transgenic) or tumour suppressor genes that are deleted (knocked out). If appropriate genes are selected, these assay systems may provide supplementary information on mechanisms of tumour formation or serve as selective tests for carcinogens. The modified transgene is expected to accelerate carcinogen-induced cancer development without interfering with other relevant genetic and/or epigenetic steps. High

spontaneous tumour incidence in control animals is a major confounding factor of the conventional bioassay; the presence of the transgene itself does not induce high spontaneous tumour incidence in the short time span of the assay. These assays have been extensively reviewed in publications, including a single-theme issue of Toxicological Pathology (26 (4), 1998) and others (Tennant *et al.*, 1995; Contrera & DeGeorge, 1998; Eastin, 1998; Bucher, 1998; Eastin &Tennant, 1998; Santos *et al.*, 2008).

166. Transgenic mouse models are still under investigation for their utility in carcinogenicity testing, and have not yet been fully validated. Several of these models are, however, generally accepted as an alternative for 2-year mouse studies for pharmaceuticals. Transgenic mouse models may be useful as hazard identification / characterization screening models as part of an initial phase of the risk assessment process. However, they are not definitive proof of potential human carcinogenicity, and they are not proof of a specific mechanism of action. Like the 2-year bioassay, the results from tests in these models need to be incorporated into an overall integrated, weight of evidence evaluation for a given compound that takes into account genotoxicity, particularly DNA reactivity, structure activity relationships, results from other bioassays, and the results of other investigations including toxicokinetics, metabolism, and mechanistic information (ICH, 1997; Meyer, 2003; NAS, 2007; EFSA, 2009).

REFERENCES

Baetcke, K.P., W. Phang and V. Dellarco (2005), *A Comparison of the Results of Studies on Pesticides from 12- or 24-Month Dog Studies with Dog Studies of Shorter Duration,* US EPA Office of Pesticide Programs, Washington, DC.

Billington, R., R. Lewis, J. Mehta, I. Dewhurst (2010), "The Mouse Carcinogenicity Study is no Longer a Scientifically Justifiable Core Data Requirement for the Safety Assessment of Pesticides", *Critical Reviews in Toxicology,* Vol 40, 35-49.

Box, R.J. and H. Spielmann (2005), "Use of the Dog as Non-rodent Test Species in the Safety Testing Schedule Associated with the Registration of Crop and Plant Protection Products (Pesticides): Present Status", *Archives of Toxicology,* 79, 615-626.

Britton, L.D., N. Downes, P. Mullins and D. Mitchell (2004), "The Most Appropriate Strain for Rat Carcinogenicity Bioassay", *Toxicol Sci,* 78:560.

Bucher, J.R. (1998), "Update on National Toxicology Program (NTP) Assays with Genetically Altered or "Transgenic" Mice", *Environ Health Perspect,* 106: 619–621.

Carmichael, N.G., H. Enzmann, I. Pate and F. Waechter (1997), "The Significance of Mouse Liver Tumor Formation for Carcinogenic Risk Assessment: Results and Conclusions from a Survey of Ten Years of Testing by the Agrochemical Industry", *Environ Health Perspect*, 105:1196-1203.

Cohen, S.M. (2004), "Human Carcinogenic Risk Evaluation: an Alternative Approach to the Two-year Rodent Bioassay", *Toxicological Sciences*, 80:225-9.

Contrera, J.F. and J.J. DeGeorge (1998), "*In vivo* Transgenic Bioassays and Assessment of the Carcinogenic Potential of Pharmaceuticals", *Environ Health Perspect,* 106 (Sup. 1): 71–80.

Doe, J.E., A.R. Boobis, A. Blacker, V. Dellarco, N.G. Doerrer, C. Franklin, J.I. Goodman, J.M. Kronenberg, *et al.* (2006), "A Tiered Approach to Systemic Toxicity Testing for Agricultural Chemical Safety Assessment", *Critical Reviews in Toxicology,* 36, 37-68.

Eastin, W.C. and R.W. Tennant (1998), *Carcinogenicity: Testing Predicting & Interpreting Chemical Effects*, In: Kitchin K.T. (Ed), pp. 395–410, Marcel Dekker Inc., New York.

Eastin, W.C. (1998), "The U.S. National Toxicology Program Evaluation of Transgenic Mice as Predictive Models for Identifying Carcinogens", *Environ Health Perspect,* 106(Sup. 1): 81–84.

ECHA (2009), *Guidance to Regulation (EC) No 1272/2008 on Classification, Labelling and Packaging (CLP) of Substances and Mixtures*", Chapter 3.6, "Carcinogenicity", European Chemicals Agency.

EFSA (2007), *Opinion of the Scientific Panel on Plant Protection Products and their Residues on a request from the Commission related to the revision of Annexes II and III to Council Directive 91/414/EEC concerning the placing of plant protection products on the market - Toxicological and metabolism studies,* Question N° EFSA-Q-2006-118, EFSA Journal 449, 1-60.

EFSA (2009), *Existing Approaches Incorporating Replacement, Reduction and Refinement of Animal Testing: Applicability in Food and Feed Risk Assessment,* Question N° EFSA-Q-2005-231, EFSA Journal 1052, 1-77.

Ennever, F.K. and L.B. Lave (2003), "Implications of the Lack of Accuracy of the Lifetime Rodent Bioassay for Predicting Human Carcinogenicity", *Reg Tox Pharm,* 38, 52-7.

Enzmann, H., E. Bomhard, M. Iatropoulos, H.J. Ahr, G. Schlueter and G.M. Williams (1998a), "Short- and Intermediate-term Carcinogenicity Testing – a Review. Part 1: The Prototypes Mouse Skin Tumour Assay and Rat Liver Focus Assay", *Food Chem Toxicol,* 36: 979–995.

Enzmann, H., M. Iatropoulos, K.D. Brunnemann, E. Bomhard, H.J. Ahr, G. Schlueter and G.M. Williams (1998b), "Short- and Intermediate-term Carcinogenicity Testing – A Review. Part 2: Available Experimental Models", *Food Chem Toxicol,* 36: 997–1013.

ESAC [ECVAM Scientific Advisory Committee] (2006), *Statement on Dog Toxicity Studies.* http://ecvam.jrc.it/publication/ESAC25_statement_DOG_20061207.pdf

Gariboldi, M., G. Manenti, F. Canzian, F.S. Falvella, M.A. Pierotti, G. Della Porta, G. Binelli and T.A. Dragani (1993), "Chromosome Mapping of Murine Susceptibility Loci to Liver Carcinogenesis", *Cancer Res.*, 53, 209–211.

Giknis, M.L.A. and C.B. Clifford (2010), *Spontaneous Neoplastic Lesions in the Crl:CD-1 (ICR) Mouse in Control Groups from 18-Month to 2 Year Studies*, Charles River Laboratories, Wilmington, MA.

Griffiths, S.A., C. Parkinson, J.A.N. McAuslane and C.E. Lumley (1994), "The Utility of the Second Rodent Species in the Carcinogenicity Testing of Pharmaceuticals", *The Toxicologist,* 14(1):214.

Gold, L.S., T.H. Slone (1993), "Prediction of Carcinogenicity from Two Versus Four Sex-species Groups in the Carcinogenic Potency Database", *Journal of Toxicology and Environmental Health,* 39:143-57.

Goldsworthy, T.L., M.H. Hanigan and H.C. Pitot (1986), "Models of Hepatocarcinogenesis in the Rat-contrasts and Comparisons", *Crit Rev Toxicol,* 17: 61–89.

Huff, J. and J. Haseman (1991), "Long-term Chemical Carcinogenesis Experiments for Identifying Potential Human Cancer Hazards: Collective Database of the National Cancer Institute and the National Toxicology Program (1976-1991)", *Environmental Health Perspectives*, 96:23-31.

IARC (1992), *Mechanisms of Carcinogenesis in Risk Identification*, Vainio, H., P. Magee, D. McGregor and A.J. McMichael Eds, IARC Scientific Publications 116, Lyon.

IARC (1999), *The Use of Short- and Medium-term Tests for Carcinogens & Data on Genetic Effects in Carcinogenic Hazard Evaluation*, McGregor, D.B., J.M. Rice and S. Vennitt Eds, IARC Scientific Publications 146, Lyon.

ICH (International Conference on Harmonisation) (1997), *ICH Guideline - S1B Testing for Carcinogenicity of Pharmaceuticals.*

ICH (International Conference on Harmonisation) (2009), *ICH Guideline - M3R2 Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals.*

Ito, N., K. Imaida, R. Hasegawa and H. Tsuda (1989), "Rapid Bioassay Methods for Carcinogens and Modifiers of Hepatocarcinogenesis", *Crit Rev Toxicol,* 19: 385–415.

Keenan, K.P. (1996), "Commentary: the Uncontrolled Variable in Risk Assessment: ad Libitum Overfed Rodents-Fat, Facts and Fiction", *Toxicol Pathol,* 24(3): 376 – 383.

King-Herbert, A.P., R.C. Sills and J.R. Bucher (2010), "Commentary: Update on Animal Models for NTP Studies", *Toxicol Pathol*, 38, 180-181.

Kobel, W., I. Fegert, R.L. Billington, K. Bentley, W. Bomann, P. Botham, B. Stahl, B. van Ravenzwaay and H. Spielmann (2010), "A 1-year Toxicity Study in Dogs is no Longer a Scientifically Justifiable Core Data Requirement for the Safety Assessment of Pesticides", *Critical Reviews in Toxicology*, 2010; 40(1): 1–15.

Manenti, G., G. Binelli, M. Gariboldi, F. Canzian, L. De Gregorio, F.S. Falvella, T.A. Dragani and M.A. Pierotti (1994), "Multiple Loci Affect Genetic Predisposition to Hepatocarcinogenesis in Mice", *Genomics*, 23, 118–124.

Manenti, G., F. Galbiati, S. Noci and T.A. Dragani (2003), "Outbred CD-1 Mice Carry the Susceptibility Allele at the Pulmonary Adenoma Susceptibility 1 (Pas1) Locus" *Carcinogenesis*, 24(6):1143-8.

Maronpot, R.R., M.B. Shimkin, H.P. Witschi, L.H. Smith and J.M. Cline (1986), "Strain A Mouse Pulmonary Tumor Test Results for Chemicals Previously Tested in the National Cancer Institute Carcinogenicity Tests", *J Natl Cancer Inst*, 76: 1101–1112.

Marsman, D.S., and J.A. Popp (1994), "Biological Potential of Basophilic Hepatocellular Foci and Hepatic Adenoma Induced by the Peroxisome Proliferator, Wy-14,643" *Carcinogenesis*, 15: 111–117.

Meyer, O. (2003) "Testing and Assessment Strategies, Including Alternative and New Approaches", *Toxicology Letters*, 140-141:21-30.

NAS (2007), *Toxicity Testing in the 21st century: A vision and a Strategy, National Research Council of the National Academies*, The National Academies Press, Washington DC, ISBN 13: 978-0-309-10992.

Nohynek, G.J., L. Longeart, B. Geffray, J.P. Provost and A. Lodola (1993), "Fat, Frail and Dying Young: Survival, Body Weight and Pathology of the Charles River Sprague-Dawley-Derived Rat Prior to and Since the Introduction of the VAFR Variant in 1988", *Human and Experimental Toxicology*, 12(2): 87-98.

OECD (1998), *Repeated Dose 90-day Oral Toxicity Study in Non-Rodents*, OECD Guideline for the Testing of Chemicals No 409, OECD, Paris.

OECD (2000), *Guidance Document on the Recognition, Assessment, and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation*, Series on Testing and Assessment No. 19, OECD, Paris.

OECD (2009), *Detailed Review Paper on Transgenic Rodent Mutation Assays*, Series on Testing and Assessment, No. 103, OECD, Paris.

REACH (2006), *Guidance on Information Requirements and Chemical Safety Assessment, Part B: Hazard Assessment*, REACH Regulation (EC) No. 1907/2006. Available at: http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_part_b_en.pdf?vers=20_10_08 .

Santos, J., P. Fernández-Navarro, M. Villa-Morales, L. González-Sánchez, J. Fernández-Piqueras (2008), "Genetically Modified Mouse Models in Cancer Studies", *Clinical and Translational Oncology*, Volume 10, issue 12, p. 794 - 803.

SCHER (2009), *The Need for Non-human Primates (NHP) in Biomedical Research, Production and Testing of Products and Devices*, Opinion of the Scientific Committee on Health and Environmental Risks.

Storer, R.D., F.D. Sistare, M.V. Reddy and J.J. DeGeorge (2010), "An Industry Perspective on the Utility of Short-Term Carcinogenicity Testing in Transgenic Mice in Pharmaceutical Development", *Toxicol Pathol,* 38: pp. 51-61.

Todd, G.C. (1986), "Induction and Reversibility of Thyroid Proliferative Changes in Rats Given an Antithyroid Compound", *Vet Pathol* 23: 110–117.

US EPA (1998c) Health Effects Test Guidelines OPPTS 870.4300 Combined Chronic Toxicity/Carcinogenicity. US Environmental Protection Agency, EPA 712-C-98-212, August 1998.

Usui, T., S.A. Griffiths and C.E.Lumley (1996), *The Utility of the Mouse for the Assessment of the Carcinogenic Potential of Pharmaceuticals*, In D'Arcy P.O.F. and D.W.G. Harron (eds), Proceedings of the Third International Conference on Harmonisation, Queen's University Press, Belfast. pp 279-284.

WHO (1990), *Principles for the Toxicological Assessment of Pesticide Residues in Food*, IPCS Environmental Health Criteria 104, World Health Organisation, Geneva.

Williams, G.M., M.F. Laspia and V.C. Dunkel (1982), "Reliability of the Hepatocyte Primary Culture/DNA Repair Test in Testing of Coded Carcinogens and Non-carcinogens", *Mutat Res,* 97: 359–370.

3.4 TOXICOKINETICS

167. Studies examining the toxicokinetics (TK) of a chemical substance are conducted to obtain adequate information on its absorption, distribution, biotransformation (i.e., metabolism) and excretion, to aid in relating concentration or dose to the observed toxicity, and to aid in understanding its mechanism of toxicity (OECD, 2010). Basic TK parameters determined from these studies will also provide information on the potential for accumulation of the test substance in tissues and/or organs and the potential for induction of biotransformation as a result of exposure to the test substance (OECD, 2010). Toxicokinetic studies may provide useful information for determining dose levels for toxicity studies (linear vs. non-linear kinetics), route of administration effects, bioavailability including differences in single versus repeat dose, internal dose, metabolism pathways (including generation of reactive intermediates) and the existence of saturation points of uptake, metabolism and/or excretion.

168. The specific objectives of a toxicokinetic study, as an adjunct to a chronic toxicity or carcinogenicity study, include the following (ICH, 1994):

- to describe the systemic exposure achieved in animals and its relationship to dose level and the time course of the toxicity study.
- to relate the exposure achieved in toxicity studies to toxicological findings and to contribute to the assessment of the relevance of these findings for other species i.e., humans/extrapolation.
- to provide information which, in conjunction with the toxicity findings, contributes to the design of subsequent toxicity studies including studies on MOA.

169. For the purpose of dose selection, TK studies are informative in indicating whether there is a "point of saturation" or saturation kinetics evident in the dose response curve (see also Section 3.1 on Dose selection).

170. The kinetics of absorption will determine the internal exposure achieved. The absorption and clearance of the compound and its metabolites will determine the systemic and target organ exposure resulting from a single dose and can be used to design the treatment regimen required to achieve a desired internal dose for either parent compound or major human metabolites. The effect of repeated exposures on absorption, metabolism, biotransformation, and clearance of a compound will provide information on the internal dose achieved during chronic exposure under conditions of the bioassay.

171. The bioavailability of test substance is often very dependent on the matrix it is administered in, e.g., due to the fat content. If this is the feed, there may be an interaction of the test substance with food matrix. The food composition may alter bioaccessibility. Therefore, exposure via food may provide different toxicokinetics compared to exposure via drinking water. In the few cases where administration in a chronic study is via gavage, it is important to realise that the composition of the gavage administration may influence bioavailability as well. These aspects should be considered when the results are used for risk assessment purposes.

172. As indicated in OECD TG 417 on Toxicokinetics (OECD, 2010), there are numerous studies that might be performed to evaluate the TK behaviour of a chemical for regulatory purposes. However, depending on particular regulatory needs or situations, not all of these possible studies may be necessary for the evaluation of a chemical. Flexibility, taking into consideration the characteristics of the substance being investigated, is needed in the design of toxicokinetic studies. In some cases, only a certain set of questions may need to be explored in order to address chemical-associated hazard and risk concerns. In some situations, TK data can be collected as part of the evaluation in other

toxicology studies. For other situations, additional and/or more extensive TK studies may be necessary, depending on regulatory needs and/or if new questions arise as part of chemical evaluation (see also Barton *et al.*, 2006).

173. In order to be of maximum utility in planning the design of a chronic toxicity or carcinogenicity study, particularly in the selection of dose levels, TK studies should be carried out, or data should be available, in the same species used in the long-term study and should preferably be performed using the same route and, where appropriate, the same vehicle as that used in the other toxicity studies. It should be noted however that such data may not be readily available for all chemicals, as they are not required under all regulatory schemes.

174. While single dose TK studies may provide useful information on absorption, distribution, metabolism and excretion of the test substance, the information most relevant in the planning and the execution of a chronic toxicity or carcinogenicity study will come from a repeat-dose toxicokinetic study over an extended period. As noted in OECD TG 417, repeated administration of the test substance may be needed to address more fully the potential for accumulation and/or persistence or changes in TK, or as required by a competent authority.

175. In addition to data from dedicated toxicokinetic studies such as OECD TG 417, useful information on repeat-dose toxicokinetics may be generated as part of a chronic toxicity (TG 452) or carcinogenicity (TG 451) study, or the combined chronic toxicity/carcinogenicity study (TG 453). In many cases, it is not necessary to add additional satellite animals into the study design for the purpose of providing excreta and blood samples for toxicokinetic analysis, since a minimum required number of blood samples to calculate representative toxicokinetics can be obtained from the study animals without compromising the outcome of the toxicity study (Saghir *et al.*, 2006). In the case of investigations carried out using the main study animals, the volume and number of blood samples which can be obtained per animal may, however, be limited by the stress imposed on the animals and the potential effects of repeated sampling on animal health and/or physiology. For more extensive TK investigations satellite groups or a separate TK study will be necessary. In the case of investigations using satellite groups, the quantity of test substance and in some cases metabolites excreted in urine, feces, and expired air should be measured on at least two time points on day 1 of collection (one of which should be at 24 hours post-dose), one at 48 hours, one at 7 days, one at 3 months, one at 12 months and one at termination. Blood samples should be taken from the satellite animals (and also in the case of an independent TK study) at suitable time points. Comparison of the area-under-curve (AUC) on Day 1 and the last day is used to assess issues such as accumulation and induction or inhibition of biotransformation, affecting AUC.

176. Guidance on toxicokinetic investigations following administration of test substance by the oral route(s) is given in the OECD TG 417. Information how to assess absorption following administration of test substance by the dermal route is given in the OECD TG 427 (*in vitro*) and TG428 (*in vivo*).

177. With respect to plasma levels of the test chemical measured in toxicity studies, an important point to note is that in rats there is a marked influence of sex hormones on liver biotransformation processes (see e.g., Chhabra & Fouts, 1974). In general, male rats metabolise xenobiotics (as well as endogenous substrates) faster than females, a finding not generally seen in other species. Thus rat studies may exhibit sex differences in plasma kinetics and in clinical and toxicological effects of the test chemical. Due to these differences in CYP3A expression pattern between males and females in rat and other species, these findings may not be relevant to human exposure.

178. As indicated in OECD TG 417 on Toxicokinetics (OECD, 2010), all available information on the test substance and relevant metabolites and analogs should be considered by the testing laboratory prior to conducting an additional toxicokinetic study in order to enhance study quality and

minimise animal usage. This could include data from other relevant test methods (*in vivo* studies, *in vitro* studies, and/or *in silico* evaluations). Physicochemical properties, such as octanol-water partition coefficient (expressed as log P_{OW}), pKa, water solubility, vapour pressure, and molecular weight of a chemical may be useful for study planning and interpretation of results. They can be determined using appropriate methods as described in the relevant OECD Test Guidelines.

179. The revised TG 417 also provides guidance on use of supplemental approaches in addition to the *in vivo* studies described in the preceding paragraphs that can provide useful information on absorption, distribution, metabolism and excretion (OECD, 2010). For example, use of freshly isolated or cultured hepatocytes and subcellular fractions (e.g., microsomes and cytosol or S9 fraction) from liver can provide useful information on metabolism of the test substance. Local metabolism in the target organ, e.g., lung, may be of interest for risk assessment. For these purposes, microsomal fractions of target tissues may be useful. Studies with microsomes may be useful to address potential gender and life-stage differences and characterize metabolic rates (Km and Vmax) which can aid in the assessment of dose dependency of metabolism in relation to exposure levels. In addition microsomes may be useful to identify the specific microsomal enzymes involved in the metabolism of the substance which can be relevant in species extrapolation.

180. In certain circumstances and under appropriate conditions, subcellular fractions coming from human tissues might be considered for use in determining potential species differences in biotransformation. Primary cell cultures from liver cells and fresh tissue slices may be used to address similar questions as with liver microsomes. In certain cases, it may be possible to answer specific questions using cell lines with defined expression of the relevant enzyme or engineered cell lines. It may also be useful to study the inhibition and induction of specific cytochrome P450 isozymes (e.g., CYP1A2, 2A1, and others) and/or phase II enzymes by the parent compound using *in vitro* studies. Information obtained may have utility for similarly structured compounds (OECD, 2010). The potential for induction of biotransformation can be examined by using liver subcellular fractions (e.g., microsomes and cytosol) of animals pretreated with the substance of interest, *in vitro* via hepatocyte induction studies or from specific cell lines expressing relevant enzymes (OECD, 2010).

181. The results from *in vitro* investigations may also have utility in the development of PBTK (physiologically-based toxicokinetic) models (Loizou *et al.*, 2008). *In vitro* dermal absorption studies may provide supplemental information to characterize absorption (OECD, 2004b, 2004c).

182. Toxicokinetic models such as PBTK modelling may have utility for various aspects of hazard and risk assessment as for example in the prediction of systemic exposure and internal tissue dose. A PBTK model comprises an independent structural mathematical model, comprising the tissues and organs of the body with each perfused by, and connected via, the blood circulatory system. PBTK modelling may be used to predict the target tissue dose of the parent chemical or its reactive metabolite. Information derived from PBTK modelling experiments may aid in the comparison of biotransformation and toxicokinetics of a test substance and/or its metabolites and may provide a basis for extrapolation across species or dosing patterns. Such experiments may also provide estimates of relevant internal tissue dose which might be important to the hazard or risk assessment process (Andersen, 2003; US EPA 2006; Nielsen *et al.*, 2008, Clewell and Clewell, 2008). Furthermore, specific questions on mode of action (see Chapter 2) may be addressed.

183. Data useful for developing PBTK models for a chemical in any given species include 1) partition coefficients, 2) metabolic rate, 3) route-specific absorption parameters and 4) *in vivo* kinetic data for model evaluation (e.g., clearance parameters for relevant (> 10 %) excretion pathways, Km and Vmax for metabolism) (OECD, 2010). The experimental data used in model development should be generated with scientifically sound methods. The model predictions should be evaluated using experimental data as much as possible (IPCS, 2010). Chemical- and species-specific parameters such

as absorption rates, blood-tissue partitioning and metabolic rate constants are often determined to facilitate development of physiologically-based models (IPCS, 2010).

184. The ICH note for guidance on the Assessment of Systemic Exposure in Toxicity Studies provides additional guidance on the value of TK data in dose selection in carcinogenicity studies (ICH, 1994). The ICH note for guidance emphasises the need to estimate systemic exposure to the parent compound and/or metabolite(s) at appropriate dose levels via TK studies and at various stages of a carcinogenicity study, in order to ensure that the findings of the study can be interpreted in relation to the comparative exposure for the animal model and humans. The note for guidance notes that increases in exposure may arise unexpectedly as a result of non-linear kinetics due to saturation of a clearance or absorption process. Increasing systemic/internal exposure may also occur during the course of a study for those compounds which have a particularly long plasma half-life. With particular reference to administration of the test compound by oral gavage careful attention should also be paid to compounds which achieve high Cmax values over comparatively short time periods within the dosing interval. Conversely, unexpectedly low internal doses may occur during a study as a result of enzyme induction over time.

REFERENCES

Andersen, M.E. (2003), "Toxicokinetic Modelling and its Application in Chemical Risk Assessment", *Toxicol Lett*, 138, 9-27.

Barton, H.A., T.P. Pastoor, K. Baetcke, J.E. Chambers, J. Diliberto, N.G. Doerrer, J.H. Driver, C.E. Hastings, S. Iyengar, R. Krieger, B. Stahl, C. Timchalk (2006), "The Acquisition and Application of Absorption, Distribution, Metabolism, and Excretion (ADME) Data in Agricultural Chemical Safety Assessments", *Crit Rev Toxicol,* Vol. 36: 9-35.

Chhabra, R.S. and J.R. Fouts (1974), "Sex Differences in the Metabolism of Xenobiotics by Extra-Hepatic Tissues in Rats", *Drug Metab Dispos*, 6: 375-379.

Clewell, R.A. and H.J. Clewell III (2008), "Development and Specification of Physiologically Based Pharmacokinetic Models for Use in Risk Assessment", *Regul Toxicol Pharmacol* 50, 129-143.

ICH (International Conference on Harmonisation) (1994), *ICH Harmonised Tripartite Guideline: Note for Guidance on Toxicokinetics: Assessment of Systemic Exposure in Toxicity Studies. S3A* Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S3A/Step4/S3A_Guideline.pdf

IPCS (2010), *Characterizing and Applying Physiologically-based Pharmacokinetic Models in Risk Assessment*, World Health Organization, International Programme on Chemical Safety, Geneva. http://www.who.int/ipcs/methods/harmonization/areas/pbpk_models.pdf

Loizou, G., M. Spendiff, H.A. Barton, J. Bessems, F.Y. Bois, M. Bouvier d'Yvoire, H. Buist, H.J. Clewell 3rd, B. Meek, U. Gundert-Remy, G Goerlitz and W. Schmitt (2008), "Development of Good Modelling Practice for Physiologically Based Pharmacokinetic Models for Use in Risk Assessment: the First Steps", *Regulatory Toxicology and Pharmacology*, 50, 400 – 411.

Nielsen, E., G. Østergaard and J.C. Larsen (2008), *Toxicological Risk Assessment of Chemicals*, Informa Healthcare, US, ISBN-13: 978-0-8493-7265-0.

OECD (2004a), *Skin absorption: in vivo method,* OECD Guideline for the Testing of Chemicals No 427, OECD, Paris.

OECD (2004b), *Skin Absorption: In Vitro Method,* OECD Guideline for Testing of Chemicals No 428, OECD, Paris.

OECD (2004c), *Guidance Document for the conduct of skin absorption studies,* Series on Testing and Assessment, OECD, Paris.

OECD (2010), *Toxicokinetics*, OECD Guideline for the Testing of Chemicals No 417, OECD, Paris.

Saghir, S.A., A.L. Mendrala, M.J. Bartels, S.J. Day, S.C. Hansen, J.M. Sushynski and J.S. Bus (2006), "Strategies to Assess Systemic Exposure of Chemicals in Subchronic/Chronic Diet and Drinking Water Studies", *Toxicology and Applied Pharmacology*, 211, 245 – 260.

US EPA (2006), *Approaches for the Application of Physiologically-Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment*.
Available at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=157668 .

3.5 HOUSING, FEEDING, HANDLING OF ANIMALS AND EXPERIMENTAL PROCEDURES

185. Many national and international standards have been developed for animal care including housing, feeding health and handling, e.g., NRC (1995, 1996), Council of Europe (2006), the European Community (EEC, 1986) the Society for Laboratory Animal Science (GV-SOLAS, 1988), the Victorian Government Department of Primary Industries (2004). The general principles outlined in these guidelines are similar, and in conducting a chronic toxicity or carcinogenicity study, those guidelines applicable at a national level should be followed.

186. An overarching principle is that the particular needs of given species and strains must take precedence and that adherence to guidelines should never replace close observation of the particular animals involved, continued throughout their lives (Council of Europe, 1997). Provision of exhaustive guidelines for all species and strains is difficult to achieve and local initiatives for improving housing conditions should be taken whenever possible. Appendix A of the Council of Europe Convention "Guidelines for accommodation and care of animals", does however provide detailed guidance on these issues, including aspects such as design and maintenance of the test facilities (Appendix A, Council of Europe, 2006). It should be consulted for in-depth information.

3.5.1 Housing

187. Taken as an example, the Council of Europe Convention (Appendix A, 2006) states on housing that special relevance should be given to the enrichment of the environment of the respective species according to their needs for social interaction, activity-related use of the space, appropriate stimuli and materials. In a review of laboratory environments and rodents' behavioural needs, Balcombe (2006) notes that there is growing recognition of the inherent problems of depriving rodents the space and resources to carry out natural behaviours, such as exploring, foraging, running, escaping hiding and hygiene maintenance. The author reports a recent survey of animal facilities at the US National Institutes of Health which indicates that a slight majority of rats and mice at these facilities are now being provided with nesting and structural (shelter) enrichment (Hutchinson *et al.*, 2005). Other indicators that rodent housing conditions are improving include the availability of commercially produced resources for nesting, shelter, gnawing and play (Key, 2004), and a sharp rise since the late 1980s in the number of citations using keywords 'environmental enrichment' and 'rodent' (Hutchinson *et al.*, 2005). Considering that two decades ago environmental rodent enrichment was scarcely being discussed, the author notes that these are laudable trends (Balcombe, 2006).

188. The Council of Europe Convention Appendix A (2006) recommendations on housing for rodents are as follows:

- Rodent species other than guinea pigs should be kept in cages made of easy to clean material and their design should allow proper inspection of the animals without unnecessarily disturbing them;
- The cages should be provided with solid floors with bedding instead of grid floors, unless there is good reason to have alternatives;
- Gregarious species should normally be group-housed, although it may be difficult to achieve stable and harmonious groups of male mice, and also female hamsters;

- Where the experimental procedures or welfare requirements make group-housing impossible, consideration should be given to accommodating animals of the same species within sight, sound or smell of one another;

- Encouragement should be given to break up the interior space of a cage by introducing objects such as platforms, tubes, boxes, etc. and attempts should be made to provide environmental enrichment with objects to explore, carry or transform, unless negative effects are observed on welfare or on the intended scientific use;

- High hygiene standards should be maintained. However, it may be advisable to maintain odour patterns left by the animals;

- Special attention should be paid to ensuring that the lighting intensity particularly on the top row of cages is not too high. Maximum light intensity should not exceed 350 Lux measured 1 metre from the floor. Provision should be made for shaded areas within the cage to allow the animals to withdraw.

189. The Convention makes specific recommendations for size of caging and stocking densities, dependent on the size/weight of the animals. In relation to environmental conditions, the Convention provides specific recommendations for temperature, humidity and ventilation for each species of laboratory animal covered in the guidelines. Those for rodents are in line with those indicated in the OECD Test Guidelines, as outlined in the next paragraph.

190. The Test Guidelines make some specific recommendations for housing of rodents only, including the recommendation (in line with that of the Convention) that animals may be housed individually e.g., when the test substance is administered via the dermal route, in order to prevent grooming behaviours and oral ingestion, or be caged in small groups of the same sex; individual housing should be considered only if scientifically justified. Animals may be group-caged by sex, but the number of animals per cage must not interfere with clear observation of each animal. The biological properties of the test substance or toxic effects may indicate a need for individual caging. Rodents should be housed individually in dermal studies and during exposure in inhalation studies.

191. The Test Guidelines also specify that cages should be arranged in such a way that possible effects due to cage placement are minimised. For rodents, the temperature in the experimental animal room should be 22°C (± 3°C). Although the relative humidity should be at least 30% and preferably not exceed 70% other than during room cleaning, the aim should be 50-60%. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. These recommendations, together with those of the Convention should be applied in any rodent chronic toxicity or carcinogenicity study conducted according to the Test Guidelines.

192. As indicated in Section 3.3, although rodents (rats or mice) are the main species used in chronic toxicity or carcinogenicity studies, other species, dog in particular, may be used on some occasions. As for rodents, specific guidelines for the care of dogs including housing, feeding health and handling have been developed, e.g., Council of Europe (2006). In relation to the housing of dogs, the Council of Europe Convention recommends that:

- Dogs should be housed in socially harmonious groups, unless the experimental procedures or welfare requirements make this impossible;

- Dogs should be exercised at least daily. Under no circumstances should dogs be caged without exercise for more than 14 days. Preferably, dogs should be exercised with other dogs.

- Dog pens should allow some privacy for the animals. They should include playthings and structures, including elevated platforms.
- Solid floors should be used for dogs. The materials, design and construction of slatted or perforated floors should provide surfaces which do not produce welfare problems such as irritation or injury of the feet or toes, blistering, etc. (these must be prevented at all times), and should supply a solid resting area.
- Temperature in dog studies should be held within a range of 15-21°C, light period between 10 and 12 hours a day, and humidity 40-70%.

193. In relation to the housing of minipigs, pigs should be housed in socially harmonious groups, unless the experimental procedures or welfare requirements make this impossible (Ellegaard *et al.*, 2010).

- Solid floors should be used for pigs. To satisfy rooting behaviour of pigs the use of bedding is recommended. The use of bedding serves as a nutritional substrate, as well as providing environmental enrichment, especially in experimental units where the possibility to move around is limited. Where slatted or perforated floors are used, the materials, design and construction should provide surfaces which do not produce welfare problems such as irritation or injury of the feet or toes, blistering, etc. (these must be prevented at all times), and should supply a solid resting area.
- Pigs are highly sensitive to environmental temperature. The recommended temperature in pig studies should be held within a range of 17.5 to 24°C when 14-16 weeks old and 34-36 weeks of age respectively (dependent on use of bedding; higher temperature for piglets), light period at a minimum of 8 hours a day, and humidity 45-75%.

3.5.2 Feeding

194. In both humans and laboratory animals, diet has a direct bearing on health, and many neoplastic and non-neoplastic diseases are caused (or prevented) by dietary factors, including variations in the composition and amount of feed consumed. The association in rats of caloric consumption, the spontaneous formation of tumours and life span is well established. Although the zero-dose group may be expected to control for the influence of diet, dietary constituents may still profoundly affect the outcome of an experiment (OECD, 2002).

195. A nutritionally-balanced diet is important both for the welfare of laboratory animals and to ensure that experimental results are not biased by unintentional nutritional factors (NRC, 1995). The US National Research Council provides detailed guidance on the nutritional requirements of a wide range of laboratory animals, with detailed information on essential nutrients and other considerations for each species (NRC, 1995). The NRC guidance emphasizes that feed palatability and intake, nutrient absorption and utilization, and excretion can be affected by physicochemical characteristics of feeds such as physical form, sensory properties, naturally-occurring refractory or anti-nutritive substances, chemical contaminants, and conditions of storage (NRC, 1995). Many biological factors also affect nutritional requirements, including genetic differences between species and strains, stage of life of the animals, environmental influences (e.g., diurnal rhythms, temperature etc.), housing and microbiological status (NRC, 1995). Detailed information is also given on diet formulation for natural-ingredient diets, purified and chemically-defined diets, and on manufacture and storage procedures and other considerations (NRC, 1995).

196. The Test Guidelines state that rodents should be fed and watered ad libitum with food replaced at least weekly. Conventional laboratory diets should normally be used/ are normally used. The diet should meet all the nutritional requirements of the species tested and the content of dietary contaminants, including but not limited to pesticide residues, persistent organic pollutants, phytoestrogens, heavy metals and mycotoxins, that might influence the outcome of the test, should be as low as possible. Control and test animals should be fed from the same batch and lot. Analytical information on the nutrient and dietary contaminant levels should be generated periodically, at the beginning of the study and whenever there is a change in the batch used, and should be included in the final report. Analytical information on the drinking water used in the study should similarly be provided. The choice of diet may be influenced by the need to ensure a suitable admixture of a test substance in the diet and to meet the nutritional requirements of the animals when the test substance is administered by the dietary route.

197. As noted in section 3.2.1, the concentration of the test substance in the feed should not normally exceed an upper limit of 5% of the total diet (FDA, 1982, Borzelleca, 1992), although higher levels are feasible (e.g., when testing carbohydrates or proteins) as long as the diet is adapted to be nutritionally adequately, e.g., incorporated, at the expense of other components, in a purified diet. Section 3.2.1 also discusses the problems associated with the palatability of diet (or drinking water) containing test substances affecting the taste and/or smell of the food. If this is marked, it may be necessary to introduce into the study design an additional control group, pair fed (i.e., having matched food intake) in parallel with the high dietary level test group.

198. An important aspect of the feeding regime used in chronic toxicity and carcinogenicity is the recognized effect on study outcome of feeding *ad libitum*. Traditionally, maximal growth and reproduction have been used as criteria for the evaluation of laboratory animal diets (NRC, 1995). However, evidence from a number of studies indicates that restricting the caloric intake of laboratory animals may have beneficial effects on life span, the incidence and severity of degenerative diseases, and the onset and incidence of neoplasia (Weindruch and Walford, 1988; Yu, 1994; Keenan *et al.*, 1997). Based on these results, allowing animals to eat *ad libitum* to produce maximum growth and reproduction may not be consistent with objectives of long-term toxicological and aging studies (NRC, 1995). Overfeeding by *ad libitum* food consumption is generally considered to be the most significant, uncontrolled variable affecting the outcome of the current rodent bioassay, and in particular, the correlation of food consumption, the resultant adult body weight and the 2-year survival in Sprague-Dawley rats is highly significant (Keenan *et al.*, 1997). However, it will probably take years to introduce dietary restriction into national and international test guidelines for toxicity testing because of concern that the delayed occurrence of, for example, cancer reflects a decrease in the sensitivity of the carcinogenicity test in detecting the carcinogenic potential of tested chemical and because the considerable database on historical control is based on data from *ad libitum* feeding studies (Meyer *et al.*, 2003). Species and strain differences in survival are discussed further in Section 3.3.

199. At a practical (experimental) level however, restriction of the caloric intake of laboratory animals is not straightforward. It may disrupt normal diurnal eating rhythms and is not compatible with group housing. It is also important to achieve caloric restriction of test animals without producing unintended nutrient deficiencies (NRC, 1995). Elevation of nutrient concentrations in the diet may be necessary to ensure that the nutrient intake of animals whose eating is restricted is comparable to that of animals allowed to eat *ad libitum*. There is, however, relatively little information available about the extent to which caloric restriction affects nutrient requirements (NRC, 1995). Since rats regulate their food intake according to caloric intake, the mineral and vitamin etc. content of the diet should be adjusted to "caloric density".

200. As already noted, the US National Research Council provides detailed guidance on the nutritional requirements of a wide range of species other than laboratory rodents, including dogs and

rabbits (NRC, 1995). In the case of a chronic toxicity or carcinogenicity study involving animals other than rodents, this guidance should be consulted for information regarding feeding.

3.5.3 Handling, Health Surveillance and Experimental Procedures

201. The quality of care provided in the laboratory may influence not only growth rate and welfare, but also the quality and outcome of experimental procedures (Council of Europe Convention, 2006). The animals should be accustomed to competent and confident handling during routine husbandry and procedures; this will reduce stress both to animals and personnel. Hurst and West (2010) note that *"consistent use of handling methods that do not induce strong anxiety responses will minimise confounding responses due to routine handling before and during experiments, reducing the need to standardise handling experience and timing"*. These authors demonstrated that picking up mice by the tail induced high anxiety and aversion, while use of tunnels or open hand led to voluntary approach, low anxiety and acceptance of physical restraint. Non-rodent species such as dogs, animals should be handled or be in social contact with humans on a regular basis. The behaviour of an animal during handling and the performance of experimental procedures depend to a considerable extent on the confidence and competence of its handler. Good technique should be unhurried, sympathetic and gentle but firm and safe for the animal and operator. All personnel should be appropriately educated and trained, and records of training maintained.

202. A strategy should be in place in all establishments to ensure that an appropriate health status is maintained, which safeguards animal welfare and meets scientific requirements (Council of Europe Convention, 2006). This strategy should include a microbiological surveillance programme, plans for dealing with health breakdowns, and should define health parameters and procedures for the introduction of new animals, e.g., quarantining. Supervision of the accommodation and care by a veterinarian or other competent person is essential.

203. In relation to the experimental phase of a chronic toxicity or carcinogenicity study, as indicated in the Test Guidelines, the animals selected for the study should have been acclimated to laboratory conditions for at least 7 days and should not have been subjected to previous experimental procedures. A period of acclimatisation is needed to allow animals to recover from transport stress, to become accustomed to a new environment and to husbandry and care practices, and to ensure that their health status is sound. The test animals should be characterized as to species, strain, source, sex, weight and age. Each animal should be assigned a unique identification number, and permanently marked with this number by tattooing, microchip implant, or other suitable method. The method chosen should be reliable and cause the minimum pain and discomfort to the animal when applied and in the long-term. Staff should be trained in carrying out the identification and marking techniques, and sedatives or local anaesthetics and analgesics should be used if necessary.

204. At the commencement of the study, the weight variation for each sex of animal used should be minimal and not exceed ± 20 % of the mean weight of all the animals within the study, separately for each sex. Animals should be randomly assigned to the control and treatment groups. After randomisation, there should be no significant differences in mean body weights between groups within each sex. If there are statistically significant differences, then the randomisation step should be repeated, if possible.

205. The animals should be inspected regularly throughout the study, at least daily by a trained person, to ensure that all sick or injured animals are identified and appropriate action taken. Regular health monitoring should be carried out. The Test Guidelines specify that all animals should be checked for morbidity or mortality, usually at the beginning and the end of each day. Animals should additionally be checked once a day following dosing in the case of gavage studies, for specific signs

of toxicological relevance, taking into consideration the peak period of anticipated effects after dosing in the case of gavage administration. Particular attention should be paid to tumour development. The time of tumour onset, location, dimensions, appearance, and progression of each grossly visible or palpable tumour should be recorded. Body weights and food/water consumption and food efficiency should be assessed and recorded at the intervals specified in the guidelines.

206. At the end of the study, for interim kills and in the case of animals found sick or moribund during the study, the animals should be humanely killed. For non-scheduled killing i.e., for animals showing clinical sign of pain, suffering or distress, OECD Guidance Document 19 on the Recognition, Assessment, and Use of Clinical Signs as Humane Endpoint for Experimental Animals Used in Safety Evaluations should be followed (OECD, 2000). All humane methods of killing animals require expertise, which can only be attained by appropriate training. Animals should be killed using a method that adheres to the principles set by the European Commission Recommendations for the euthanasia of experimental animals (Part 1 and Part 2) (EEC, 1986). A deeply unconscious animal can be exsanguinated, but drugs which paralyse muscles before unconsciousness occurs, drugs with curariform effects and electrocution without passage of current through the brain, should not be used without prior anaesthesia. Disposal should not be allowed until death has been confirmed.

207. Records of source, use and final disposal of all animals bred, kept for breeding, or for subsequent supply for use in scientific procedures should be used not only for statistical purposes but, in conjunction with health and breeding records, as indicators of animal welfare and for husbandry and planning purposes.

REFERENCES

Balcome, J.P. (2006), "Laboratory Environments and Rodents' Behavioural Needs, a Review", *Laboratory Animals*, 40, 217–235.

Borzelleca, J. F. (1992), "Macronutrient Substitutes: Safety Evaluation", *Regulatory Toxicology and Pharmacology*, 16, 253-264.

Council of Europe (1986, adopted 2005), *European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes*, CETS No.123. Appendix A of the Convention: Guidelines for accommodation and care of animals (Article 5 of the Convention) approved by the Multilateral Consultation, 15th June, 2006. Available at: http://conventions.coe.int/Treaty/EN/Treaties/PDF/123-Arev.pdf

EEC (1986), *Council Directive 86/609/EEC on the approximation of laws, regulations, and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes*, Official Journal 29, L358, 18 December 1986.

Ellegaard, L., A. Cunningham, S. Edwards, N. Grand, T. Nevalainen, M. Prescott and T. Schuurman (2010), "Welfare of the Minipig with Special Reference to Use in Regulatory Toxicology Studies, *Journal of Pharmacological and Toxicological Methods*, 62, 167-183.

FDA (1982), *Toxicological principles for the safety assessment of direct food additives and color additives used in food*, US Food and Drug Administration, Washington, DC.

GV-SOLAS (Society for Laboratory Animal Science, Gesellschaft für Versuchstierkunde) (1988), Publication *on the Planning and Structure of Animal Facilities for Institutes Performing Animal Experiments*, ISBN 3-906255-04-2. http://www.gv-solas.de/publ/heft1_1988.pdf

Hurst, J.L. and R.S. West (2010), "Taming Anxiety in Laboratory Mice", *Nature Methods*, 7(10), 825-826.

Hutchinson, E., A. Avery and S. VandeWoude (2005), "Environmental Enrichment for Laboratory Rodents", *ILAR Journal*, 46, 148–61.

Keenan, K.P., G.C. Ballam, R. Dixit, K.A. Soper, P. Laroque, B.A. Mattson, S.P. Adams, J.B. Coleman (1997), "The Effects of Diet, Overfeeding and Moderate Dietary Restriction on Sprague-Dawley Rat Survival, Disease and Toxicology", *J Nutr*, 127, 851S-856S.

Key, D. (2004), "Environmental Enrichment Options for Laboratory Rats and Mice", *Lab Animal*, 33, 39–44.

Meyer, O. and O. Svendsen (2003), *Animal Models in Pharmacology and Toxicology*, Chapter 2 in Handbook of Laboratory Animal Science, second edition Volume II, Animal Models, J. Hau and G.L. Van Hoosier Jr, (eds), CRC Press.

NRC (National Research Council) (1995), *Nutrition Requirements of Laboratory Animals*, 4th revised edition, National Academy Press, Washington, D.C.

NRC (National Research Council) (1996), *Guide for Care and Use of Laboratory Animals*, National Academy Press, Washington, D.C.

OECD (2000), *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints*, Series on Testing and Assessment, No. 19, OECD, Paris.

OECD (2002), *Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies*, Series on Testing and Assessment No. 35, OECD, Paris.

Victorian Government Department of Primary Industries, Australia (2004), *Code of Practice for the Housing and Care of Laboratory Mice, Rats, Guinea Pigs and Rabbits*. Available at: www.anu.edu.au/ro/ORI/Animal/code_of_practice_lab_animals.pdf

Weindruch, R. and R.L. Walford (1988), *The Retardation of Aging and Disease by Dietary Restriction*, Charles C Thomas, Springfield, IL.

Yu, B.P., ed. (1994), *Modulation of Aging Processes by Dietary Restriction*. CRC Press, Boca Raton, FL.

3.6 INVESTIGATIONS (INCLUDING HISTOPATHOLOGICAL GUIDANCE)

3.6.1 Introduction

208. This Section on investigations includes guidance on the design and conduct of pathological, histopathological and ophthalmoscopic investigations together with more general advice on the avoidance of bias during investigations. Sponsors and study directors should consider this guidance when planning and performing studies to reduce the likelihood of misleading or ambivalent findings that could result in inappropriate conclusions or the need to repeat studies and to optimize the use of animals for generation of data. It complements existing Guidance contained in OECD Guidance Document No. 35 on the analysis and evaluation of chronic toxicity and carcinogenicity studies. Guidance Document 35 should be consulted for detailed guidance on mortality, clinical observations, body weight changes, food and water consumption, absolute and relative organ weights, haematological, clinical and urinary measurements, post mortem observations and analysis of toxicokinetic and metabolism data. The Society of Toxicologic Pathology recommendations for organ weights have also been published (Sellers *et al.*, 2007).

209. Ophthalmoscopy is an extremely useful clinical technique that allows examination of the anterior and posterior of the eyeball (fundus), including the retina, optic disc, choroid, and blood vessels in the eye. Ophthalmoscopy can help to detect diseases of the eye, to diagnose other conditions or diseases that damage the eye, and can detect other diseases such as some brain tumours. In life ophthalmoscopy allows the progression of the disease to be followed in time and can indicate that additional sampling at necropsy would be informative.

210. Histopathology evaluation is an important part of the assessment of the adverse effects of chemicals on the whole organism. Conventional histological, histochemical and special staining techniques, together with electron microscopy can be used both to define the identity and morphology of tissue, cellular and subcellular structures and to indicate the chemical characteristics of their constituents. In addition to conventional special staining techniques and electron microscopy, new techniques involving immunohistochemistry, molecular biology and novel visualisation procedures can be applied to provide additional more objective methods. Information from such investigations can be used to obtain better functional and morphological characterization of induced alterations in tissues of the body, when needed.

3.6.2 Ophthalmoscopy

211. In toxicity studies, ophthalmoscopic evaluation should be conducted by a suitably trained and experienced individual, preferably using indirect fundoscopic examination *and* a slit-lamp evaluation. A topical mydryatic agent will normally be employed to facilitate ophthalmoscopic examination.

212. Animals should be examined pre-treatment and pre-terminally. Preferably, they will also be examined at least once during the course of the study. Dependent upon the nature of any abnormalities observed, additional ophthalmoscopic examination or other clinical investigations may be warranted. Ocular abnormalities may indicate additional tissue sampling at postmortem to enable histological examination of the ocular adnexa. When ocular abnormalities are detected by ophthalmoscopy and treatment-related causes cannot be confidently excluded, histopathological examination of the eye should attempt to identify the morphological correlate of the abnormality. If ocular abnormalities are focal, it may be appropriate to perform histological examination on additional sections. Appropriate sections from all groups should be taken to provide adequate controls.

213. The study report should contain an integrated interpretation of all treatment-related ocular findings (ophthalmoscopic, macroscopic and microscopic examinations), along with pertinent individual animal data.

214. It should be noted that rodent models are not suitable for the detection of disturbances of ocular pressure by tonometry or for monitoring for certain visual disturbances, such as dyschromatopsia.

3.6.3 Pathology and Histopathology

3.6.3.1 General Considerations

215. Pathology has an important role in toxicology since it provides information on the differences in tissue and organ morphology that establish the presence or absence of lesions and whether or not there are dose–effect relationships. Pathology data can facilitate the interpretation of other data, such as organ weight changes, clinical biochemistry or haematology findings (e.g., Krinke *et al.*, 1991), and evaluators should always make it clear whether there are any associations between pathological abnormalities and other findings of physiological significance. Nevertheless, not all changes in tissue morphology are accompanied by abnormalities in other parameters, and perturbation in organ biochemistry will not necessarily be accompanied by changes in the histological appearance of the affected organ(s). An overview of physiological and environmental factors that can complicate the interpretation of findings in a toxicity study may be found in the *Handbook of Toxicologic Pathology* (Bucci, 1991).

3.6.3.2 Sampling

216. The pathologist should ensure that standard sections of all appropriate tissues and organs are present on the slides to be evaluated. The use of standard sections helps to ensure comparable samples across all animals, thus reducing inconsistency. If flawed or incomplete specimens (eg. missing medulla of adrenal, pars distalis of pituitary, mucosa of intestine, the parathyroid), impair the pathologist's ability to detect, or evaluate treatment-related effects, it is the pathologist's responsibility to obtain recuts, to the extent possible, of those missing or inadequate tissues. Any irretrievable omissions should be taken into account in the interpretation of the data and discussed in the pathology narrative.

217. Procedures for tissue sampling and trimming to ensure optimal fixation, vary between laboratories. However, more standardised approaches to the selection of blocks, orientation of tissues and number of slides examined for rat and mouse organs and tissues in regulatory type toxicity studies are now available (Ruel-Fehlert *et al.*, 2003; Kittel *et al.*, 2004 and Morawietz *et al.*, 2004) and are recommended. These publications are based on the experience of the European Registry of Industrial Toxicology Animal-Data (RITA) and North American Control Animal Database (NACAD). They are an extended revision of the trimming guidelines published by Bahnemann *et al.* (1995). The articles describe in detail the optimum localisation for tissue preparation, the sample size, the direction of sectioning and the number of sections to be prepared, organ by organ. However, existing information on the substances tested and gross findings at autopsy may indicate that samples from non standard locations should also be taken. If a test substance is administered by the dermal route or by inhalation route for example, samples should be taken from the site of application, in addition to the standard specimens.

3.6.3.3 Histopathology specimen processing and quality

218. High quality tissue specimens are necessary for histopathologic evaluation. For routine purposes, fixation in a suitable concentration of a formalin solution followed by processing and embedding in paraffin wax, and cutting of a suitable thickness of histological section is adequate for most tissues. Exceptions are the eyes and testes for which formalin is generally regarded as an inadequate fixative. In these instances, Bouin's or Davidson's fixatives are preferable, especially if evidence from shorter term studies indicated treatment-related testicular or ocular toxicity. Bouin's fluid is generally considered the best fixative for testis although there has been a move away from this because of the safety concerns of its picric acid content (Latendresse *et al.*, 2002). Formalin, when used under stringent conditions is usually also a good fixative for many immunohistochemical and molecular biological techniques on tissue sections. Frozen sections are however usually needed for studies of enzyme activity or those requiring intact RNA. Tissues should be processed in a manner that reduces the potential for variation to be introduced among groups, as indicated in Section 3.6.5.

219. Special stains should be used where appropriate. Haematoxylin and eosin (H&E) remains the most widely used stain, supplemented, where appropriate by a Romanovsky stain for haemopoietic cells, Periodic-acid Schiff (PAS) stain for hepatic glycogen, glomerular basement membrane and the acrosome on testicular germ cells, trichrome and elastic stains for the myocardium, and blood vessels and oil red O applied to frozen sections for neutral lipids. While still not standard, immunohistochemical techniques are now widely used in special cases and the advantages and pitfalls of the techniques are discussed in detail by Greaves (2009). All specially stained tissues should be accompanied by a specimen known to be positive for the particular test to confirm the proper functioning of the technique. Where procedures other than the standard H & E staining are carried out, the specific fixation procedures appropriate to that method should be followed.

220. The use of larger semi-thin (1 to 3 µm thick) plastic or resin embedded sections is a technically demanding but cost effective compromise between electron microscopy and conventional light microscopy. It requires specialist equipment frequently not present in routine histology laboratories. Sometimes termed 'high resolution light microscopy', light microscopic evaluation of semi-thin sections provides a means of avoiding extensive use of the electron microscope because it can locate cytoplasmic organelles in a way sometimes not possible in a paraffin wax embedded material.

221. In view of the hazard to health of formalin, used for fixation, and xylene, used during processing, an increasing number of laboratories are adopting technologies which are free of both formalin and xylene. Nassiri *et al.* (2008) describes the utilization of a formalin-free fixation and processing system for tissue detection of two important biomarkers in breast cancer at the RNA and protein levels. Falkeholm *et al.* (2001) demonstrated that xylene-free histological sections are qualitatively on a par with conventional paraffin sections for routine diagnostic work. Overall, these techniques appear to have good safety profiles, provide excellent histology quality and are suitable for further analysis by immunohistochemistry and nucleic acid extraction. There can however be subtle differences between xylene-free and xylene sections thus the pathologist should be aware of the potential for these differences especially when reading between studies on the same compound where one study uses sections using xylene and the other is xylene-free.

222. In addition to histopathology and immunohistochemistry, tissue may be required for molecular genomic or transcriptomic analysis of biomarkers. This should be borne in mind when collecting and studying tissue. It is important therefore to be prepared to extract RNA, DNA and even protein from paraffin embedded material. This can be achieved using both formalin-containing and formalin-free fixed tissues. Hewitt *et al.* (2008) summarizes the current state-of-the-art of preanalytic

factors in tissue handling and processing as they impact the quality of RNA obtainable from formalin fixed paraffin-embedded tissue.

223. Artefacts may be produced at each of the following stages in the processing of tissue sections: before death, at postmortem or necropsy, during the fixation of tissues, during processing, paraffin embedding and microtomy, during the mounting of tissue sections onto glass slides, staining procedures and coverslipping. Some artefacts are easily distinguishable from normal or diseased tissue components but these render the task of the pathologist more difficult and great care should be taken to avoid their introduction. Some artefacts are difficult to distinguish from pathological changes and are thus of particular concern. McInnes (2005) sets out some of the more common artefacts that are most frequently encountered as a result of inadequacies in the preparation of microscopic tissue sections.

224. All slides evaluated should be identified with a unique code from which identifies the study number, animal number, slide (block) number and the tissue[s] present on the slide.

3.6.3.4 Histopathologic evaluation

225. Details of the nature of the test substance, results of any previous toxicity studies and known activities of this class of compounds should be made available to the pathologist before evaluation of the tissue slides begins. Knowledge of target organs and tissues and the types of changes previously encountered, even in different species, facilitates the evaluation of tissues and provides for the consistent use of terminology. Previous knowledge of target tissues should be utilized during the protocol development to determine whether special pathology procedures should be used in obtaining, fixing, processing, or staining of sections.

226. The study pathologist should have access to in-life clinical observations, organ and body weight data, haematology and clinical biochemistry data, and macroscopic findings of the postmortem examination for each animal in addition to complete information about the experimental design, characteristics of the animal (age, sex and strain) and husbandry of the study population. These include, but are not necessarily limited to: study protocol, including amendments and relevant deviations, species, strain, and age of animals, route, doses, and duration of dosing.

227. Metabolic, pharmacokinetic, or toxicokinetic information may be necessary for understanding patterns of change and interpreting differences in species responses.

228. In-life data (i.e., clinical signs, body weight changes, food consumption, ophthalmoscope findings etc.) from animals may help greatly in the identification of target organs and in understanding mechanisms of toxicity. Haematology, clinical chemistry, and urinalysis results also aid the identification of target organs and may contribute to an understanding of the mechanism of action. Results of special assays, such as hormone concentrations or enzyme induction, are equally important in locating morphologic changes and in the understanding of their significance.

229. Necropsy (gross) findings for individual animals must be available to the pathologist for lesion tracking and correlation with histopathology findings. The pathologist should also be aware of organ weight changes. Often histomorphologic correlates of altered weights can be identified.

230. All of the above-mentioned data, if available, should be provided to the pathologist at the time of the initial slide evaluation. Provision of information on previous findings for the individual animals may render the treatment status of the animal obvious. The pathologist may wish to examine samples from animals of the same treatment group for the same histopathological changes. This may

improve the efficiency of the process but comes at the price of inevitably introducing the potential for bias into the process. Careful consideration needs to be given to the balance between ensuring that subtle or rare changes are detected and the possible introduction of bias. This is discussed further in section 3.6.5.

231. Tissues may be evaluated animal by animal or organ by organ as the preference of the pathologist dictates. The animal by animal technique affords an encompassing overview of an animal's complete health status. The organ by organ technique allows more focused attention to changes and aids in the consistent grading of changes in a particular organ. Where concise terminology is inadequate to convey lesion complexity, detailed free text descriptions should be used to define the diagnostic term used for tabulation.

232. For common lesions in a species or strain of animal, it is important to know if the experimental treatment alters severity. The pathologist should use a severity grading system that allows for an appropriate severity classification, as treatment may affect the incidence (number of animals showing the pathology) or the severity of a particular pathology lesion. Toxicological lesions are frequently found in a continuous spectrum of severity. Therefore, severity grading systems should be: 1) definable, 2) reproducible, and 3) meaningful. A description of each of the various grades should be included in the narrative for target lesions where severity is critical to interpretation of the data. Photomicrographs may be helpful in conveying the severity differences for the grading system used. Well-defined severity grading systems greatly aid the pathology peer-review process. For carcinogenicity studies, it is the pathologist's responsibility to distinguish between hyperplasia, dysplasia, neoplasia and to classify tumours, where applicable, as 1) benign or malignant, and 2) primary or metastatic.

233. Computerised systems of recording findings help to organise the process of evaluation and ensure that all animals and all tissues have been examined, and that gross and microscopical correlation is followed. They also allow rapid, and reproducible, formation of the pathology tables needed to ensure appropriate interpretation of treatment-related effects.

3.6.3.5 Procedures to enhance the accuracy and consistency of histopathology

234. Histopathology is a descriptive and interpretive science and therefore includes an element of subjectivity. However, the pathologist should evaluate tissues as consistently as possible to avoid the introduction of artificial differences or between-group bias. Evaluation of all tissues within a study by one pathologist with consistent standards for detecting, naming, and grading tissue changes, facilitates the detection of differences induced by treatment. However, two or more pathologists are on rare occasions involved in evaluation of a study. In such circumstances, the utmost care must be taken to ensure that nomenclature and severity grading systems used by the contributing pathologists are harmonized to limit variation and that steps are taken to avoid bias (see section 3.6.5). For example, the situation where all of the samples in one dose group are evaluated by one pathologist and those in another dose group by another pathologist should be avoided. Standardized criteria and consistent terminology should be agreed upon for grading systems of common spontaneous and treatment-related findings. The use of a shared computer system that can define a study-specific lexicon for capturing data facilitates this process. Guidance on standard nomenclature is available in OECD Guidance Document No. 35.

235. Pathologists are aware of the phenomenon of "diagnostic drift." Drift refers to a gradual change in nomenclature or severity grading of lesions within a single study. Diagnostic drift usually develops from the increased awareness of a lesion by the pathologist and it is more of a problem in large studies with many animals and tissues requiring evaluation over a prolonged period of time. It is

a source of inconsistency that can negatively impact detection of treatment-related lesions or artefactually introduce apparent treatment-related effects where none exist and can erroneously affect the determination of no-effect-levels. When a pathologist becomes aware of drift in his/her selection of terminology or severity grading, he/she must re-evaluate the tissue(s) involved. Any suspected treatment-related effects should normally be re-evaluated through the use of a "blinding" or "masking" technique, where appropriate.

3.6.3.6 Image Capture

236. It is useful to capture images of key features, both as an aide memoire but also as a means of sharing with other pathologists, thus reducing problems of diagnostic drift. Rather than relying solely upon subjective opinions consideration should be given to use of quantitative morphometry for morphological features, where possible, for example in metabolic bone studies, and the accurate quantitation of protein expression detected by immunofluorescence.

3.6.3.7 Peer Review of histopathology

237. Peer review increases confidence in the accuracy of the histopathology findings from a study. Peer review by an independent pathologist is essential to ensure consistency of the histopathological findings in any studies but in particular those evaluated by more than one pathologist. The objectives of a formal histopathology peer review are several: 1) determine accuracy and consistency of nomenclature, i.e., survey for the presence of incorrectly diagnosed or inaccurately described treatment-related lesions, 2) determine completeness, i.e., survey for the presence of undiagnosed treatment-related lesions, 3) determine the appropriateness of the NOEL, NOAEL or other point of departure, e.g., BMDLx, by reviewing all target tissues and organs, and 4) review the correctness of the textual interpretations derived from those data. The methods employed may vary depending on the purpose of the peer review. For a routine peer review, tissues from a sufficient number of treated animals need to be evaluated to assure that significant lesions were not missed and that a "no effect level" can be verified.

238. A number of procedures can be used for peer review (Eighmy, 1996; Peters, 1996; The Society of Toxicologic Pathologists, 1991 and 1997). Peer reviews are generally included in the study protocol and are conducted prior to the issuance of the study report (Ward *et al.*, 1995). It is important that the original plan for the review process include a joint review by the original and reviewing pathologist of the two sets of results to explain or resolve apparent differences, should they occur. The results of the peer review should be documented and archived and any differences of opinion resolved through consensus. This procedure is then part of the process that leads to finalizing diagnoses and interpretations.

239. If any differences cannot be resolved through a joint review of the data, then arbitration through a third pathologist should take place or the problem referred to a "pathology working group" (PWG) or to a panel of expert consultants for resolution. Contingency plans to allow for this should also be included in the original study plan. A Pathology Working Group (PWG) may be formed to review, revise and/or interpret diagnoses (Peters, 1996; The Society of Toxicologic Pathologists, 1991) and this should be done prior to finalising the report. The PWG should be composed of individuals having expertise both in the pathology of the test species and with the specific lesion in question. The PWG may include both the original and the reviewing pathologist but the diagnoses of both the original and the reviewing pathologist should not be known to the other PWG members. Prior to beginning work, a PWG chairman is appointed who directs the focus of the PWG and orchestrates

agreement of the criteria used to make the diagnoses of the lesions in question. For impartiality reasons, neither the original pathologist nor the reviewer pathologist should be nominated as chairman (Mann P.C., 1996). This process should be clearly defined and documented. The PWG may request additional sections from tissues to aid with the diagnostic process of the peer review. Following slide examination by the PWG in a blinded fashion and tabulation of the results; the findings of the PWG are compared to those of the original pathologist. When the consensus of the PWG is clearly different from that of the original pathologist, the diagnosis for an individual lesion should be changed. When there is a close split on a vote of a diagnosis, however, the diagnosis of the original pathologist should be allowed to stand. The records of the PWG should indicate the final diagnoses for each lesion and the degree of certainty of the diagnoses.

240. Regardless of its exact form, the peer review process should encourage direct interaction between the original and the reviewing pathologist and should result in the production of a single, scientifically robust pathology report that appropriately and accurately summarises the results and any uncertainties. The process should be constructive, meet the needs and objectives of the review and have an inbuilt procedure for resolving differences, should they arise.

241. Retrospective peer reviews by a single pathologist or a PWG may also be undertaken after the completion of a study. This type of peer review is necessary when unexpected issues are highlighted after a study is finalised and additional work is needed to clarify the issue. In these exceptional circumstances, where the conclusions of the study may change for example, it is necessary to take steps to ensure that the original report is marked to indicate that a revision of the final conclusions has occurred. A separate report or a report amendment may be appropriate and any changes must be documented as required by GLP.

3.6.3.8 Early Death Investigations

242. Guidance on the interpretation of early deaths is provided in OECD Guidance Document 35. In this Section further guidance on investigations that are valuable is provided.

243. In accordance with OECD Guidance Document 19 (GD 19), the earliest possible endpoints that are indicators of distress, severe pain, or impending death should be used as indications for humanely killing the animals prior to them reaching a moribund state or dying. GD 19 was developed largely for animal welfare reasons. However, this practice also has the advantage or reducing the number of animals that are found dead with the accompanying risk of loss of tissues from autolysis or cannibalism. It also avoids confusing direct effects of the test substance with secondary effects resulting from post-mortem or moribund changes.

244. Reasonable efforts should be made to determine the cause, or likely cause, of individual deaths or severe toxicity leading to euthanasia. It is important to identify causes unrelated to exposure to the test agent (e.g., acute or chronic infections, age or disease-related degenerative processes, anatomical abnormalities, mishandling or accident) from toxicity-induced effects. All in-life data, as well as the results of the post mortem, of euthanized or dead animals in a study, should be used in an attempt to make this distinction.

245. Assignment of the cause of death or of severe toxicity, leading to euthanasia is a component of the assessment of carcinogenicity studies in rodents, which is widely practised to aid statistical evaluation (Kodell *et al*., 1995). The US National Centre for Toxicological Research (NCTR) requests routine assignment of cause of death for all dead and moribund animals examined histologically in carcinogenicity studies. A determination as to whether a particular tumour under consideration was fatal to an animal or was simply an incidental finding at necropsy is important for determining the

type of information that an individual animal contributes to age-adjusted statistical tests for carcinogenic effects (Peto *et al.*, 1995).

246. Organ weights from animals that die or are euthanized prior to scheduled necropsy are of little value in most cases to the overall study because of differences in nutritional status and tissue congestion and oedema secondary to the health status of the animal, and may not need to be recorded (Sellers *et al.*, 2007). The absence of matched concurrent control data further hinders interpretation of organ weights from animals that die or are euthanized prior to termination of the study. Gross pathologic and/or histopathologic changes of the organ can however be useful in indicating a potential cause of death in the individual animal.

247. An experienced pathologist should, under appropriate conditions, be able to distinguish peri- and post-mortem changes in tissues from treatment-related effects on tissues. While cannibalism can inhibit accurate pathology evaluation it is almost always useful to fix and block tissues from all animals, even those found dead, since different tissues show different degrees of autolytic change in death and conclusions can often be made even when some time has elapsed between death and tissue sampling.

3.6.3.9 Pathology Final Conclusions

248. A final pathology/histopathology conclusion is reached when the following have been achieved: 1) histopathologic diagnosis and description of all findings for each animal on the study, 2) the diagnoses and descriptions reflecting the consensus of a peer review, if a peer review was done prospectively, 3) tabular summary data providing an accurate representation of the findings, 4) a NOEL, NOAEL, or other point of departure, e.g., BMDLx, for pathology/histopathological findings in the study, if appropriate. Final conclusions and diagnoses may often be the result of an iterative process in which the diagnosis is continually refined and modified until a final conclusion can be reached. This process may include taking additional samples for evaluation. Details of areas of disagreement and/or uncertainty that arose during the study and how these were resolved should be documented. The final pathology report must accurately and completely present the data and their interpretation.

249. Issues arise, particularly in carcinogenicity studies, as to whether findings are split into separate categories such as adenoma and carcinoma, grouped as benign and malignant, combined across multiple tumour sites and the inclusion or exclusion of metastatic tumours. Information on related metastases and its site(s) should be reported for the primary tumour. The general convention is that metastases from primary tumours should not be included separately in statistical analysis, except where no primary tumours could be identified. Benign and malignant lesions should always be presented separately. Combination may result in a loss of sensitivity if the substance tested accelerates the progression to malignancy, since the increase in malignant neoplasms is balanced by a decrease in benign neoplasms. However, benign and malignant neoplasms of the same cell type are sometimes combined because one is seen as the progression of the other. Therefore, whilst combined data may eventually be useful, the pathology narrative should be full and precise to prevent misinterpretation or doubts over the interpretation.

250. McConnell *et al.* (1986) recommended that the statistical analysis of each tumour type be carried out separately and, if it is considered scientifically defensible, further statistical analysis may be performed on the combined tumours (benign and malignant) of the same histogenic origin, even when those tumours are in different tissues: so for example, the incidence of lipoma & liposarcoma should be combined across all tissues, as should the overall incidence of vascular tumours and tumours of connective tissue types. More detailed guidance on combining neoplasms is given in

McConnell *et al.* (1986) and further guidance on statistical analysis is given in Chapter 4 of this Guidance Document.

251. The narrative of the final pathology report should address all the significant pathology findings, report any group differences and address the issue of disease progression. When the disease is progressive, then the identification in the pathological examination of the presence of the more severe form of a disease should automatically mean that the earlier (precursor) form is or has been present. Background pathology and other underlying conditions such as infections or the presence of incidental lesions in both control and treated animals should also be reported. There should be precise descriptions of any treatment-related findings and their likely importance. Clear reasons for concluding that any observations are unrelated to treatment or are not biologically or toxicologically significant should also be provided (Morton *et al.*, 2006).

3.6.4 Supplementary Investigations

252. Previous knowledge of the effects of the substance may indicate that non-standard investigations would be of value in the interpretation of the results. Samples for such investigations should be taken and stored appropriately, provided that this does not jeopardise the basic requirements of the study. Such supplementary sampling would not thus invalidate the study as an OECD-compliant study according to the Test Guideline. Findings in the initial standard investigations may raise further questions that may be addressed by supplementary studies if appropriate stored samples are available.

3.6.5 Avoiding bias

253. Bias can arise if samples or observations in one group differ from those in another group for any reason other than treatment with the test substance. Therefore it is important that steps are taken to prevent the introduction of bias throughout the whole study and not restricted to the randomisation of animals to the treatment groups at the beginning of the study. Bias can be introduced whenever the operator or observer, equipment used or the timing of investigations are confounded with treatment group. The study design should minimise such effects by use of randomisation or a well balanced investigation scheme.

254. Randomisation should be done formally, not semi-subjectively. There are several methods for performing the randomisation process. The most commonly used is a computerised algorithm, but card assignment and the use of a random number table may also be used. Randomisation is normally used for the sequence of necropsy and can also be used for the sequence of other investigations. Similarly, if different operators or observers and equipment are used for the same investigation, randomisation should be used to prevent confounding with the treatment group. A well balanced investigation scheme can be used instead of randomisation for investigations.

255. Re-randomisation of samples is not necessary at each stage of the study and an initial randomisation or well balanced investigation scheme can be re-used throughout the study investigations. Slides are not normally randomised for histopathologic evaluation; slides from the control and highest dose groups are usually evaluated before other groups. The benefits and drawbacks of blinding or masking samples for histopathological evaluation are discussed below.

256. Subjective evaluations, such as clinical signs and histopathology are susceptible to bias if the evaluator is aware of the treatment group to which the individual animal or slide belongs. Blinded

or masked evaluation is carried out without prior knowledge of treatment group, which could include untreated or other control groups as a mechanism to minimize the introduction of observational bias. Blinded/masked evaluation is a more rigorous scientific approach to assessment.

257. For the evaluation of clinical signs, the absolute masking of the treatment group for individual animals is practically difficult as animals within one cage are usually from the same treatment group. However, blinding of the evaluator should be maintained to the extent possible without jeopardising other aspects of the study, particularly by increasing the risk of errors in dosing or animal identification.

258. Appropriate blinding procedures should be considered by the pathologist before, during, and after the microscopic evaluation of tissues. Blinding need not prevent the pathologist being provided with detailed information of other observations in the study, including animal-specific findings (see Section 3.6.3.4). The foremost objection to blinded slide evaluation in the initial histopathological examination is loss of knowledge of the range of normal that exists in known controls. This may be particularly pertinent for studies on new chemical entities or in new species, where little is known about potential treatment-related effects and their background frequency in control animals. Without this baseline, subtle differences between treated and control groups may be difficult to detect. Knowledge of the treatment group also allows the pathologist to assess the spectrum of related morphological changes and determine the most appropriate diagnostic terminology, including combining related diagnoses where indicated. However, these advantages must be weighed against the inevitable bias introduced by non-blinded evaluations. If an initial non-blinded examination is carried out to allow the recognition of treatment-related findings, subsequent masked evaluations of target tissues can be very helpful because it ensures an unbiased determination of a NOEL (no-observable-effect-level), or other point of departure.

259. Randomisation, well balanced investigation scheme and blinding or masking procedures employed to minimise bias in investigations should be described in the protocol and study report.

3.6.6 Data Archives

260. Archiving of documents and specimens generated during a non-clinical laboratory study is a fundamental Good Laboratory Practice (GLP) requirement. The records and material that should be archived as well as the characteristics and the organisation of archive facilities are addressed in the OECD series on Principles of Good Laboratory Practice No. 1 (1997) and the OECD GLP document 15 (2007). Each kind of material has its own individual requirements for proper storage and may have different retention times. Materials should be retained for the period specified by the appropriate authorities.

261. Raw data are defined in the OECD series on Principles of Good Laboratory Practice No. 1 (1997) as "all original test facility records and documentation, or verified copies thereof, which are the results of the original observations and activities in a study. Raw data also may include, for example, photographs, microfilm or microfiche copies, computer readable media, dictated observations, records data from automated instruments, or any other data storage medium that has been recognised as capable of providing secure storage of information for a time period as stated in section 10".

262. Any computerised system used to collect or store histopathology data should be GLP compliant. A variety of security measures will have been incorporated into each of these computer systems to ensure both data integrity throughout the process of histopathology data collection and reporting and to ensure that any changes made after the pathology contribution has been locked are

recorded, and that an audit trail for the changes, including who made them and when, is kept. These security controls include: limited user access, single and/or multiple password requirements, procedural controls and technical controls built into the systems. A critical requirement of an audit trail is a record of any change to the data.

REFERENCES

Abdo, K.M. and F.W. Kari (1997), "The Sensitivity of the NTP Bioassay for Carcinogen Hazard Evaluation can be Modulated by Dietary Restriction", *Exp Toxic Pathol,* 48: 129–137.

Bahnemann, R., M. Jacobs, E. Karbe, W. Kaufmann, G. Morawietz, T. Nolte and S. Rittinghausen (1995), "RITA – Registry of Industrial Toxicology Animal-data – Guides for Organ Sampling and Trimming Procedures in Rats", *Exp Toxic Pathol,* 47: 247–266.

Boorman G.A. and S.L. Eustis (1986), *The Pathology working group as a means for assuring pathology quality in toxicological studies*, In: Hoover B.K., Baldwin J.K., Uelner A.F., Whitmire C.E., Davies C.L. and Bristol D.W. (Eds) Managing Conduct and Data Quality of Toxicology Studies, pp. 271–275, Princeton Scientific Publishing Co., Princeton, N.J.

Bucci, T.J. (1991), *Evaluation of Altered Morphology*, In: Hascheck W.M. and Rousseaux C.G. (Eds) Handbook of Toxicologic Pathology, ch. 3, pp. 23–25, Academic Press, San Diego, Cal.

Cammermeyer, J. (1978), "Is the Solitary Dark Neuron a Manifestation of Postmortem Trauma to the Brain Inadequately Fixed by Perfusion?" *Histochemistry,* 56: 97-115.

Chhabra, R.S. and J.R. Fouts (1974), "Sex Differences in the Metabolism of Xenobiotics by Extrahepatic Tissues in Rats", *Drug Metab Dispos,* 6: 375–379.

Christian, M.S., A.M. Hoberman, M.D. Johnson, W.R. Brown and T.J. Bucci (1998), "Effect of Dietary Optimization on Growth, Survival, Tumor Incidences and Clinical Pathology Parameters in CD Sprague–Dawley and Fischer 344 Rats: a 104-week Study", *Drug Chem Toxicol,* 21: 97–117.

Darke, P.G.G., J.D. Donagura and D.F. Kelly (1996), *Color Atlas of Veterinary Cardiology,* Mosby, St Louis.

Dayton, P.G. and J.E. Sanders (1983), "Dose-dependent Pharmacokinetics: Emphasis on Phase I Metabolism", *Drug Metab Rev*, 14: 347–405.

Eastin, W.C. and R.W. Tennant (1998), *Carcinogenicity: Testing Predicting & Interpreting Chemical Effects,* In: Kitchin KT (Ed) pp. 395–410, Marcel Dekker Inc., New York.

Eastin, W.C. (1998), "The U.S. National Toxicology Program Evaluation of Transgenic Mice as Predictive Models for Identifying Carcinogens", *Environ Health Perspect*, 106(Sup. 1): 81–84.

Eighmy, J. J. (1996), "Study Pathologist Perspective of Pathology Peer Review." *Toxicol Pathol,* 24(5): 647–649.

Falkeholm, L., C.A. Grant, A. Magnusson and E. Möller (2001), "Xylene-free Method for Histological Preparation: a Multicentre Evaluation", *Lab Invest*, 81:1213-21.

Fix, A.S. and R.H. Garman (2000), "Practical Aspects of Neuropathology: A Technical Guide for Working with the Nervous System", *Toxicol Pathol* 28; 122.

Greaves, P. (2009), *General and Applied Toxicology*, 3rd Edition, Wiley Press, Editors Ballantyne, B, Marrs, T, Syversen, T. ISBN: 978-0-470-72327-2.

Hart, R.W., D.A. Newmann, R.T. Robertson (1995), *Dietary Restriction: Implications for the Design and Interpretation of Toxicity and Carcinogenicity Studies*, ILSI Press, Washington, DC.

Hausmann, R., H. Bock, T. Biermann and P. Betz (2004), "Influence of Lung Fixation Technique on the State of Alveolar Expansion - a Histomorphometrical Study", *Leg Med (Tokyo)* 6:61–65.

Hayward, J.J., B.S. Shane, K.R. Tindall and M.L. Cunningham (1995), "Differential *in vivo* Mutagenicity of the Carcinogen–noncarcinogen Pair 2,4- and 2,6-diaminotoluene", *Carcinogenesis,* 16: 2429–2433.

Hewitt, S.M., F.A. Lewis, Y Cao, R.C. Conrad, M. Cronin, K.D. Danenberg, T.J. Goralski, J.P. Langmore, R.G. Raja, P.M. Williams, J.F. Palma and J.A. Warrington (2008), "Tissue Handling and Specimen Preparation in Surgical Pathology: Issues Concerning the Recovery of Nucleic Acids from Formalin-fixed, Paraffin-embedded Tissue", *Arch Pathol Lab Med*, 132:1929-35.

House, D. E., E. Berman, J. C. Seely, and J. E. Simmons (1992), "Comparison of Open and Blind Histopathologic Evaluation of Hepatic Lesions" *Toxicol Lett,* 63: 127–133.

IARC (1992–1997) International Classification of Rodent Tumours, Part 1. The Rat (set of 10 fascicles), Mohr U (Ed), IARC Scientific Publications 122, Lyon.

Iatropoulos, M.J. (1984), "Editorial", *Toxicol Pathol,* 12(4): 305–306.

Keenan, K.P., G.C. Ballam, K.A. Soper, P. Laroque, J.B. Coleman and R. Dixit (1999), "Diet, Caloric Restriction, and the Rodent Bioassay", *Toxicol Sci,* 52 (Suppl): 24–34.

Keenan, K.P., J.B. Coleman, C.L. Mc Coy, C.-M. Hoe, K.A. Soper and P. Laroque (2000), "Chronic Nephropathy in ad libitum Overfed Sprague-Dawley Rats and its Early Attenuation by Increasing Degrees of Dietary (Caloric) Restriction to Cntrol Growth", *Toxicol Pathol*, 28: 788–798.

Keenan, K.P., P. Laroque, G.C. Ballam, K.A. Soper, R. Dixit, B.A. Mattson, S.P. Adams and J.B. Coleman (1996), "The Effects of Diet, ad libitum Overfeeding, and Moderate Dietary Restriction on Rodent Bioassay: The Uncontrolled Variable in Safety Assessment", *Toxicol Pathol*, 24: 757–768.

Kemi, M., K.P. Keenan, C. McCoy, C.-M. Hoe, K.A. Soper, G.C. Ballam and M.J. vanZwieten (2000), "The Relative Protective Effects of Moderate Dietary Restriction versus Dietary Modification on Spontaneous Cardiomyopathy in Male Sprague-Dawley Rats", *Toxicol Pathol*, 28: 285–296.

Kittel, B., C. Ruehl-Fehlert, G. Morawietz, J. Klapwijk, M.R. Elwell, B. Lenz, M.G. O'Sullivan, D.R. Roth and P.F. Wadsworth (2004), "Revised Guides for Organ Sampling and Trimming in Rats and Mice - Part 2 - A Joint Publication of the RITA and NACAD Groups", *Experimental and Toxicologic Pathology,* 55, 413-431.

Kodell, R. L., B.-N. Blackwell, T.J. Bucci, and D.L. Greenman (1995), "Cause-of-death Assignment at the National Center for Toxicological Research", *Toxicologic Pathology,* 23, 241-247.

Krinke, G.J., L.P.A. Perrin and R. Hess (1991), "Assessment of Toxicopathological Effects in Ageing Laboratory Rodents", *Arch Toxicol*, (Sup. 14) 43–49.

Latendresse, J.R., A.R. Warbrittion, H. Jonassen and D.M. Creasy (2002), "Fixation of Testes and Eyes Using a Modified Davidson's Fluid: Comparison with Bouin's Fluid and Conventional Davidson's Fluid", *Toxicologic Pathology*, 30, 524-533.

Levy, G. (1968), *Dose Dependent Effects in Pharmacokinetics*, In: Important Fundamental Principles in Drug Evaluation, pp. 141–172.

Mann, P.C. (1996), "Pathology Peer Review from the Perspective of an External Peer Review Pathologist" *Toxicol Pathol,* 24(5): 650–653.

McConnell, E.E., H.A. Solleveld, J.A. Swenberg and G.A. Boorman (1986), "Guidelines for Combining Neoplasms for Evaluation of Rodent Carcinogenicity Studies", *J. Natl.Cancer Inst,* 76 283-289.

McInnes, E. (2005), "Artefacts in Histopathology", *Comp Clin Path*, 13: 100–108.

Mohr, U. and G. Morawietz (1995), "Recent Progress in the Field of Toxicologic Pathology in Europe and the Need for Standardisation", *J Toxicol Pathol*, 8: 219–226.

Mohr, U., (1999), "RITA – Registry of Industrial Toxicology Animal-data: Optimization of carcinogenicity bioassays", *Exp Toxic Pathol*, 51: 461–475 (editorial).

Morawietz, G., S. Rittinghausen, U. Mohr (1992), "RITA – Registry of Industrial Toxicology Animal-data – Progress of the Working Group", *Exp Toxic Pathol*, 44: 301–309.

Morawietz, G., C. Ruehl-Fehlert, B. Kittel, A. Bube, K. Keane, S. Halm, A. Heuser andJ. Hellmann (2004), "Revised Guides for Organ Sampling and Trimming in Rats and Mice - Part 3 - A Joint Publication of the RITA and NACAD Groups", *Experimental and Toxicologic Pathology*, 55, 433-449.

Morton, D., R.K. Kemp, S. Francke-Carroll, K. Jensen, J. McCartney, T.M. Monticello, R. Perry, O. Pulido, N. Roome, K. Schafer, R. Sellers and P.W. Snyder (2006), "Best Practices for Reporting Pathology Interpretations within GLP Toxicology Studies", *Toxicologic Pathology*, 34, 806-809.

Munro, I.C. (1977), "Considerations in Chronic Toxicity Testing: The Chemical, the Dose, the Design", *J Environ Path Toxicol*, 1: 183–197.

Nassiri, M., S. Ramos, H. Zohourian, V. Vincek, A.R. Morales and M. Nadji (2008), "Preservation of Biomolecules in Breast Cancer Tissue by a Formalin-free Histology System", *BMC Clinical Pathology*, 8:1doi:10.1186/1472-6890-8-1.

National Academy of Sciences/National Research Council (1983), *Risk Assessment in the Federal Government: Managing the Process*, National Academy Press, Washington, D.C.

National Academy of Sciences/National Research Council (1994), *Science and Judgement in Risk Assessment*, National Academy Press, Washington, D.C.

Newberne, P.M. and F.A. de la Iglesia (1985), "Editorial: Philosophy of Blind Slide Reading", *Toxicol Pathol*, 13(4): 255.

OECD (2002), *Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies*, Series on Testing and Assessment No. 35 and Series on Pesticides No. 14, OECD Paris.

OECD (2000) *Guidance Document on the Recognition, Assessment, and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation*, Series on Testing and Assessment No 19, OECD Paris.

OECD (2007), *Establishment and control of archives that operate in compliance with the principles of GLP*, Series on Principles of Good Laboratory Practice and Compliance Monitoring, Number 15, OECD Paris.

OECD (1998), *OECD Principles on Good Laboratory Practice*, Series on Principles of Good Laboratory Practice and Compliance Monitoring, Number 1, OECD Paris.

Pearson, G.R., E.F. Logan (1978), "The Rate of Development of Post-mortem Artefact in the Small Intestine of Neonatal Calves", *Br J Exp Pathol*, 59:178–182.

Peters, T.S. (1996), "Pathology Peer Review—A Concept for Consideration" *Toxicol Pathol*, 24(5): 654–656.

Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, S. Richards and J. Wahrendorf (1980), *Guidelines for Simple, Sensitive Significance Tests for Carcinogenic Effects in Long-term Animal Experiments*, In: Long-Term and Short-Term Screening Assays for Carcinogens: A Critical Appraisal (IARC Monographs, Supplement 2), R. Montesano, H. Bertsch and L. Tomatis (eds). International Agency for Research on Cancer, Lyon, pp. 311-426.

Prasse, K., P. Hildebrandt and D. Dodd (1986), "Letter to the Editor" *Vet Pathol* 23: 540–541.

Rao, G.N. (1997), "New Nonpurified Diet (NTP-2000) for Rodents in the NationalToxicology Program's Toxicology and Carcinogenesis Studies", *J Nutr*, 127: 842S–846S.

Rao, G.N., R.W. Morris, J.C. Seely (2001), "Beneficial Effects at NTP-2000 Diet on Growth, Survival and Kidney and Heart Diseases of Fischer 344 Rats in Chronic Studies", *Toxicol Sci*, 63: 245-255.

Ruehl-Fehlert, C., B. Kittel, G. Morawietz, P. Deslex, C. Keenan, C.R. Mahrt, T. Nolte, M. Robinson, B.P. Stuart and U. Deschl (2003), "Revised Guides for Organ Sampling and Trimming in Rats and Mice - Part 1", *Experimental and Toxicologic Pathology*, 55, 91-106.

Sellers, R.S., D. Morton, B. Michael, N. Roome, J.K. Johnson, B.L. Yano, R. Perry and K. Schafer (2007), "Society of Toxicologic Pathology position paper: organ weight recommendations for toxicology studies", *Toxicol Pathol* 35 751-755.

Sharma R.P. and R.A. Coulombe, Jr (1996), *Pharmacokinetics and Risk Assessment*, In: A.M. Fan and L.W. Chang (Eds) Toxicology and Risk Assessment, ch. 7, pp. 81–99. Marcel Dekker Inc., New York, N.Y.

Society of Toxicologic Pathologists' Position Paper on Blinded Slide Reading (1986), *Toxicologic Pathology* 14(4): 493–494.

Society of Toxicologic Pathologists Position Paper: Documentation of Pathology Peer Review (1997), *Toxicol Pathol* 25(6): 655.

Society of Toxicologic Pathologists' Recommendations on Statistical Analysis of Rodent Carcinogenicity Studies (2002), *Toxicol Pathol* 30(3), 415–418.

Standardized System of Nomenclature and Diagnostic Criteria. *Guides For Toxicologic Pathology*. STP/ARP/AFIP, Washington, DC.

Summers, B.A., J.F. Cummings and A. De Lahunta (1995), *Veterinary Neuropathology*, Mosby Year Book, pp 34–35.

Thompson, S.W. and L.G. Luna (1978), "An Atlas of Artifacts Encountered in the Preparation of Microscopic Tissue Sections", Charles Louis Davis DVM Foundation, Gurnee.

Travlos, G.S., R.W. Morris, M.R. Elwell, A. Duke, S. Rosenblum and M.B. Thompson (1996), "Frequency and Relationships of Clinical Chemistry and Liver and Kidney Histopathology Findings in 13-week Toxicity Studies in Rats", *Toxicol*, 107: 17–29.

WHO (1987), *Principles for the Safety Assessment of Food Additives and Contaminants in Food*, IPCS Environmental Health Criteria 70, World Health Organisation, Geneva.

4. STATISTICAL AND DOSE RESPONSE ANALYSIS, INCLUDING BENCHMARK DOSE AND LINEAR EXTRAPOLATION, NOAELS AND NOELS, LOAELS AND LOELS

4.1 Preamble

263. This document is intended to provide guidance on the statistical issues associated with the design and analysis of chronic toxicity and carcinogenicity bioassays, the analysis and interpretation of resulting data and the use of these data in the identification of Benchmark doses, linear extrapolation and various no observed adverse effect levels (NOAELs) and lowest observed adverse effect levels (LOAELs) measures.

264. The fundamental point that this document aims to convey is:

"The statistical methods most appropriate for the analysis of results, given the experimental design and objectives, should be established before commencing the study" (OECD 2009, paragraph 9 of Test Guideline (TG) 451/453 and paragraph 8 of TG 452).

265. Statistical analysis of biological data is intertwined with the experimental design of studies so this document also includes discussion of issues related to study design. Consequently there is some overlap with topics discussed in other sections but this can allow relevant linkages and cross-references to be made to other sections of the document.

4.2 Introduction

266. The central concept of this document is that the experimental design represents the strategy for answering the question of interest and that the specific statistical analyses are tactical methods used to help answer the questions. Therefore, the statistical methods most appropriate for the analysis of the data collected should be established at the time of designing the experiment and before the study starts.

267. Many of the standard methods in long-term animal experimentation have been in use for a considerable time. There is, therefore, considerable experience of the statistical properties and strengths/weaknesses of these designs and in the integration of the biological importance of findings with the results of statistical analyses. There is, though, a need to ensure that studies are planned before beginning so as to optimize the use of the resources available and to avoid the time and costs associated with the unnecessary repeat of poorly designed or the conduct of duplicate studies. From an ethical perspective, there is a requirement to avoid the use of more animals than is absolutely necessary. Statistical analysis plans should be in place before the start of the study and staff with appropriate statistical expertise should be involved in all stages of the study, especially the design stage.

268. The principles of the formal study design developed over many years should continue to underpin the subsequent pragmatic interpretation often needed in the interpretation of data. It is important to ensure that the quality of studies does not decline because of familiarity with the methods. Long-term studies require a high standard of conduct to ensure unbiased statistical analysis. These requirements are well known and have not undergone any significant changes in recent years. (See paragraph 275 – Current statements on statistical methods.)

269. Much of this Chapter will concentrate on the statistical methods proposed in various documents and guidelines specifically developed for the analysis of chronic toxicity data. Particular

attention will be given to OECD Guidance Document N° 35 (2002) which made some general points about the strengths and weaknesses of statistical methods and listed some common statistical tests (reproduced and augmented here as Figure 1) and provided references to a number of sources.

270. It is important at this point to stress that there is no single approach to the statistical analysis of data. There are different schools of thought about statistical methodologies. These can raise fundamental, almost philosophical, issues about the role of statistical analysis. An example is the long-running debate between 'Frequentists' and 'Bayesians'. Statistical methods also continue to develop so that new and modified approaches may continue to be proposed. As a consequence there can be alternative approaches to those suggested in various guidance documents. Such methods may, in practice, satisfy the requirements of a regulatory authority but it is always recommended that such approaches are discussed in advance with the relevant regulatory authority.

4.3 Objectives

271. In the 1960s and 1970s the primary objective of a long-term rodent carcinogenicity bioassay was qualitative hazard identification (i.e., identification of chronic toxicity and the evaluation of the carcinogenic potential of a chemical administered to rodents for most of their lives).

272. The purpose of the long-term rodent carcinogenicity bioassay has, however, widened to extend to a number of other objectives (e.g., paragraph 6 of TG 451 or TG 453). These include objectives relating to hazard characterization, describing the dose-response relationship and the derivation of an estimate of a Point of Departure (POD) such as the Benchmark Dose (BMD) or a no observed adverse effect level (NOAEL) which can then be used to establish an acceptable level of human exposure.

273. As a consequence of multiple potential objectives study design can become a compromise with a trade-off in the ability to answer competing questions: hazard identification/characterization on the one hand and characterization of the dose-response on the other (see paragraph 278).

274. The objective of the chronic toxicity bioassay (TG 452) is to characterize the toxicological response of a substance in a mammalian species following prolonged and repeated exposure. Determining the carcinogenic potential is also the objective of the combined chronic toxicity/carcinogenicity study (TG 453), but is the primary objective of the carcinogenicity study (TG 451). In this guidance the term "long-term bioassay" will refer to both the carninogenicity and the chronic toxicity bioassay.

4.4 Current statements on statistical methods

275. OECD TGs 451, 452 and 453 contain sections relating to the statistical analysis of the data. All three guidelines state:

"When applicable, numerical results should be evaluated by an appropriate and generally acceptable statistical method. The statistical methods and the data to be analysed should be selected during the design of the study (paragraph 8 [TG452] or 9 [TG451] and [TG453]). Selection should make provision for survival adjustments, if needed."

276. In TGs 451 and 453, paragraph 9 states:

"The statistical methods most appropriate for the analysis of results, given the experimental design and objectives, should be established before commencing the study. Issues to consider include

whether the statistics should include adjustment for survival, analysis of cumulative tumour risks relative to survival duration, analysis of the time to tumour and analysis in the event of premature termination of one or more groups. Guidance on the appropriate statistical analyses and key references to internationally accepted statistical methods are given in a Guidance Document on the design and conduct of chronic toxicity and carcinogenicity studies , available on the OECD public website on Test Guidelines, and also in Guidance Document No.35 on the analysis and evaluation of chronic toxicity and carcinogenicity studies).''

This passage is the same as paragraph 8 in TG 452, with the exception of the second sentence which states:

''Issues to consider include whether the statistics should include adjustment for survival and analysis in the event of premature termination of one or more groups.''

277. Other organizations have made suggestions for the statistical methods to be used. For instance, the US EPA's Guidelines for Carcinogen Risk Assessment (EPA, 2005) advises that:

"Statistical analysis of a long-term study should be performed for each tumor type separately. The incidence of benign and malignant lesions of the same cell type, usually within a single tissue or organ, are considered separately and are then combined when scientifically defensible (McConnell et al., 1986). Trend tests and pairwise comparison tests are the recommended tests for determining whether chance, rather than a treatment-related effect, is a plausible explanation for an apparent increase in tumor incidence. A trend test such as the Cochran-Armitage test (Snedecor & Cochran, 1967) asks whether the results in all dose groups together increase as the dose increases. A pairwise comparison test such as the Fisher exact test (Fisher, 1950) asks whether an incidence in one dose group is increased over the control group. By convention, for both tests a statistically significant comparison is one for which p is less than 0.05 that the increased incidence is due to chance. Significance in either kind of test is sufficient to reject the hypothesis that chance accounts for the result. A statistically significant response may or may not be biologically significant and vice versa. The selection of a significance level is a policy choice based on a trade-off between the risks of false positives and false negatives. A significance level of greater or less than 5% (the most common significance level) is examined to see if it confirms other scientific information. When the assessment departs from a simple 5% level, this should be highlighted in the risk characterization. A two-tailed test or a one-tailed test may be used. In either case a rationale is provided. "

4.5 Study designs

278. Strategic issues relevant to study design include the number of doses, their spacing, the choice of the top dose, the group sizes, the length of study and the choice of control groups. In particular, the choice of the number of dose levels and the dose spacing is crucial to achieving the objectives of the study (e.g., hazard identification or dose-response/risk assessment) and is important for subsequent statistical analysis. An important consideration in the design of studies is the statistical power of the study. This is the probability of detecting an effect of a certain size as statistically significant if it really exists. Most toxicological studies evolved from designs developed before formal power calculations were incorporated in to experimental design.

279. OECD TGs 451, 452 and 453 have a core design consisting of three treatment groups (with different dose levels) and one or more negative control groups for each sex. Each group should be at least 50 animals of each sex for TG 451 and 453 and at least 20 animals for each group for TG 452.

280. The OECD TG 453 recommends, for the chronic phase of the study, at least three dose groups and a control group, each group containing at least 10 males and 10 females per group.

281. There are different strategies for the allocation of resources (i.e., the animals) to the groups in the design depending upon the objective of a study. These strategies could range from the equal allocation between a single negative control and high dose group to maximize the power to detect a difference (hazard identification) to the allocation of single animals to a large number of different doses across the whole dose range. The first case would require an ANOVA-style analysis while the latter would use a regression analysis with a test for the lack of fit of the dose-response relationship. Intermediate designs could be different numbers of animals allocated to each of a number of different dose groups. This design can be analysed by the ANOVA methodology where the between group comparison can be broken down into linear, quadratic, other components and a lack of fit component. This above strategy reflects the continuum that exists between the ANOVA and regression modelling approaches.

282. There has been debate over whether a 4 group 50 animal/sex/group design should be replaced with an 8 group 25 animals/sex/group design. Unpublished work examined the power of the different designs, using three different scenarios to detect a linear trend in the proportions together with other dose selection issues using the nQuery Advisor software. These calculations made no assumptions about differential survival. This analysis showed that the power of the 8 group design was between 5 and 22% lower than the 4 group design. To achieve comparable power with the 8 group design, the 8 group sample sizes would need to increase by about 35 to 40%. However, both TG 451 and 453 recommend 50 animals/sex/group so that the alternative 8 group design is not recommended.

4.6 Control groups and length of study

283. The control groups can be either untreated or a vehicle control group. The animals in these groups are expected to be treated in an identical fashion to those in the test groups. A discussion of the implications for the statistical analysis of the inclusion of more than one control group can be found in the section 4.21 'Use of control data and dual control groups' (paragraphs 393-397). A control group of pair-fed animals may be included if the palatability of the substance administered in the diet is of concern (i.e., a continuous reduction of 20% of more in food intake).

284. All animals should be treated identically throughout the length of the study (see section 3.3.3 on study duration). At the end of the study a full detailed gross necropsy should be carried out on all the animals from the control and test groups.

285. Planned interim kills may be part of the design. It is suggested in the OECD TG 453 that a group of 10 male and 10 females per group (reduced from the original 20 per sex per group) should be included and that the terminal kill of these animals could act as an interim kill for the main carcinogenicity study. The justification for this reduction is that more information is available from the animals in the carcinogenicity phase. The proposed statistical analysis of the results from the interim kill should form part of the statistical analysis plan of the whole study to avoid introducing biases but the data should not be combined with the data from the main study.

286. The investigator should consider the implications of including such a small interim kill sub-group on the power of the study. Animals, such as those from the chronic portion of TG 453, should be allocated to the treatment group before starting the study based upon some randomisation process (see Bannasch *et al.*, 1986). Identifying these animals before the start of the study could, however, lead to them experiencing slightly different test environments such as this sub-study being maintained in a different room under slightly different conditions which might introduce biases into the statistical analysis. Therefore, precautions such as ensuring all animals are exposed to the same environmental factors and animal handlers being blind to which animals will be in the interim kill group should be instigated.

4.7 Purpose of statistical analyses

287. The objective of the statistical analyses of the data generated in long-term toxicity tests is to assess whether exposure to the test chemical is associated with adverse effects such as, for instance, an increased tumour incidence. The results of these analyses will then help achieve the other objectives of the study outlined in paragraph 272.

288. Statistical analyses can address this aim in two ways. On the one hand, the objective may be to test a hypothesis that one or more treated group is different from the concurrent group; alternatively, the objective may be to estimate the size of an effect in a comparison between groups and provide some indication of the precision or confidence that can be ascribed to that estimate.

289. As mentioned previously there are different 'schools of thought' about the statistical analysis of data. Much of the work in toxicology has been carried out based upon a traditional frequentist approach particularly around the concept of hypothesis testing. While recognizing that alternative viewpoints exist and this is a controversial area, most of the emphasis in this document will be on the traditional approaches.

290. There is a need to remain aware of the distinction between statistical significance and biological importance. Statistical analysis involves more than the reporting of the statistical significance of a hypothesis test. This may be one, often small component, of the much larger component of the design, analysis and interpretation of an experiment with statistical analysis being a part of the interpretation of the biological importance, not an alternative. Many statisticians argue against the reporting of significance levels arguing instead that the emphasis should be on emphasising the size of effects and the confidence in them. This avoids the problem of a small biologically unimportant effect being declared statistically significant and the artificiality of trying to dichotomise a result into a positive or negative finding on the basis of a P-value of, for instance, either 0.051 or 0.049. When reporting the results of significance tests, precise P-values (e.g., P=0.051) should be reported rather than referring to specific critical values.

291. The concept of statistical significance is an important component of the hypothesis testing approach. In a test of a null hypothesis the P value is a measure of how likely a result that has been obtained, or one more extreme, might have arisen if there were no difference between, for example, the two groups. The P value is dependent upon a number of factors: endpoint, variability, sample size, experimental design and statistical method. Conventionally, certain critical values ($P < 0.05$, $P < 0.01$ and $P < 0.001$) denote specific levels of statistical significance although many statisticians dislike this approach. Denoting something as statistically significant does not mean it is biologically important. (The use of the term biologically or clinically important is an attempt to avoid misunderstandings as the word 'significant' has a specific and precise meaning for statisticians.)

292. Similarly, declaring a result non-significant (often designated as $P>0.05$ or NS, again a nomenclature not favoured by statisticians) should not be interpreted as meaning the effect is not biologically important or that the null hypothesis is correct. Rather it means that there is not sufficient evidence to reject the null hypothesis.

293. Selecting an appropriate statistical method to analyse data is dependent on both the study purpose and the type of data. The specific tests applied will have different results because they are testing different hypotheses. A trend test will be testing whether there is a linear trend with a slope greater than zero; a pairwise comparison will be comparing whether a treated group is significantly different from the controls. In general, testing a trend which is a more specific hypothesis has greater power than a pair-wise comparison. It is also a single test compared with the 3 pair-wise comparisons between the dose groups and the negative controls. This introduces the concept of the use of corrections for multiple comparisons. Corrections are sometimes used to address concerns that when a

large number of comparisons (e.g., between pairs of treatments) are made that there is a risk of Type 1 errors. (A Type I error is the risk of wrongly rejecting the null hypothesis in a statistical test when, in fact, it is true and thus declaring results significant when they are not).

Parametric versus non-parametric methods

294. Deciding whether to use parametric or non-parametric methods to analyze data can be an issue, for each basic parametric test has a non-parametric counterpart (e.g., a pairwise comparison via Student's t-test or Mann-Whitney U-test; multigroup mean comparison via 1-way ANOVA or Kruskal-Wallis; trend test via a linear dose-response or Jonkheere-Terpstra test.)

295. Non-parametric tests have an advantage when some of the assumptions, particularly normality, underlying parametric tests are violated. They then give exact and accurate probabilities regardless of the shape of the distribution the data were randomly sampled from. However, while non-parametric tests may be distribution free, they are not assumption free so are as vulnerable, if not more so, to differences in the distributions between the groups. Non-parametric tests aim to ensure that correct Type I errors are derived but are less suitable for more complex designs, estimation and model fitting. When the assumptions underlying parametric methods are met their non-parametric equivalents have lower power and thus waste data. In the case of small sample sizes (e.g., 4 or 5 experimental units per group) comparisons using non-parametric tests have low power to detect even quite large treatment effects.

Limitations of statistical analysis

296. It should always be appreciated that a statistical analysis has its limitations. Statistical analysis cannot rescue poor data resulting from a flawed design or a poorly conducted study. Good experimental design, again the 'strategy' (paragraph 266), is the critical role of statistical analysis in a study. An appropriate data analysis will follow directly from a correct experimental design (including the selection of statistical methods to be applied) and implementation.

4.8 Types of data: Qualitative and Quantitative endpoints

297. Different types of data are collected in the course of a long-term bioassay. Endpoints can be qualitative or quantitative. The power associated with the detection of biologically important effects can be very different between a qualitative and a quantitative endpoint.

298. Qualitative data can be binary, categorical or ordinal. Examples of binary data are where the classification can take one of two (binary) forms: an animal can be dead or alive or have a tumour or not.

299. There are important issues about how pathological findings are described (see section 3.6.3.9). Sometimes these can be considered categorical (no ordering) or ordinal (where there is some ranking or ordering of the types). Issues arise as to whether findings are split into separate categories. This can alter the background incidence of the tumour types which affects power considerations. Splitting into a number of different separate categories of tumours also raises multiple comparison issues. Benign and malignant tumours should be analysed separately (McConnell *et al.*, 1986; US EPA, 2005) and, if it is considered scientifically defensible, further analysed using the combined incidences.

300. Further guidance on reporting of tumours and other histopathological lesions is provided in Section 3.6.3.9. Any final decision on the datasets to be analyzed should be agreed with the study pathologist.

301. Pathology data are categorical but can be converted into semi-qualitative or ordinal data. For instance, histopathology grading can be separated into categories of gradation from no effect, through mild to more severe. The assumption is that there is increasing severity but there is no assumption that the differences between classes are on a linear scale. (For instance, a change from a severity Category I to Category II may not be directly comparable to a change from Category II to Category III). Dichotomisation or other categorisation of continuous or ordinal variables may sometime be desirable, but categorisation will result in a loss of information.

302. A considerable amount of quantitative data is collected during the course of a long-term bioassay. Much of this is continuous data such as body and organ weights; clinical chemistry and haematological data (e.g., white blood cell counts) are also collected in chronic toxicity studies. In the case of carcinogenicity studies, the length of time either to the death or the identification of a tumour is a quantitative measure (used in, some cases, as a surrogate measure of the time until a tumour arises). There can also be quasi-continuous data where, although the data are discrete counts (e.g., the numbers of various types of blood cells), these counts are such large numbers that they can be analysed as if they are continuous data,

303. The specific statistical methods or tests used to analyse qualitative and quantitative endpoints are different. These have been represented as various times in a decision tree format with different flowcharts. An example developed by the OECD (2002, Guidance Note 35) previously will be discussed below (See pargraph 309 onwards). It is important to appreciate that while the statistical methodologies and algorithms differ, the underlying statistical concepts associated with the interpretation of the tests are basically the same. However, the power associated with the different endpoints within the context of the same basic experimental design can be very different.

4.9 Sample size and power considerations

304. The power of a study is the probability of detecting a true effect of a specific size or larger using a particular statistical test at a specific probability level. The power is $(1-\beta)$ where β is the Type II error associated with a hypothesis test. (The Type II error is the probability of wrongly accepting the null hypothesis as true when it is actually false while the Type I error is the probability of rejecting the null hypothesis when it is actually true.)

305. The power of a study for a qualitative trait depends upon 5 factors: the sample size (n), the significance level (α), whether the test is one- or two-sided, the size of effect of interest ($d = q-p$) where q is the incidence in the treated group and p is the control incidence and the type of statistical test. In the case of quantitative data; the proportions are replaced by the size of effect of interest and a measure of the inter-individual variability such as the standard deviation. Numerous software packages and programs are available for carrying out these calculations.

306. The OECD Test Guidelines indicate the appropriate sample sizes for each group. In the carcinogenicity study, the sample size is usually at least 50 animals of each sex at each dose level. This group size reflects a trade-off between the statistical power of the design and economic practicalities of the design. In practice, the carcinogenicity study has low power in the sense that treatment effects that might be considered biologically important cannot be detected routinely as statistically significant.

307. It is recognized that the power of the study can only be increased modestly by increasing sample sizes. Similarly, allocating animals differently between the test groups can increase the statistical power of detecting, for instance, an effect in a particular dose group (Portier & Hoel, 1983, 1984).

308. A common feature is that the power associated with qualitative data is less than that associated with quantitative data. The carcinogenicity study design, for instance, has low power for comparisons focussed on the low dose group and is only able consistently to detect large increases over the negative control incidence in tumour incidence with the power being reduced (further) if there is a high control incidence (Haseman, 1984).

4.10 Selection of Statistical Methods using Flowcharts

309. Flowcharts which provide a decision tree for the choice of statistical tests have been developed and used extensively in the analysis of statistical data in the biological sciences. There are obvious practical advantages in having a set of standard methodologies with the choice of particular methods made at key decision points based upon data. A number of examples can be found in textbooks. Gad (2006) for example has developed some for the analysis of toxicological data. The OECD produced a flowchart (Guidance Note 35, OECD 2002) in a previous document which is reproduced here (Figure 1).

310. The choice of the statistical method to use is based upon whether the data are qualitative or quantitative and upon the assumptions required by the test being met. The choice of route through the flow chart is based upon the results of answers to queries higher in the chart. For example, in the event of a test for non-normality of the data being statistically significant a non-parametric test may be chosen in preference to a parametric one. This methodology speeds the analysis and reduces the amount of valuable time a statistician spends on an analysis.

311. Gad (2000), for instance, suggested that statistical analysis of endpoints such as body and organ weight data are "universally best analysed by ANOVA followed, if called for, by a post hoc test." He suggests Bartlett's test is performed first to ensure homogeneity of variances. With smaller sample sizes he suggests that Kruskal-Wallis test may be more appropriate. In the case of clinical chemistry he points to the limitations of univariate analyses when the analysis must consider a battery of biochemical parameters to determine the overall effect.

312. Critics point out that, although there are efficiency gains in the application of flow charts, there is a 'de-skilling' of the task, an over-emphasis on significance testing for decision making and vulnerability to artefactual results. There is also the methodological problem of the use of a multiple testing procedure where one hypothesis test is used to choose another test which can complicate quantifying the true probability values associated with various comparisons.

313. Some concern has been expressed over whether tests for normality or for heterogeneity of variances are over-sensitive and, as a consequence, unnecessarily rule out the use of robust statistical methods such as the analysis of variance and, thus, potentially reduce the power of the design.

314. In the case of tests for normality, the null hypothesis is that the data are normally distributed. In practice, this can mean that a small sample which may deviate from a normal distribution may fail to 'trigger' a significant result, while a large sample with a slight deviation from a normal distribution may be considered significant and lead to a switch in the subsequent statistical test. Such test behaviour is counter-intuitive to what in practice is required in the selection of statistical tests and the switch between different statistical tests should not be automatic.

315. Treatment effects can result in heterogeneity in response between individuals leading to violations to the assumptions such as non-normality and heterogeneity of variances. Attempts to remove these violations by the use of transformations such as the logarithmic transformation are not always successful. One approach is then to apply both parametric and non-parametric tests to such data. It is important to realise that different statistical methods will produce different results (i.e.,

probability values) when they are applied to the same data sets. Concordant results then give increased confidence in the findings while discordant results 'flag' up anomalous data.

316. Other decision rules based upon trying to find an optimal transformation can lead to a very heterogeneous set of analyses based upon decision rules which lead on to another different transformation. The US FDA Redbook (FDA 2007), for instance, indicated that unnecessary transformation should be avoided as should the use of a mixture of parametric and non-parametric methods on the same endpoint.

317. The use of decision rule-based statistical analysis differs from the more current way of carrying out analyses using statistical models. In this there is an iterative process of fitting and testing models and checking assumptions.

318. Visual representation of data is also an important aspect of the analysis, relying on inspection of the data for outliers, trends, goodness of fit and checks of assumptions. Care should, therefore, be taken in carrying out statistical analyses using flowcharts. The usual considerations for the interpretation of statistical analyses should always be kept in mind.

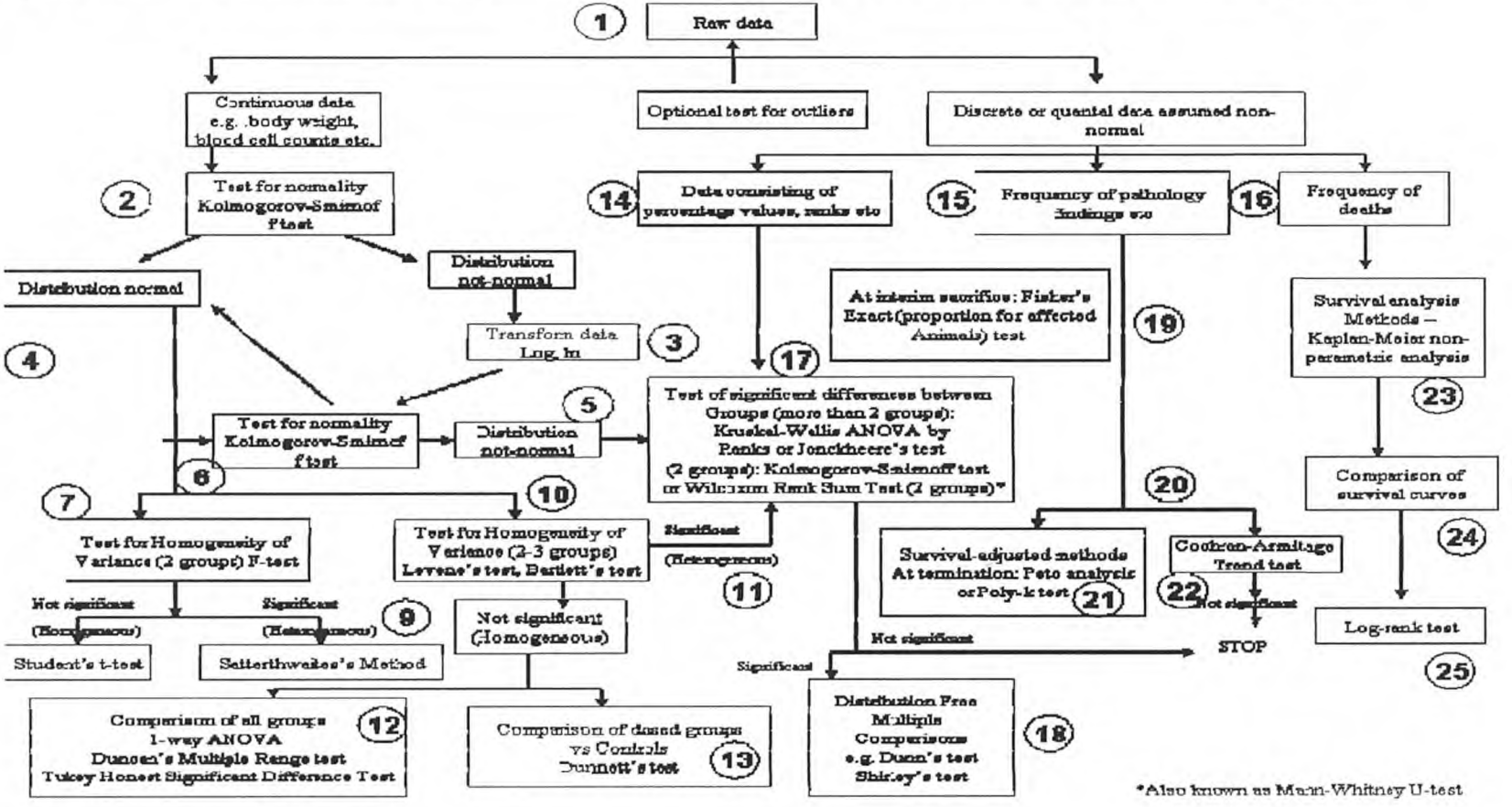
4.11 Description of the OECD flowchart

319. In this section Fig 1, a slightly modified version of the OECD (Guidance Note No 35; OECD, 2002, Fig 1) flowchart, will be worked through to illustrate the points that arise in its use. This will then be followed with some of the issues that arise if statistical analyses are totally dependent upon such an algorithmic approach. The objective, here, is to provide a brief overview of the methods used. The individual tests are described briefly in the glossary. This flowchart is similar to an approach used by the US National Toxicology Program (NTP).

320. For simplicity the description of the flowchart will be by working from left to right. This is for convenience and there is no priority implicit in this ordering. When an option is reached, the left choice initially will be described. Once the end of the 'tree' has been reached the description will move back to the previous node/decision point and continue down that. The same 'zig zag' procedure will be carried until the tests at the far right of the flowchart are reached.

321. In general, (but not precisely) the boxes at the top are concerned with aspects of the nature of the data. Moving down there are tests of the assumptions underlying the methods followed by 'omnibus' tests (which test for overall differences between groups irrespective of the specific design). These are followed by tests of whether there is a linear trend across the treatment groups and/or between treatment groups and the negative control group. In some cases there is a circular/interactive path when, for instance, following rejection of the assumption of normality the data are transformed to logarithms (natural or to the base 10) and the test for normality made again. The numbers in brackets refer to the circled numbers representing specific points on the flowchart. The assumption is that a decision is made to choose one or other route if the P value associated with the test statistic is <0.05. (Note that the use of hypothesis tests, the specific choice of P values, the choice of parametric or non-parametric and the use of multiple testing procedures are controversial issues and the particular routes outlined in this particular flowchart would not necessarily be agreed on by all statisticians.) Note that some of the statistical methods, particularly those related to the analysis of tumour data are discussed in more detail in paragraph 346 onwards.

Figure 1 A statistical decision tree, summarising common statistical procedures used for the analysis of data in long-term toxicology studies.



322. The top level of the flowchart (1) relates to a check of the data for overall quality and the identification of the type of data, whether quantitative continuous or qualitative or discrete. As part of the data checking an optional test for outliers such as the Dixon & Massey test is suggested. Moving to the methods included for continuous data, a test for the assumption of normality (either the Kolmogorov-Smirnov test or the Shapiro-Wilk) test is identified (2). If the test is significant indicating that the data are not normally distributed, the option is that the data are logarithmically transformed (3) and the test for normality carried out again as well as perhaps testing for outliers using the Extreme Studentized Deviate statistic (4). If neither transformation or identifying outliers results in normality the suggestion is to assume the distribution of the data is not normal and move (5) to the use of non-parametric methods (in the centre of the flowchart). If the data are assumed to be normal then a further test for homogeneity of variances (6) is suggested. In the case of a two group comparison this is an F-test (7). If the F test is not significant then the two groups are analysed by a standard Student's t-test (8); if the F test is significant then the comparison is by the modified t-test using Satterthwaite's method for unequal variances (9). Returning to point (6): either Levene's test or Bartlett's test are used to test for homogeneity of variances (10) when there are three or more groups. If the variances are considered heterogeneous the flowchart directs the analysis to non-parametric methods (11). If the variances are considered homogeneous then the comparison of all groups is suggested to be firstly by a one-way analysis of variance (1-way ANOVA) (12) then followed by a decision whether to proceed to a multiple comparison procedure such as Duncan's multiple range test or Tukey's Honest Significant Difference test. A number of other multiple comparison methods are available such as the Williams multiple comparison test (Williams, 1971). In the case of pair-wise comparisons between the control and the dosed groups, the flowchart suggests Dunnett's test (13) following the one-way analysis of variance.

323. Returning to the qualitative data (1) a distinction is made between data which are ranked or are discrete counts of an endpoint within the experimental unit (14) and those which are pathology findings (15) and those related to death or survival (16) The pathology findings and number of deaths may be summarized as frequencies or percentages for the different experimental groups. It is suggested that ranked data, together with quantitative data which were either determined to be non-normally distributed (5) or have heterogeneous variances (11) are analysed by non-parametric methods (17). The suggested methods for comparisons are the Kruskal-Wallis test for comparisons between the groups and Jonckheere's test for a trend in the data. Methods identified for comparisons between the control and test groups are the Kolmogorov-Smirnoff test and the Wilcoxon Rank Sum Test (which is equivalent to the Mann-Whitney U-test). If these tests are significant then further testing using distribution free multiple comparison tests can be carried out using tests such as Dunn's or Shirley's tests (18). Returning to pathology findings (15), if an interim kill is carried out comparisons between the proportions of animals with pathological findings in a treated group can be compared with the proportion in the control group using Fisher's exact test (19). An overall test of differences in proportions between the groups can be carried out using the chi-square test of heterogeneity outlined in paragraph 341. At the end of the study a choice is made (20) between tests which take into account information on how long the animals lived without a tumour (21) and those that do not such as the Cochran-Armitage test for a trend in proportions (22). The survival adjusted tests (21) are the Peto analysis (which requires information about whether the tumour is 'incidental' or 'fatal') and the poly-k test which does not need this information. (These methods are described in more detail in the survival adjusted analyses section of this Chapter, paragraph 346 onwards.)

324. Returning to qualitative data (1) and to data on survival/death (16), the flowchart identifies survival analysis approaches such as the Kaplan-Meier non-parametric methods (23) followed by comparison of the graphs of the survival curves (24) followed by analysis using the log rank test (25).

4.12 Intercurrent mortality

325. Inter-current mortality is death that arises during the course of a study from anything other than a tumour. Chronic studies can last up to two years. Animals can and do die in the course of the studies for a variety of reasons both related and unrelated to the treatment they have received. Inter-current mortality is different from the planned interim sacrifices or kills specifically included in some designs. Inter-current mortality complicates the statistical analysis of comparisons between test groups. For instance, older animals are more likely to develop a tumour than younger ones. The risk of getting a tumour and of dying because of a tumour increases with age. Consequently, the probability that an animal that dies unexpectedly during the course of a study also has a tumour will depend upon the animal's age at death. The test chemical may also affect the survival of different groups by causing either more deaths (through non-tumour related toxicity) or fewer deaths (by, for instance, reducing food intake and making the animals less susceptible to obesity-related morbidity and mortality). Peto *et al.* (1980) point out that comparisons based upon crude tumour rates in the presence of differential mortality can produce serious errors in the analysis.

326. A simple statistical analysis which does not account for inter-current mortality (described in paragraph 341) can underestimate the carcinogenic effects if the treatment decreases survival. Conversely, if the treatment increases survival then the tests may overestimate the carcinogenic effects. Failure to take intercurrent mortality into effect can, therefore, produce serious biases in the interpretation of results. Peto *et al.* (1980) state "the effects of differences in longevity on numbers of tumour-bearing animals can be very substantial, and so, whether or not they (the effects) appear to be, they should routinely be corrected when presenting experimental results" They argued that to avoid this problem from occurring, adjustments are needed for differences in survival between the groups and this correction should be routinely used.

327. Inter-current mortality, if not related to the tumours, is a serious problem for the interpretation of chronic studies because a high number of animals in the groups need to survive for a successful analysis (see paragraphs 158-159 for a more detailed discussion).

328. There are occasions when inter-current mortality may not be a major problem. However, it is important that any organization carrying out such an analysis should contact the relevant regulatory authority for advice in the event of problems with survival so as to try to ensure that the optimum amount of information is obtained from any study which has to be stopped early.

4.13 Cause of death (COD) and Context of observation (COO)

329. The context of observation (COO) relates to whether the tumour is considered fatal or incidental. Cause of death (COD) relates to the specific cause of death of the animal.

330. Information on the cause of death (COD) of the animal and the 'context of observation' of a tumour (COO) (Peto, 1974)) may be important for subsequent statistical analysis of a study (step 21 in the flowchart). Peto *et al.* (1980) argue that a distinction needs to be made between whether a tumour is called 'fatal' or 'incidental' for a correct statistical analysis to be carried out. The 'context of observation' (COO) distinguishes between whether a 'fatal' tumour is one that is considered to have caused the death of the animal; an 'incidental' tumour is found when the animal died from an unconnected reason or was found at the terminal kill at the end of the experiment. Distinguishing between these two COOs can be both difficult and controversial. However, this distinction is a critical feature of the 'Peto' analysis described in the 1980 IARC monograph (Peto *et al.*, 1980).

331. Complications arise because this dichotomy is not always simple. Information on the COO may not be available because it is not provided by the pathologist or is unreliable because it is based upon assumptions about lethality which may be contentious.

332. A pathologist may not be able to determine the COD of an animal. Some pathologists have argued that it is difficult retrospectively to diagnose accurately if a tumour is the true cause of death of an animal. There may not be a single factor which solely determines the death; there may be multiple causes of death with the presence of a tumour being just one of them. Alternatively, more than one tumour may contribute to the cause of death.

333. Peto *et al.*'s (1980) approach to this problem was to suggest four categories of tumour COO - (1) probably fatal, (2) possibly fatal, (3) possibly incidental and (4) probably incidental - with the suggestion of combining the categories in different ways using different cut-offs to produce a binary endpoint: e.g., combining 1+2 versus 3+4; 1+2+3 versus 4 and 1 versus 2+3+4 and analyzing each combination as a form of sensitivity analysis. This has proved controversial because it is argued that it is a device to subsequently reduce the categories back to two: 'fatal' and 'incidental'.

334. Lee & Fry (cited in Lee *et al.*, 2002) point out that the 'fatal' definition is often misunderstood by pathologists. The key issue is that a 'fatal' tumour is wholly responsible for the death of the animal not that the tumour had the potential to kill the animal at some time in the future. Lee and Haseman (cited in Lee *et al.*, 2002) also disagree on how easy it is for the pathologist to make a distinction between the COOs. Lee quoted Peto's example of the high proportion of definitive diagnoses in the BIBRA nitrosamine study where 94% of over 4500 tumours could be classified as either "definitely incidental" or "definitely fatal," even though the pathologists had initially expressed reservations about whether such classifications could be made reliably. Haseman (cited in Lee *et al.*, 2002) argues that there were instances where it was difficult to make the distinction and when made it was often incorrect. There was also a tendency for the over-designation of fatal tumours (Abu *et al.*, 2001; Kodell *et al.*, 1982). Haseman argues that pathologists should be allowed the freedom to make their own judgements. Soper and Kodell (cited in Lee *et al.*, 2002) argue for a more objective classification based upon the large historical data base available.

4.14 Time to tumour onset

335. The time until an event such as death occurs can be analyzed by the statistical technique of survival analysis. There is much widely available statistical software that can be used to carry out the standard analyses. Such methods are used to analyze the survival data from the carcinogenicity study (see paragraph 346 onwards).

336. Analysis of time to tumour data, however, is less straightforward. Ideally, the time when a tumour first occurred is needed so that these times can be compared between treatment groups. There are some tumours which can be observed during routine observation of an animal such as some skin and some palpable mammary tumours. In practice, the time, when a tumour is identified is an arbitrary but objective endpoint such as recognizing when the tumour reaches a certain size. These tumours are called 'mortality independent' and can be analysed by standard life-table survival analysis methods.

337. However, these are the exception. Most tumours are internal (and, therefore, hidden or 'occult') and will usually be detected only during a post mortem examination of the animal. The specific time (the age of the animal or the tumour onset time) when the tumour initially arose is consequently unknown. It cannot simply be replaced by the time when the tumour was first identified as an 'occult' tumour (in other words the time of identification of an incidental tumour is not a surrogate for the time of onset). This lack of information complicates analysis. Statistical methods

which take into account the time until an event occurs (such as survival analysis methods) need to make a number of assumptions for analyses of such 'occult' tumours.

338. Dinse (1994), for instance stated: "without direct observations of the tumour onset times, the desired survival adjustment usually is accomplished by making assumptions concerning tumour lethality, cause of death, multiple sacrifices or parametric models" (Dinse, 1994). Approaches to get around this problem are to model and impute the onset times, to include extra data (context of observation) or to make more assumptions (poly-k test)" (see paragraph 361).

339. One suggested approach to the problem is to increase the sample size and have planned interim kills/'sacrifices'. In practice, this is rarely done and there are no universal guidelines for the analysis of such studies.

4.15 Standard (simple) statistical analysis of qualitative data

340. Tests for pair-wise comparisons and trends are recommended for the statistical analysis of tumour data to test for treatment-related effects (EPA, 2005). There are a set of standard statistical methods for comparing proportions, such as tumour incidence, between one or more groups. In these tests the basic information is the number of animals at risk as the denominator and the number of animals with the tumour (or pathology) as the numerator of interest. These tests make no assumptions or corrections to take into account 'COO' or 'COD' information or the time to tumour or death. They are described in many standard statistical textbooks and software for analysis is widely available.

341. The three main tests are:

1) A pair-wise test between the negative control and the treated group using the Fisher exact test (Fisher, 1950).

2) A chi-square test for heterogeneity of proportions between groups. This is an 'omnibus' test of differences between a series of groups (with no ordering to the groups).

3) A test for a linear trend such as the Cochran-Armitage trend test (Snedecor & Cochran, 1980).

342. Pair wise comparisons are carried out using Fisher's exact test or its Chi-square approximation. Fisher's exact test is now preferred because of the availability of software for carrying it out.

343. The Cochran-Armitage trend test aims to detect a linear trend. The null hypothesis is that all animals are at equal risk of developing a tumour during the study. Problems arise if there are differences in mortality between the groups. The test is sensitive to increases in treatment related lethality and this leads to an incorrect level of the Type 1 error (the risk of falsely rejecting the null hypothesis).

344. The selection of the dose metric (the values representing the actual doses used) for use in the analysis is important. The doses could be the 'applied', logarithmic, some equally spaced rank or some measure of the effective dose at the target organ (based upon pharmacokinetics). A practical issue arises using a logarithmic scale with the need to choose a value to substitute for the zero dose level.

345. A trend test is more powerful than the pair-wise test. A complication is that a trend test may fail to detect curvi-linear responses such as might arise from non-linear effects such as complications from saturation. In such situations the pair-wise tests may give more appropriate results.

4.16 Survival adjusted analyses

346. For the reasons discussed above, survival adjusted methods are strongly advocated for comparisons of tumour incidences among groups.

347. Some exact probability test methods are available for use instead of the approximations based upon the normal distribution used in the standard methods. The standard methods may underestimate p values when the numbers of tumours in the groups are small. In these situations exact permutation tests which are extensions of the Fisher exact test are suggested for use (Lin, 2010).

348. Three different types of statistical procedures have been developed depending upon the type of tumour.

1) The prevalence method for non-lethal (incidental) tumours (Hoel & Walburg, 1972; Peto *et al.*, 1980)
2) The death rate method for lethal (fatal) tumours (Tarone, 1975; Peto *et al.*, 1980)
3) The onset–rate for mortality independent (observable) tumours (Peto *et al.*, 1980)

The prevalence method

349. The prevalence method, also called the incidental method, is, effectively the Hoel-Walburg procedure for nonlethal tumours which makes no assumption about the COO of the tumour (Hoel & Walburg, 1972).

350. The procedure involves carrying out a life-table analysis, a method for analysing censored observations that have been grouped into intervals, for tumours which were found incidental (incidental context) to the death of the animal. The experimental period is split into a set of intervals (including any interim or terminal sacrifices).

351. It has been argued that the choice of partitions is not critical. Peto *et al.*, (1980) suggest the partitions should not be so short that the prevalence of incidental tumours is unstable nor so large that the prevalence could differ markedly from one half of an interval to the other and suggest a partition based upon 'ad hoc' runs. However, there have been concerns about constructions of ad hoc intervals in the Peto analysis and attempts made to standardize them. The US National Toxicology Program (NTP) for instance previously used intervals 0 – 52, 53 – 78, 79 – 92, 93 – 104 plus the terminal kills.

352. In the analysis the denominator is the number of animals dying within the specific partition and the numerator is the number of animals dying with an incidental tumour. The analyses for each individual partition are combined using the Mantel-Haenszel method. This compares two groups for proportions with an adjustment for control variables. A series of k 2 x 2 contingency tables are produced with k being the number of strata of the different control variables such as age or sex. The stratification increases the power of the design to detect an association. The test statistic is a chi-square. Full details of the equations used to conduct it are found in Lin (2010).

353. The methodology uses normal approximations for the tests but the approximation may be unreliable if the number of tumours in a group is small. A permutation test involving the hypergeometric distribution may be used (Lin, 2010).

The death rate method (for comparing rapidly fatal tumours)

354. This test is also referred to as the log rank test assuming that all the tumours cause death (Peto, 1974). It is used when tumours are observed in a fatal context. In the analysis the stratification is based upon a partition into intervals (often a week) where one or more animals died. In each

stratum the number of animals entering the partition or strata who are tumour free is the denominator and the number of animals dying during the partition with a fatal tumour is the numerator. The analyses for each individual partition are combined using the Mantel-Haenszel methods. Full details of the method and equations are found in Lin (2010).

Mortality independent analysis (onset rate method)

355. In those tissues such as skin and the mammary gland where tumours can be observed in the live animal, the tumours are described as mortality independent. The onset rate method (the same basic statistical method as used for the death rate or fatal analysis) is used for these mortality independent tumours. The endpoint in this case is the occurrence of a tumour based upon it reaching some predefined size rather than the death of the animal. Once a tumour has been identified using this criterion, the animal makes no further contribution to the analysis. It is no longer 'at risk' of developing a tumour because it now has a tumour even though it may live on for some time. The calculations used to produce the statistics for the onset rate methods are comparable to those produced by the death rate method.

Peto / IARC analysis

356. The Peto analysis, as described in an IARC monograph (Peto *et al.*, 1980), is a combination of the prevalence and death rate based upon the COO of a tumour as either being called non-lethal ('incidental') or the cause of the animal's death ('fatal'). It is a joint test of age-adjusted tumour lethality and age-adjusted tumour prevalence. The assumptions underlying the method are that the control and treated animals are equally likely to be killed at any particular stage in a tumour's development, the animals dying of other causes are representative of all animals surviving in that interval and that the pathologist is prepared to make the 'fatal'/'incidental' classification.

357. The analysis combines the two separate approaches. Animals whose tumours are termed 'fatal' are analysed by Peto's death rate method while those called 'incidental' are analysed using the prevalence method. The two analyses are then combined using the Mantel-Haenszel method to provide a test for trend. A problem is that the analysis will be biased if the assumption on the nature of the tumour being either called 'fatal' or 'incidental' is inaccurate (Dinse, 1994). Although the Peto test is described as robust, there is some debate as to how big a problem misclassification is.

358. Peto *et al.* (1980) provide an illustration of the implications of wrongly defining the COO. For example, in a study of pituitary tumours in animals treated with N–nitrosodimethylamine (NDMA) different conclusions are drawn depending upon the COO. If all the tumours were considered fatal then NDMA was wrongly considered carcinogenic, whereas the tumours were considered incidental then NDMA was wrongly considered health protective. Taking the COO into account produced what Peto *et al.* considered to be the correct interpretation that there was no carcinogenic effect in the pituitary.

Multivariate regression methods

359. Statistical methods used for survival analysis such as the Cox proportional hazard method have been used for the analysis of carcinogenesis data. This is a regression method which takes into account the time until the binary event of interest (death) occurs. The modelling also allows the effect of an exposure/treatment to be investigated after adjusting for confounding effects. The method was applied when the tumours were considered lethal.

360. Dinse and Haseman (1986) suggested a logistic regression approach which was applicable when the tumours were considered incidental, Logistic regression is a method used where the

outcome is binary such as whether an animal has a tumour or not. The effect of an exposure on this binary outcome can be adjusted for confounding factors in a study.

Poly-k test

361. More recent approaches have been the development of the poly-k tests (Bailer & Portier, 1988; Dinse, 1994). The poly-k test does not need arbitrary partitions of time periods or COO information. The test is based upon the assumption that the time to tumour onset can be modelled based upon the tumour onset times raised to the power k. Initially, the test was proposed without identifying how to derive k but now it is suggested that k should be 3 because of observations that tumours can be modelled by a polynomial of order 3 from an analysis of NTP historical control data for F344 rat and B6C3F1 mice (Portier *et al.*, 1986.) The poly-3 test is then a special case of the poly-k test. The power of 6 (or k=6) can be used when the tumour onset times are close to a polynomial of order 6. The value of k need not be critical as poly-k tests are reported to give valid results if the true value lies between 1 and 5 (Bailer and Portier, 1988).

362. The tests are, in effect, modified Cochran-Armitage tests which adjust for differences in mortality in the treated groups by a modification of the number of animals in the denominator to reflect the less than whole animal contributions because of reduced survival. The approach gives a value (w) from 0 to 1 for each animal based upon a weight which relates to the time of death or the time of the final sacrifice. Thus w relates to the fraction of the length of time the animal survived in the study over the total length of the study to the power k. The value w is <1 if the animal died early without developing a tumour and w = 1 if the animal died with a tumour or survived until the study was completed. The number of animals at risk is replaced by a new estimate in the Cochran Armitage test. The method tests for a dose-related trend in the mortality-adjusted lifetime tumour incidence rate.

363. The Bieler-Williams variance (1993) is used in the poly-3 test (which is also sometimes referred to the Bieler-Williams method/test) where the test is modified by using the delta method and weighted least squares techniques in order to adjust the variance estimation of the test statistic and to improve the performance of the test.

Comparison between Peto and Poly-k methods

364. Rahman & Lin (2008) compare the false positive rates of the Peto and poly-k tests using a simulation study. Kodell, as cited in Lee *et al.* (2002), compared the properties of the Peto and poly-k tests and concluded that both are valid for adjusting for differential mortality. The problem remains that the comparison is ideally based on tumour onset data but because most tumours are occult, the tumour onset cannot be actually observed. In addition, there is a need to adjust the analysis for any differences in inter-current mortality.

365. The NTP routinely used to carry out two trend tests. One assumed that all tumours in dead or moribund animals were 'fatal'; the other assumed all the tumours were non-fatal ('incidental'). The current NTP approach is to no longer use life-table tests or prevalence tests. Instead, the poly-3 test with Bieler-Williams variance with a trend test and pair-wise tests with controls is used. Sometimes this test is used with k=1.5 and/or k=6.

366. Debate over the methods for the analysis of data on the COO of tumours using the Peto analysis has generated controversy. The Society of Toxicologic Pathology (STP) set up the STP Peto Analysis Working Group and produced draft recommendations for the classification of rodent tumours for the Peto analysis. Instead of the Peto analysis the STP Working Group recommended that the poly-3 methodology be used (STP 2001).

367. STP concluded:

1) Pathologists cannot determine the time of onset accurately from post mortems.

2) If the Peto analysis uses death as a surrogate for time of onset then the method seems inappropriate

3) Better to use other methods which do not require COO information

368. Lee & Fry (cited in Lee *et al.*, 2002) responded to the STP Peto Analysis Working Group recommendations and their comments together with those of other statisticians were published as a collection of comments in Toxicologic Pathology (cited in Lee *et al.*, 2002).

369. Kodell (cited in Lee *et al.*, 2002) concluded that both the Peto and poly-3 tests are valid for adjusting for differential mortality. Both methods are fairly robust to deviations from their assumptions although both could be improved by modifications and further development.

370. Based upon these comments particularly relating to issues around the use of the Peto method and the onset of fatal tumours, the STP withdrew its criticism of the Peto approach while maintaining that the poly-3 test is appropriate in certain circumstances (STP, 2002).

371. The STP (2002) put forward new recommendations that:

- The Peto test should be performed whenever study pathologists and peer review pathologists can consistently classify neoplasms as fatal or incidental

- If fatal and incidental classifications are not applied, the Poly-3 or another alternative to the Peto test should be employed

4.17 Tests of difference in survival

372. As discussed earlier (see paragraph 326), differences in survival can affect the conclusions drawn from the simple analyses. A number of methods can be used to test for differences in survival and for significant dose-response relationships or trends in studies. These include: the Cox test (Cox 1972; Thomas, Breslow, and Gart, 1977; Gart *et al.*, 1986); the generalized Wilcoxon or Kruskal-Wallis test (Breslow 1970; Gehan 1965; Thomas, Breslow, and Gart 1977); and the Tarone trend tests (Cox 1959; Peto *et al.*, 1980; Tarone 1975).

373. The number of animals surviving until the scheduled terminal kill can be compared by what is termed the product-limit or Kaplan-Meier method (Kaplan and Meier, 1958) and plotted graphically. This analysis compares length of survival in the groups (and does not involve any pathology findings.) Animals that are found dead other than from natural causes are usually treated as censored and excluded from the survival analysis. There is a pair-wise comparison method (Cox, 1972; Tarone & Ware, 1977) and a test for linear trend, Tarone's life-table test (Tarone, 1975).

374. The Kaplan-Meier analysis involves calculating the ratios of the surviving animals divided by the numbers of animals at risk of dying. Every time an animal dies the ratio is recalculated. These ratios can be plotted to show a curve which displays the probability of survival. When there are different treatment groups a curve can be generated for each group. Formal statistical tests such as the log rank test can be used to test difference between groups. (Curves which are close together may indicate that a difference is not statistically significant.)

4.18 Assumptions for statistical analysis

375. In the context of the theme of this document, it should be reiterated that the design of the experiment is fundamental to the choice of statistical methods.

376. Statisticians have long had a specific interest in the methods used to allocate animals to treatment, sample size determination, the control of possible confounding effects such as cage location and their rotation, the dose selection and the length of the study (Haseman, 1984).

377. An important consideration is the need to balance the considerable experience built up in the use and interpretation of the data from these experimental studies over many years with the appreciable empirical knowledge that exists on the implication of violating the assumptions associated with the statistical methods.

Randomisation

378. One assumption underlying the design and subsequent statistical analysis such as the ANOVA is that the animals have been assigned at random to the treatment groups. Each animal entering the study should have the same chance as any other of being allocated to one of the experimental groups (including the interim kills and control groups). Randomisation can be carried out by a number of different methods. The primary objectives are to prevent bias and to ensure that uncontrolled covariates do not affect the results of the analysis.

379. Stratified randomisation with groupings based upon body weights may be used to reduce bias and ensure compatibility of the various treatment groups with respect to uncontrolled variables. Ideally the animals from all groups should be placed into the study at the same time. If this is not practical, strata or blocks can be included as factors in the subsequent statistical analyses and can be created by starting subgroups from the control and each of the treatment groups over several days.

380. An assumption in the statistical analysis is that randomisation occurs at all points in a study so that biases are not introduced. If, for instance, the animals are initially randomised into groups but subsequent procedures are carried out in a pre-defined systematic order there is a risk that biases may be introduced.

Independent experimental units

381. The animal is often both the experimental and observational unit. If a cage is assigned to a treatment it becomes the experimental unit. However, the common statistical methods used in the analyses make the assumption that the experimental units are independent. In some cases this is clearly not so. In practice, although the assumption of independence of experimental units is an important one (more important than normality and homogeneity of variances; van Belle, 2009) it is not always taken fully into account in toxicological studies.

382. Individual housing is preferable to meet statistical assumptions but may have implication for the welfare and the representativeness of the animal. The trade off is between a theoretical optimum design, the avoidance of cage/confounding effects and practical husbandry issues, the 'pathology' of single housing and the possible increased variability of isolated animals. Litter and caging effects should be taken into account in the statistical analysis. If this cannot be done this should be noted with an explanation of why doing this is not possible, and the potential implications for the study.

383. Lack of independence may also arise from contamination of doses within and between cages. In the case of airborne contamination the random assignment of cages will mean that some of the lower dose and control groups will be exposed to some or higher concentrations than is implicit in

the experimental design. There is clearly a need to balance the potential limitations resulting from contaminations arising from randomisation in a room (but which also provides some protection from this uncontrolled variable) with the potential confounding effect of separate handling and housing of dosed groups necessary to prevent cross treatment group contamination.

4.19 One- or two-sided tests

384. The choice of whether to use a one- or two-side test should be made at the design rather than the analysis stage. A two-sided statistical hypothesis test tests for a difference from the negative control (in a pair-wise comparison) in either direction. A one-sided comparison tests for a difference in only one pre-specified direction, but as a consequence has more power. In a carcinogenicity study, the expectation is often that the change will be an increase in tumours in the treated group so a one-sided test may be considered more appropriate, although this can be controversial. If the treatment could also be protective (i.e., reduce tumour incidence or delay it) then a two-sided comparison may be more appropriate. Regulatory authorities may have specific opinions. For instance, the US EPA (2005) notes that either "a two-tailed test or a one-tailed test may be used".

Equal information per unit

385. There are implications for the assumptions underlying the statistical analysis in the non-random reading of histopathology slides. In some studies more effort has been put into reading slides from the control and top dose groups with less emphasis on the examinations at the intermediate doses. This approach was more common when qualitative hazard identification was the objective. Although such studies have been accepted by some regulatory authorities, this method can create problems in the statistical analysis of dose-response trends and cannot be recommended if dose-response characterization is an objective of a study. Similar considerations arise if, because of uncertainty in a diagnosis, more slides are read for some animals than others.

Blind or unblind reading of slides

386. Blinding procedures such as for the reading of slides is discussed in paragraph 258 (section 3.6.5). Many statisticians expect blinding to treatment group to be included, as in clinical trials as a protection against biases. Temple *et al.* (1988) have argued that blinding should be used to guard against biases that might arise (inadvertently) from the pathologist knowing the treatment group of the animal from which the tissue was derived. However, Haseman, in the STP discussions (cited in Lee *et al.*, 2002) argues from a statistical point of view, that blinding is not necessary given the long experience of using the assay. Blinding is, however, often used when re-reading slides in disputed cases or when the results are close.

Confounding variables

387. One of the purposes of randomised and blinded studies is to minimize the effects of uncontrolled covariates and to prevent the introduction of biases by preventing confounding factors from distorting the results. A confounding variable is one which is so closely related to both treatments (or another factor in the study design) and effects such that the individual contributions to the effect of interest cannot be separated. Haseman (1984) discusses a number of confounding effects such as cage location and litter mates although the latter should be controlled by randomisation. Butterworth *et al.* (2004) discuss another form of confounding where, rather than the chemical under study, a contaminant may be responsible for the carcinogenicity detected.

388. In the case of long-term bioassays, early differences in body weight between control and treated animals which persist through the study create a potential confounding effect between body

weight, life span and tumour incidence. Ad libitum overfeeding is the most important uncontrolled variable which affects the results of a long-term bioassay. Keenan *et al.* (1996) report, for instance, that there is a highly significant correlation between food consumption, body weight / obesity and shortened life-span in rodents. Kodell *et al.* (2000) notes that the reduced survival of Sprague Dawley rats questions the continued use of this stock in the carcinogenicity study.

389. If the tumour profile of lean and obese rodents is different, then there is a possibility that apparent treatment related differences in tumour incidence may, in fact, be wholly or partially caused by the body weight differences. Confounding may then make it difficult to identify if an effect was a direct result of the treatment or an indirect effect of the treatment through affecting food intake and consequently body weight. Ibrahim *et al.* (1998) noted that there is usually not enough information to provide a regression based adjustment for differences in body weight.

4.20 Interpretation of statistical analyses

390. Interpreting the results of a long-term bioassay is complex. A critical issue is the practical problem of the low power of the design when the tumour incidence is rare together with the multiple comparisons issues arising from the investigation of 20 or more tissues from both sexes of two species. As a result, there is a risk of both Type I (false positives) and Type II (false negative) errors. It is also necessary to integrate the results of the full battery of statistical tests, significant or otherwise, and the importance of a series of biological issues in the assessment of the result. Factors which add importance to findings include: uncommon tumours, multiple sites, positive findings using more than one route of administration, effects in multiple species/ strains/ sexes, effects at the same site in both sexes and/or species, tumour/disease progression, increased preneoplastic lesions, reduced latency, metastases, unusually large responses, dose-related responses and a high proportion of malignant tumours. In addition, a comparison of the observed tumour incidence in the study with the historical control rates should be made.

391. Various approaches are taken to control the false positive rate. Haseman (1983) developed a decision rule. This approach is based upon comparisons between the top dose and the negative control using data from the standard NTP experimental design studies using F344 rats and B6C3F1 mice. The rule uses a criterion of a statistically significant difference at $P<0.01$ for common and <0.05 for rare tumours (Lin & Ali, 1994). The definition of a rare tumour is an incidence of <1%, based on historical controls. Above 1%, the tumours are considered common. It has been argued that Haseman's rule produces too high a false positive rate because all treatment groups rather than just the top dose are, in fact, used in the comparisons. Any decision rule used to control the false positive rate should be justified.

392. A study where no treatment–related effects are found should be reviewed to ensure it is valid by checking that sufficient numbers of animals lived long enough to ensure that adequate exposure had been achieved and so were 'at risk' of developing 'late in life' tumours. A check should also be made that the objectives of the study design in terms of dosage have been achieved (see section 3.1 on dose selection).

4.21 Use of control data and dual control groups

393. OECD TGs 451/452/453 specify that a concurrent control group should be included in the study design. Study designs where there are more than one group of concurrent controls are rarely necessary, but when used, come in two basic types: firstly, where the two control groups are treated differently, e.g., vehicle contol and an untreated control group or a pair fed control vs ad libitum (see paragraph 283); secondly, where there are two identical control groups (Haseman *et al.*, 1986).

394. With the first type, the objective is to see whether differences between the 2 control groups have any effects on tumour incidence or any influence on any other toxicological effects in the control animals. These different control groups should not be combined for use in statistical analyses. The concurrent control group that only differs from the test groups by the absence of the test substance is the most appropriate for the comparison with the test groups.

395. The argument made for the second type, the identical dual controls (C1 and C2) is that they provide a way of identifying the degree of variability in the negative control animals. This gives a better basis for addressing the biological importance of any increase found in the treated groups. They can be considered as 'contemporary' historical control data.

396. Data from the two dual control groups (C1 and C2) can usually be combined (Haseman *et al.*, 1990) but if differences are found either in mortality or tumour incidence then it has been suggested that three tests with the treatment group (T) could be carried out (C1 v. T) (C2 v. T) and ((C1+C2) v T).

397. Some have argued that if the C1 and C2 groups differ then comparisons with the test groups are only considered positive when comparisons with each control group individually is significant on the grounds that findings should be reproducible. However, if it is accepted the carcinogenicity study is underpowered, then the comparison would be considered positive if any of the comparisons of the treated group with the control groups were significant. The first approach risks more false negatives, the second approach more false positives.

4.22 Historical control considerations

398. Historical control data can help interpret results in a number of situations (see GD 35). In any discussion about historical control data, it should be stressed that the concurrent control group is always the most important consideration in the testing for increased tumour rates. The historical control data can, though, be useful provided that the data chosen are from studies that are comparable with the study being investigated. It is widely recognized that large differences can result from disparities in factors such as pathology nomenclature, strain, husbandry, pathologists.

399. It has been suggested that historical control data should only be used if the concurrent control data are appreciably 'out of line' with recent previous studies and that only historical data collected over the last 5 years should be used. Such historical control data can be helpful in evaluating how 'normal' the data from the concurrent control groups are, for evaluating differing results from the dual control groups and as a form of quality control for carcinogenicity studies. Any concerns over the appropriateness of the control groups need to be evaluated and discussed.

400. Elmore & Peddada (2009) discuss how to incorporate historical control data into the statistical analysis of the carcinogenicity study. The mean and SD can be affected by a 'rogue' outlier while the median and interquartile range (IQR) is not. They argue outliers should be identified and not discounted but considered alongside other relevant data in the assessment of the results. They suggest the use of exploratory methods such as box and whisker plots (with their associated 5 number summaries) to give graphical presentations of historical control data (when there are more than 15 studies) together with the results of the treated groups and the concurrent negative control in a study. They illustrate the advantages of using the median and quartiles over the mean, SD and range especially to identify potentially misleading outlying results. They recommend using Bailer & Portier (1988) survival adjusted tumour incidence rates.

401. There are a number of statistical methods developed for incorporating historical control data into formal statistical analysis. Although this is an area of active statistical research, these methods have not been used in a regulatory context.

402. A number of methods have been described for incorporating historical control data into the statistical analysis of a trend in the data. One suggestion has been to use the upper confidence limit on the binomial proportion to help in interpretation. Tarone (1982) developed a method using the beta-binomial distribution (a binomial distribution where the value of p is a random variable rather than a single fixed value) to account for the variability between studies to model historical control data and to derive both exact and approximate tests. Including the historical control data increases the power of tests, especially for comparisons with rare tumours but the method does not take into account differential survival. Ibrahim & Ryan (1996) have developed a test which uses historical control data in survival-adjusted tests. The study period is split into intervals and in each of these the multinomial distribution is used to model the number of animals dying with tumours. The prior distribution for the historical control rate is based upon the Dirichelet distributions. The method, though, can only be applied to fatal tumours. Ibrahim *et al.* (1998) developed methods for incorporating historical control data into age-adjusted tests. This approach, though, is limited because it makes strong assumptions about the tumour lethality.

403. Ibrahim and co-workers have also developed a method that assumes all the tumours are lethal (Ibrahim & Ryan, 1996) or all tumours are non-lethal (Ibrahim, Ryan and Chen 1998). The two tests, therefore, represent the extreme events and may not be accurate in practice. Fung *et al.* (1996) developed methods for incorporating historical control data but this approach is also not an age-adjusted test.

404. A Bayesian approach has been suggested by Dempster *et al.* (1983) making the assumption that the logits of the historical control rate are normally distributed. Another Bayesian approach has been developed by Dunson & Dinse (2001) which relaxes some of the assumptions regarding the nature of the tumours. The prior probabilities for the parameter in the model, however, have to be chosen carefully and this requires a consensus between the pathologist and the toxicologist.

405. Peddada *et al.* (2007) have proposed a non-parametric approach which provides separate comparisons of the dose groups with the concurrent and historical control data. A third comparison can be made between the two control data sets. The poly-3 correction is made to sample sizes to adjust for survival rate differences between the groups. Consequently, individual data rather than summary data for a group are needed. The three p-values obtained are compared using a 'weight of evidence' approach. Without the survival data there is a possibility of bias.

4.23 Dose-response modelling

406. Previously, most of the emphasis on the carcinogenicity study had been on hypothesis testing and the identification of a carcinogenic hazard. The tests have been about the identification of a significant effect or trend. An alternative approach is the estimation of the size of any dose-response relationship. However, the evolution of the carcinogenicity study towards greater emphasis on quantitative dose response modelling has implication for the study design. The traditional design is limited because of some of the other objectives described earlier (paragraph 272).

407. The current approach to experimental design is a trade-off between the optimum allocation of equal group sizes between a control and treated group to maximize the power to detect an effect (i.e., test a null hypotheses) and the need to describe in detail the shape of a dose response relationship. Studies need to be designed to identify which parts of the relationship are important and, consequently, may require a two-stage approach of identifying a dose range of interest in a

preliminary range finding study and then moving to investigate this in more detail by concentrating resources there.

4.24 Extrapolation to low doses

408. In the current paradigm of risk assessment, the objective is to identify the risks associated with levels of potential human exposure. Risk assessment has traditionally been carried out differently for genotoxic and non-genotoxic carcinogens and non-carcinogens. A no-threshold model has usually been assumed for genotoxic carcinogens which has been associated with low dose extrapolation. On the other hand, a threshold model has been assumed for some non-genotoxic carcinogens and non-carcinogenic endpoints, provided there are adequate supporting mode of action data. The threshold model, below which there are no observed toxic effects, has been linked to the concept of a point of departure (POD) e.g., a no observed (adverse) effect level (NO(A)EL). In the absence of sufficient mode of action data, the risk assessment approach usually reverts to the non-threshold model.

409. Low dose extrapolation has been carried out by some regulatory agencies by fitting mathematical models to the observational data for carcinogens and then extrapolating the models to the low doses/exposures that might be expected to occur in the human population (EPA, 1986). For a long time, The US EPA used the Linearized Multistage procedure as the default approach for such extrapolations. These approaches aimed to identify, for instance, Virtually Safe Doses (VSDs) where it has been estimated that such life time exposure would lead to an upper bound increase of 1 extra lifetime cancer death in 1 million exposed individuals (or a 10-6 lifetime risk) or some similar low risk. The low dose extrapolations are, however, highly dependent upon the mathematical function assumed for the dose-response relationship and could give very different estimates of, for instance, the VSDs. Such low dose extrapolation may be conservative and alternative approaches have been proposed. Modelling may be confined to the observed experimental ranges and a POD identified such as, for example, an estimate of the Benchmark Dose.

410. Some authorities now propose linear extrapolation from the POD by drawing a straight line from the POD to the zero extra/additional risk and reading off the VSD associated with the 10^{-6} excess/additional risk (US EPA, 2005). Others propose using the ratio of the POD to the human exposure to derive a Margin of Exposures (MOE).

4.25 NOEL, NOAEL, LOEL, LOAEL approach

411. A NOEL (no observed effect level) is obtained by identification of the highest dose level of a test substance that does not cause a significant increase in any treatment-related effects compared with the negative/vehicle control. The LOEL (lowest observed effect level) is the lowest dose where there is a statistically significant effect.

412. The NOAEL (the no observed adverse effect level, sometimes designated NO(A)EL) is the highest level of a test substance that does not cause any observed and statistically significant adverse effects compared with the controls. The NOAEL distinguishes between changes which are adverse rather than any treatment related effect which may in some case not be adverse. Similarly the LOAEL (sometimes referred to as LO(A)EL) is the lowest dose where there is a significant adverse effect compared with the controls. Other terms used are NOEC, NOAEC, LOEC and LOAEC where the C refers to concentration rather than level, e.g., in inhalation studies.

413. The NOAEL identified in a long-term toxicity or carcinogenicity study may be used to derive a health-based guidance value such as an Acceptable or Tolerable Daily Intake (ADI and TDI), by applying a Safety or Uncertainty Factor (SF or UF) to the NOAEL derived from the study. Application of Safety or Uncertainty Factors depends of regulatory context. It is then considered that

there are no appreciable health risks below this health-based guidance value (WHO, 1999). The SF/UFs are used to account for inter- and intra-species variability, and an extra SF is applied if no NOAEL can be found and the LOAEL is used to derive health-based guidance values. The NOAEL approach is only relevant in the case of those non-genotoxic carcinogens or non-cancer endpoints where there is believed to be a threshold dose below which no toxic effects occur.

414. Although widely used, the following limitations of the NOAEL approach have been suggested:

- The NOAEL is inherently dependent on the study design, sample sizes, dose spacing, background levels, the power of the design (and statistical test)
- It is also based upon a hypothesis testing approach where a failure to reject the null hypothesis is taken as evidence of no difference. This differs from how a failure to reject a null hypothesis is interpreted by statisticians,
- The statistical comparisons are pair-wise comparisons and do not take into account information on the dose-response relationship or the more powerful trend test.
- The smaller the experiment or the more variable the endpoint the less chance of detecting a real effect
- The NOAELs tend to be higher for endpoints with a high control/background level because it is more difficult to demonstrate a statistically significant difference than for endpoints with low control/background levels
- It is a single dose level and does not reflect the other dose-response information e.g., the shape of the dose-response curve
- The NOAEL is not a measurement as such so cannot be presented with a measure of precision such as a confidence interval (CI).
- The NOAEL is a dose at which no adverse effect was detected. This does not preclude effects that could not be detected by the techniques used and also does not indicate that higher doses below the LOAEL may also have no detectable adverse effects. It is a single dose level dependent on the study design and is not necessarily representative of the true 'threshold'

4.26 Benchmark dose approach

415. An alternative approach to the NOAEL approach, the Benchmark dose (BMD) approach, was first proposed by Crump (1984) as an alternative for the identification of estimates of dose levels helpful for risk assessment. It had a slow acceptance despite early work to generalize the concept but has gradually become more widely accepted.

416. The BMD approach is based upon fitting a mathematical model to dose-response data collected for either qualitative or quantitative endpoints from experiments with 3 or more dose groups plus a negative control group. Mathematical models are fitted to the observed data. The dose associated with some predefined increase over background is chosen as the BMD. The lower confidence limit on the dose is then chosen as the POD for further investigation of the likely response at lower dose depending upon the choice of method to be used for extrapolation to lower doses.

417. The modelling approach involves the choice of endpoints to model, the choice of an appropriate mathematical model to fit to the data, the selection of model(s) which are considered to provide a satisfactory fit to the observed data and the selection of an appropriate BMD for determination of the POD.

418. The BMD is the dose associated with a pre-specified change in response. This response, the BMR, is a change in the endpoint of interest above the control response. In its original form, the response was based upon an increase over the control incidence of a quantal measure (such as tumour incidence). Values such as a 5% or 10% additional or extra incidence have conventionally been used to define the BMD for quantal data.

419. The use of confidence limits in the BMDL approach provides an estimate of uncertainty, that the uncertainty is reduced in large studies with better designs usually leading to higher PODs. It uses all the data in a dose-response experiment. The POD does not have to be one of the experimental dose levels and a POD can be calculated even if there is no NOAEL derived from a study.

420. Acceptance of the BMD approach is not universal. Travis *et al.* (2005), for example, argue against its routine use and that there are issues about how to apply it in cases where there is either no LOAEL identified (i.e., no obvious dose-response) or no NOAEL is identified (significant effect at all dose levels). Travis *et al.* (2005) argue that the NOAEL is best for routine use in toxicology studies but that the BMD may have a role in the interpretation of the most influential/critical studies in a regulatory package. The BMD approach also has potential limitations in that the selection of the model type and the parameters to include are chosen by the 'assessor'. This means that the same data may produce different BMD/BMDLs. There may also be difficulty in arriving at a consensus in defining a BMR level of a measure. The BMD, just as the NOEL, is also not a risk- or response-free exposure level.

Mathematical modelling for the BMD

421. Those endpoints which show 'visual' trends are analysed further to identify if the dose-response data are suitable for further analysis by fitting dose-response models to the experimental data. After suitable models have been identified, the BMD and the BMDL are determined for each suitable model. The specific BMD value is determined either by interpolation within the experimental data or by extrapolation beyond the experimental data. The BMDL is the lower one-sided 95% confidence limit on the BMD value. The BMDL can be interpreted as meaning that there is 95% confidence that the true effect at this dose would be less than the effect associated with the BMDL.

422. Values are obtained using software (BMDS developed by the US EPA (http://www.epa.gov/ncea/bmds/index.html) or the S-plus and R based PROAST software developed by RIVM) specifically developed for the purposes or using statistical packages such as SAS. (It is important that the various assumptions and defaults underlying these approaches and, in some cases, incorporated in the software are acknowledged.) One assumption in the modelling approach is the distribution of the data. In the case of quantal data the assumption is a binomial distribution. In the case of quantitative data, either a normal or a lognormal distribution is assumed.

423. The modelling approaches use algorithms to identify the optimal values of the parameters which specify the mathematical model. These values are derived by minimizing the difference between the fitted values and the observed values. One approach is the maximum likelihood method where likelihood is a measure of how likely the parameters have these specific values given the observed data; parameter values that maximize the likelihood measures are considered the 'best' estimates'.

424. The BMDL is usually estimated using a likelihood ratio test, a method also used for comparing the fits of different models. Twice the absolute difference between the log likelihoods of two different (comparable) models follows a chi-square distribution with degrees of freedom equal to the difference in number of parameters in each model. A chi-square value significant at $P<0.05$ is taken by some as evidence that the two models are considered significantly different.

425. The range of models potentially suitable for quantal data includes the 'standard' tolerance functions: probit, logistic and Weibull and log based versions. Others include the multistage and gamma multi-hit models. More complex models are provided in the software to take into account the more complex multilevel data from developmental/teratology data where there are, for instance, intra-litter correlations.

426. A large number of models can be used to describe a dose-response relationship in the case of quantitative (continuous) data. Examples include the polynomial, power, exponential and Hill function models (Slob, 2002). The Hill and exponential models have the properties of being sigmoid (S-shaped) and bounded (levelling of at a particular maximum and minimum response value) and parameters that can be easily related to the shape of the dose-response relationship. (On a log scale the Hill model is symmetrical but not on the normal dose scale.). Models can also be specified to take into account different variances within the dose groups.

427. A sufficient number of dose groups (with different response levels) are needed to assess whether a dose-response is linear, sublinear, sigmoidal or another form. In the case of a sigmoid/S-shaped dose-response, the modelling requires at least five dose levels (including controls) to avoid over-parameterization (perfect fit) and allows a test of model fit. A family of a specific model, e.g., polynomial, linear, power, Hill and exponential, is often used in the BMD approach for continuous data. . Testing within the families can be carried out with a likelihood ratio test for whether extra parameters improve the fit. If they do not, then by parsimony, they are left out. If doses were not high enough to measure where a sigmoid dose-response relationship levels off, the approach of nested models would fit a simpler model, without the need of identifying the maximum response level. Including high doses (where the response levels off) would result in more precise BMD estimates, but they are not crucial.

428. There is complexity in comparing models across classes. The US EPA has suggested the use of Akaike's Information Criteria (AIC). The AIC is a measure of the fit of the model weighted by the number of parameters fitted with the model, with the lowest AIC selected. A complication is that, when models are similar, the relative ranking in terms of AIC may be somewhat arbitrary. In practice, however, the objective is not to find the "best model" but rather to identify models which are plausible.

429. In practice, different models with the same number of parameters can often be found to give a satisfactory fit to the same dataset. The approach then is to calculate the BMDs and BMDLs from these various mathematical models and compare the range of values for all acceptable models for their similarity and consistency. The choice of which model is used for subsequent calculations may be based upon criteria such as which approach gives the lowest or most conservative BMDL or gives the best visual fit to data. Such choices, therefore, have potentially appreciable input from the risk assessor.

430. For the selection of the BMDL, the case is made for the selection of the lowest (i.e., most conservative) BMDL from those models which fit the data satisfactorily. Model averaging has also been proposed using, for instance, Bayesian model averaging where the averaging is a weighting derived from the support for a particular model taking into account the data (Bailer *et al.*, 2005). The BMDL which is considered most appropriate is identified as the RP which is used by the different approaches to defining a guidance level for risk assessment (ADI, TDI, MOE etc).

431. In the case of quantal data the BMR may be expressed either as 'additional' or 'extra' risk. Extra risk is an adjusted rate which includes an adjustment for the background incidence rate and is based only on the fraction who are expected not to have a background incidence with BMR = P(BMD) – P(0) / 1 - P(0). Additional risk is an absolute rate: P(BMD) – P(0). The two terms become the same when P(0), the background frequency, is zero.

432. The extra or additional risk (BMR) used to derive default BMDL values has usually been 5 or 10% for quantal data based upon the similarity between the BMDL derived from them and the NOAEL derived from developmental toxicity studies. The BMDL for a 10% risk level initially seemed most similar to the NOAEL derived from the same studies. Using more complex models taking into account the intra-litter correlations suggested the BMDL for a 5% risk level were most similar to the NOAEL.

433. One limitation of modelling for carcinogenicity study data is that it is generally done solely on incidence (quantal) data and does not make use of time-to-tumour (continuous) data. Including these data should be both more informative and provide a more accurate assessment. However, modelling methods using them have not been developed and validated and this limitation is particularly important when there is uncertainty over the cause of an animal's death. Other complexities for such methods are the consequence of early termination of studies or of some groups.

434. In the case of quantitative data the percentage change of effects relative to control levels that are considered biologically important need to be defined. There have been a number of different definitions of the BMR based upon different underlying approaches (IPCS, 2009)

435. In one approach the BMR is related to the CES (critical effect size) (Slob & Pieters, 1988). The CES is a BMR defined as a percentage change in the average response compared with the average response level in the controls: e.g., a 10% change, in the mean body weight over the control values for an adult animal or some fold change in an enzyme level or clinical chemistry value. There remains a debate over what the CES should be for each toxicological endpoint and the potential to use within animal variability to define it (Dekkers *et al.*, 2001; 2006). A 5% CES has been suggested in the absence of other information (Woutersen *et al.*, 2001; EFSA, 2009) in part because such a CES seemed close to the NOAEL found in some studies. However, it could be argued that BMD modelling may not be appropriate in the absence of an understanding of an endpoint.

436. In another approach the BMD related to a change in response equal to one standard deviation above the negative control mean has also been suggested (Crump 1984, 1995; Kavlock *et al.*, 1995). The US EPA had suggested using a change equivalent to one standard deviation (1 SD) in the endpoint. A 5% decrease in BMR for foetal weight and a 10% decrease for brain cholinesterase have been suggested as adverse.

437. Another approach, based on change relative to the dynamic range (the maximum to minimum values which links to the Hill or exponential model parameters), has been suggested (Murrell *et al.*, 1998). Change relative to the dynamic response may make comparison across endpoints possible. Other options include identifying doses where there are 'change points' on the dose response curve or the steepest point on the sigmoid curve.

438. One consideration is whether quantitative data should be converted into either binary or ordinal data so that it can be handled as if it were quantal data. Quantal data can in this approach be thought of as being on a continuous scale with various classes being defined as specific break points. Such dichotomization, however, results in a loss of information.

439. In conclusion, although modelling can appear a precise and a formal process, there is the opportunity for an appreciable amount of expert but subjective input into, for instance, the choice of options and the inclusion or exclusion of outliers and anomalous curve fits. Similar opportunities for expert but subjective judgement also apply to the NOAEL approach.

4.27 Other statistical methods for identifying change points in a dose-response relationship

440. Other statistical approaches have been developed for identifying "change points" in a dose-response study as an alternative to the NOAEL for continuous data. The change point is defined as the largest dose level which has the same response as in the negative control group. West & Kodell (2005) propose a method that investigates the profile of the least squares criterion over each of the intervals between the dose points in an experiment. They carried out simulation studies to show that the 95% lower confidence interval of the estimate of the change point had better statistical properties than the NOAEL. West & Kodell suggest linking the approach to the BMD methodology and that the method will need to be developed for a range of relevant change point model but note that conventional toxicology studies may have too few doses to estimate the parameters that explain more complex dose-response relationships.

441. The NOSTASOT method (no statistical significance of trend test) identifies the maximal dose which is not significantly different from the negative control group (Tukey *et al.*, 1985). In general the NOSTASOT dose is higher than the no-effect

Appendix 1 Common statistical methods used in the analysis of data

The following glossary gives a brief description of the various tests in the flowchart (Figure 1) plus other tests that are commonly encountered in chronic studies. The definitions below are based upon a number of widely available glossaries and in particular Everitt (1995).

Tests for outliers

Dixon/Massey test: A test for outliers in a sample

Extreme Studentized Deviate (ESD) Statistic: A method used for identifying outliers; also known as Grubbs' test

Tests for non-normality

Chi-square test: A goodness of fit test that a set of data come from a hypothesized distribution such as the normal.

Kolmogorov–Smirnov one-sample test: A method that tests for goodness of fit of the data to a defined distribution

Shapiro–Wilk test: A method that tests that a set of random variable arise for a specified probability distribution used to test for departure from normality.

Tests for homogeneity of variance

Bartlett's test: A test for the equality (homogeneity) of the variances of a number of samples. The test is sensitive to departures from normality

Levene's test: A test for the equality (homogeneity) of the variances of a number of samples. The test is less sensitive to departures from normality

F test of variances: A test for a difference in the size of two variances.

Assumed normally distributed data

1. Overall tests

Analysis of variance (ANOVA): Statistical methodology which partition variability attributable to various causes. Family of modelling approaches simplest and commonest of which is the fixed effects one-way ANOVA which compares means across a set of samples.

Analysis of covariance (ANCOVA): Extension of ANOVA which allows for possible effects of covariates on endpoint in to effects of treatments which may reduce error mean squares associated with analysis.

Pearson's correlation coefficient: A test of the association between two variables

Linear regression: A test of the relationship between the two variables: one the independent like the dose the other the dependent i.e., the response. Used to examine trends in dose effects and to test the significance of the regression slopes.

2. Pair-wise comparisons

Duncan's multiple range test: A modified version of the Newman-Keuls multiple comparison test used to test for multiple comparisons when the initial ANOVA between groups is significant.

Dunnett's t-test : A multiple comparison test which compares each of a number of treatments to a single control.

Scheffe's test: A multiple comparison test with less power than Newman–Keuls multiple range test.

Williams' t-test: A multiple comparisons method for comparing each of a number of treatments with a single control.

Student's t-test: A number of different tests but here the independent two sample t-test assuming equal variance in the two groups and testing for a difference between two means.

Satterthwaite test: An alternative to the pooled-variance t test, and is used when the assumption that the two populations have equal variances seems unreasonable. It provides a t statistic that asymptotically (that is, as the sample sizes become large) approaches a t distribution, allowing for an approximate t test to be calculated when the population variances are not equal. Also known as Welch's test,

Fisher's least significant difference (LSD) test: A pair-wise test equivalent to the independent two-sample t-test except that the estimate of error is based upon the within group error of an ANOVA.

Tukey's Honest Significant (HSD) Difference test: A single step multiple comparison method used after the initial ANOVA between groups is significant.

Non-parametric procedures (percentage values, ranks, etc.)

Kendall's coefficient of rank correlation: A non-parametric test of the association between two variables based upon ranks

Pearson's rank correlation: Another non-parametric test of the association between two variables based upon ranks

Mann–Whitney U-test: A non-parametric alternative to the independent two-sample t-test. Also called the Wilcoxon Rank Sum Test.

Kolmogorov–Smirnov two-sample test: A distribution-free method that tests for any difference between two population distributions.

Wilcoxon signed-rank test: A non-parametric alternative to the paired t-test for matched or paired data.

Kruskal–Wallis ANOVA test: A distribution-free method that is the analogue of the one-way analysis of variance. Which tests whether the groups to be compared have the same population mean.

Jonckheere-Terpstra test: A test for detecting departures from independence where both the rows and columns of a contingency table have a natural order.

Distribution-free multiple comparisons tests

Dunn's test: A multiple comparison test based upon the Bonferroni test

Shirley's test: A non-parametric equivalent of Williams' test.

Quantal data (mortalities, pathology findings, etc.)

Fisher's exact test: Test for independence of two variables forming a 2 x 2 contingency table, based upon the hypergeometric distribution.

R x C chi-square test: A measure of association between the row and column classification or a r x c contingency table of variables

Litchfield & Wilcoxon test: Graphical method of probit analysis for calculation ED50 and confidence intervals

Cochran-Armitage linear trend test: Chi-square test for linear trend in counts and proportions.

Multivariate methods

Hotellings T^2: A generalization of the t-test to multivariate data

MANOVA: A multivariate analysis of variance to test the equality of the means of more than 2 populations.

Survival-adjusted procedures for analysis of carcinogenicity data

Log-rank test: Compares the survival distributions of two or more samples, sometimes called the Mantel-Cox test.

Peto analysis: A test in IARC monograph combining a life table test for fatal tumours with a prevalence analysis for incidental tumours.

Life table test: A survival adjusted test for fatal cancers or cancers with observable onset times.

Hoel–Walberg procedure: A survival adjusted test for incidental tumours. Also called the prevalence method.

Logistic regression: A form of regression analysis used when the response is binary: i.e., tumour/no tumour.

Poly-k test: A survival-adjusted Cochran-Armitage test for testing for a dose-related trend and/or a pair-wise difference in the incidence of tumours.

REFERENCES

Ahn, H., R.L. Kodell and H. Moon (2000), "Attribution of Tumor Lethality and Estimation of Time to Onset of Occult Tumors in the Absence of Cause-of-death Information", *Applied Statistics,* 49 157-169.

Bailer, A.J., R.B. Noble and M.W. Wheeler (2005), "Model Uncertainty and Risk Estimation for Quantal Responses", *Risk Analysis*, 25 291-299.

Bailer, A.J. and C.J. Portier (1988), "Effects of Treatment-induced Mortality and Tumor-induced Mortality on Tests for Carcinogenicity in Small Samples", *Biometrics,* 44 417-431.

Bannasch, P., R.A. Griesemer, F. Anders, B. Becker, J.R. Cabral, G.D. Porta, V.J. Feron, D. Henschler, N. Ito, R. Kroes, P.N. Magee, B. McKnight, U. Mohr, R. Montesano, N.P. Napalkov, S. Nesnow, A.E. Pegg, G.N. Rao, V.S. Turusov, J. Wahrdrendorf and J. Wilbourn (1986), *Long-term Assays for Carcinogenicity in Animals, in Long-Term and Short-Term Assays for Carcinogens: A Critical Appraisal*, Editors: Montesano, R., Bartsch, H., Vainio, H., Wilbourn, J. & Yamasaki, H. IARC Scientific Publications No. 83, Lyon, France.

Bieler, G.S. and R.L.Williams (1993), "Ratio Estimates, the Delta Method, and Quantal Response Tests for Increased Carcinogenicity", *Biometrics,* 49 793-801.

Breslow, N. (1970), "A generalized Kruskal–Wallis Test for Comparing K Samples Subject to Unequal Patterns of Censorship", *Biometrics*, 57 579–594.

Butterworth, B.E., O.B. Mathre, K.E. Ballinger and O. Adalsteinsson (2004), "Contamination is a Frequent Confounding Factor in Toxicology Studies with Anthraquinone and Related Compounds", *Int J Toxicol,* 23 335-344.

Cox, D.R. (1959), "The Analysis of Exponentially Distributed Life-times with Two Types of Failures", *Journal of Royal Statistical Society*, Series B, 21 4121-421.

Cox, D.R. (1972), "Regression Models and Life Tables (with Discussion)", *Journal of Royal Statistical Society*, Series B, 34, 187-220.

Crump, K.S. (1984), "A New Method for Determining Allowable Daily Intakes", *Fund Appl Toxicol*, 4 854-871.

Crump, K.S. (1995), "Calculation of Benchmark Doses from Continuous Data", *Risk Analysis,* 15 79-89.

Dekkers, S., C. de Heer and M. Rennen. (2001), "Critical Effect Sizes in Toxicological Risk Assessment: A Comprehensive and Critical Evaluation", *Environ Toxicol Pharmacol*, 10 33–52.

Dekkers, S., J. Telman, M.A.J. Rennen, M.J. Appel and C. de Heer (2006), "Within-animal Variation as an Indication of the Minimal Magnitude of the Critical Effect Size for Continuous Toxicological Parameters Applicable in the Benchmark Dose Approach", *Risk Analysis,* 26 867–880.

Dempster, A.P., M.R. Selwyn and B.J. Weeks (1983), "Combining Historical and Randomized Controls for Assessing Trends in Proportions", *Journal of the American Statistical Association*, 78 221-227.

Dinse, G.E. (1994), "A comparison of Tumor Incidence Analyses Applicable in Single-sacrifice Animal Experiments", *Statistics in Medicine,* 13 689-708.

Dinse, G.E. and J.K. Haseman (1986), "Logistic Regression Analysis of Incidental-tumor Data from Animal Carcinogenicity Experiments", *Fundamental and Applied Toxicology* 6 751-770.

Dunson, D. B. and G.E. Dinse (2001), "Bayesian Incidence Analysis of Animal Tumorigenicity Data", *Appl Stat,* 50 125–41.

Elmore, S. and S.D. Peddada (2009), "Points to Consider on the Statistical Analysis of Rodent Cancer Bioassay Data when Incorporating Historical Control Data", *Toxicologic Pathology,* 37 672-676.

EFSA (2005), *Opinion of the Scientific Committee on a request from EFSA related to a harmonised approach for risk assessment of substances which are both genotoxic and carcinogenic,* The EFSA Journal, 282 1–31.

EFSA (2009), *Guidance of the Scientific Committee on a request from EFSA on the use of the benchmark dose approach in risk assessment*, The EFSA Journal, 1150: 1-72.

EPA (1986), *Guidelines for Carcinogen Risk Assessment*, U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, DC.

EPA (2005), *Guidelines for Carcinogen Risk Assessment,* U.S. Environmental Protection Agency, Washington, DC.

Everitt, B.S. (1995), *The Cambridge Dictionary of Statistics in the Medical Sciences,* Cambridge University Press.

FDA Redbook (2007), *Toxicological Principles for the Safety Assessment of Food Ingredients*, Food and Drug Administration, Washington, DC.

Fisher, R.A. (1932), *Statistical Methods for Research Workers*, (3rd edition) Oliver and Boyd, Edinburgh.

Fisher, R.A. (1950), *Statistical Methods for Research Workers,* (11th edition) Oliver and Boyd, Edinburgh.

Fung, K.Y., D. Krewski and R.T. Smythe (1996), "A Comparison of Tests for Trend with Historical Controls in Carcinogen Bioassay", *Can. J. Statist.,* 24 431-454.

Gad, S.D. (2000), *Statistics for Toxicologists*, In Principles and Methods of Toxicology, Ed: A. Wallace Hayes, 4th edition, Taylor and Francis: Boston, MA, pp 369-451.

Gad, S. (2006), *Statistics and Experimental Design for Toxicologists and Pharmacologists*, 4th Edition, Taylor and Francis, Boca Raton, FL.

Gart, J.J., D. Krewski, P.N. Lee, R.E. Tarone and J. Wahrendorf (1986), *Statistical Methods in Cancer Research, III The Design and Analysis of Long-term Animal Experiments*, IARC Scientific Publications 79. vol. III pp. 1–219.

Gehan, E.A. (1965), *A Generalized Wilcoxon Test for Comparing K Samples Subject to Unequal Patterns of Censorship*, Biometrika, 52 203-223.

Graves, T.S. (1994), "A Comparison of Test Procedures for Multidose Animal Carcinogenicity Assays under Competing Risks", *Journal of Biopharmaceutical Statistics*, 4 289 - 320.

Haseman, J.K. (1983), "A Reexamination of False-positive Rates for Carcinogenesis Studies". *Fundamental and Applied Toxicology,* 3 334-339.

Haseman, J.K. (1984), "Statistical Issues in the Design, Analysis and Interpretation of Animal Carcinogenicity Studies", *Environmental Health Perspective*, 58 385-392.

Haseman, J.K., G. Hajian, K.S. Crump, M.R. Selwyn and K.E. Peace (1990), *Dual Control Groups in Rodent Carcinogenicity Studies,* In Statistical Issues in Drug Research and Development, Ed, Peace, K. E. Marcel Dekker, New York.

Haseman, J.K., J. Huff and G.A. Boorman (1984), "Use of Historical Control Data in Carcinogenicity Studies in Rodents", *Toxicologic Pathology,* 12 126-135.

Haseman, J.K., J.E. Huff, G.N. Rao and S.L. Eustis (1989), "Sources of Variability in Rodent Carcinogenicity Studies", *Fundam Appl Toxicol*, 12 793-804.

Haseman, J.K., J.S. Winbush, M.W. O'Donnell (1986), "Use of Dual Control Groups to Estimate False Positive Rates in Laboratory Animal Carcinogenicity Studies", *Fundamental and Applied Toxicology*, 7 573-584.

Hoel, D.G. and H.E. Walburg (1972), "Statistical Analysis of Survival Experiments", *J Natl Cancer Inst,* 49 361-372.

Iatropoulos, M.J. (1984), "Appropriateness of Methods for Slide Evaluation in the Practice of Toxicologic Pathology", *Toxicol Pathol*, 12 305-306.

Iatropoulos, M.J. (1988), *Society of Toxicologic Pathologists Position Paper: "Blinded" Microscopic Examination of Tissues from Toxicologic or Oncogenic Studies*, in Carcinogenicity, The Design, Analysis, and Interpretation of Tong-Term Animal Studies, edited by H.C. Grice and J.L. Ciminera, ILSI Monographs, Spring-Verlag, New York.

Ibrahim, J.G. and L.M. Ryan (1996), "Use of Historical Controls in Time-adjusted Trend Tests for Carcinogenicity", *Biometrics,* 52 1478-1485.

Ibrahim, J.G., L.M. Ryan and M.-H. Chen (1998), "Use of Historical Controls to Adjust for Covariates in Trend Tests for Binary Data", *J. Amer. Statist. Assoc.*, 93 1282-1293.

IPCS (2009), *Principles for Modelling Dose–Response for the Risk Assessment of Chemicals,* Environmental Health Criteria 48, International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland.

JECFA (2005), *Joint FAO/WHO Expert Committee on Food Additives Sixty-fourth Meeting, Rome - Summary and Conclusions*, FAO/WHO.

Kaplan, E.L. and P. Meier (1958), "Nonparametric Estimation from Incomplete Observations", *J. Am. Stat. Assoc.,* 53 457-481.

Kavlock, R. J., B. C. Allen, E. M. Faustman and C. A. Kimmel (1995), "Dose–response Assessments for Developmental Toxicity. IV. Benchmark Doses for Fetal Weight Changes", *Fundam. Appl. Toxicol.,* 26 211–222.

Keenan, K.P., P. Laroque, G.C. Ballam, K.A. Soper, R. Dixit, B.A. Mattson, S. Adams and J.B. Coleman (1996), "The Effects of Diet, ad libitum Overfeeding, and Moderate Dietary Restriction on the Rodent Bioassay: the Uncontrolled Variable in Safety Assessment", *Toxicologic Pathology,* 24 757-768.

Kodell, R.L, J.H. Farmer, D.W. Gaylor and A.M. Cameron (1982), "Influence of Cause-of-death Assignment on Time-to-tumor Analyses in Animal Carcinogenesis Studies", *J Natl Cancer Inst.*, 69 659-64.

Kodell, R.L., K.K. Lin, B.T Thorn and J.J Chen (2000), "Bioassays of Shortened Duration for Drugs: Statistical Implications", *Toxicological Sciences,* 55 415-432.

Lee, P.N., J.S. Fry, W.R. Fairweather, J.K. Haseman, R.L. Kodell, J.J. Chen, A.J. Roth, K. Soper and D. Morton (2002), "Current Issues: Statistical Methods for Carcinogenicity Studies" *Toxicological Pathology*, 30 403–414.

Lin, K.K. and M.W. Ali (1994), *Statistical Review and Evaluation of Animal Tumorigenicity Studies*, In Statistics in the Pharmaceutical Industry, 2nd Ed., Ed. Buncher, C.R & Tsay J.Z. Marcel Dekker, New York.

Lin, K.K. and M.A. Rahman (1998), "Overall False Positive Rates in Tests for Linear Trend in Tumor Incidence in Animal Carcinogenicity Studies of New Drugs (with discussions)", *Journal of Pharmaceutical Statistics*, 81 1-22.

Lin, K.K. (2010), *Carcinogenicity Studies of Pharmaceuticals*, in Encyclopedia of Biopharmaceutical Statistics, Third edition, Revised and Expanded, Volume 1, edited by Shein-Chung Chow, Informa Healthcare, New York and London.

McConnell, E.E., H.A. Solleveld, J.A. Swenberg and G.A. Boorman (1986), "Guidelines for Combining Neoplasms for Evaluation of Rodent Carcinogenesis Studies", *J. Natl. Cancer Inst.,* 76 283-289.

Murrell, J.A., C.J. Portier and R.W. Morris (1998), "Characterizing Dose-response I: Critical Assessment for the Benchmark Dose Concept", *Risk Anal*, 18 1326.

NTP (2005), *Description of Study Types: Statistical Procedures*, US National Toxicology Program (NTP) Website: http://www.niehs.ntp.gov.

OECD (2002) *Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies,* Series on Testing and Assessment N° 35, OECD, Paris.

OECD (2009), *Carcinogenicity Studies*, OECD Guideline for the Testing of Chemicals, No. 451, OECD, Paris.

OECD (2009), *Chronic Toxicity Studies,* OECD Guideline for the Testing of Chemicals, No. 452, OECD, Paris.

OECD (2009), *Combined Chronic Toxicity/Carcinogenicity Studies,* OECD Guideline for the Testing of Chemicals, No. 453, OECD, Paris.

Peddada, S.D., G. Dinse and G. Kissling (2007), "Incorporating Historical Control Data when Comparing Tumor Incidence Rates", *J. Amer. Stat. Assoc.,* 102 1212-1220.

Peto R. (1974), "Guidelines on the Analysis of Tumour Rates and Death Rates in Experimental Animals", *British J Cancer*, 29 101-105.

Peto, R., M.C. Pike, N.E Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, S. Richards and J. Wahrendorf (1980), *Guidelines for Simple Sensitive Significance Tests for Carcinogenic Effects in Long-Term Animal Experiments,* In: IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans.

supplement 2: Long-term and Short-term Screening Assays for Carcinogens: A Critical Appraisal, Lyon: International Agency for Research on Cancer 311-346.

Portier, C.J., J.C. Hedges and D.G. Hoel (1986), "Age-specific Models of Mortality and Tumor Onset for Historical Control Animals in the National Toxicology Program's Carcinogenicity Experiments", *Cancer Res*, 46 4372−4378.

Portier, C. and D.G. Hoel (1983), "Optimal Bioassay Design Under the Armitage-Doll Multistage Model", *J Toxicol Environ Health,* 12 1−19.

Portier, C. and D.G. Hoel (1984), "Design of Animal Carcinogenicity Studies for Goodness-of-fit to Multistage Models", *Fundam Appl Toxicol.,* 4 949−959.

Prasse, K. (1986), Letter to the Editor (on blinded microscopic evaluation of slides from toxicity and carcinogenicity studies), *Toxicology and Applied Pharmacology,* 83 184-185.

Rahman, M.A. and K.K. Lin (2008), "A Comparison of False Positive Rates of Peto and Poly-3 Methods for Long-term Carcinogenicity Data analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman", *J Biopharm Statist*, 18 949-958.

Sand, S., K. Victorin and A.F. Filipsson (2008), "The Current State of Knowledge on the Use of the Benchmark Dose Concept in Risk, *J Appl Toxicol*., 28 405-21.

Slob, W. (2002), "Dose–response Modeling of Continuous Endpoints", *Toxicol. Sci.* 66 298–312.

Slob, W. and M.N. Pieters (1998), "A Probabilistic Approach for Deriving Acceptable Human Intake Limits and Human Health Risks from Toxicological Studies: General Framework", *Risk Analysis* 18 787–798.

Snedecor, G.W. and W.G. Cochran (1967), *Statistical Methods* (6th edition), The Iowa State University Press, Ames, Iowa.

Snedecor, G.W. and W.G. Cochran (1980), *Statistical Methods,* (7th edition), The Iowa State University Press, Ames, Iowa.

STP Working Group (2001), "Draft Recommendations on Classification of Rodent Neoplasms for Peto Analysis", *Toxicologic Pathology*, 29 265-268.

STP Peto Analysis Working Group (2002), "The Society of Toxicologic Pathology's Recommendations on Statistical Analysis of Rodent Carcinogenicity Studies", *Toxicologic Pathology,* 30 415–418.

Tarone, R.E. (1975), "Tests for Trend in Life Table Analysis", *Biometrika*, 62 679-682.

Tarone, R.E. (1982), "The Use of Historical Control Information in Testing for a Trend in Proportions" *Biometrics*, 38 215-220.

Tarone, R.E. and J. Ware (1977), "On distribution-free Tests for Equality of Survival Distributions", *Biometrika*, 64 165-60.

Temple, R.T., W.R. Fairweather, V.C. Glocklin, and R.T. O'Neill (1988), "The Case for Blinded Slide Reading", *Comments on Toxicology*, 2 99-109.

Thomas, D.G., N. Breslow and J.J. Gart (1977), "Trend and Homogeneity Analyses of Proportions and Life Table Data", *Computer and Biomedical Research*, 10 373-381.

Travis, K., I. Pate and Z. Welsh (2005), "The Role of the Benchmark Dose in a Regulatory Context", *Regulatory Toxicology and Pharmacology,* 43 280-291.

Tukey, J.W., J.L. Ciminera and J.F. Heyse (1985), "Testing the Statistical Certainty of a Response to Increasing Doses of a Drug", *Biometrics,* 41 295-301.

van Belle, G. (2009), *Statistical Rules of Thumb*, 2nd Edition, Wiley Hobeken.

WHO (1999), *Evaluation of Certain Food Additives and Contaminants,* Forty-Ninth Report of the Joint FAO/WHO Expert Committee on Food Additives, WHO Technical Report Series, No. 884, WHO, Geneva.

West, R.W. and R.L. Kodell (2005), "Changepoint Alternatives to the NOAEL", *Journal of Agricultural, Biological and Environmental Statistics*, 10 197-211.

Williams, D.A. (1971), "A test for Differences between Treatment Means when Several Dose Levels are Compared with a Zero Dose Control", *Biometrics*, 27 103-117.

Woutersen, R.A., D. Jonker, H. Stevenson, J.D. te Biesebeek and W. Slob (2001), "The Benchmark Dose Approach Applied to a 28-day Toxicity Study with Rhodorsil Silane in rats: the Impact of Increasing the Number of Dose Groups", *Food Chem. Toxicol.*, 39 697-707.

5 DEFINITIONS/GLOSSARY

ADI/TDI: Acceptable daily intake/Tolerable daily intake: the amount of a test article in food or drinking water that can be ingested (orally) over a lifetime without an appreciable health risk.

Adverse Effect: Change in the morphology, physiology, growth, development, reproduction or life span of an organism, system, or (sub) population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences.

AIC Akaike's Information Criteria

AUC: Area Under the Curve (Area under the plasma concentration-time curve): Area under the curve in a plot of concentration of substance in plasma over time.

Autolysis: The destruction of a cell through the action of its own enzymes.

Bayesian: Relating to statistical methods based on Bayes' theorem. A statistical approach that assesses the probability of a hypothesis being correct (for example, whether an association is valid) by incorporating the prior probability of the hypothesis and the experimental data supporting the hypothesis.

Benchmark dose (BMD): The dose corresponding to a small specified increase in effect over the background level.

BMDLx: is the lower confidence limit of BMD (see above) Typically, BMDL1 (with a response rate set at 1%) or BMDL10 (with a response rate set at 10%) is selected).

BMR: Benchmark Response

Bioaccumulation: the accumulation of the test article in an exposed organism. Bioaccumulation occurs when an organism absorbs a test article at a rate greater than that at which it is excreted.

Bias: A systematic error occurring in a measurement that is inherent in the sample itself or caused by the operator.

Bioavailability: Fraction of an administered dose that reaches the systemic circulation or is made available at the site of physiological activity.

Biomarker: A characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.

Carcinogenicity: Substances are defined as carcinogenic if they induce tumours (benign or malignant), increase its incidence or shorten the time of tumour, when inhaled, ingested, dermally applied or injected.

CES: Critical Effect Size

Chronic Toxicity: Toxicity (adverse effect) after an exposure period of 12 months or longer due to a test article that has been ingested inhaled, dermally applied or injected.

COD: Cause of Death

COO: Context of Observation

Detoxification pathways: Series of steps leading to the elimination of toxic substances from the body, either by metabolic change or excretion.

Dose: Total amount of a test article administered to, taken up by, or absorbed by an organism, system, or (sub) population.

Dose-response: Relationship between the amount of an agent administered to, taken up by, or absorbed by an organism, system, or (sub) population and the change developed in that organism, system, or (sub) population in reaction to the agent.

Dyschromatopsia: Deficiencies of colour vision

Dysplasia: refers to any disordered growth and maturation of an epithelium.

Electron microscope: A type of microscope that uses a beam of electrons, rather than light to produce an image of a specimen for detailed observation.

Extrapolation: Inference of one or more unknown values on the basis of that which is known or has been observed.

Exposure: Concentration or amount of the test article that reaches a target organism, system, or (sub) population in a specific frequency for a defined duration.

Fixation: A chemical process by which biological tissues are preserved from decay, to as close to its natural state as possible in the process of preparing tissue for examination.

Frequentist: An advocate of frequency probability. This is the inference framework in which the well-established methodologies of statistical hypothesis testing and confidence intervals are based

Fundoscopy (See ophthalmascopy)

Genomic(s): The study of all of the nucleotide sequences, including structural genes, regulatory sequences, and noncoding DNA segments, in the chromosomes of an organism.

Genotoxic/genotoxicity: A deleterious action on a cell's genetic material affecting its integrity.

GLP: Good Laboratory Practice

Hazard: The inherent property of a test article to cause adverse effects when an organism, system, or (sub) population is exposed to that test article.

Hazard identification: The identification of the type and nature of adverse effects that an agent has as inherent capacity to cause in an organism, system or (sub) population.

Haematology: The study of blood, the blood-forming organs, and blood diseases.

Histochemistry: The branch of histology dealing with the identification of chemical components in cells and tissues.

Histology: The study of the microscopic anatomy of cells and tissues of animals.

Histopathology: The study of the microscopic anatomical changes in diseased tissue

Hyperplasia: The proliferation of cells within an organ or tissue beyond that which is ordinarily seen.

Induction/Enzyme induction: Enzyme synthesis in response to an environmental stimulus or inducer molecule;

Immunohistochemistry: The process of localizing antigens (e.g., proteins) in cells of a tissue section exploiting the principle of antibodies binding specifically to antigens in biological tissues.

Local effect: Adverse effect at the site of first contact (e.g., skin, eye, mucous membrane/gastro-intestinal tract, or mucous membrane/respiratory tract).

LO(A)EL, also **LOAEL**: Lowest Observed (Adverse) Effect Level: The lowest level of a test substance that causes an observed and significant adverse effect on the test species compared with the controls.

LO(A)EC, also LOAEC: Lowest Observed (Adverse) Effect Concentration: The lowest concentration of a test article that causea an observed and significant adverse effect on the test species compared with the controls.

LOEC: Lowest Observed Effect Concentration. The lowest concentration of a test article that causes an observed and significant effect on the test species compared with the controls.

LOEL: Lowest Observed Effect Level: The lowest level of a test substance that causes an observed and significant effect on the test species compared with the controls.

Macroscopic/macroscopy: An observation that is large enough to be perceived or examined by the unaided eye.

Mechanism of Action (MOA): The individual biochemical and physiological events leading to a toxic effect).

Metabolism: In the context of this Guidance, refers to all the chemical processes in the body that alter the structure of the substanve administered.

Microscopic/microscopy: An observation of something extremely small in size; visible only with the aid of a microscope.

Mode of Action (MOA): Mode of Action: the processes by which a chemical induces toxicity. A MOA can inform about relevance of observed effects in laboratory animals to humans and the variability of response within the human population.

Morphology: The form and structure of organs and tissues and their specific structural features.

MTD: Maximum Tolerated Dose

MTC: Maximum Tolerated Concentration

Mydryatic agent: A chemical agent that induces dilation of the pupil.

Neoplasia: Refers to tumour growth. A neoplasm may be benign or malignant.

NO(A)EC, also **NOAEC**: No-Observed- Adverse-Effect-Concentration. The highest concentration of a test substance that does not cause any observed and statistically significant adverse effect on the test species compared with the controls.

NO(A)EL, also **NOAEL**: No-Observed- Adverse-Effect-Level. The highest level of a test substance that does not cause any observed and statistically significant adverse effect on the test species compared with the controls.

NOEC: No-Observed-Effect-Concentration. The highest concentration of a test substance that does not cause any observed and statistically significant effect on the test species compared with the controls.

NOEL: No-Observed-Effect-Level. The highest level of a test substance that does not cause any observed and statistically significant effect on the test species compared with the controls.

NOSTASOT: No statistical significance of trend (test)

Ocular: Relating to the eye or the sense of sight.

Ocular adnexa: Structures adjacent to the eye such as the lacrimal apparatus, the extraocular muscles and the eyelids, and the conjunctiva

Ophthalmoscopy (funduscopy or fundoscopy): Visual non-invasive examination the interior of the eye, including the lens, retina and optic nerve using a specialised instrument (ophthalmoscope or funduscope) containing a concave mirror and a battery-powered light

Palatability/ decrease in palatability: Acceptable to the taste, sufficient agreeable to be eaten/reduced taste / makes the food less aggreable due to organoleptic changes e.g., change in odor/flavor leading to reduced food intake.

Pathology: The study of disease.

Perimortem: At or around the time of death.

POD: Point of Departure: The dose-response point that marks the beginning of a low-dose extrapolation. This point is most often the upper bound on an observed incidence or on an estimated incidence from a dose-response model.

Postmortem: Done, occurring, or collected after death

(Q)SAR: Quantitative Structure Activity Relationship

Randomisation: A term used to describe the selection of samples for each 'arm' of a study or experiment based on chance alone– i.e., a theoretical coin toss, which is intended to minimize the influence of irrelevant details and selection bias, and produce statistically valid data.

Read-across: The endpoint information for one or more chemicals is used to make prediction of the endpoint for the target chemical i.e., use of structure similarity analysis to identify analogous compounds to determine whether there may be opportunities to bridge databases for one chemically structurally similar compound to another.

Reference dose (RfD): An estimate of a daily exposure to a chemical that is unlikely to cause harmful effects during a lifetime.

RP: Reference Point

Route of administration (oral, IV, dermal, inhalation, etc.): Refers to the means by which substances are administered to the body (e.g., orally by gavage, orally by diet, dermal, by inhalation, intravenously, etc).

Route-to-route extrapolation: The prediction of an equivalent dose and dosing regime that produces the same toxic endpoint or response as that obtained for a given dose and dosing regime by another route.

SF: Safety Factor

SAR: Structure Activity Relationship

Systemic effect: A toxicological effect that affects the entire body or many organs.

Systems Modeling: (Pharmacokinetic-based, Physiologically-based Pharmacokinetic, Biologically-based, etc.): Abstract model that uses mathematical language to describe the behaviour of a system.

Target tissue: Tissue in which the principal adverse effect of a toxicant is manifested.

TDI: Tolerable Daily Intake

Threshold: Dose or exposure concentration of an agent below which a stated effect is not observed or expected to occur.

Toxicity: Inherent property of an agent to cause an adverse biological effect.

Toxicodynamics: The processes of interaction of toxicologically active substances with target sites, and the biochemical and physiological consequences leading to adverse effects i.e., the way in which the chemical substance behaves/interacts within the biological system to cause toxicity

Toxicokinetics (Pharmacokinetics): A term describing the processes of chemical absorption, distribution, metabolism, and excretion in the organism (ADME) i.e., the way in which the chemical substance is absorbed, moves within the body and are being excreted.

Tonometry: The measurement of tension or pressure. In ophthalmoscopy, tonometry measures intraocular pressure by recording the resistance of the cornea to pressure (indentation).

Transcriptomics: The study of the complete set of RNA transcripts produced by the genome at any one time.

UF: Uncertainty Factor

VSD: Virtually Safe Dose

Elwell MR, Mahler JF and Ruecker FA (2004). Proliferative and non-proliferative Lesions in the Heart and Vasculature in Mice. MCV-1. In: **Guides for Toxicologic Pathology**. STP/ARP/AFIP. Washington, DC.

Proliferative and Non-proliferative Lesions in the Heart and Vasculature in Mice

M. R. Elwell [1], J.F. Mahler [2], F.A. Ruecker [3]*

[1] Pfizer Global Research & Development, Groton, Connecticut
[2] National Institute of Environmental Health Sciences, RTP, North Carolina
[3] Monsanto, St. Louis, Missouri

INTRODUCTION

This document outlines nomenclature and diagnostic criteria for lesions of the heart and blood vessels of the mouse. A longitudinal section of the heart, ideally including both ventricles, portions of both atria, the interventricular septum, and major blood vessels at the base, is routinely examined in toxicity and carcinogenicity studies. As an essential histologic component of tissues, blood vessels are evaluated during the course of microscopic examination of all histologic sections; abnormalities are typically given the topography of the organ or tissue site affected. Both non-proliferative and proliferative lesions of the cardiovascular system are outlined below. The majority of lesions that occur in the hearts of mice are degenerative or inflammatory; the presence of primary cardiac neoplasms is extremely rare in the strains of mice commonly used for carcinogenicity studies. Proliferative and non-proliferative lesions of the blood vessels are not uncommon in mice. Given the relatively common occurrence of spontaneous, background lesions and the potential for treatment induced lesions in the cardiovascular system, careful microscopic examination and appropriate use of terminology is crucial for the assessment of this system.

* Deceased.

Address correspondence to: Society of Toxicologic Pathology, 1821 Michael Faraday Dr., Suite 300, Reston, VA 20190-5332.

MORPHOLOGY

HEART

NON-PROLIFERATIVE LESIONS

Murine Cardiomyopathy (degenerative or progressive cardiomyopathy) (Figure 1)

The term cardiomyopathy is commonly used for the mouse, as well as the rat, to include a spectrum of spontaneous, progressive, age-related morphologic changes that occur in the heart. The modifier murine may be included in the diagnostic terminology to avoid potential misinterpretation or confusion of this common change in mice with the unique types of cardiomyopathy that occur in specific rodent models of cardiotoxicity, or the unique types of cardiomyopathy that occur in humans. Morphologic changes in murine cardiomyopathy include myocardial cell (fiber) degeneration, interstitial fibrosis, and frequently a minimal to mild infiltrate of mononuclear inflammatory cells. These degenerative and inflammatory components that comprise murine cardiomyopathy have been reported for different strains of older laboratory mice (8, 30) and in practice these age-related changes have been combined under the general term of cardiomyopathy (24). Although murine cardiomyopathy is typically present in strains of laboratory mice used in carcinogenicity studies, the incidence is not known. This appears to be the result of differences in terminology and possible combinations of terms that may be applied to this condition (i.e., cardiomy-

opathy, spontaneous cardiomyopathy, degenerative cardiomyopathy, fibrosis, chronic inflammation, myocardial degeneration, etc.). In addition, there are apparent differences among pathologists' definitions for the threshold of the morphologic change required for this diagnosis (30). Murine cardiomyopathy is often initially evident in the ventricular muscle but may also involve the atria. It should be noted that in the ventricular muscle at the base of the heart there is normally a slight prominence of interstitial connective tissue compared to other areas. In the development of murine cardiomyopathy the interstitial tissue in this area often becomes more prominent (fibrosis). As the severity of this condition progresses, there is evidence of muscle fiber degeneration (cytoplasmic swelling, shrinkage, and/or vacuolation), mononuclear cell infiltrate, and fibrosis throughout the myocardium.

Spontaneous, age-related murine cardiomyopathy has histological and ultrastructural myocardial changes that are distinguishable from the strain-specific hereditary muscular distrophies of mice (24, 29). It is also important to note that one or more of the morphologic components of spontaneous murine cardiomyopathy may also occur separately as incidental or treatment-related cardiac lesions (29), for example, degeneration and/or necrosis, and should be diagnosed separately when they are not considered to be components of cardiomyopathy. As discussed below, specific terminology (e.g., cytoplasmic vacuolation, hemorrhage, necrosis, pigmentation, etc.) should be applied when these changes are not considered to represent components of cardiomyopathy. There may be instances in which a treatment-related effect in the myocardium may be morphologically indistinguishable from spontaneous murine cardiomyopathy and detected only as an apparent increased severity or incidence of this change.

Amyloidosis (amyloid deposition) (Figure 2–3)

Amyloid occurs in the heart as well as in other tissues, but varies markedly in incidence between strains of laboratory mice. Amyloidosis is extremely rare in the B6C3F1 strain (12), but is a common age-related systemic disease in the CD-1 strain (18, 31). Amyloid is usually more prominent in other tissues (small intestine, renal glomerulus). The early change with amyloidosis in the heart consists of an interstitial accumulation of pale-staining amyloid material; large areas of the myocardium may be replaced by typical homogeneous masses amyloid in more severe cases.

Inflammation (Figure 4)

Inflammatory lesions in the heart may involve the endocardium, cardiac valves, ventricular or atrial myocardium, or epicardium. Inflammation of the myocardium should be distinguished from the minimal mononuclear cell infiltrate that is typically associated with the spectrum of morphologic changes observed in cardiomyopathy. The cellular components of inflammation may consist primarily of polymorphonuclear leukocytes (acute), mononuclear inflammatory cells (chronic), or a mixture of both types of inflammatory cells with or without fibrosis (chronic-active). Myocardial necrosis and fibrosis associated with chronic-active inflammation may be a component of the inflammatory process. An inflammatory cell infiltrate in the myocardium that appears to be a localized extension of arterial inflammation (see below) is generally only diagnosed with the vascular topography (e.g., heart-artery, inflammation). Inflammation in the heart may occur with various bacterial infections of mice, particularly in immunosuppressed mice with streptococcal septicemia (22). Tyzzer's disease may also cause myocardial inflammation and necrosis (28). Inflammation of the heart has been reported with murine cytomegalovirus, Coxsackie virus, and other experimental virus infections (13, 15, 22).

Mineralization (calcification) (Figure 5)

Mineralization in the heart is a genetically determined trait that varies markedly between mouse strains with respect to incidence, severity, and anatomic distribution. Within a given laboratory mouse strain there is also variation related to age, sex, and diet (8). Mineralization in the heart of mice has generally been considered to represent a dystrophic calcification within the myocardium or epicardium (4, 6, 8). Experimentally, calcification (dystrophic) has been shown to be increased in mice with freeze-thaw injury to the heart (5), and with cytomegalovirus or other viral infections (11). Mineralization may be observed as basophilic granular material in degenerative myofibers in hematoxylin & eosin (H&E) stained sections; larger crystalline aggregates of mineralized material may be present in the myocardium. Small vessels in the myocardium as well as in the aorta may also be mineralized. This basophilic material is positive with special stains for calcium salts. Inflammation is usually not a prominent component of myocardial mineralization; minimal fibrosis may be present, particularly around epicardial foci of mineralization.

Pigmentation

Pigmentation is generally present as a brown or yellow-brown, granular to globular material within myocytes or the interstitium. If it is necessary to specifically identify the nature of a pigment, special staining methods may be required to determine the composition (e.g., hemosiderin, lipofuscin, test article or metabolite). It should be noted that a small amount of melanin pigment may be found in the valve leaflets of the heart in non-albino strains of mice.

Degeneration (Figure 6)

Degeneration is a term usually applied to various morphologic changes in the heart that are interpreted to be the

opathy, spontaneous cardiomyopathy, degenerative cardiomyopathy, fibrosis, chronic inflammation, myocardial degeneration, etc.). In addition, there are apparent differences among pathologists' definitions for the threshold of the morphologic change required for this diagnosis (30). Murine cardiomyopathy is often initially evident in the ventricular muscle but may also involve the atria. It should be noted that in the ventricular muscle at the base of the heart there is normally a slight prominence of interstitial connective tissue compared to other areas. In the development of murine cardiomyopathy the interstitial tissue in this area often becomes more prominent (fibrosis). As the severity of this condition progresses, there is evidence of muscle fiber degeneration (cytoplasmic swelling, shrinkage, and/or vacuolation), mononuclear cell infiltrate, and fibrosis throughout the myocardium.

Spontaneous, age-related murine cardiomyopathy has histological and ultrastructural myocardial changes that are distinguishable from the strain-specific hereditary muscular distrophies of mice (24, 29). It is also important to note that one or more of the morphologic components of spontaneous murine cardiomyopathy may also occur separately as incidental or treatment-related cardiac lesions (29), for example, degeneration and/or necrosis, and should be diagnosed separately when they are not considered to be components of cardiomyopathy. As discussed below, specific terminology (e.g., cytoplasmic vacuolation, hemorrhage, necrosis, pigmentation, etc.) should be applied when these changes are not considered to represent components of cardiomyopathy. There may be instances in which a treatment-related effect in the myocardium may be morphologically indistinguishable from spontaneous murine cardiomyopathy and detected only as an apparent increased severity or incidence of this change.

Amyloidosis (amyloid deposition) (Figure 2–3)

Amyloid occurs in the heart as well as in other tissues, but varies markedly in incidence between strains of laboratory mice. Amyloidosis is extremely rare in the B6C3F1 strain (12), but is a common age-related systemic disease in the CD-1 strain (18, 31). Amyloid is usually more prominent in other tissues (small intestine, renal glomerulus). The early change with amyloidosis in the heart consists of an interstitial accumulation of pale-staining amyloid material; large areas of the myocardium may be replaced by typical homogeneous masses amyloid in more severe cases.

Inflammation (Figure 4)

Inflammatory lesions in the heart may involve the endocardium, cardiac valves, ventricular or atrial myocardium, or epicardium. Inflammation of the myocardium should be distinguished from the minimal mononuclear cell infiltrate that is typically associated with the spectrum of morphologic changes observed in cardiomyopathy. The cellular components of inflammation may consist primarily of polymorphonuclear leukocytes (acute), mononuclear inflammatory cells (chronic), or a mixture of both types of inflammatory cells with or without fibrosis (chronic-active). Myocardial necrosis and fibrosis associated with chronic-active inflammation may be a component of the inflammatory process. An inflammatory cell infiltrate in the myocardium that appears to be a localized extension of arterial inflammation (see below) is generally only diagnosed with the vascular topography (e.g., heart-artery, inflammation). Inflammation in the heart may occur with various bacterial infections of mice, particularly in immunosuppressed mice with streptococcal septicemia (22). Tyzzer's disease may also cause myocardial inflammation and necrosis (28). Inflammation of the heart has been reported with murine cytomegalovirus, Coxsackie virus, and other experimental virus infections (13, 15, 22).

Mineralization (calcification) (Figure 5)

Mineralization in the heart is a genetically determined trait that varies markedly between mouse strains with respect to incidence, severity, and anatomic distribution. Within a given laboratory mouse strain there is also variation related to age, sex, and diet (8). Mineralization in the heart of mice has generally been considered to represent a dystrophic calcification within the myocardium or epicardium (4, 6, 8). Experimentally, calcification (dystrophic) has been shown to be increased in mice with freeze-thaw injury to the heart (5), and with cytomegalovirus or other viral infections (11). Mineralization may be observed as basophilic granular material in degenerative myofibers in hematoxylin & eosin (H&E) stained sections; larger crystalline aggregates of mineralized material may be present in the myocardium. Small vessels in the myocardium as well as in the aorta may also be mineralized. This basophilic material is positive with special stains for calcium salts. Inflammation is usually not a prominent component of myocardial mineralization; minimal fibrosis may be present, particularly around epicardial foci of mineralization.

Pigmentation

Pigmentation is generally present as a brown or yellow-brown, granular to globular material within myocytes or the interstitium. If it is necessary to specifically identify the nature of a pigment, special staining methods may be required to determine the composition (e.g., hemosiderin, lipofuscin, test article or metabolite). It should be noted that a small amount of melanin pigment may be found in the valve leaflets of the heart in non-albino strains of mice.

Degeneration (Figure 6)

Degeneration is a term usually applied to various morphologic changes in the heart that are interpreted to be the result of specific cardiotoxicities and therefore distinct from the changes of spontaneous cardiomyopathy. Degeneration may be characterized by the presence of myocyte vacuolation, swelling, atrophy, or loss of cross striations (hyalinization). Vacuolation may be the result of dilation of sarcoplasmic reticulum or lipid accumulation and may require special staining or ultrastructural examination to differentiate the type of vacuolation present (29). Loss of cross striations is generally due to myofibrillar degeneration and likewise may require ultrastructural examination for verification.

It should be noted that the nuclei of some myocardial or interstitial cells in the heart may be enlarged (karyomegaly) and contain a linear segment or bar of densely stained chromatin that resembles cells characterized as "Aschkow-like" myocytes. The incidence or prominence of these cells is variable, but often a few can be observed in the heart of control and treated mice from chronic studies. A few cells with these enlarged nuclei in the heart of older mice is generally not considered to represent a specific degenerative change and is usually not diagnosed.

Necrosis

Focal areas of necrosis are characterized by lysis of myocytes and may represent an incidental finding or a treatment effect (21, 29). Necrosis may be accompanied by acute hemorrhage and a mild inflammatory cell infiltrate (21).

Hemorrhage

Hemorrhage consists of accumulation of erythrocytes beneath the epicardial surface or between myocytes. When hemorrhage is present within the myocardium there are usually other morphologic changes (degeneration/necrosis) present within the heart (21).

Thrombus (atrial thrombosis) (Figure 7)

Thrombus formation typically occurs in one or both atria. It has been reported to be more common in the left atrium (9). In the more acute lesions the lumen of the thrombosed atrium is distended with an accumulation of laminated layers of fibrin, leukocytes, and red blood cells. Lesions of longer duration have features typical of an organized thrombus, with attachment to the atrial wall and variable amounts of fibrous connective tissue, inflammation and occasionally cartilaginous metaplasia. In some instances there is a severe murine cardiomyopathy or amyloid deposition in the ventricular/atrial myocardium that may have preceded the thrombosis. Dietary factors (2, 17) and treatment-related degeneration of myocytes (10) have been shown to affect the incidence of atrial thrombosis.

Ectopic Tissue (Figure 8)

Embryonic rests of tissues such as thyroid may be present as an incidental finding in the epicardium or mediastinum/pericardium adjacent to the base of the heart. Epithelial lined cysts containing eosinophilic colloid-like material, possibly of thyroglossal duct origin have been found in the wall of the heart.

Bacteria

Septicemic animals, particularly immunodeficient mice, may have bacterial colonies. The colonies are usually composed of basophilic coccoid organisms in cardiac muscle with little or no associated inflammatory response.

PROLIFERATIVE LESIONS

Rhabdomyosarcoma

Rhabdomyosarcoma occurred in the left ventricle of the heart of two mice of the TM strain (27). These malignant neoplasms consisted of irregular sized spindle shaped cells with large nuclei. Cross striations were present in the cytoplasm of the neoplastic cells; both neoplasms metastasized to the kidney and/or lung. One transgenic mouse strain developed atrial neoplasms of the heart which are comprised of well-differentiated cardiac muscle fibers (25).

Sarcoma, Not Otherwise Specified (NOS)

Primary mesenchymal neoplasms of the heart have rarely been reported. One neoplasm diagnosed as mesenchymoma occurred at the base of the heart in a B6C3F1 mouse (14) that had been exposed to gamma irradiation. This neoplasm consisted of bone, cartilage and fibrous connective tissue. In the same report, three cardiac mesenchymal cell sarcomas that metastasized to several sites were also described in mice. These malignant neoplasms were described as consisting of variably sized and shaped cells, multinucleated cells and collagen. Sarcoma, NOS is to be used to designate a poorly differentiated malignant neoplasm considered to be of mesenchymal origin.

Mesothelioma

Mesothelioma of the epicardial surface of the heart, pericardium and other serous surfaces (lung, abdominal cavity) was reported in mice infected with polyoma virus (26). This neoplasm consisted of finger-like projections of mesothelial cells on a fine, connective tissue stroma.

Endothelial Hyperplasia, Hemangioma, Hemangiosarcoma

Vascular hyperplasia and neoplasia occasionally occur in the heart and may be treatment-related. These are described in the subsequent section on vascular proliferative lesions.

Systemic and Metastatic Neoplasms (Figure 9)

Systemic neoplasia and metastatic or locally invasive

neoplasia may occur in the heart. For example, hematopoietic neoplasms including lymphoma and histiocytic sarcoma may infiltrate the myocardium, particularly in more advanced systemic neoplasms where there is extensive involvement of multiple tissue sites. In disseminated lymphoma, focal aggregates of neoplastic lymphocytic cells may be present in the myocardium or immediately beneath the epicardial surface. Locally invasive or metastatic alveolar bronchiolar carcinoma occurs in the epicardium/myocardium and is much more commonly observed than various types of metastatic neoplasms from other sites (e.g., subcutaneous sarcoma, hepatocellular carcinoma, squamous cell carcinoma, etc.).

BLOOD VASCULATURE

NON-PROLIFERATIVE LESIONS

Arteritis (periarteritis, panarteritis, polyarteritis, systemic arteritis, vasculitis, necrotizing polyarteritis) (Figures 10-12)

This inflammatory process may involve one or multiple small- or medium-sized muscular arteries, and is more frequently observed in arteries of the uterus, ovaries, testes, kidney, urinary bladder, mesentery, and heart. Variable components of this lesion include fibrinoid degeneration, necrosis of the tunica media, and an inflammatory cell infiltrate that may be predominately polymorphonuclear leukocytes, mononuclear inflammatory cells, or a mixture of these cell types. Fibrosis and thickening of the muscular wall is also a common feature of inflammation of the arteries or veins. This lesion is more common in older mice but the incidence varies with strain and sex. Although the etiology of this spontaneous vascular lesion is generally not known, in most instances it is considered to be immune-mediated (1, 23).

Inflammation

Inflammatory lesions in blood vessels considered to represent changes that should be distinguished from the spontaneous arteritis (described above) are diagnosed with the appropriate terminology to indicate the morphologic changes. Inflammation may consist of infiltrates of primarily polymorphonuclear leukocytes (acute), mononuclear inflammatory cells (chronic) with or without fibrosis, or a mixture of the two types with or without fibrosis (chronic-active).

Amyloidosis (amyloid deposition) (Figures 13–14)

Vascular deposition of amyloid is typically diagnosed with the topography of organ involvement (e.g., kidney, glomerulus-amyloid deposition) and also as a systemic lesion. Amyloid deposition occurs in many organs: kidneys, gastrointestinal tract, lung, liver, testes, tongue, thyroid gland, and heart are more often affected. The histological appearance of amyloid is similar to that described for the heart.

Angiectasis (telangiectasis, dilatation) (Figure 15)

Angiectasis is characterized by dilatation of the small endothelial lined channels which form the sinusoidal network of some organs (e.g., liver, adrenal gland, pituitary, lymph node medulla), or by dilatation of larger vessels (e.g., uterine veins). The absence of endothelial cell proliferation and/or nuclear atypia are important features to differentiate angiectasis from hemangioma.

Congestion (hyperemia)

Congestion consists of an increased prominence of erythrocytes in the capillary bed or larger blood vessels of an organ. There is generally no appreciable distention (angiectasis) or other morphological alteration of congested vessels. Congestion is typically diagnosed as a microscopic correlation for certain gross observations (i.e., darkened areas on a mucosal surface, reddened lymph nodes).

Thrombus (Figure 16)

Thrombus formation may occur at almost any site, and generally, the morphologic diagnosis (e.g., uterus-thrombosis) is not usually indicated by a vascular topography.

Mineralization (calcification) (Figure 17)

Foci of mineral deposition may be within the tunic media of arteries or in veins. In some strains of mice the mineralization of small vessels in the brain is common, but the morphologic diagnosis (e.g., brain, thalamus-mineralization) generally does not include the blood vessel topography. Mineralization of vessels consists of basophilic granular to crystalline aggregates in H&E stained sections which are positive with special stains for calcium salts.

Medial Hypertrophy (Figure 18)

Medial hypertrophy of the muscular arteries has been reported as a vascular finding of uncertain significance and etiology (9, 23). Thickened walls of muscular arteries having little or no apparent inflammatory component occur in sections of lung. While it has been suggested that this change may sometimes be due to a contraction and thickening of the muscle fibers in the tunica media rather than hypertrophy, a possible relationship of this finding to venous congestion and pulmonary hypertension has been proposed (8). It is important to note that hypertrophy of the tunica media is also a frequent component of arteritis (described above) which occurs at many sites throughout the body. The hypertrophy, present in arteries with inflammation, is not generally diagnosed as a separate lesion.

neoplasia may occur in the heart. For example, hematopoietic neoplasms including lymphoma and histiocytic sarcoma may infiltrate the myocardium, particularly in more advanced systemic neoplasms where there is extensive involvement of multiple tissue sites. In disseminated lymphoma, focal aggregates of neoplastic lymphocytic cells may be present in the myocardium or immediately beneath the epicardial surface. Locally invasive or metastatic alveolar bronchiolar carcinoma occurs in the epicardium/myocardium and is much more commonly observed than various types of metastatic neoplasms from other sites (e.g., subcutaneous sarcoma, hepatocellular carcinoma, squamous cell carcinoma, etc.).

BLOOD VASCULATURE

NON-PROLIFERATIVE LESIONS

Arteritis (periarteritis, panarteritis, polyarteritis, systemic arteritis, vasculitis, necrotizing polyarteritis) (Figures 10-12)

This inflammatory process may involve one or multiple small- or medium-sized muscular arteries, and is more frequently observed in arteries of the uterus, ovaries, testes, kidney, urinary bladder, mesentery, and heart. Variable components of this lesion include fibrinoid degeneration, necrosis of the tunica media, and an inflammatory cell infiltrate that may be predominately polymorphonuclear leukocytes, mononuclear inflammatory cells, or a mixture of these cell types. Fibrosis and thickening of the muscular wall is also a common feature of inflammation of the arteries or veins. This lesion is more common in older mice but the incidence varies with strain and sex. Although the etiology of this spontaneous vascular lesion is generally not known, in most instances it is considered to be immune-mediated (1, 23).

Inflammation

Inflammatory lesions in blood vessels considered to represent changes that should be distinguished from the spontaneous arteritis (described above) are diagnosed with the appropriate terminology to indicate the morphologic changes. Inflammation may consist of infiltrates of primarily polymorphonuclear leukocytes (acute), mononuclear inflammatory cells (chronic) with or without fibrosis, or a mixture of the two types with or without fibrosis (chronic-active).

Amyloidosis (amyloid deposition) (Figures 13–14)

Vascular deposition of amyloid is typically diagnosed with the topography of organ involvement (e.g., kidney, glomerulus-amyloid deposition) and also as a systemic lesion. Amyloid deposition occurs in many organs: kidneys, gastrointestinal tract, lung, liver, testes, tongue, thyroid gland, and heart are more often affected. The histological appearance of amyloid is similar to that described for the heart.

Angiectasis (telangiectasis, dilatation) (Figure 15)

Angiectasis is characterized by dilatation of the small endothelial lined channels which form the sinusoidal network of some organs (e.g., liver, adrenal gland, pituitary, lymph node medulla), or by dilatation of larger vessels (e.g., uterine veins). The absence of endothelial cell proliferation and/or nuclear atypia are important features to differentiate angiectasis from hemangioma.

Congestion (hyperemia)

Congestion consists of an increased prominence of erythrocytes in the capillary bed or larger blood vessels of an organ. There is generally no appreciable distention (angiectasis) or other morphological alteration of congested vessels. Congestion is typically diagnosed as a microscopic correlation for certain gross observations (i.e., darkened areas on a mucosal surface, reddened lymph nodes).

Thrombus (Figure 16)

Thrombus formation may occur at almost any site, and generally, the morphologic diagnosis (e.g., uterus-thrombosis) is not usually indicated by a vascular topography.

Mineralization (calcification) (Figure 17)

Foci of mineral deposition may be within the tunic media of arteries or in veins. In some strains of mice the mineralization of small vessels in the brain is common, but the morphologic diagnosis (e.g., brain, thalamus-mineralization) generally does not include the blood vessel topography. Mineralization of vessels consists of basophilic granular to crystalline aggregates in H&E stained sections which are positive with special stains for calcium salts.

Medial Hypertrophy (Figure 18)

Medial hypertrophy of the muscular arteries has been reported as a vascular finding of uncertain significance and etiology (9, 23). Thickened walls of muscular arteries having little or no apparent inflammatory component occur in sections of lung. While it has been suggested that this change may sometimes be due to a contraction and thickening of the muscle fibers in the tunica media rather than hypertrophy, a possible relationship of this finding to venous congestion and pulmonary hypertension has been proposed (8). It is important to note that hypertrophy of the tunica media is also a frequent component of arteritis (described above) which occurs at many sites throughout the body. The hypertrophy, present in arteries with inflammation, is not generally diagnosed as a separate lesion.

Embolus

An embolus is any detached intravascular mass which may be carried by the blood to a site distant from its point of origin. Bacterial or foreign body emboli (e.g., hairshaft or epidermis introduced with intravenous injection procedure) are the most common type in mice. The presence of malignant tumor cell masses within the lumen of a blood vessel may be diagnosed as embolus; however, the topography of the blood vessel and the term embolus are generally not included for diagnoses of metastatic tumors (e.g., lung-hepatocellular carcinoma, metastatic).

Leukocytosis

Leukocytosis is a diagnostic term that may occasionally be used to distinguish an increased number of non-neoplastic leukocytes within the vasculature from a neoplastic leukemic change. Leukocytosis often occurs in mice with severe inflammatory lesions (e.g., abscesses, gastrointestinal ulceration) and is typically associated with increased hematopoiesis in the spleen, bone marrow and other sites. Leukocytosis is often most prominent in the liver sinusoids and alveolar wall capillaries of the lung. The morphologic diagnosis (e.g., liver-leukocytosis) generally does not include the blood vessel topography.

Atherosclerosis

Vacuolation and plaque-like thickening of the tunica intima occurs very rarely in commonly used strains of laboratory mice. Only two cases, both involving the thoracic aorta, have been described (9). The apolipoprotein E deficient mouse spontaneously develops atherosclerosis in the aorta and coronary arteries (32).

Aneurysm

Dilatation or dissection of the wall or larger arteries rarely occurs in common strains of laboratory mice. The blotchy mouse is a strain that is genetically predisposed to develop an aneurysm of the aorta and its abdominal branches by 6 months of age (3).

Vascular Malformation

Vascular malformations are congenital/developmental lesions that on rare occasions have been observed as an incidental microscopic finding and characterized by a focal increase in the number of blood vessels within a particular tissue, or a hamartomatous mass comprised of blood vessels. The blood vessels within the focus or mass have normal morphology, distinguishing the malformation from vascular tumors.

PROLIFERATIVE LESIONS

Endothelial Cell Hyperplasia (Figure 19)

Hyperplasia of the capillary endothelium is not a common finding in mice but has been reported as a preneoplastic lesion with treatment-related hemangiosarcoma in the heart (20). Hyperplasia consists of a focal or multifocal increase in the number and size of endothelial cell nuclei; nuclear atypia and mitotic figures may be present. Some foci of endothelial cell hyperplasia in the heart may represent a section through the periphery of a hemangiosarcoma. A different form of endothelial hyperplasia has been reported in female mice (B6C3F1 strain) with mammary gland adenocarcinoma (16). This is a diffuse change in the endothelium throughout the heart; mitoses are not a feature of this hyperplasia that is possibly a secondary effect of mammary adenocarcinoma in this mouse strain. In that report, the hyperplastic endothelium stained positive for WGA and ConA lectins. Angiogenesis or proliferation of capillary structures in granulation tissue is not usually diagnosed separately but considered part of a regenerative/inflammatory process.

Hemangioma (Figures 20–21)

Hemangioma is a focal, expansive, non-invasive expansion of capillary or sinusoidal structures. Variably sized cavernous spaces are lined by a single layer of well-differentiated endothelium. Endothelial nuclei are usually increased in number and may be slightly enlarged in hemangioma; mitoses are sometimes present but are not common. Hemangioma may distort surrounding tissues, especially within lymph nodes or liver. This can also be observed with dilated, blood filled spaces resulting from angiectasis.

Hemangiosarcoma (Figures 22–23)

Hemangiosarcoma is a poorly delineated, invasive focus of capillary or sinusoidal structures. Typically, these neoplasms have solid areas of collapsed vascular channels and areas with poorly formed, irregular-shaped vascular spaces. There is an increase in the number of endothelial cells, many of which have pleomorphic, enlarged, hyperchromatic nuclei. Mitotic figures are usually present. Metastases may occur, but in the absence of a large highly malignant neoplasm it may be difficult to distinguish a metastatic focus from a vascular neoplasm of multicentric origin. The diagnosis of vascular neoplasia is based primarily on the morphologic features in H&E stained sections. Although ultrastructural features consistent with Weibel-Palade bodies have rarely been observed, they have been inconsistently present (7) and are not considered to be a reliable finding to characterize these neoplasms (19).

RECOMMENDED NOMENCLATURE AND DIAGNOSTIC CRITERIA

HEART

NON-PROLIFERATIVE LESIONS

Murine Cardiomyopathy

1. Myocardial fiber degeneration (vacuolization, necrosis), interstitial fibrosis, and/or small number of mononuclear inflammatory cells are typical features of spontaneously occurring murine cardiomyopathy.

Amyloidosis

1. Deposition of characteristic hyaline material around small vessels and between myofibers. Larger deposits of amyloid may replace focally extensive areas of the myocardium.

Inflammation

1. Acute—Infiltrates composed primarily of polymorphonuclear leukocytes in the myocardium, epicardium, or endocardium.
2. Chronic—Infiltrates composed primarily of mononuclear inflammatory cells in the myocardium, epicardium, or endocardium; fibrosis may be present.
3. Chronic-active—A mixture of inflammatory cells in the myocardium, epicardium, or endocardium; includes abscess formation and may have an associated myofiber degeneration/necrosis and fibrosis.

Mineralization

1. Focal aggregates of basophilic granular to crystalline material in myocardium.
2. Minimal fibrosis and a scant inflammatory cell infiltrate may be associated with some foci of mineralization.

Pigmentation

1. Generally a brown or yellow-brown, granular to globular material within myocytes or the interstitium.
2. If necessary to identify pigment (e.g., hemosiderin, lipofuscin, exogenous material), special staining of selected sections may be required.

Degeneration

1. Presence of myocyte vacuolation, swelling, atrophy, or loss of cross striations (hyalinization).

Necrosis

1. Lysis of myocytes with loss of nuclear and cytoplasmic detail.
2. May be accompanied by acute hemorrhage and a mild inflammatory cell infiltrate.

Hemorrhage

1. Accumulation of erythrocytes beneath the epicardial surface or between myocytes.
2. Other changes (degeneration/necrosis) may be present with hemorrhage.

Thrombus

1. Most often occur in the atria and observed grossly as masses in dilated or distended atria.
2. Laminated layers of fibrin, leukocytes, and red blood cells.
3. Lesions of longer duration may show evidence of organization with attachment to the endocardial surface, fibrosis, mineralization, or cartilaginous metaplasia.

Ectopic Tissue

1. Presence of thyroid gland or other tissue within the heart (e.g., thyroid gland, cyst) in the endocardium or myocardium or adjacent to the base of the heart.

Bacteria

1. Presence of aggregates of bacterial organisms within the myocardium, epicardium, or endocardium.

PROLIFERATIVE LESIONS

Rhabdomyosarcoma

1. Irregular sized spindle shaped cells and well-differentiated cardiac muscle fiber with large nuclei and cytoplasmic cross striations.

Sarcoma, NOS

1. Poorly differentiated mesenchymal neoplasms with a variety of histologic features to include bone, cartilage and fibrous connective tissue.

Mesothelioma

1. Finger-like projections of mesothelial cells on a fine, connective tissue stroma.
2. Arises from epicardial or pericardial surfaces.

Systemic Neoplasms

1. Any of several hematopoietic neoplasms (e.g., lymphoma, histiocytic sarcoma) that form masses or cellular infiltrates within the myocardium.
2. Characteristic morphologic features are described in Hematopoietic System.

Metastatic and Invasive Neoplasms

1. Variety of metastatic (e.g., hepatocellular carcinoma) and locally invasive (most commonly bronchoalveolar carcinoma) neoplasms.
2. Identified by the morphologic features described in the chapters that discuss these proliferative lesions in their primary sites.

RECOMMENDED NOMENCLATURE AND DIAGNOSTIC CRITERIA

HEART

NON-PROLIFERATIVE LESIONS

Murine Cardiomyopathy

1. Myocardial fiber degeneration (vacuolization, necrosis), interstitial fibrosis, and/or small number of mononuclear inflammatory cells are typical features of spontaneously occurring murine cardiomyopathy.

Amyloidosis

1. Deposition of characteristic hyaline material around small vessels and between myofibers. Larger deposits of amyloid may replace focally extensive areas of the myocardium.

Inflammation

1. Acute—Infiltrates composed primarily of polymorphonuclear leukocytes in the myocardium, epicardium, or endocardium.
2. Chronic—Infiltrates composed primarily of mononuclear inflammatory cells in the myocardium, epicardium, or endocardium; fibrosis may be present.
3. Chronic-active—A mixture of inflammatory cells in the myocardium, epicardium, or endocardium; includes abscess formation and may have an associated myofiber degeneration/necrosis and fibrosis.

Mineralization

1. Focal aggregates of basophilic granular to crystalline material in myocardium.
2. Minimal fibrosis and a scant inflammatory cell infiltrate may be associated with some foci of mineralization.

Pigmentation

1. Generally a brown or yellow-brown, granular to globular material within myocytes or the interstitium.
2. If necessary to identify pigment (e.g., hemosiderin, lipofuscin, exogenous material), special staining of selected sections may be required.

Degeneration

1. Presence of myocyte vacuolation, swelling, atrophy, or loss of cross striations (hyalinization).

Necrosis

1. Lysis of myocytes with loss of nuclear and cytoplasmic detail.
2. May be accompanied by acute hemorrhage and a mild inflammatory cell infiltrate.

Hemorrhage

1. Accumulation of erythrocytes beneath the epicardial surface or between myocytes.
2. Other changes (degeneration/necrosis) may be present with hemorrhage.

Thrombus

1. Most often occur in the atria and observed grossly as masses in dilated or distended atria.
2. Laminated layers of fibrin, leukocytes, and red blood cells.
3. Lesions of longer duration may show evidence of organization with attachment to the endocardial surface, fibrosis, mineralization, or cartilaginous metaplasia.

Ectopic Tissue

1. Presence of thyroid gland or other tissue within the heart (e.g., thyroid gland, cyst) in the endocardium or myocardium or adjacent to the base of the heart.

Bacteria

1. Presence of aggregates of bacterial organisms within the myocardium, epicardium, or endocardium.

PROLIFERATIVE LESIONS

Rhabdomyosarcoma

1. Irregular sized spindle shaped cells and well-differentiated cardiac muscle fiber with large nuclei and cytoplasmic cross striations.

Sarcoma, NOS

1. Poorly differentiated mesenchymal neoplasms with a variety of histologic features to include bone, cartilage and fibrous connective tissue.

Mesothelioma

1. Finger-like projections of mesothelial cells on a fine, connective tissue stroma.
2. Arises from epicardial or pericardial surfaces.

Systemic Neoplasms

1. Any of several hematopoietic neoplasms (e.g., lymphoma, histiocytic sarcoma) that form masses or cellular infiltrates within the myocardium.
2. Characteristic morphologic features are described in Hematopoietic System.

Metastatic and Invasive Neoplasms

1. Variety of metastatic (e.g., hepatocellular carcinoma) and locally invasive (most commonly bronchoalveolar carcinoma) neoplasms.
2. Identified by the morphologic features described in the chapters that discuss these proliferative lesions in their primary sites.

BLOOD VASCULATURE

NON-PROLIFERATIVE LESIONS

Arteritis (panarteritis, periarteritis, polyarteritis, systemic arteritis, necrotizing polyarteritis)

1. A mixed (mononuclear and polymorphonuclear leukocyte) inflammatory cell infiltrate within the tunica adventitia that may extend through the tunica media to the endothelium.
2. Thickness of vascular wall is often increased and fibrinoid degeneration and necrosis of the tunica media may be present.
3. Often involves arteries at multiple sites (i.e., kidney, uterus, ovaries, testes, heart, etc.).
4. Generally the term applies to the spontaneous (immune-mediated) lesion.

Inflammation

1. Acute—Infiltrates composed primarily of polymorphonuclear leukocytes within or adjacent to vessel wall.
2. Chronic—Infiltrates composed primarily of mononuclear inflammatory cells within or adjacent to vessel wall; fibrosis may be present.
3. Chronic-active—A mixture of inflammatory cells and fibrosis within or adjacent to vessel wall.

Amyloidosis

1. Deposition of hyaline amyloid substance appearing in capillaries and larger blood vessels.
2. Most common in the glomerulus of the kidney.

Angiectasis

1. Dilatation (distention) of endothelial-lined channels forming delicate sinusoidal network (e.g., liver, adrenal gland, pituitary gland) or larger vessels (e.g., uterine veins).
2. Absence of endothelial alteration (increased mitoses, nuclear atypia) distinguishes angiectasis from hemangioma.

Congestion

1. Increased prominence (number) of erythrocytes in the capillary bed or larger vessels of an organ.
2. No appreciable distention (angiectasis) of the vessel lumen.
3. Often diagnosed to correlate with gross observation (e.g., reddened, darkened focus).

Thrombus

1. May occur in any vessel but more commonly seen in the uterus.
2. Often not diagnosed separately when it is a component of another lesion (i.e., hepatocellular carcinoma).

Mineralization

1. Foci of mineral deposits in vascular wall of blood vessels.
2. May occur at any site but is more often observed in larger vessels of the heart and capillaries of the brain (thalamus).

Medial Hypertrophy

1. Thickening of the tunica media of arteries in the absence of inflammatory changes (arteritis).
2. Has only been described in arteries of the lung.

Embolus

1. Presence of an intravascular mass (e.g., bacterial, foreign body, malignant tumor cells).

Leukocytosis

1. Prominent increase in the number of intravascular non-neoplastic leukocytes.
2. More commonly observed in the liver and lungs.

Atherosclerosis

1. Vacuolation and plaque-like thickening of the tunica intima.

Aneurysm

1. Localized abnormal dilatation of a blood vessel or dissection and expansion of vascular wall by entry and accumulation of blood between layers of the vascular tunics.

Vascular Malformation

1. A focus or mass composed of an increased number of blood vessels with normal morphology.

PROLIFERATIVE LESIONS

Endothelial Cell Hyperplasia

1. Focal or multifocal increase in the number and size of endothelial cell nuclei; nuclear atypia and mitotic figures may be present.
2. Diffuse change in the endothelium throughout the heart in female mice; possible secondary effect of mammary adenocarcinoma in the B6C3F1 mouse strain.

Hemangioma

1. Focal, non-invasive expansion of a capillary or sinusoidal structures.
2. Variably sized cavernous spaces are lined by a single layer of well-differentiated endothelium.
3. Endothelial nuclei increased in number and may be slightly enlarged; mitoses are sometimes present.

Hemangiosarcoma

1. Focal, poorly delineated, expansive mass with local invasion by capillary or sinusoidal structures.
2. Solid areas of collapsed vascular channels and areas with poorly formed, irregular-shaped vascular spaces.
3. An increase in the number of endothelial cells, many of which have pleomorphic, enlarged, hyperchromatic nuclei. Mitotic figures are usually present.
4. Metastases may occur.

REFERENCES

1. Andrews AG, Dysko RC, Spilman SC, Kunkel RG, Brammer DW, Johnson KJ. Immune complex vasculitis with secondary ulcerative dermatitis in aged C57BL/6NNia mice. *Vet Pathol.* 1994;31:293-300.
2. Ball CR, Clower BR, Williams WL. Dietary-induced atrial thrombosis in mice. *Arch Path.* 1965;80:391-396.
3. Brophy CM, Tilson JE, Braverman IM, Tilson MD. Age of onset, pattern of distribution, and histology of aneurysm development in a genetically predisposed mouse model. *J Vasc Surg.* 1988;8:45-48.
4. Brownstein DG. Genetics of dystrophic epicardial mineralization in DBA/2 mice. *Lab Anim Sci.* 1983;33:247-248.
5. Brunnert SR, Altman NH. Dystrophic cardiac calcinosis in mice: abnormal myocardial response to freeze-thaw injury. *Lab Anim Sci.* 1990;40:616-619.
6. Eaton GJ, Custer RP, Johnson FN, Stabenow KT. Dystrophic cardiac calcinosis in mice: genetic, hormonal, and dietary influences. *Am J Pathol.* 1978;90:173-186.
7. Elwell MR, Mahler JF. Heart, blood and lymphatic vessels. In: Maronpot RR, ed. *Pathology of the Mouse.* Clearwater, FL: Cache River Press; 1999.
8. Faccini JM, Abbott DP, Paulus GJJ. *Mouse Histopathology.* New York, NY: Elsevier Science Pub Co, Inc; 1990:64-71.
9. Frith CH, Ward JM. *Color Atlas of Neoplastic and Non-neoplastic Lesions in Aging Mice.* New York, NY: Elsevier Science Pub Co, Inc; 1988:1-6.
10. Fujihira S, Yamamoto T, Matsumoto M, et al. The high incidence of atrial thrombosis in mice given doxorubicin. *Toxicol Pathol.* 1993;21:362-368.
11. Gang DL, Barrett LV, Wilson EJ, Rubin RH, Medearis DN. Myopericarditis and enhanced dystrophic cardiac calcification in murine cytomegalovirus infection. *Am J Pathol.* 1986;124:207-215.
12. Hagiwara A, Tamano S, Hirose M. Changes in the heart. In: Mohr U, Dungworth DL, Capen CC, Carlton WW, Sundberg JP, Ward JM, eds. *Pathobiology of the Aging Mouse.* Washington, DC: ILSI Press; 1996:361-372.
13. Hashimoto I. Myocarditis in mice infected with Coxsackie virus B3. *Br J Exp Pathol.* 1986;67:737-745.
14. Hoch-Ligeti C, Restrepo C, Stewart HL. Comparative pathology of cardiac neoplasms in humans and in laboratory rodents: a review. *J Natl Cancer Inst.* 1986;76:127-142.
15. Huber SA. Coxsackievirus-induced myocarditis is dependent on distinct immunopathogenic responses in different strains of mice. *Lab Invest.* 1997;76:691-701.
16. Iwata H, Nomura Y, and Enomoto M. Spontaneous hemangioendothelial cell hyperplasia of the heart in B6C3F1 mice. *Toxicol Pathol.* 1994;22:423-429.
17. Klevay LM. Atrial thrombosis, abnormal electrocardiograms and sudden death in mice due to copper deficiency. *Atherosclerosis.* 1985;54:213-224.
18. Maita K, Hirano M, Harada T, et al. Mortality, major cause of moribundity, and spontaneous tumors in CD-1 mice. *Toxicol Pathol.* 1988;16:340-349.
19. Murray AB. Weibel-Palade bodies are not reliable ultrastructural markers for mouse endothelial cells. *Lab Anim Sci.* 1987;37:483-485.
20. National Toxicology Program (NTP). Toxicology and carcinogenesis studies of 1,3-butadiene in B6C3F1 mice. *NTP Tech Rep.* 1984;No. 288.
21. National Toxicology Program (NTP). Toxicity studies of urethane in F344 rats and B6C3F1 mice. *NTP Toxicity Rep.* 1996;No. 52.
22. Percy DH, Barthold SW. *Pathology of Laboratory Rodents and Rabbits.* Ames, Iowa: Iowa State University Press; 1993.
23. Plendl J, Kolle S, Sinowatz F, Schmahl W. Nonneoplastic lesions of blood vessels. In: Mohr U, Dungworth DL, Capen CC, Carlton WW, Sundberg JP, Ward JM, eds. *Pathobiology of the Aging Mouse.* Washington, DC: ILSI Press; 1996:385-392.
24. Price P, Papadimitriou JM. Genetic and immunologic determinants of age-related and virus-induced cardiopathy. In: Mohr U, Dungworth DL, Capen CC, Carlton WW, Sundberg JP, Ward JM, eds. *Pathobiology of the Aging Mouse.* Washington, DC: ILSI Press; 1996:373-384.
25. Steinhelper ME, Field LJ. Cardiac tumors and dysrhythmias in transgenic mice. *Toxicol Pathol.* 1990;18:464-469.
26. Stewart SE, Eddy BE, Borgese NG. Neoplasms in mice inoculated with a tumor agent carried in tissue culture. *J Natl Cancer Inst.* 1958;20:1223-1243.
27. Szepsenwol J, Boschetti NV. Primary and secondary heart tumors in mice maintained on various diets. *Oncology.* 1975;32:58-72.
28. Tsuchitani M, Umemura T, Narama I, Yanabe M. Naturally occurring Tyzzer's disease in a clean mouse colony: high mortality with coincidental cardiac lesions. *J Comp Pathol.* 1983;93:499-507.
29. Van Vleet JF, Ferrans VJ. Myocardial diseases of animals. *Am J Pathol.* 1986;124:98-178.
30. Ward JM, Goodman DG, Squire RA, Chu KC, Linhart MS. Neoplastic and nonneoplastic lesions in aging (C57BL/6N x C3H/HeN)F1 (B6C3F1) mice. *J Natl Cancer Inst.* 1979;63:849-854.
31. Yamate J, Tajima M, Kudow S. Comparison of age-related changes between long-lived CRJ:CD-1 (ICR) and CRJ:B6C3F1 mice. *Nippon Juigaku Zasshi.* 1986;48:273-284.
32. Zhang SH, Reddick RL, Piedrahita JA, Maeda N. Spontaneous hypercholesterolemia and arterial lesions in mice lacking apolipoprotein E. *Science.* 1992;258:468-471.

REFERENCES

1. Andrews AG, Dysko RC, Spilman SC, Kunkel RG, Brammer DW, Johnson KJ. Immune complex vasculitis with secondary ulcerative dermatitis in aged C57BL/6NNia mice. *Vet Pathol*. 1994;31:293-300.
2. Ball CR, Clower BR, Williams WL. Dietary-induced atrial thrombosis in mice. *Arch Path*. 1965;80:391-396.
3. Brophy CM, Tilson JE, Braverman IM, Tilson MD. Age of onset, pattern of distribution, and histology of aneurysm development in a genetically predisposed mouse model. *J Vasc Surg*. 1988;8:45-48.
4. Brownstein DG. Genetics of dystrophic epicardial mineralization in DBA/2 mice. *Lab Anim Sci*. 1983;33:247-248.
5. Brommer SR, Altman NH. Dystrophic cardiac calcinosis in mice: abnormal myocardial response to freeze-thaw injury. *Lab Anim Sci*. 1990;40:616-619.
6. Eaton GJ, Custer RP, Johnson FN, Stabenow KT. Dystrophic cardiac calcinosis in mice: genetic, hormonal, and dietary influences. *Am J Pathol*. 1978;90:173-186.
7. Elwell MR, Mahler JF. Heart, blood and lymphatic vessels. In: Maronpot RR, ed. *Pathology of the Mouse*. Clearwater, FL: Cache River Press; 1999.
8. Faccini JM, Abbott DP, Paulus GJJ. *Mouse Histopathology*. New York, NY: Elsevier Science Pub Co, Inc; 1990:64-71.
9. Frith CH, Ward JM. *Color Atlas of Neoplastic and Non-neoplastic Lesions in Aging Mice*. New York, NY: Elsevier Science Pub Co, Inc; 1988:1-6.
10. Fujihira S, Yamamoto T, Matsumoto M, et al. The high incidence of atrial thrombosis in mice given doxorubicin. *Toxicol Pathol*. 1993;21:362-368.
11. Gang DL, Barrett LV, Wilson EJ, Rubin RH, Medearis DN. Myopericarditis and enhanced dystrophic cardiac calcification in murine cytomegalovirus infection. *Am J Pathol*. 1986;124:207-215.
12. Hagiwara A, Tamano S, Hirose M. Changes in the heart. In: Mohr U, Dungworth DL, Capen CC, Carlton WW, Sundberg JP, Ward JM, eds. *Pathobiology of the Aging Mouse*. Washington, DC: ILSI Press; 1996:361-372.
13. Hashimoto I. Myocarditis in mice infected with Coxsackie virus B3. *Br J Exp Pathol*. 1986;67:737-745.
14. Hoch-Ligeti C, Restrepo C, Stewart HL. Comparative pathology of cardiac neoplasms in humans and in laboratory rodents: a review. *J Natl Cancer Inst*. 1986;76:127-142.
15. Huber SA. Coxsackievirus-induced myocarditis is dependent on distinct immunopathogenic responses in different strains of mice. *Lab Invest*. 1997;76:691-701.
16. Iwata H, Nomura Y, and Enomoto M. Spontaneous hemangioendothelial cell hyperplasia of the heart in B6C3F1 mice. *Toxicol Pathol*. 1994;22:423-429.
17. Klevay LM. Atrial thrombosis, abnormal electrocardiograms and sudden death in mice due to copper deficiency. *Atherosclerosis*. 1985;54:213-224.
18. Maita K, Hirano M, Harada T, et al. Mortality, major cause of moribundity, and spontaneous tumors in CD-1 mice. *Toxicol Pathol*. 1988;16:340-349.
19. Murray AB. Weibel-Palade bodies are not reliable ultrastructural markers for mouse endothelial cells. *Lab Anim Sci*. 1987;37:483-485.
20. National Toxicology Program (NTP). Toxicology and carcinogenesis studies of 1,3-butadiene in B6C3F1 mice. *NTP Tech Rep*. 1984;No. 288.
21. National Toxicology Program (NTP). Toxicity studies of urethane in F344 rats and B6C3F1 mice. *NTP Toxicity Rep*. 1996;No. 52.
22. Percy DH, Barthold SW. *Pathology of Laboratory Rodents and Rabbits*. Ames, Iowa: Iowa State University Press; 1993.
23. Plendl J, Kolle S, Sinowatz F, Schmahl W. Nonneoplastic lesions of blood vessels. In: Mohr U, Dungworth DL, Capen CC, Carlton WW, Sundberg JP, Ward JM, eds. *Pathobiology of the Aging Mouse*. Washington, DC: ILSI Press; 1996:385-392.
24. Price P, Papadimitriou JM. Genetic and immunologic determinants of age-related and virus- induced cardiopathy. In: Mohr U, Dungworth DL, Capen CC, Carlton WW, Sundberg JP, Ward JM, eds. *Pathobiology of the Aging Mouse*. Washington, DC: ILSI Press; 1996:373-384.
25. Steinhelper ME, Field LJ. Cardiac tumors and dysrhythmias in transgenic mice. *Toxicol Pathol*. 1990;18:464-469.
26. Stewart SE, Eddy BE, Borgese NG. Neoplasms in mice inoculated with a tumor agent carried in tissue culture. *J Natl Cancer Inst*. 1958;20:1223-1243.
27. Szepsenwol J, Boschetti NV. Primary and secondary heart tumors in mice maintained on various diets. *Oncology*. 1975;32:58-72.
28. Tsuchitani M, Umemura T, Nozima I, Yanabe M. Naturally occurring Tyzzer's disease in a clean mouse colony: high mortality with coincidental cardiac lesions. *J Comp Pathol*. 1983;93:499-507.
29. Van Vleet JF, Ferrans VJ. Myocardial diseases of animals. *Am J Pathol*. 1986;124:98-178.
30. Ward JM, Goodman DG, Squire RA, Chu KC, Linhart MS. Neoplastic and nonneoplastic lesions in aging (C57BL/6N x C3H/HeN)F1 (B6C3F1) mice. *J Natl Cancer Inst*. 1979;63:849-854.
31. Yamate J, Tajima M, Kudow S. Comparison of age-related changes between long-lived CRJ:CD-1 (ICR) and CRJ:B6C3F1 mice. *Nippon Juigaku Zasshi*. 1986;48:273-284.
32. Zhang SH, Reddick RL, Piedrahita JA, Maeda N. Spontaneous hypercholesterolemia and arterial lesions in mice lacking apolipoprotein E. *Science*. 1992;258:468-471.

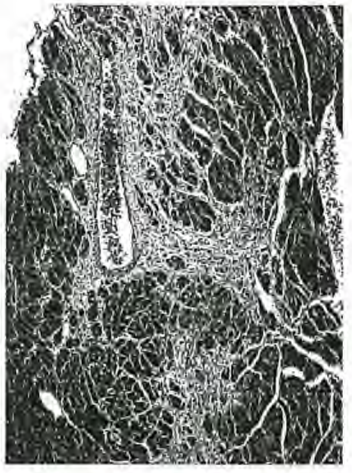


Fig. 1 - Murine cardiomyopathy, heart.

Fig. 2 - Amyloidosis, heart.

Fig. 3 - Amyloidosis, heart.

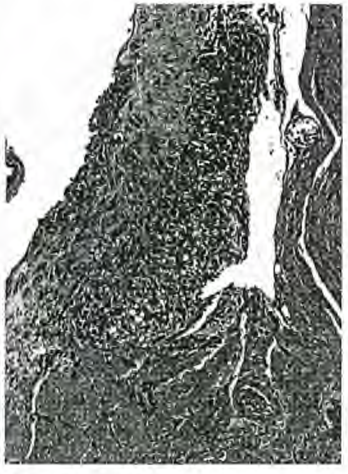


Fig. 4 - Inflammation, heart.

Fig. 5 - Mineralization, heart.

Fig. 6 - Degeneration, heart.

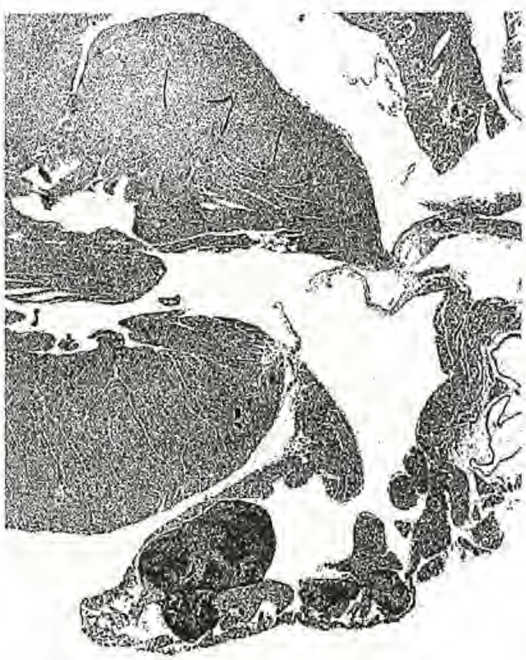


Fig. 7 - Thrombus, atrium, heart.

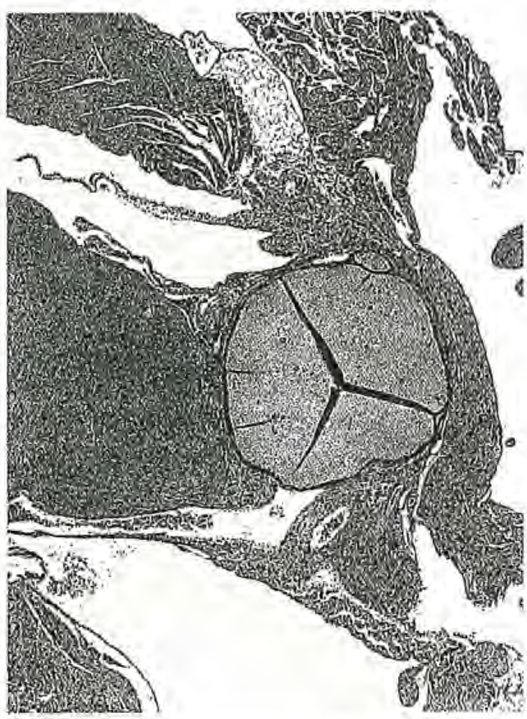


Fig. 8 - Ectopic tissue (embryonic duct cyst), heart.

Fig. 5 - Mineralization, heart.

Fig. 6 - Degeneration, heart.

Fig. 7 - Thrombus, atrium, heart.

Fig. 8 - Ectopic tissue (embryonic duct cyst), heart.

Fig. 9 - Lymphoma, heart.

Fig. 10 - Arteritis, meninges.

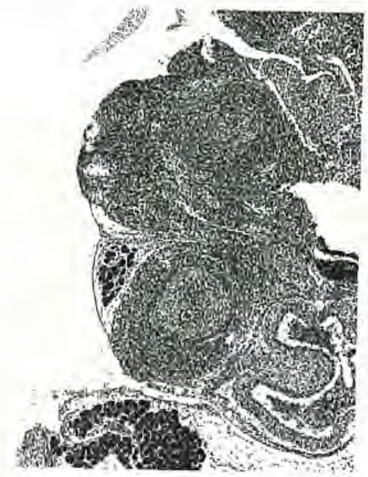


Fig. 11 - Arteritis, mesentery.

Fig. 12 - Arteritis, renal artery.

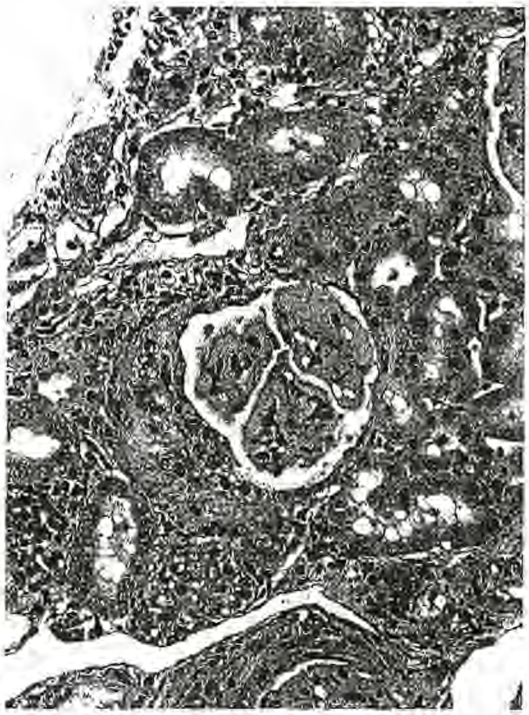


Fig. 13 - Amyloidosis, kidney.

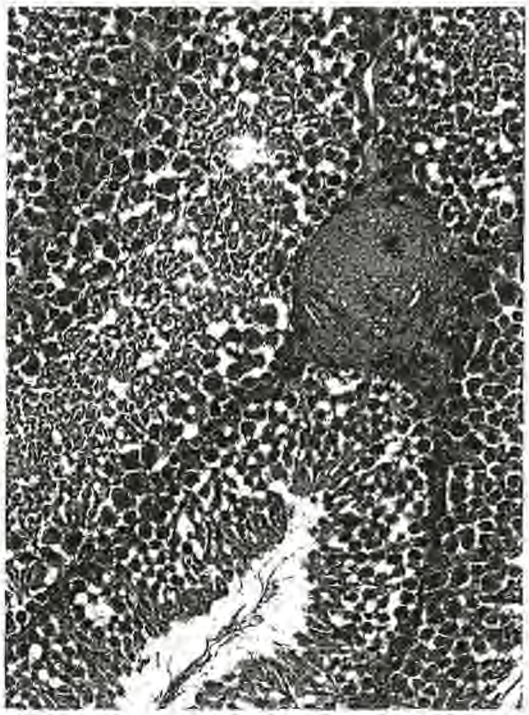


Fig. 14 - Amyloidosis, testis.

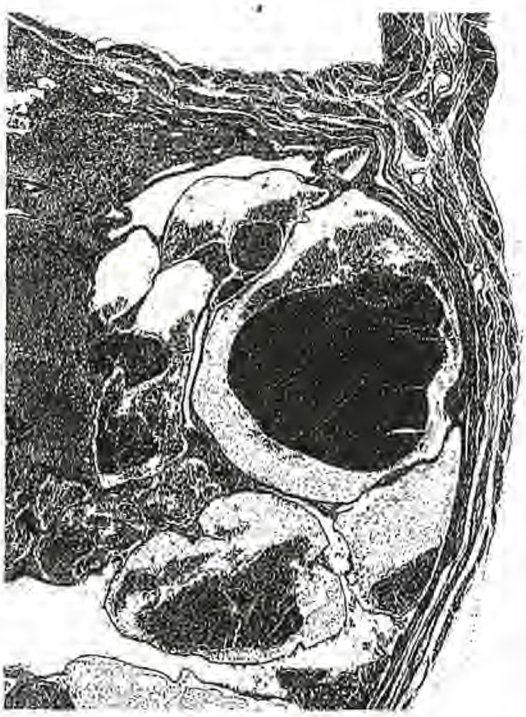


Fig. 15 - Angiectasis, uterus.

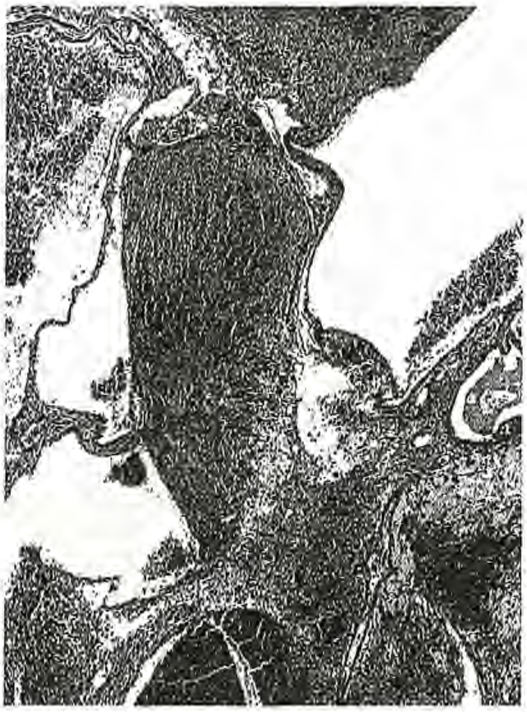


Fig. 16 - Thrombus (in angiectatic vessel), uterus.

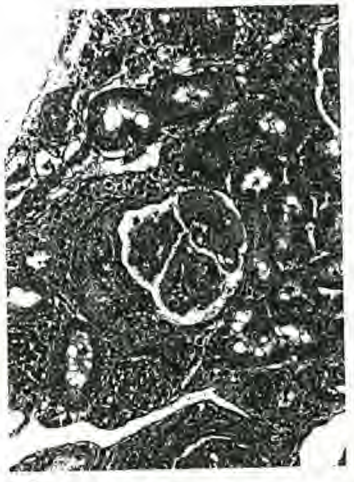


Fig. 13 - Amyloidosis, kidney.

Fig. 14 - Amyloidosis, testis.

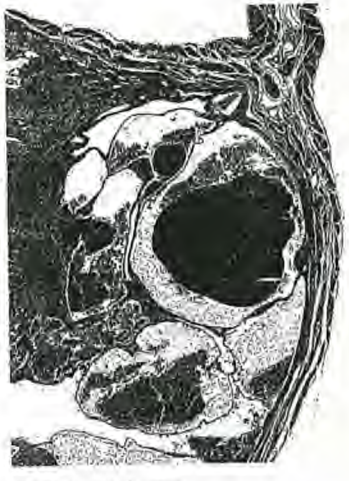


Fig. 15 - Angiectasis, uterus.

Fig. 16 - Thrombus (in angiectatic vessel), uterus.

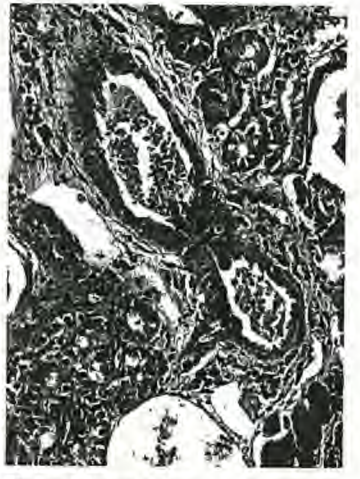


Fig. 17 - Mineralisation, kidney.

Fig. 18 - Medial hypertrophy, artery, lung.

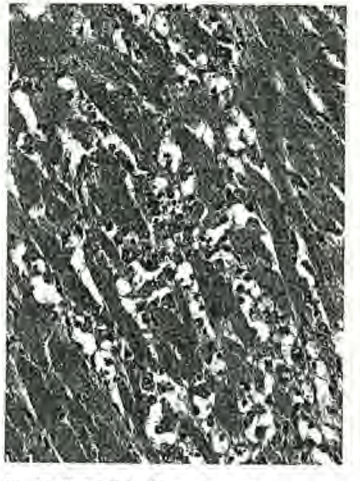


Fig. 19 - Endothelial cell hyperplasia, heart.

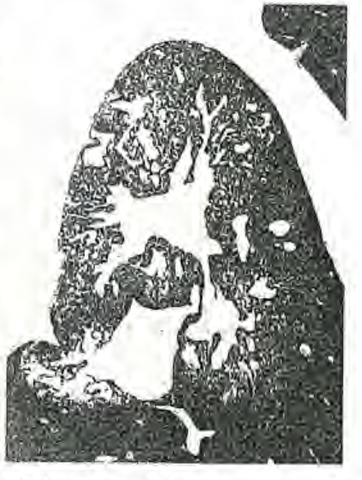


Fig. 20 - Hemangioma, liver.

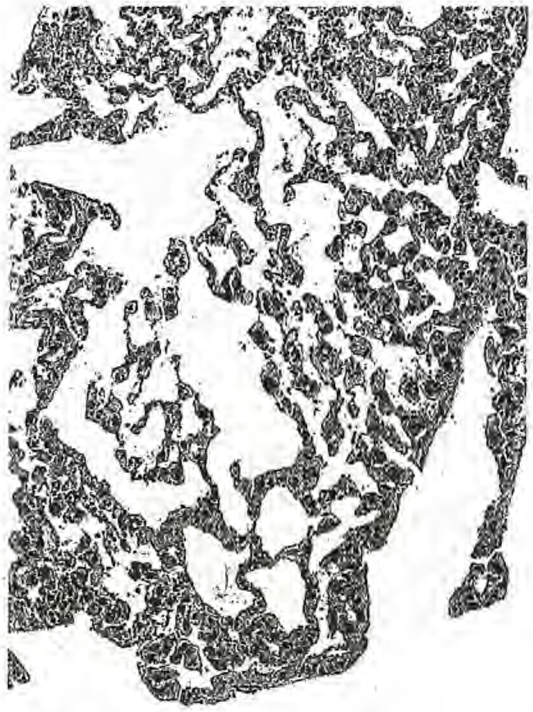


Fig. 21 - Hemangioma, liver.

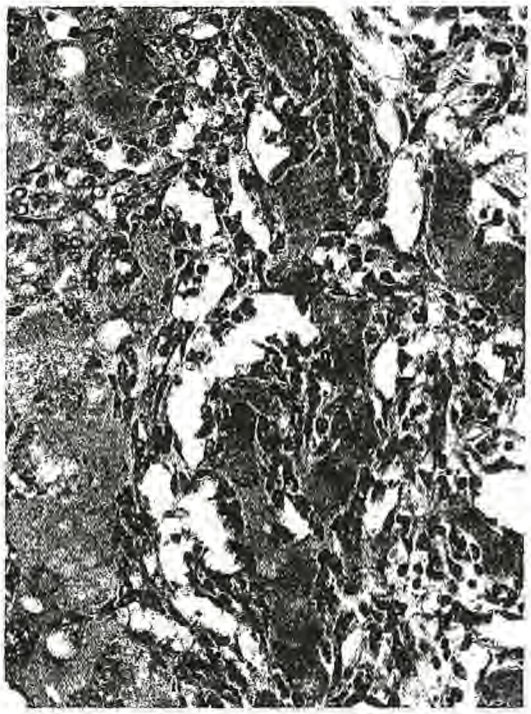


Fig. 22 - Hemangiosarcoma, liver.

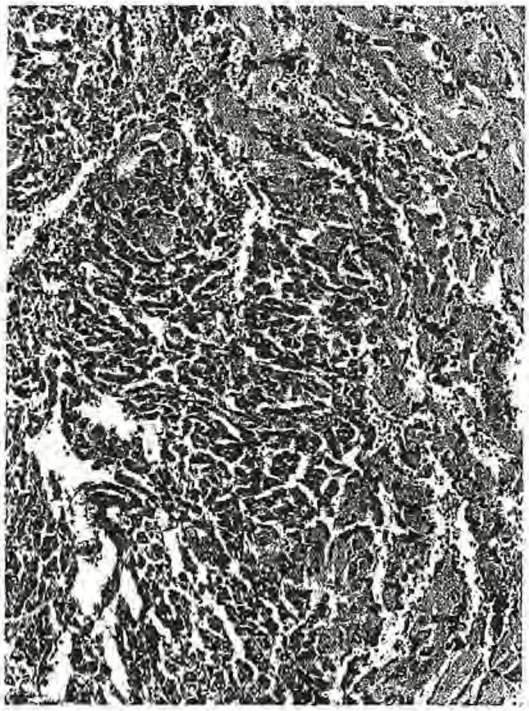


Fig. 23 - Hemangiosarcoma, heart.