

EXHIBIT 108

Regulatory Forum

Toxicologic Pathology, 31: 679-693, 2009
 Copyright © 2009 by The Author(s)
 ISSN: 0192-6233 print / 1533-1601 online
 DOI: 10.1177/0192623309336154

Best Practices for Use of Historical Control Data of Proliferative Rodent Lesions

CHARLOTTE KEENAN,¹ SUSAN ELMORE,² SABINE FRANCKE-CARROLL,³ RAMON KEMP,⁴ ROY KERLIN,⁵ SHYAMAL PEDDADA,⁶
 JOHNN PLETCHER,⁷ MATTHIAS RINKE,⁸ STEPHEN PETER SCHMIDT,⁵ IAN TAYLOR,⁹ AND DOUGLAS C. WOLF¹⁰

¹GlaxoSmithKline, King of Prussia, Pennsylvania, USA

²National Toxicology Program (NTP), National Institute of Environmental Health Sciences (NIEHS),
 Research Triangle Park, North Carolina, USA

³Center for Food Safety and Applied Nutrition (CFSAN), U.S. Food and Drug Administration, College Park, Maryland, USA

⁴Merck Research Laboratories, Kenilworth, New Jersey, USA

⁵Pfizer Inc., Groton, Connecticut, USA

⁶NIEHS, Research Triangle Park, North Carolina, USA

⁷Charles River, Frederick, Maryland, USA

⁸Bayer Schering Pharma AG, Wuppertal, Germany

⁹Huntingdon Life Sciences, Eye, United Kingdom

¹⁰U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA

Keywords: historical control data; rodent tumors; carcinogenicity studies; best practices.

SUMMARY AND BEST PRACTICES RECOMMENDATION

The Historical Control Data Working Group, under the direction of the Scientific and Regulatory Policy Committee (SRPC) of the Society of Toxicologic Pathology (STP) was tasked with reviewing the current scientific practices, regulatory guidance, and relevant literature pertaining to rodent microscopic historical control data (HCD) of proliferative lesions to provide best practice recommendations for locating, generating, and applying such data. The Working Group focused exclusively on HCD of proliferative lesions from nonclinical rodent carcinogenicity studies. The HCD Working Group recommends the following consensus principles to guide the use of HCD of proliferative lesions from chronic rodent (rats/mice) bioassays:

- The concurrent control group is the most relevant comparator for determining treatment-related effects in a study.

Address correspondence to: Charlotte Keenan, GlaxoSmithKline, 709 Swedeland Rd, UE0360, King of Prussia, PA 19406, USA; e-mail: charlotte.m.keenan@gsk.com.

The views expressed in this article are those of the authors and do not necessarily represent the policies, positions, or opinions of their respective agencies and organizations. This research was supported (in part) by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. The recommendations in this article are endorsed by the European Society of Toxicologic Pathology and the British Society of Toxicologic Pathologists.

Conflict of interest: The authors have not declared any conflict of interest.

- HCD may be useful in the interpretation of rare tumors, marginally greater incidences and/or severity of proliferative changes in treated animals compared to controls, and unexpected increases or decreases of tumor incidences in study control animals. HCD can be used as a tool to provide scientific perspective of disparate findings in dual concurrent control groups and review trends in tumor biology and behavior that may evolve over time in these rodent models.
- Study design-related parameters such as laboratory, species/strain, route of administration, vehicle, feed, feeding practices, study duration, and housing have a potential to impact study outcomes and control findings. These parameters should be considered when selecting the appropriate studies for the HCD.
- Pathology practices, including necropsy and trimming procedures and application of diagnostic criteria, can impact study data and HCD. HCD are best if these factors are standardized.
- HCD from the laboratory that conducted the study under review will likely be more comparable than HCD compiled from several laboratories.
- Similarly, HCD that underwent a peer-review process are generally more reliable than those that did not.
- Published HCD should be evaluated carefully. It may provide guidance in evaluating data associated with particular effects, but difficulties in assessing the

quality of published data should be considered along with the "weight of evidence" for determining its relevance to study findings.

- HCD may be presented as a range of incidences or percentages, mean, and standard deviation for a given change. Reporting of incidences per study will allow both presentation and use of a broad range of observations and provide transparency of potential influences of outlier populations.
- Although a limited time span of two to seven years for collection of HCD is proposed in the guidance documents of several agencies, wider intervals may be appropriate if tumor types are stable over a longer period.
- HCD should be considered as one of many sources of information that add to the "weight of evidence" approach when assessing the potential carcinogenic effect of a compound.

INTRODUCTION

The Historical Control Data Working Group, under the direction of the SRPC of the STP, was tasked with reviewing current scientific practices and the relevant literature and regulatory guidances to provide best practice recommendations for locating, accessing, generating, and applying rodent microscopic HCD for proliferative lesions from rodent carcinogenicity studies.

The usefulness of HCD has been an ongoing topic of vigorous discussion within the field of toxicologic pathology. Authors have either dismissed or supported the use of HCD as a reliable comparator to concurrent study-specific control data. However, there is general consensus that for any particular study, the incidence of proliferative lesions in the concurrent control group or groups is the most appropriate and accurate comparison to that of the treated groups. When HCD are used as a comparator, the design of studies comprising them should be similar to the study under review (Boorman et al. 2002; Deschl et al. 2002; Gopinath 1994; Greirn et al. 2003; Haseman, Huff, and Boorman 1984; Haseman 1995a, 1995b; Keenan et al. 2002; van Zwieten et al. 1988).

The use of HCD can be valuable when, within a specific study, the concurrent control results give equivocal comparisons and interpretation to the treated groups or when there is a need to provide quality control for intercurrent factors that may have compromised the survival of the control or treated animals (Deschl et al. 2002; van Zwieten et al. 1988). A typical example of a result that may be difficult to explain would be the presence or high incidence of a rare tumor or other uncommon findings in the treated animals compared to the control animals. The reverse situation would be when an unusually variable or high incidence of a tumor type is found in the control animals, but at a lower incidence in the treated animals, thereby possibly masking a compound-related effect (Deschl et al. 2002; van Zwieten et al. 1988).

The potential limitation for the use of HCD generally focuses on the variability and drift over time in animal and

study-related factors such as animal genetics, the experimental environment, and the macroscopic and microscopic pathological interpretations (Wolf and Mann 2005). It is generally believed that HCD from within a laboratory are more homogeneous. In contrast, there is a concern that data from multiple laboratories may be of limited interpretive value because of increased variability of diagnostic interpretation (Roe 1994; Yoshimura and Matsumoto 1994).

Progress has been made in recent years by various groups such as the Registry of Industrial Toxicology Animal Data (RITA), North American Control Animal Database (NACAD) International Agency for Research on Cancer (IARC), National Toxicology Program (NTP), STP, and others to reduce some of the sources of variability with international efforts to improve harmonization and standardization of terminology, trimming procedures, and study designs. Examples include the efforts by the International Conference on Harmonization (ICH; <http://www.ich.org/cache/compo/502-272-1.html>); guides produced by the International Agency for Research on Cancer (IARC) (*International Classification of Rodent Tumours*, part 1, *Rat*, ed. U. Mohr [Lyon, France: IARC Scientific Publications, 1992]; complete listing provided at http://reni.item.fraunhofer.de/reni/public/rita/icrt_rathtrn [accessed March 17, 2009]); the Revised Guides for Organ Sampling and Trimming in Rats and Mice: Parts 1-3 provided by the RITA group (<http://www.item.fraunhofer.de/reni/trimming/index.php> [accessed March 13, 2009]); and the Society of Toxicologic Pathology Standardized System of Nomenclature and Diagnostic Criteria (SSNDC) Guides (<http://www.toxpath.org/ssndc.asp>). The U.S. Food and Drug Administration (FDA) has recognized the efforts to harmonize, and this is reflected in the Redhook (<http://www.cfsan.fda.gov/~redbook/red-ivc6.html>). The STP also regularly publishes position papers outlining Best Practice Guidelines (Crissman et al. 2004; Morton et al. 2006). In 1994, a project to harmonize rat nomenclature was established by the Joint STPs and International Life Sciences Institute (ILSI) Committee on International Harmonization of Nomenclature and Diagnostic Criteria in Toxicologic Pathology. The results of discussions among the Rat Nomenclature Reconciliation Subcommittees, a presentation during the 1999 annual STP meeting in Washington, D.C., and comments that have been sent to the committee are available on Web sites (http://www.item.fraunhofer.de/reni/rat_nomenclature/index.htm [accessed March 13, 2009]; <http://www.toxpath.org/normen/index.htm> [accessed March 13, 2009]). This project evolved into the current International Harmonization of Nomenclature and Diagnostic Criteria for Lesions in Rats and Mice (INHAND) initiative. This project is sponsored by several Societies of Toxicologic Pathology: Japanese Society of Toxicologic Pathology (JSTP), British Society of Toxicologic Pathologists (BSTP) and European Society of Toxicologic Pathology (ESTP) and has been organized to describe and publish uniform nomenclature for both proliferative and nonproliferative lesions in laboratory rodents (http://www.goren.org/iback_inhand.php [accessed March 13, 2009]). This harmonization of nomenclature should improve uniformity and accuracy of histopathological diagnoses and further increase the value of HCD.

At present, detailed procedures that assist with the use of HCD in interpreting the significance of lesions within a rodent study submitted for regulatory review are not available. To gain a better understanding of how HCD are currently used in the professional community, a survey was undertaken. The survey questions focused on current practices and use of historical data, the source of data, factors affecting the model system, presentation of the data, and demographic information. Respondents represented the pharmaceutical and chemical industry as well as government and regulatory bodies.

This article summarizes existing regulatory guidance pertaining to HCD; provides a report of current practice by interpreting the results of the survey; presents a review of the relevant literature; and compiles recommendations for generating, locating, accessing, and applying HCD.

REGULATORY GUIDANCE

A short summary of guidances from a variety of regulatory agencies addressing HCD is provided.

U.S. Environmental Protection Agency (EPA)

The U.S. EPA Cancer Guidelines consider that the standard for determining statistical significance of tumor incidences comes from a comparison of tumors in dosed animals with those in concurrent control animals; however, additional insights concerning statistical and biological significance can come from appropriate HCD. Historical control data can provide information on the biological significance of uncommon tumor types or when there is a high spontaneous tumor incidence. The guidance also suggests that caution be exercised in using ranges because the range does not account for survival and is dependent on the number of animals in the database. It is also suggested that the most relevant historical data come from the same laboratory and the same animal supplier, and that the data be gathered within two or three years prior to or following the study under review. In evaluating the data from historical controls, statistically significant increases in tumors based on the concurrent control should not be discounted simply because incidence rates in the treated groups are within the range of historical controls or because incidence rates in the concurrent controls are low. Proper study design, including appropriate randomization and statistical procedures, should provide confidence that statistically significant results are not due to chance (U.S. EPA 2005).

U.S. Food and Drug Administration (FDA)

The information most often requested by the FDA for a follow-up pathology review is clarification of diagnostic criteria and HCD for a lesion in question (U.S. FDA Center for Food Safety and Applied Nutrition [CFSAN] 2000). In the FDA's draft guidance on the statistical aspects of analysis, design, and interpretation of rodent carcinogenicity studies (U.S. FDA Center for Drug Evaluation and Research [CDER] 2001), it is stated that "the concurrent control group is always the most

appropriate and important in testing drug related increases in tumor rates in a carcinogenicity experiment." However, HCD, as long as the data are chosen from studies comparable to the study in question, are considered useful in classifying tumors as common or rare, and as a quality control tool for establishing the reasonableness of the spontaneous tumor rates in the concurrent control groups. As an example, CFSAN FDA recognizes the beneficial effects of caloric restriction on the survivability of certain species. However, the FDA requires "sufficient" HCD from studies using such methods before they can accept results from studies using caloric restriction (U.S. FDA CFSAN 2000, 2006, 2007). HCD can help investigators reduce false positive and false negative results and may be incorporated into formal statistical tests for analyzing study tumor rates. The validity of these tests, however, can be dependent on the availability of a large database of appropriate HCD (U.S. FDA CDER 2001).

European Medicines Agency (EMA)

The note for guidance on carcinogenic potential, produced by the EMA (2002), states that the concurrent control group should be regarded as the primary reference when considering treatment-related carcinogenicity. When HCD are used, they should be from the same strain of animal and the same testing laboratory, performed during the five years prior to the study in question. If data from the literature are considered to be informative, they may also be added. Any increase in tumor incidence should be interpreted in the light of the historical incidence of that tumor, and if HCD are used to support the interpretation of that study, those data should also be included in the study report.

Although the 2004 guidance states that a factor limiting the regulatory acceptance of transgenic models for carcinogenicity assessment is the relatively small database for neoplastic HCD available for these models (EMA 2004), the body of information on these models continues to grow.

Organization for Economic Cooperation and Development (OECD)

The OECD describes HCD from the same strain of animals, derived from studies run under the same laboratory conditions, as "desirable" or "indispensable" for correctly assessing the significance of changes in the numbers of tumors (and other lesions normally occurring in the strain of animals used) when interpreting the results of chronic toxicity or carcinogenicity studies (OECD 1981a, 1981b, 1981c).

In the OECD guidance for the evaluation and analysis of chronic toxicity and carcinogenicity studies (OECD 2002), it is stated that concurrent control groups should always be used and that statistical comparisons with HCD are generally not appropriate because of the many variables affecting the incidence of spontaneous tumors. The guidance does, however, note that HCD can be useful for establishing the acceptability of the "normal" data from the control groups and for judging the biological significance of the occurrence of rare or unusual tumors and nonneoplastic abnormalities. However, the

guidance stresses that any HCD used should be comparable with the study in question. Ideally, HCD should be generated by the same laboratory at which the study being assessed was performed. Furthermore, the HCD should come from studies conducted within five years prior to or two to three years after the study being evaluated. The guidance recommends that parameters that could affect the occurrence of spontaneous tumors in HCD are identified. The guidance refers the reader to the European requirements for the submission of HCD in the European Commission (EC) Directive 91/414/EEC (European Economic Community 1991). Within this directive, Amendment M4-Section 5.5 Long Term Toxicity and Carcinogenicity, Test Conditions, pp. 61-62, states that while the standard reference point for evaluating treatment responses in long-term toxicity and carcinogenicity studies is the concurrent control group, HCD may be helpful in the interpretation of particular studies. The HCD should, however, come from animals of contemporaneous studies in the same species and strain, maintained under similar conditions.

Japan

There is no specific Japanese guideline; the Ministry of Health, Labour and Welfare follows the ICH guideline for carcinogenicity studies-<http://www.pmda.go.jp/ich/safety.htm> (accessed March 13, 2009). The ICH guideline does not specifically address the use of HCD. The Expert Committee of Food Safety Commission, which evaluates the safety of pesticides, often requests HCD when study data are difficult to interpret (Matthias Rinke, February 18, 2008, Bayer Schering Pharma AG, Wuppertal, Germany). Use of HCD is recognized for evaluation of chemical carcinogenicity but is not specifically described in the guidelines for pesticides.

In summary, all regulatory guidance suggests that the concurrent control should always be the principal comparator and that properly selected HCD can be informative in specific cases where there may be questions on whether the treated groups show an effect or not.

SURVEY RESULTS

An informal survey (Table 1) was sent to more than one hundred contacts covering a wide range of industry and regulatory agencies throughout the world, and fifty-five responses were returned. Responses consisted of individual opinions of contact representatives as well as of single, harmonized responses from organizations with multiple operational sites. Because of the relatively small sample size, no formal statistical analysis was possible. The responses were suitably proportional to the groups surveyed with the majority from pharmaceutical industry, followed by chemical industry, contract research organizations (CROs), government research organizations (e.g., NIH), and regulatory authorities. Regulatory respondents included staff from U.S. FDA, U.S. EPA, Health Canada, the European Union (EU), and Japan. Approximately 443 of respondents were from the United States, 333

from Europe, 73 from Japan, 43 from Canada, and 123 indicated a worldwide location.

For most survey questions, respondents had more than one option to answer the question (Table 1); therefore, the total of all percentage rates given is greater than 1003.

In general, 493 of respondents indicated reviewing between one and five carcinogenicity studies per year. Approximately 243 of respondents review more than five studies a year. This suggests a substantial experience base among respondents as well as a substantial need for availability of high-quality HCD on a regular basis.

The majority of respondents use HCD whenever there are statistically significant increased incidences of neoplastic or hyperplastic lesions (913). Most respondents replied that HCD are particularly useful for the interpretation of rare proliferative lesions (933) or borderline differences from concurrent control groups (733). Frequently respondents stated they consult HCD for biological trends regardless of statistical significance (603), incorporate HCD as part of their interpretation on the request of regulatory agencies (533), and/or use them to evaluate the consistency of the concurrent control data (453).

The majority of respondents (603) prefer to use HCD from animal models similar to those of the study being evaluated. Respondents felt that rodent strain, feeding practices, and route of administration are the three most important comparators when evaluating HCD.

Typically, respondents use up to five years of collected HCD if available (533), while 333 of respondents include more than five years, 133 include the most recent three years, and 13 include only the most recent two years.

On how data are routinely presented, most respondents chose more than one of the five possible answers; therefore, the total percentage rate is greater than 1003. However, most respondents chose to report a historical control range of incidence (743) followed by presentation of HCD as percent (513) and range or mean with scientific opinion on relevance (443). A majority of respondents (913) report HCD without performing statistical analyses.

There is no consistent preference for the source of the data as respondents appeared to routinely use with equal frequency published literature (713), internal databases (703), public and nonpublic sources (e.g., the public Charles River [563] and NTP [543] databases and the proprietary database of the RITA [343]) as well as data from the laboratory that conducted the study (503). About two-thirds of the respondents using RITA as a source for HCD are located in Europe or are part of globally operating organizations.

An important concern expressed by 673 of respondents is the lack of consistency and robustness of the databases available to obtain HCD. As a consequence, some respondents commented that HCD are not applied appropriately in evaluating the relevance of proliferative lesions. Specifically, the respondents' concerns centered on the variability in recording background lesions due to the lack of a standard lexicon; variability in the experience of the study pathologist; and differences between studies as well as dissimilar parameters used

TABLE 1.-Survey questions.

-
1. How many carcinogenicity studies on average do you evaluate per year (select one)?
 - a. <1
 - b. 1-5
 - c. >5
 2. When do you use historical control data (select all appropriate)?
 - a. Presence of statistically significant neoplastic or hyperplastic data
 - b. Biological trend without statistical significance
 - c. Presence of rare proliferative lesion
 - d. Proliferative/neoplastic lesion that shows a borderline increase relative to concurrent controls
 - e. For quality control aspects to determine if concurrent control tumor or proliferative lesion incidences are consistent with previously reported tumor/proliferative lesion rates
 - f. At the request of a regulatory agency
 - g. Do not use historical data
 - h. Other
 3. Where do you obtain historical data (select all appropriate)?
 - a. Internal database
 - b. Registry of Industrial Toxicology Animal-data (RITA)
 - c. External database
 - i. Contract organization conducting your study
 - ii. Contract organization at location other than study conduct
 - iii. Published literature (peer reviewed journals)
 - iv. Public database
 1. NTP
 2. Charles River
 - d. Other
 4. How similar is the animal model system from which historical control data were derived to the animal model system of the Study being conducted and evaluated (select one)?
 - a. Very Similar-ensure that both model systems use similar feeding practices, vehicles, route of administration, vendor sources, strains, caging
 - b. Somewhat similar-go to question 5 below
 - c. Don't Know (data are not available from my source)
 - d. Don't Know (generally this has not been a consideration for use of historical control data)
 5. If the model system is somewhat similar what factors are considered most critical (select top three choices)?
 - a. Feeding practice
 - b. Caging
 - c. Vehicles
 - d. Route of administration
 - e. Vendor source
 - f. Strain
 - g. Other
 6. How many years do you routinely include (select one)?
 - a. Most recent 2 yrs
 - b. Most recent 3 yrs
 - c. Most recent 5 yrs
 - d. >5 yrs
 7. How are the data presented (select all appropriate)?
 - a. Range of incidence
 - b. Percent
 - c. Mean with standard deviation
 - d. Range or mean with scientific opinion on relevance
 - e. Other
 8. Do you conduct statistical analysis on historical data (select one)?
 - a. No
 - b. Yes
 - c. If yes, what statistical tools are used
 9. Demographic information: your geographic location
 - a. Canada
 - b. Europe
 - c. Japan
 - d. United States
 - e. Worldwide (multiple locations)
-

(continued)

TABLE I -(continued)

-
10. Demographic information: Industry
- Pharmaceutical/biologics
 - Chemical
 - CRO
 - Independent consultant
11. Demographic information: Government agency:
- FDA
 - EPA
 - NIH
 - EMEA
 - Health Canada
 - Other, please specify
12. What is your major concern with the use of historical control data?
13. Please tell us ways in which the use of historical data can be improved
-

with regard to animal age, feed, strain, caging, and dose regimen.

The consistency of the application of diagnostic criteria (263); the fact that HCD are overvalued, overinterpreted, or abused (233); and the lack of a reliable source for robust HCD (233) were additional concerns expressed by respondents. Some presented concerns with genetic drift (193) and the need for considering the biological significance of study findings and HCD (163).

A common theme in the "comments" section of the survey for ways to improve HCD was the desire for a readily accessible, centralized, independent, and standardized database for HCD. Many respondents suggested that such a database for HCD would be best served by the use of harmonized diagnostic criteria, terminology, and enhanced statistical methods.

FACTORS TO CONSIDER WHEN USING HISTORICAL DATA

When using HCD, several factors must be considered that can be consolidated into two categories: the *in-life* (strain, age, study duration, body weight, housing, route of administration, diet, vehicle, and test article cross-contamination of controls) and the *postmortem* factors (necropsy, trimming, histopathologic criteria, and terminology).

Strain (Source/Vendor, Age-Matched)

Among the *in-life* factors, the strain and source of the animal model play an important role. It is well known that there are major differences in the incidences of certain lesions in different strains of rats and mice. This has resulted in publications presenting lists of findings from large laboratories or CROs. Frequently these are the only HCD references that regulatory organizations and small companies can access. The pattern of background proliferative lesions may be different between outbred stocks and inbred strains. Therefore, it may not be desirable to compare frequently diagnosed proliferative lesions across stocks and strains. However, this may be the only option in cases of exceedingly rare tumors. Also, there might be considerable variation within strains depending on the breeder and,

among stocks, there might be further variation due to housing conditions (van Zwieten et al. 1988; Deschl et al. 2002). There may also be a lack of genetic quality control, which in one reported case resulted in an error in supply, and animals with a wrong strain declaration were employed in a study (Roe 1994). However, most vendors perform genetic quality control tests periodically and have the necessary documentation on file. For example, some large suppliers of laboratory animals perform genetic tests of inbred strains quarterly by a panel of markers that can distinguish among all inbred strains produced by the company, primarily for evidence of accidental breeding with other strains. Outbred stocks are monitored for evidence of loss of heterozygosity or dramatic shifts in allele frequency. Because of the large breeding populations and different focus of the genetic testing, monitoring is conducted less frequently than for inbred strains. Outbred stocks are genetically monitored every five years or approximately every ten generations (Charles Clifford, March 14, 2008, Charles River, Wilmington, MA).

There is a general presumption that genetic drift occurs over time, resulting in changes in the reported frequency of certain lesions leading to the recommendation of using only HCD of certain time intervals (U.S. EPA, <http://www.epa.gov/iris/backgr-d.htm>; EMEA 2002). However, longitudinal analyses (Eiben and Bombard 1999; Eiben 2001; Haseman et al. 1989; Tennekens et al. 2004a, 2004b) have shown that some parameters may change while others remain stable for periods of up to twenty years (Deschl et al. 2002).

As the species and strain of the animals used in carcinogenicity testing of chemicals influence the spontaneous tumor profile, likewise the sex of the test animal also influences the profile. Gender differences in the spontaneous tumor incidences of rats and mice are well documented (Attia 1996; Baldrick 2005; Eiben 2001).

Other factors may influence the variety and incidence rates of proliferative lesions within HCD, such as litter size and birth weight (Roe 1994). In addition, since most neoplasms increase in incidence with age, it is important to consider the age of the animal when comparing historical controls to concurrent controls and treatment groups. It may not be clear in reports or

publications of tumor incidence if tumors identified in decedents (unscheduled deaths) have been combined with those at scheduled termination. Most rodent studies are conducted for 18 to 24 months in mice and 24 months in rats. However, some individual investigators extend the study to full lifetime or "natural" death of the animal. These studies can last significantly longer than most standard bioassays and can be difficult to interpret relative to HCD. As an example, in one study, there were clear trends of higher incidence rates over time in pituitary tumors, adrenal pheochromocytomas, and mammary gland tumors, which also showed a shift towards malignancy from 2/31 malignant tumors after 24 months to 34/101 after 30 months in study durations (Bombard 1992; Bombard and Rinke 1994). A study investigating strain differences between Sprague-Dawley and Wistar rats showed that pancreatic islet cell carcinomas occurred late in the study (Germann et al. 1999). In 36 male Wistar rats, the first carcinoma was observed at 649 days of age (mean age 847 days) and on day 575 in females (mean age 781 days of 7 cases). In Sprague Dawley rats, which had more severe nephropathy and did not survive as long, the first male with carcinoma was 591 days old (mean age 709 days of 25 cases) and the first female was 682 days old (mean age 761 days). These data highlight the importance of considering age effects when comparing historical controls, concurrent controls, and treatment groups.

Housing (Caging, Identification Methods, Temperature, Humidity, Lighting)

The structural and social aspects of an environmental system can influence the physiology and behavior of animals occupying that system. A study examining the physiological effects of environmental enrichment (EE) with toys and nestlets on stress-responsive hormones of the hypothalamic-pituitary-adrenal axis under basal conditions and mild stress conditions showed that individually housed male and female rats with EE had significantly lower baseline adrenocorticotropic hormone and corticosterone concentrations compared to those housed without EE (Belz et al. 2003). Although valid tumor data from studies with and without EE are lacking, different housing methods may potentially influence the incidence rates of certain hormone-dependent or stress-related proliferative responses.

The type of caging and number of animals per cage can influence proliferative responses. A decrease in testicular interstitial cell tumors and increase in pituitary tumors in F344 rats was reported in studies where rats were housed individually compared to studies with group housing (Haseman et al. 1997; Nyska et al. 1998 2002). The survival rate for male B6C3F1 mice housed individually was significantly lower (66.53 vs. 823) when housed in wire mesh instead of polycarbonate cages (Rao and Crockett 2003). However, in this evaluation, mice kept in wire mesh cages were used in inhalation studies and compared to animals from feeding studies.

The unique identification of animals is critical for every toxicological study. The implantation of electronically readable transponder microchips is a reliable and frequently used identification method. However, induction of proliferative lesions associated with these implants have been reported as an additional variable impacting HCD (Rao and Edmondson 1990; Tillmann et al. 1997; Elcock et al. 2001; Le Calvez, Perron-Lepage, and Burnett 2006; Blanchard et al. 1999).

There are also physiological effects from different levels of noise, lighting, temperature, and humidity (http://www.nal.usda.gov/awic/pubs/Rodents/noise_light_temp.htm [accessed March 13, 2009]). Recent experiments have evaluated light cycle disruption mimicking jet-lag in mouse tumor models. Prior to inoculation with tumor cells, mice were synchronized with 12 hours of light and 12 hours of darkness or underwent repeat 8-hour advances of the light/dark cycle every 2 days to simulate jet-lag. The 24-hour rest/activity cycle was ablated, and the rhythms of body temperature, serum corticosterone, and clock protein expression were markedly altered in jet-lagged mice as compared with controls. Tumors grew faster in the jet-lagged animals, compared with controls, suggesting that altered environmental conditions can disrupt circadian clock molecular coordination in peripheral organs including tumors and could play a role in malignant progression (Filipski et al. 2004).

Route of Exposure

The route of test article administration can affect the tumor yield at a given site because it dictates pathways of internal distribution and metabolism and affects the concentration in the tissues (Amdur, Doull, and Klaassen 1991). A study evaluating two study designs showed that F344 rats had consistently improved survival in feed studies compared to inhalation studies (Haseman et al. 2003). Interestingly, tumors of endocrine active organs, especially of the pituitary gland, occurred far more frequently in inhalation studies than in feeding studies, while the inverse situation was noted for testicular tumors and leukemias (Haseman et al. 2003). These examples underscore the need to consider route of exposure as well as other parameters when evaluating tumor incidence rates.

Diet (Type and Feeding Practice)

One of the major factors affecting the rodent lifespan has been *ad libitum* feeding. The scope of this article is not to discuss all aspects of overfeeding or diet restriction but rather to address the impact of feeding practices on HCD.

Extensive research has consistently shown the longevity benefits of moderate caloric restriction or diet modification by reducing protein and/or increasing fiber compared to the *ad libitum* feeding of a standard nutritionally balanced rodent diet. Dietary modifications that contribute to a reduction in body weight gain and a consequent reduced incidence and severity of chronic degenerative conditions lead to improved survival (Keenan et al. 1999; Rao, Edmondson, and Elwell 1993; Masoro, Shimokawa, and Yu 1991). Improved survival allows for the detection of tumors that occur later in life. Some

debate has occurred regarding whether the restricted food consumption results in a decreased sensitivity to the development of tumors (Allaben et al. 1996); therefore, a distinction must be made between moderate caloric restriction and more severe caloric restriction, which has been shown to delay the onset of tumors in rodents (Kritchevsky 1993). In studies where rats were given an amount of nutritionally balanced feed that provided for initial growth and then maintenance over the life span of rodents, the incidence of spontaneous tumors observed at two years was in the same range as that from *ad libitum* fed rodents, but the trend was for tumors to be identified at the terminal necropsy rather than in early unscheduled necropsies (Keenan et al. 1999).

Because of the impact of diet on the development of tumors and degenerative disease, diet is a major factor in comparing tumor data, either within the same laboratory or between laboratories. Information on the type of diet (standard, reduced protein, or increased fiber) and the amount provided is important to a sound comparative review. Additionally, depending on the type of feeder used, or whether animals are individually or group housed, partial food restriction may occur unintentionally even in studies with an *ad libitum* feeding regimen.

Another important consideration is the potential for vehicle effects on study results. Of the many commonly used vehicles, only corn oil has been associated with increased body weights, an increased incidence of pancreatic acinar cell adenomas, and with a decreased incidence of mononuclear cell leukemia in male F344 rats (Haseman et al. 1985; Haseman and Rao 1992; NTP 1994). This observation underscores that any novel vehicle used in long-term studies needs to be evaluated for potential effects before their study data are used as HCD for comparative purposes.

The potential contamination of control animals with test article also needs to be considered. Trace levels of contamination that are below the lower limit of quantification are, in principle, considered to be nonrelevant (EMEA 2005).

Tissue Sampling and Trimming Procedures

Postmortem factors such as necropsy technique, accurate description of macroscopic observations using consistent nomenclature, trimming procedures, and correct labeling play an important role in establishing valid HCD. The importance of standardized necropsy and trimming techniques can be especially critical in tissues such as the mammary gland because different topographic areas contain structures that differ in their morphology, cell kinetic characteristics, hormone responsiveness, and carcinogenic potential (Russo et al. 1990). Perhaps less obvious is the importance of the direction in which the organ is trimmed either sagittally or transversely and of the number of sections and slides prepared and examined by the pathologist (Eustis et al. 1994). The results of other investigations have shown that it is possible to increase the number of proliferative lesions identified in the thyroid 55 tumors in serial sections of 140 thyroids compared to 9 tumors in single random sections of 177 thyroids (Thompson and Hunt

1963). To support harmonized and standardized organ processing, guides were published for organ sampling and trimming in rats and mice through an international collaboration between pathologists and histotechnicians from various European countries and the United States (Ruehl-Fehlert et al. 2003; Kittel et al. 2004; Morawietz et al. 2004; <http://www.item.fraunhofer.de/reni/trimming/index.php> [accessed March 13, 2009]).

Diagnostic Criteria and Terminology (Use of Standardized Nomenclature)

The histopathologic diagnosis of proliferative lesions at the end of a carcinogenicity study is the most important aspect of the study, with the principal endpoint being tumor incidence rates (Fitzgerald 1985). Pathology has been considered a subjective science, and diagnoses made by different pathologists may vary according to the pathologist's experience, education and training, and geographic location (Ward and Reznik 1983; Hardisty 1985). The familiarity a pathologist has with the typical spectrum of background lesions associated with a particular species or strain of laboratory rodent can also influence diagnoses (Goodman 1988). Consequently, there is much attention paid to the diagnostic criteria used in the histological interpretation of tissue changes. Standardization of diagnostic criteria is considered essential for the consistent and appropriate interpretation of a study (Ettlin and Prentice 2002; Greim et al. 2003; Haseman, Huff, and Boorman 1984). Differences in diagnostic criteria and their interpretation are considered to be one of the main causes of interlaboratory variability in the incidences of spontaneous tumors (Deschl et al. 2002; Gopinath 1994; Haseman 1990; van Zwieten et al. 1988) and perhaps the most important source of variability in tumor rates (Haseman 1992; Haseman et al. 1997; Roe 1994; Ward 1983). Changes in diagnostic criteria over time as a result of a greater scientific understanding of the processes involved in the development of a particular tumor can also play a part in interlaboratory variability (Greim et al. 2003; Poteracki and Walsh 1998; Wolf and Mann 2005). Reevaluation of lesions from earlier studies using criteria established at a later date has shown that the incidence of selected tumors can change considerably (Rao et al. 1990a, 1990b).

When considering published HCD, the use of different nomenclature for the same lesion can be misleading to those unaware that terminology may be synonymous. This can be a problem when different terminology for the same lesion is used in the same study (Hardisty 1985). There may be a need for different terminologies when describing a lesion in morphological descriptive terms or to address differences in topography (Goodman 1988). The use of synonymous terms within a study can impact the statistical interpretation of that study. Considering HCD, the use of different terminologies for the same change in different studies can be confusing and may lead to erroneous conclusions unless the synonyms are defined and understood (Haseman, Huff, and Boorman 1984; U.S. FDA CDER 2001; U.S. EPA 2005).

A complete record of the diagnostic criteria by which the data were compiled is critical for appropriate use of HCD. When there are multiple instances of the same tumor in a tissue or an animal, it should be clear if they were counted separately or combined. In addition, there are neoplasms where classification as hyperplastic or benign or malignant can be difficult and somewhat subjective. For example, the diagnosis of thyroid follicular tumors can be challenging and differentiation of hyperplasia from adenoma is difficult; often the key feature separating these two entities is compression (Society of Toxicologic Pathology, SSNDG Guides, <http://www.toxpath.org/ssndc/ThyroidParathyroidPro.pdf> [accessed March 13, 2009]). Other complexities include when tumors that are related ontogenically are combined according to the predominant differentiated cell type, though they may have variable differentiation, such as basal cell tumors in the epidermis, adnexa, or Zymbal's glands. In all of these cases, information on how these decisions were made within a particular study(ies) used as part of HCD would assist interpretation.

When utilizing published HCD, it is helpful to know which diagnostic criteria were used; however, this information is not typically available. A number of papers have been published presenting and evaluating HCD from studies carried out by the NTP (Haseman, Hailey, and Morris 1998; Haseman and Rao 1992; Brix et al. 2005). Consistent diagnostic criteria were applied in the interpretation of these findings. Some authors did publish the diagnostic criteria used in their studies but reported that the criteria changed for studies included in the published historical data (Baldrick and Reeve 2007; McMartin et al. 1992; Tennekes et al. 2004a, 2004b). Other authors have compiled HCD tables where the same pathologists were involved in all aspects of the evaluation and peer review process, to ensure consistency between the studies used in the tabulation of data (Brix et al. 2005). This information should be taken into consideration when published data are used for comparison with study data.

An important aim in the development of the RITA historical control database was the establishment of harmonized nomenclature and standard diagnostic criteria for proliferative lesions in rodents to ensure consistency between pathologists and laboratories (Mohr et al. 1990; Morawietz, Rittinghausen, and Mohr 1992). This effort resulted in the International Classification of Rodent Tumors published by IARC (1992-1997). As noted in the introduction, there is a current effort under way with INHAND to further increase global harmonization of diagnostic criteria and terminology.

Peer Review

As noted above, there are several areas in the histopathological evaluation of tissues that can account for differences in the reporting of tumor incidences. Many of these sources of variability can be countered by the use of appropriate quality assurance procedures, such as the peer-review process.

Histopathologic diagnoses are one of the most important sources of variability in tumor rates (Haseman 1993). A peer-

review procedure conducted at the laboratory of origin ensures that consistent criteria are used for the diagnoses of all tumors in studies conducted at that laboratory and increases the reliability of study data (Gopinath 1994; Greim et al. 2003; Hardisty 1985; Haseman 1990, 1992, 1993, 1995a, 1995b; van Zwieten et al. 1988; Ward et al. 1995). The STP has published recommendations specifically on the peer-review process (STP 1991). The peer-review process is described within the Best Practices Guidelines for Toxicologic Histopathology (Crissman et al. 2004).

A database for control animal pathology data, therefore, must have established peer-review procedures to ensure comparability of histopathology diagnoses for all studies entered into such a database (Haseman 1992). The peer-review process is a pivotal procedure completed prior to incorporation of data into the NTP and RITA historical control databases (Boorman et al. 2002; Deschl et al. 2002; Haseman et al. 1997; Morawietz, Rittinghausen, and Mohr 1992; Ward and Reznik 1983).

Innate Biological Variability

The intrinsic variation in the incidence of neoplasia can be considerable, even in the absence of confounding factors discussed above (Tarone, Chu, and Ward 1981). Dual control groups can be and have been used to assess intrinsic variability in groups of animals on study. For example, a review of dual control groups in CD-1 mice highlighted the wide variability in tumor incidence rates within a study (Baldrick and Reeve 2007). In one study, the lymphoma incidence was 3/60 in control group I and 11/59 in control group II. An evaluation of a large series of dual controls (18 studies) found 23 significant ($p < .05$) differences among the two control groups; importantly, this total number of significant differences was actually slightly less than the number of significant differences expected by chance alone, which was 24.4 (Haseman, Winbush, and O'Donnell 1986). One of the more striking examples was an increased incidence of prostate tumors observed (0/67 vs. 11/67; $p < .001$). Prostate tumors are uncommon, occurring in less than 13% of the male control rats in the other studies reviewed in this article. Thus, the increased incidence would have likely been regarded as a biologically significant response if it occurred in a test article-dosed group. Understanding the range of variability of control incidences under even the best of conditions is necessary to aid interpretation of the impact of other variables on the data within historical control databases.

WHEN AND HOW TO USE HISTORICAL DATA

HCD may be useful in the interpretation of rare tumors, marginally greater incidences and/or severity of proliferative changes in treated animals compared to controls, and unexpected increases or decreases of tumor incidences in study control animals. HCD can be used as a tool, including the possibility of statistical evaluation, to provide scientific perspective of disparate findings in dual concurrent control groups and review trends in tumor biology and behavior that may evolve over time in these rodent models (Haseman, Huff, and Boorman 1984). There are a

TABLE 2.-Key parameters included in the National Toxicology Program (NTP), Charles Rivers (CR), and Registry of Industrial Toxicology Animal Data (RITA) databases.

Data available	CR	NTP	RITA
Animal systems			
Species/strains	CR species/strains	Rats: Fischer 344/N, SD Mice: B6C3F1	Rats: multiple strains (e.g. Wistar, SD) Mice: multiple strains (e.g. B6C3F1, CD) p53 (few)
Genetically modified mouse models	N/A	Tg.Ac, p53 models	p53 (few)
Database content			
Details of study protocol (e.g., route of administration, vehicles)	Limited information provided	Full study protocols provided	Full study protocols provided
Data included	Group animal data only	Group and individual animal data	Group and individual animal data
Type of histopathological lesions	Neoplastic	Neoplastic	Neoplastic and preneoplastic
Peer review of histopathology results	No. Results presented as received	Yes. Three-tier independent review and expert panel	Yes. Three-tier independent review and expert panel
Database logistics			
Location of database	Printed monographs/web site	Web site	Server at Fraunhofer ITEM
Access requirements	None	None	Membership fee
Database searchable for historical control data (HCD)	N/A	Yes. Formatted tables are provided by species, sex, route of exposure, and type of vehicle and are searchable by a number of factors	Yes. Parameter searching across all studies by finding, strain, breeder, study duration, etc.

N/A = not applicable.

number of situations in which HCD may be used to assist in interpretation of study data. It is beyond the scope of this article to address all potential situations.

Occasionally, the concurrent controls may have an incidence that is at the high end of the HCD range, which may mask a treatment effect; or the concurrent controls are at the low end of a normal range for the laboratory, which could result in an overinterpretation of a possible treatment related effect. There is also the possibility that the study is flawed for design or technical reasons that may bias a study toward an apparent increase, or decrease, in a specific tumor incidence unrelated to a test article effect (Ettlin and Prentice 2002). For example, apparently simple differences in housing-group versus individual--can lead to significant differences in the occurrence of certain tumor types, such as testicular and pituitary tumors, in otherwise identical studies (Haseman, Hailey, and Morris 1998).

An important consideration impacting selection of HCD for comparison with study data is whether to combine neoplasms from different anatomic locations and similar histologic ontogeny but different histomorphology. Guidelines for combining neoplasms have shown that some neoplasms can be combined and, in some cases, preneoplastic lesions such as hyperplasia could be included as part of the weight of evidence in support of carcinogenicity (McConnell et al. 1986). Two examples are the transition from hyperplasia or dysplasia to malignancy for tumors of the nasal cavity and of the glandular stomach in rats. The guidelines also list neoplasms for which combining is inappropriate, such as combining malignant lymphomas and

histiocytic sarcomas, which are of different cellular lineage (McConnell et al. 1986). Appropriate combination of lesions such as benign and malignant hepatocellular tumors (similar cellular lineage) can provide evidence in determining a mode of action of a compound. In contrast, inappropriate combinations can result in overinterpretation of an effect or masking of a response when one is actually present (McConnell et al. 1986; Linkov et al. 2000).

WHERE TO OBTAIN HISTORICAL DATA

HCD from the laboratory that conducted the study under review will generally be more comparable than HCD collected from several laboratories, but the laboratory may not always have an adequate number of studies for compilation of HCD. HCD are widely available in organized databases and in published literature, apart from company-owned or CRO databases. Organized historical control databases have been, and continue to be, compiled by large institutions and organizations such as the NTP, Charles River (CR), and RITA. Each of these databases differs in the way the data are collected, handled and presented (Table 2).

The NTP was established in 1978 as a cooperative effort to coordinate toxicology testing programs within the federal government. Other goals were to strengthen the science base in toxicology; develop and validate improved testing methods; and provide information about potential adverse human health effects of chemicals to health, regulatory, and research agencies, the scientific and medical communities, and the public. As a way to

follow changes in the biology of the test species and to evaluate test results, a database of HCD of neoplastic lesions from untreated or control groups was established and is available electronically through their Web site at <http://ntp.niehs.nih.gov/?objectid=92E61F1B-F1F6-975E-7D3BED551F07DCOA> (accessed March 17, 2009). To ensure that current HCD are presented, the NTP database is maintained as a five-year window of the most recent NTP data and updated annually. Costs to maintain this database are provided for in the conduct of the studies. To date, the NTP has published technical reports from more than 500 two-year, two species, toxicology and carcinogenicity bioassays in F344/N rats, Sprague Dawley rats, and B6C3F1 mice; and all corresponding raw data are provided at <http://ntp.niehs.nih.gov/ntpweb> (accessed March 17, 2009). The rigorous pathology peer-review process includes three independent pathology reviews and a final pathology working group of nine to eleven pathologists who meet to review and decide on diagnostic or terminology discrepancies. The publicly available historical control database includes tumor incidences and growth and survival curves. These data are summarized by species, strain, sex, route of administration, and vehicle.

Over the past several decades, at their own cost, CR has compiled HCD on CR-produced rodent strains used in chronic studies conducted in various laboratories in the United States and Europe and published them as strain-specific monographs. Some are limited to specific organs (e.g., ophthalmic) and specific parameters (e.g., caloric restriction). Although there has been some attempt to standardize the diagnostic nomenclature, the data are, for the most part, presented as received from the laboratories where the studies were performed. Each publication specifies common study parameters such as time frame, rodent strain production site(s), diet versus gavage, untreated versus vehicle controls, and various husbandry/environmental factors dependent on sponsor disclosure. Each publication's focus is on a specific CR rodent and contains data from multiple studies composed of thousands of control animals. These published compilations can be obtained from CR (<http://www.criver.com/en-US/ProdServ/ByType/ResModOver/Pages/On-lineLiterature.aspx> [accessed March 17, 2009]).

RITA is a proprietary pathology database for historical control data founded in 1988 in Hannover, Germany, as a cooperative venture between the Fraunhofer Institute of Toxicology and Experimental Medicine (Fraunhofer ITEM) and thirteen pharmaceutical and chemical companies from Germany and Switzerland (Morawietz, Rittinghausen, and Mohr 1992). The objective of this coalition was to establish a centralized European database providing standardized and valid historical background data in specific rodents to be used for carcinogenic risk assessment (Deschl et al. 2002). From the very beginning, the development of standardized nomenclature and diagnostic criteria was considered as key for success (Mohr et al. 1990) and resulted in the previously mentioned publication of ten IARC/WHO fascicles, *International Classification of Rodent Tumours*, part 1, *The Rat* (1992-1997), each dealing with an organ system, and later the *International Classification of Rodent Tumors: The Mouse*, which was edited as a book (Mohr

2001). The RITA data collection has been ongoing since 1988 and contains data on animals from more than two hundred carcinogenicity studies of the major rat and mouse strains used in Europe including the Sprague Dawley and Wistar Han rats and the CD1 and B6C3F1 mice. The animals are from a variety of breeders and vendors. Detailed information is available for each individual animal and includes environmental factors such as housing conditions, feeding, group size, and others as described previously in this article. Companies throughout Europe and North America are currently participating in the RITA Group effort via a membership fee. The RITA project adheres to a rigorous peer-review process in which every preneoplastic and neoplastic lesion entered into the database is confirmed by an actual examination of the respective tissue section by an experienced independent pathologist, with all questionable findings submitted to a panel of experienced pathologists to establish a final diagnosis. By using systemized trimming procedures, nomenclature, and diagnostic criteria, the group adheres to standardized data acquisition and data validation procedures. Photomicrographs representing typical and equivocal histopathologic diagnoses are available for group members and to a large extent to users of the password-protected Web program "goRENI," which is accessible to members of any Toxicologic Pathology Society on request (http://www.goreni.org/back_inhand.php [accessed March 13, 2009]). Findings are searchable in the database by various criteria such as strain, breeder, time period, and study duration.

There are many published reports of proliferative changes in rodents, and these often present results from individual toxicology studies conducted by investigators in industry or academia and represent a variety of rodent strains, study designs, and feeding and husbandry practices. Individually, the reports do not provide the comprehensive incidence of proliferative lesions that are found in larger databases such as those discussed above; however, they do provide useful information regarding tumor occurrences. These reports often summarize current literature on specific tumor types and/or incidence rates in specific organ systems and may include consideration of variables in study design, feeding practices, and strain among others. Published literature may serve as a useful supplement to the larger organized databases when searching for HCD. An organized listing of publications is available on the STP Web site, <http://www.toxpath.org/positions.asp> (accessed April 10, 2009). This compilation provides a listing of references in regard to strain, tumor types, and other factors.

PRESENTATION OF HCD

When HCD are used, descriptions of the statistical analyses and all relevant data (including adjustments for survival) should be a component of study reports and available for review. A few examples of data presentation are provided on the STP Web site. As described above in the HCD survey results, HCD are commonly presented as a mean, standard deviation, and range of tumor incidence from a historical control database or published literature from studies with same

route and all exposure routes combined. This method provides a representation of reported observations of a given tumor incidence. When using this approach, it should be noted that the range can be influenced by extreme outliers even from a single study. For example, consider mammary gland carcinoma in F344/N rats. In the NTP's database for the twenty studies conducted during the period January 6, 1997, and August 6, 2001 (based on NTP 2000 feed), there were 64 rats out of 1,050 diagnosed with mammary gland carcinoma. These studies had a mean of 6.03, a standard deviation of 4.393, and a range of 03 to 203 for all routes and all vehicles. However, the 203 incidence (10 out of 50 rats) was found in only one inhalation study in a total of twenty studies. The next largest incidence was 103. Without the control from this single study, the range would have been 03 to 103. To address this concern, incidences per study may also be reported as part of the HCD dataset. Reporting of incidences per study will allow both presentation and use of a broad range of observations and provide transparency of potential influences of outlier populations.

Since the early 1980s, a number of attempts have been made to develop statistical procedures for analyzing concurrent experimental data by formally making use of the HCD (Tarone 1982; Dempster, Selwyn, and Weeks 1983; Hoel 1983; Hoel and Yanagawa 1986; Tamura and Young 1986, 1987; Prentice et al. 1992; Ibrahim and Ryan 1996; Ibrahim, Rynn, and Chen 1998; Dunson and Dinse 2001). Each of these methods has strengths and limitations. There may, however, be value, in some situations, in analyzing the concurrent data by applying additional informal or formal statistical methods to historical controls and evaluating HCD in the context of a specific data set or study (U.S. FDA CDER 2001; Elmore and Peddada 2009). HCD may aid in the interpretation of data for the assessment of a xenobiotic-induced proliferative change, but its use and application should be presented in the context of sound biological principles with regard to the pooling of findings, combining ontogenetically similar tissues, and related criteria.

HCD AND WEIGHT OF EVIDENCE APPROACH

While the concurrent control group provides the most relevant control data, one should consider HCD as one of many sources of information that add to the "weight of evidence" approach when assessing the potential carcinogenic effect of a compound. HCD can be used as a tool to assess the spontaneous tumor rates in the concurrent control group and to evaluate disparate findings in dual concurrent control groups (Haseman 1984; Haseman, Huff, and Boorman 1984). The HCD can help to determine whether marginally significant trends in common tumors are likely real or false positives. If the tumor rates in the treated groups are within the range of reliable HCD, then a marginally significant trend for a common tumor could be discounted due to a random occurrence of a low concurrent control rate. However, the incidences of other lesions of similar cell lineage (hypertrophy, hyperplasia, papilloma, adenoma, etc.) may also be considered in this weight of evidence approach. Such a weight of evidence approach is

particularly helpful when the proliferative lesions are considered to be on a biological continuum as in the forestomach of the mouse where the lesions progress from focal hyperplasia to papilloma to squamous cell carcinoma (Leininger and Jokinen 1994; Leininger et al. 1999). In data on a pesticide submitted to the U.S. EPA, the incidence of thyroid follicular cell tumors was increased in treated female rats (0/60, 0/60, 2/60, 2/60, and 4/60 at 0, 200, 1,000, 4,000, and 20,000 ppm, respectively). The increased incidence of thyroid adenomas was statistically significant in the female 20,000 ppm group by the Cochran-Armitage trend test but not by the Fisher's pair-wise exact test. There was a lack of hyperplasia or other preneoplastic morphologic changes and no progression of the thyroid follicular adenomas to carcinomas at the high dose. There was a lack of evidence for the typical progression of thyroid follicular adenomas, which typically includes perturbation of the thyroid pituitary axis and increased follicular cell hyperplasia. The effect on the thyroid was in female rats, but there was no associated thyroid follicular effect in male rats, which tend to be the more sensitive sex for follicular cell hypertrophy and neoplasia. In addition, an evaluation of the HCD from the source of the study animals showed that the spontaneous incidence of thyroid follicular adenomas in the source population was 1.13 to 6.13. The incidence of thyroid follicular adenomas from the testing facility was 03 to 33. The incidence of thyroid follicular adenomas in the high-dose treated female rats (6.73) was near the upper end of the source historical control range and greater than the laboratory historical range. Based on the weight of the biological evidence, the slight increase in incidence of thyroid follicular cell adenomas was interpreted to be unrelated to treatment (Douglas Wolf, March 11, 2008, Environmental Protection Agency, Research Triangle Park, NC).

One could also consider if there is histological evidence that any of the malignant lesions (e.g., carcinoma) arise within the benign counterparts (e.g., adenoma) and if there is multiplicity in site-specific tumors. Other issues to consider include body weight, survival, plasma concentration of test compound, time of tumor onset, if the neoplastic lesion occurs in both males and females (although there may be differences due to sex steroids), if it occurs in both rodent species, if there is a positive dose-related response, or if there are bilateral lesions in paired organs. Combining of benign and malignant neoplasms of the same histogenesis in the same or different organs for statistical analyses (i.e., hepatocellular adenomas and carcinomas; cardiovascular system, vascular endothelium, hemangiomas, hemangiosarcomas) may also add to the weight of evidence.

SUMMARY

In conclusion, when evaluating proliferative lesions from nonclinical rodent carcinogenicity studies, the concurrent control group is the most relevant. However, when the biological significance of a change in incidence of proliferative lesions in compound-treated groups relative to concurrent controls is uncertain, historical control data can aid in the overall

evaluation. When using HCD, several issues should be taken into consideration such as source and quality of HCD, in-life and postmortem factors associated with the origination of the HCD, and type of statistical tools used to present the HCD, which reflect the consensus principles summarized by the HCD Working Group.

ACKNOWLEDGMENTS

The Working Group would like to thank Drs. C. Gopinath, S. Rittinghausen, C. Clifford, T. Peters, and J. Haseman and members of the various STPs for their review and critical comments.

REFERENCES

- Allaben, W. T., Turturro, A., Leakey, J. E., Seng, J. E., and Hart, R. W. (1996). FDA points-to-consider documents: The need for dietary control for the reduction of experimental variability within animal assays and the use of dietary restriction to achieve dietary control. *Toxicol Pathol* 24, 77-81.
- Amdur, M. O., Doull, J., and Klaassen, C. D., eds. (1991). *Casareu and Doull's Toxicology: The Basic Science of Poisons* 4th ed., p. 165. Pergamon Press, New York.
- 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, 461-73.
- Baldrick, P. (2005). Carcinogenicity evaluation: Comparison of tumor data from dual control groups in the Sprague-Dawley rat. *Toxicol Pathol* 33, 283-91.
- Baldrick, P., and Reeve, L. (2007). Carcinogenicity evaluation: Comparison of tumor data from dual control groups in the CD-1 mouse. *Toxicol Pathol* 35, 562-69.
- Belz, E. E., Kennell, J. S., Czambel, R. K., Rubin, R. T., and Rhodes, M. E. (2003). Environmental enrichment lowers stress-responsive hormones in singly housed male and female rats. *Pharmacol Biochem Behav* 76, 481-86.
- Blanchard, K. T., Barthel, C., French, J. E., Holden, H. E., Moretz, R., Pack, F. D., Tennant, R. W., and Stoll, R. E. (1999). Transponder-induced sarcoma in the heterozygous p53^{+/−} mouse. *Toxicol Pathol* 27, 519-27.
- Bombard, E. (1992). Frequency of spontaneous tumors in Wistar rats in 30-months studies. *Exp Toxicol Pathol* 44, 381-92.
- Bombard, E., and Rinke, M. (1994). Frequency of spontaneous tumors in Wistar rats in 2-year studies. *Exp Toxicol Pathol* 46, 17-29.
- Boorman, G. A., Haseman, J. K., Waters, M. D., Hardisty, J. F., and Sills, R. C. (2002). Quality review procedures necessary for rodent pathology databases and toxicogenomic studies: The National Toxicology Program experience. *Toxicol Pathol* 30, 88-92.
- Brix, A. E., Nyska, A., Haseman, J. K., Sells, D. M., Jokinen, M. P., and Walker, N. J. (2005). Incidences of selected lesions in control female Harlan Sprague-Dawley rats from two-year studies performed by the National Toxicology Program. *Toxicol Pathol* 33, 477-83.
- Crissman, J. W., Goodman, D. G., Hildebrandt, P. K., Maronpot, R. R., Prater, D. A., Riley, J. H., Seaman, W. J., and Thake, D. C. (2004). Best practices guideline: Toxicologic histopathology. *Toxicol Pathol* 32, 126-31.
- Dempster, A. P., Selwyn, M. R., and Weeks, B. J. (1983). Combining historical and randomized controls for assessing trends in proportions. *J Amer Stat Assoc* 78, 221-27.
- Desch, U., Kittel, B., Rittinghausen, S., Morawietz, G., Kohler, M., Mohr, U., and Keenan, C. (2002). The value of historical control data—scientific advantages for pathologists, industry and agencies. *Toxicol Pathol* 30, 80-87.
- Dunson, D. B., and Dinse, G. E. (2001). Bayesian incidence analysis of animal tumorigenicity data. *Appl Stat* 50, 125-41.
- Eiben, R. (2001). Frequency and time trends of spontaneous tumors found in B6C3F1 mice oncogenicity studies over 10 years. *Exp Toxicol Pathol* 53, 399-408.
- Eiben, R., and Bombard, E. M. (1999). Trends in mortality, body weights and tumor incidences of Wistar rats over 20 years. *Exp Toxicol Pathol* 51, 523-36.
- Elcock, L. E., Stuart, B. P., Wable, B. S., Hoss, H. E., Crabb, K., Millard, D. M., Mueller, R. E., Hastings, T. F., and Lake, S. G. (2001). Tumors in long-term rat studies associated with microchip animal identification devices. *Exp Toxicol Pathol* 52, 483-91.
- Elmore, S. A., and Peddada, S. (2009). Points to consider on the statistical analysis of rodent cancer bioassay data when incorporating historical control data. *Toxicol Pathol* 37 (5), 672-76.
- Ettlin, R. A., and Prentice, D. E. (2002). Unexpected tumour findings in lifetime rodent bioassay studies—What to do? *Toxicol Pathol* 30, 1733.
- European Economic Community (EEC). (1991). COUNCIL DIRECTIVE of 15 July 1991 concerning the placing of plant protection products on the market (91/414/EEC). http://europa.eu/eur-lex/en/consleg/pdf/1991/en_1991L0414_do_001.pdf (accessed March 13, 2009).
- European Medicines Agency (EMA). (2002). CPMP, Note for guidance on carcinogenic potential, EMA, CPMP/SWP/2877/00, London, 25 July 2002. <http://www.emea.europa.eu/pdfs/human/swp/287700en.pdf> (accessed March 13, 2009).
- European Medicines Agency (EMA). (2004). CHMP SWP conclusions and recommendations on the use of genetically modified animal models for carcinogenicity assessment, EMA, CPMP/SWP/2592102 Rev I, London, 23 June 2004. <http://www.emea.europa.eu/pdfs/human/swp/259202en.pdf> (accessed March 13, 2009).
- European Medicines Agency (EMA). (2005). Guideline on the evaluation of control samples in nonclinical safety studies: Checking for contamination with the test substance, London, 17 March 2005, CPMP/SWP/1094/04. <http://www.emea.europa.eu/pdfs/human/swp/109404en.pdf> (accessed April 8, 2009).
- Eustis, S. L., Hailey, J. R., Boorman, G. A., and Haseman, J. K. (1994). The utility of multiple-section sampling in the histopathological evaluation of the kidney for carcinogenicity studies. *Toxicol Pathol* 22, 457-72.
- Filipski, E., Delaunay, F., King, V. M., Wu, M., Claustrat, B., Gnichez-Cassiau, A., Guettier, C., Hastings, M. H., and Francis, L. (2004). Effects of chronic jet lag on tumor progression in mice. *Cancer Res* 64, 7879-85.
- Fitzgerald, J. E. (1985). Editorial: Dilemma in diagnostic tumor pathology. *Toxicol Pathol* 13, 81.
- Germann, P. G., Heinrichs, M., Piischner, H., Hellmann, J., Kohler, M., and Ernst, H. (1999). RITA/Registry of Industrial Toxicology Animal data: A comparative immunohistochemical study of 77 islet cell carcinomas in Sprague-Dawley and Wistar rats using antibodies against insulin, glucagon, somatostatin and gastrin. *Exp Toxicol Pathol* 51, 477-87.
- Goodman, D. G. (1988). Factors affecting histopathologic interpretation of toxicity-carcinogenicity studies. In: *LSI Monographs: Carcinogenicity—The Design, Analysis and Interpretation of Long-Term Animal Studies* (H. Grice and J. Ciminera, eds.), pp. 109-18. Springer-Verlag, New York.
- Gopinath, C. (1994). Spontaneous tumour rates: Their use to support rodent bioassays. *Toxicol Pathol* 22, 160-64.
- Greim, H., Gelbke, H. P., Reuter, U., Thielmann, H. W., and Edler, L. (2003). Evaluation of historical control data in carcinogenicity studies. *Hum Exp Toxicol* 22, 541-49.
- Hardisty, J. F. (1985). Factors influencing laboratory animal spontaneous tumor profiles. *Toxicol Pathol* 13, 95-104.
- Haseman, J. K. (1984). Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. *Environ Health Perspect* 58, 385-92.
- Haseman, J. K. (1990). Controlling sources of variability and use of statistical decision rules in rodent carcinogenicity studies. In: *Proceedings of the Biopharmaceutical Section of the American Statistical Association Meeting, August 6-9, 1990, Anaheim, CA, 1990*, pp. 260-9.
- Haseman, J. K. (1992). Value of historical controls in the interpretation of rodent tumor data. *Drug Inf J* 26, 191-200.

- Haseman, J. K. (1993). The value and limitations of a large source of control animal pathology data. In: *Computerized Control Animal Pathology Databases: Will They Be Used?* (J. A. N. McAuslane, C. Parkinson, and C. E. Lumley, eds.). Carshalton, Centre for Medicines Research, Carshalton, England.
- Haseman, J. K. (1995a). Data analysis: Statistical analysis and use of historical control data. *Regul Toxicol Pharmacol* 21, 52-59.
- Haseman, J. K. (1995b). Data analysis: Statistical analysis and use of historical control data. *Regul Toxicol Pharmacol* 21, 81-86.
- Haseman, J. K., Hailey, J. R., and Morris, R. W. (1998). Spontaneous neoplasm incidences in Fischer 344 rats and B6C3F1 mice in two-year carcinogenicity studies: A National Toxicology Program update. *Toxicol Pathol* 26, 428-41.
- Haseman, J. K., Huff, J. E., and Boorman, G. A. (1984). Use of historical control data in carcinogenicity studies in rodents. *Toxicol Pathol* 12, 126-35.
- Haseman, J. K., Huff, J. E., Rao, G. N., Arnold, J. E., Boorman, G. A., and McConnell, E. E. (1985). Neoplasms observed in untreated and com oil gavage control groups of F344/N rats and (C57BU6N 3/C3WHEN)F1 (B6C3F1) mice. *J Natl Cancer Inst* 75, 975-84.
- Haseman, J. K., Huff, J. E., Rao, G. N., and Eustis, S. L. (1989). Sources of variability in rodent carcinogenicity studies. *Fundam Appl Toxicol* 12, 793-804.
- 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., and Rao, G. N. (1992). Effects of com oil, time-related changes, and inter-laboratory variability on tumor occurrence in control Fischer 344 (F344/N) rats. *Toxicol Pathol* 20, 52-60.
- 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. *Fundam Appl Toxicol* 1, 573-84.
- Haseman, J. K., Young, E., Eustis, S. L., and Hailey, J. R. (1997). Body weight-tumor incidence correlations in long-term rodent carcinogenicity studies. *Toxicol Pathol* 25, 256-63.
- Hoel, D. G. (1983). Conditional two sample tests with historical controls. In: *Contributions to Statistics* (P. K. Sen, ed) North-Holland Publishing Company, Amsterdam, the Netherlands.
- Hoel, D. G., and Yanagawa, T. (1986). Incorporating historical controls in testing for a trend in proportions. *J Amer Stat Assoc* 81, 1095-99.
- Ibrahim, J. G., and Ryan, L. M. (1996). Use of historical controls in time-adjusted trend tests for carcinogenicity. *Biometrics* 52, 1478-85.
- Ibrahim, J. G., Ryan, L. M., and Chen, M. (1998). Using historical controls to adjust for covariates in trend tests for binary data. *J Amer Stat Assoc* 93, 1282-93.
- International Conference on Harmonization. (n.d). <http://www.ich.org/cache/compo/502-272-1.html> (accessed March 17, 2009).
- Keenan, K. P., Ballam, G. C., Soper, K. A., Laroque, P., Coleman, J. B., and Dixit, R. (1999). Diet, caloric restriction, and the rodent bioassay. *Toxicol Sci* 52 (Suppl.), 24-34.
- Keenan, C., Hughes-Earle, A., Case, M., Stuart, B., Lake, S., Mahrt, C. R., Halliwell, W., Westhouse, R., Elwell, M. R., Morton, D., Morawietz, G., Rittinghausen, S., Deschl, U., and Mohr, U. (2002). The North American control animal database: A resource based on standardized nomenclature and diagnostic criteria. *Toxicol Pathol* 30, 75-79.
- Kittel, B., Ruehl-Fehlert, C., Morawietz, G., Klapwijk, J., Elwell, M. R., Lenz, B., Gerard, M., O'Sullivan, M. G., Roth, D. R., and Wadsworth, P. F. (2004). Revised guides for organ sampling and trimming in rats and mice-Part 2-A joint publication of the RITA and NACAD groups. *Exp Toxicol Pathol* 55, 413-31.
- Kritchevsky, D. (1993). Energy restriction and carcinogenesis. *Food Res Internat* 26, 289-95.
- Le Calvez, S., Perron-Lepage, M. F., and Burnett, R. (2006). Subcutaneous microchip-associated tumours in B6C3F1 mice: A retrospective study to attempt to determine their histogenesis. *Exp Toxicol Pathol* 57, 255-65.
- Leininger, J. R., and Jokinen, M. P. (1994). Tumours of the oral cavity, pharynx, oesophagus and stomach. In: *Pathology of Tumours in Laboratory Animals. Volume ff-Tumours of the Mouse*, 2nd ed. (V. Turusov and U. Mohr, eds.), pp. 167-77. International Agency for Research on Cancer Scientific Publications No. 111, Lyon, France.
- Leininger, J. R., Jokinen, M. P., Dangler, C. A., and Whiteley, L. O. (1999). Oral cavity, esophagus, and stomach. In: *Pathology of the Mouse* (Robert R. Maronpot, ed.), pp. 38-42. Cache River Press, Vienna, IL.
- Linkov, I., Polkanov, M., Shagiakhrnetov, A., Wilson, R., and Gray, G. (2000). Correlations among tumor types in mouse cancer bioassays: Liver adenomas, liver carcinomas, leukemias and lymphomas. *Toxicol Ind Health* 16, 16-40.
- Masoro, E. J., Shimokawa, I., and Yu, B. P. (1991). Retardation of the aging processes in rats by food restriction. *Ann New York Acad Sci* 621, 337-52.
- McConnell, E. E., Solleveld, H. A., Swenberg, J. A., and Boorman, G. A. (1986). Guidelines for combining neoplasms for evaluation of rodent carcinogenesis studies. *J Natl Cancer Inst* 76, 283-89.
- 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, 212-25.
- Mohr, U., ed. (1992-1997). *International Classification of Rodent Tumours, Part I, The Rat*. IARC Scientific Publications, Lyon, France. Vol. 1 (1992): *Respiratory System*; Vol. 2 (1992): *Soft Tissue and Musculoskeletal System*; Vol. 3 (1992): *Urinary System*; Vol. 4 (1993): *Haematopoietic System*; Vol. 5 (1993): *Integumentary System*; Vol. 6 (1994): *Endocrine System*; Vol. 7 (1994): *Central Nervous System; Hear Eye, Mesothelium*; Vol. 8 (1997): *Male Genital System*; Vol. 9 (1997): *Female Genital System*; Vol. 10 (1997): *Digestive System*.
- Mohr, U., ed. (2001). *International Classification of Rodent Tumours. The Mouse*. Springer-Verlag, Berlin.
- Mohr, U., Bader, R., Ernst, H., Eitlin, R., Gemhardt, C., Harleman, J. H., Hartig, F., Jahn, W., Kaliner, G., Karbe, E., et al. (1990). Tumor Registry Data Base. Suggestions for a systematized nomenclature of pre-neoplastic and neoplastic lesions in rats. *Exp Pathol* 38, 1-18.
- Morawietz, G., Rittinghausen, S., and Mohr, U. (1992). RITA-Registry of Industrial Toxicology Animal-data-Progress of the working group. *Exp Toxicol Pathol* 44, 301-9.
- Morawietz, G., Ruehl-Fehlert, C., Kittel, B., Bube, A., Keane, K., Halm, S., Heuser, A., and Hellmann, J. (2004). Revised guides for organ sampling and trimming in rats and mice Part 3A joint publication of the RITA and NACAD groups. *Exp Toxicol Pathol* 55, 433-49.
- Morton, D., Kemp, R. K., Francke-Carroll, S., Jensen, K., McCartney, J., Monticello, T. M., Perry, R., Pulido, O., Roome, N., Schafer, K., Sellers, R., and Snyder, P. W. (2006). Society of Toxicologic Pathology position paper. Best practices for reporting pathology interpretations within GLP toxicology studies. *Toxicol Pathol* 34, 806-9.
- National Toxicology Program. (1994). NTP comparative toxicology studies of com oil, safflower oil, and tricaprilyn (CAS Nos. 8001-30-7, 8001-23-8, and 538-23-8) in male F344/N rats as vehicles for gavage. *Natl Toxicol Program Tech Rep Ser* 426, 1-314.
- Nyska, A., Hester, S. D., Cooper, R. L., Goldman, J. M., Stoker, T. E., House, D., and Wolf, D. C. (2002). Single or group housing altered hormonal physiology and affected pituitary and interstitial cell kinetics. *J Toxicol Sci* 27, 449-57.
- Nyska, A., Leininger, J., Maronpot, R., Haseman, J., and Hailey, J. (1998). Effect of individual versus group caging on the incidence of pituitary and Leydig cell tumours in F344 rats: Proposed mechanism. *Med Hypotheses* 50, 525-29.
- Organization for Economic Cooperation and Development (OECD). (1981a). OECD Guidelines for the Testing of Chemicals. Test No. 451: Carcinogenicity Studies. <http://browse.oecdbookshop.org/oecd/pdfs/browseit/9745101E.PDF> (accessed March 17, 2009).
- Organization for Economic Cooperation and Development (OECD). (1981b). OECD Guidelines for the Testing of Chemicals. Test No. 452: Chronic Toxicity Studies. <http://browse.oecdbookshop.org/oecd/pdfs/browseit/9745201E.PDF> (accessed March 17, 2009).
- Organization for Economic Cooperation and Development (OECD). (1981c). OECD Guidelines for the Testing of Chemicals. Test No. 453: Combined Chronic Toxicity/Carcinogenicity Studies. <http://browse.oecdbookshop.org/oecd/pdfs/browseit/9745301E.PDF> (accessed March 17, 2009).

- Organization for Economic Cooperation and Development (OECD). (2002). OECD Environment, Health and Safety Publications. Series on Testing and Assessment No. 35 and Series on Pesticides No. 14. Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies. *ENVIRONMENTAL MONITORING* 19. [http://www.olis.oecd.org/olis/2002doc.nsf/LinkTo/env-jm-mono\(2002\)19](http://www.olis.oecd.org/olis/2002doc.nsf/LinkTo/env-jm-mono(2002)19) (accessed March 17, 2009).
- Poteracki, J., and Walsh, K.M. (1998). Spontaneous neoplasms in control Wistar rats: A comparison of reviews. *Toxicol Sci* 45, 1-8.
- Prentice, R. L., Smythe, R. T., Krewski, D., and Mason, M. (1992). On the use of historical control data to estimate dose response trends in quantitative bioassay. *Biometrics* 48, 459-78.
- Rao, G. N., and Crockett, P.W. (2003). Effect of diet and housing protocols on growth, body weight, survival and tumor incidences of B6C3F1 mice in chronic studies. *Toxicol Pathol* 31, 243-50.
- Rao, G. N., and Edmondson, J. (1990). Tissue reaction to an implantable identification device in mice. *Toxicol Pathol* 18, 412-16.
- Rao, G. N., Edmondson, J., and Elwell, M. R. (1993). Influence of dietary protein concentration on severity of nephropathy in Fischer-344 (F-344/N) rats. *Toxicol Pathol* 21, 353-61.
- Rao, G. N., Haseman, J. K., Grumbein, S., Crawford, D. D., and Eustis, S. L. (1990a). Growth, body weight, survival, and tumor trends in (C57BU6 X C3H/HeN)F1 (B6C3F1) mice during a nine-year period. *Toxicol Pathol* 18, 71-77.
- Rao, G. N., Haseman, J. K., Grumbein, S., Crawford, D. D., and Eustis, S. L. (1990b). Growth, body weight, survival, and tumor trends in F344/N rats during an eleven year period. *Toxicol Pathol* 18, 61-70.
- Roe, F. J. (1994). Historical histopathological control data for laboratory rodents: Valuable treasure or worthless trash? *Lab Anim* 28, 148-54.
- Ruehl-Fehlert, C., Kittel, B., Morawietz, G., Deslex, P., Keenan, C., Mahrt, C. R., Nolte, T., Robinson, M., Stuart, B.P., and Desch, U. (2003). Revised guides for organ sampling and trimming in rats and mice-Part 1-A joint publication of the RITA and NACAD groups. *Exp Toxicol Pathol* 55, 91-106.
- Russo, J., Gusterson, B.A., Rogers, A. E., Russo, I.H., Wellings, S.R., and van Zwieten, M. J. (1990). Comparative study of human and rat mammary tumorigenesis. *Lab Invest* 62, 1-32.
- Society of Toxicologic Pathologists. (1991). Peer review in toxicologic pathology: Some recommendations. *Toxicol Pathol* 19, 290-92.
- Society of Toxicologic Pathology. (n.d.). SSNDG Guides. <http://www.tox-path.org/ssndc.asp> (accessed March 13, 2009).
- Tamura, R. N., and Young, S.S. (1986). The incorporation of historical information in tests of proportions: Simulation study of Tarone's procedure. *Biometrics* 42, 343-49.
- Tamura, R. N., and Young, S.S. (1987). A stabilized moment estimator for the beta-binomial distribution. *Biometrics* 43, 813-24.
- Tarone, R. E. (1982). The use of historical control information in testing for a trend in proportions. *Biometrics* 38, 215-20.
- Tarone, R. E., Chu, K.C., and Ward, J. M. (1981). Variability in the rates of some common naturally occurring tumors in Fischer 344 rats and (C57BU6N x C3H/HeN)F1 (B6C3F1) mice. *J Natl Cancer Inst* 66, 1175-81.
- Tennekes, H., Gembardt, C., Dammann, M., and van Ravenzwaay, B. (2004a). The stability of historical control data for common neoplasms in laboratory rats: Adrenal gland (medulla), mammary gland, liver, endocrine pancreas, and pituitary gland. *Regul Toxicol Pharmacol* 40, 18-27.
- Tennekes, H., Kaufmann, W., Dammann, M., and van Ravenzwaay, B. (2004b). The stability of historical control data for common neoplasms in laboratory rats and the implications for carcinogenic risk assessment. *Regul Toxicol Pharmacol* 40, 293-304.
- Thompson, S. W., and Hunt, R. D. (1963). Spontaneous tumors in the Sprague-Dawley rat: Incidence rates of some types of neoplasms as determined by serial section versus single section techniques. *Ann New York Acad Sci* 108, 832-45.
- Tillmann, T., Kamino, K., Dasenbrock, C., Ernst, H., Kohler, M., Morawietz, G., Campo, E., Cardesa, A., Tomatis, L., and Mohr, U. (1997). Subcutaneous soft tissue tumours at the site of implanted microchips in mice. *Exp Toxicol Pathol* 49, 197-200.
- U.S. Environmental Protection Agency (EPA). (2005). Guidelines for Carcinogen Risk Assessment. EPA/630/P-03/001B. <http://www.epa.gov/iris/backgr-d.htm> (accessed March 13, 2009).
- U.S. Food and Drug Administration (FDA). (2000). Toxicological Principles for the Safety Assessment of Food Ingredients Redbook 2000, July 2000 Section IV.B.3. Pathology Considerations in Toxicity Studies. <http://www.cfsan.fda.gov/~redbook/red-ivb3.html> (accessed March 13, 2009).
- U.S. Food and Drug Administration (FDA). (2006). Toxicological Principles for the Safety Assessment of Food Ingredients Redbook 2000, January 2006. Section IV.C.6 Carcinogenicity Studies with Rodents. <http://www.cfsan.fda.gov/~redbook/red-ivc6.html> (accessed March 13, 2009).
- U.S. Food and Drug Administration (FDA). (2007). Toxicological Principles for the Safety Assessment of Food Ingredients Redbook 2000, July 2007. Section IV.C.5a. Chronic Toxicity Studies with Rodents. <http://www.cfsan.fda.gov/~redbook/red-ivc5a.html> (accessed March 13, 2009).
- U.S. Food and Drug Administration (FDA) CDER. (2001). Guidance for Industry: Statistical Aspects of the Design, Analysis and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals. (Section C). <http://www.fda.gov/cder/guidance/815dft.htm> (accessed March 13, 2009).
- van Zwieten, M. J., Majka, J. A., Peter, C. P., and Burek, J. D. (1988). The value of historical control data. In: *LSI Monographs: Carcinogenicity-The Design, Analysis and Interpretation of Long-Term Animal Studies* (H. C. Grice and J. L. Ciminera, eds.), pp. 39-51. Springer-Verlag, New York.
- Ward, J. M. (1983). Background data and variations in tumor rates of control rats and mice. *Proc Exp Tumor Res* 26, 241-58.
- Ward, J. M., and Reznik, G. (1983). Refinements of rodent pathology and the pathologist's contribution to evaluation of carcinogenesis bioassays. *Prog Exp Tumor Res* 26, 266-91.
- Ward, J. M., Hardisty, J. F., Hailey, J. R., and Streett, C. S. (1995). Peer review in toxicologic pathology. *Toxicol Pathol* 23, 226-34.
- Wolf, D.C., and Mann, P.C. (2005). Confounders in interpreting pathology for safety and risk assessment. *Toxicol Appl Pharmacol* 202, 302-8.
- Yoshimura, I., and Matsumoto, K. (1994). Notes on the use of historical controls. *Environ Health Perspect* 102 (Suppl. 1), 19-23.