Comments of Christopher J. Portier, PhD.
USEPA (EPA-HQ-OPP-2016-0385-0094)
Glyphosate Issue Paper: Evaluation of Carcinogenic Potential
October 4, 2016

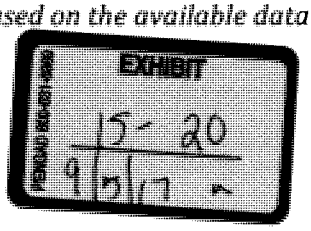
Disclaimer: This work was done with my own resources and on my own time. I have received no reimbursement for any of these comments and no other party has contributed to the drafting of these comments. These comments are solely my opinion and my responsibility.

General Comments and Overall Summary

My comments on the glyphosate review by the USEPA (EPA-HQ-OPP-2016-0385-0094) is rather long and detailed. Realizing that the time and energies of the Science Advisory Panel (SAP) are limited, I will summarize my findings here. Each summarized finding is linked to the line(s) in my more technical review for those who wish to see more details. Because of my own limited time, I have chosen to focus my comments on the human evidence and the animal carcinogenicity evidence, foregoing the review of the other evidence presented. However, I will note that after reading the review on the mechanistic evidence relating to genotoxicity and oxidative stress, I still agree with the findings from the IARC Working Group that there is *strong evidence* that these mechanisms are operable.

Human Evidence Findings

1. The meta-analyses are improperly characterized by the EPA (lines 21-33)
2. The exposure-response relationship in the Agricultural Health Study (AHS) has greater weight than in the other studies, but has problems of its own (lines 39-45)
3. It is not clear in which direction possible confounding would alter the relative risks (lines 61-66) although possible confounding is an issue (line 68).
4. Recall bias is a concern, especially with the case-control studies (lines 70-72)
5. The EPA speculates without data that the more positive studies should have had lower relative risks than other studies (lines 77-80)
6. The follow-up time in the AHS study is likely to be too short to have seen an impact of the magnitude seen in the case-control studies and EPA does a poor job of characterizing the data they used to reach an opposite conclusion (lines 85-109)
7. The EPA speculates that earlier years of exposure prior to the start of the AHS would have effectively expanded the time on study in the AHS without any solid basis (lines 111-114)
8. The Bradford-Hill criteria outlined in the 1997 Guidelines for Carcinogenic Risk Assessment (GCRA) support a conclusion that a causal association in the epidemiology data is credible, but that chance, bias and or confounding could possibly explain the results. (lines 116-127)
9. EPA's interpretation that *"the association between glyphosate exposure and risk of NHL cannot be determined based on the available data"* does not correctly



characterize the human data presented. A better interpretation is that "*a positive association has been observed between exposure to the agent and cancer for which a causal interpretation is considered to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence*". This is the interpretation given these data by the IARC Working Group (lines 129-139)

Animal Carcinogenicity Data Findings

1. EPA's QCAR sets clear guidelines on evaluating animal cancer data with regard to when a high dose is exceeded (lines 151-159), how to interpret trend tests and pairwise comparisons (lines 163-169), how to use historical control data (lines 173-176) and what constitutes a valid historical control data set (180-182)
2. EPA has misinterpreted the language in OCSPP 870.4200 and OCSPP 870.4300 by assuming that an optional highest dose in an animal carcinogenicty study is also a threshold for inclusion of doses in their evaluation. In other words, 1000 mg/kg/day is not an upper bound, 5% in diet is the upper bound (lines 184-210)
3. I have individual comments on every rat study evaluated by the EPA (lines 212-287).
4. EPA consistently dismisses significant findings in rat studies because of a lack of a preneoplastic finding (studies listed starting a lines 217, 229, 277). This presumes that all mechanisms by which chemicals induce tumors in animals will involve enough stages that there would be a histologically identifiable preneoplastic lesion from which final tumors are formed. This simply is not the case and this criteria is applied without any concern for its validity by the EPA.
5. EPA consistently dismisses significant findings in rat studies because of a lack of a significant pairwise comparison even though there is a significant trend in violation of the GCRA (studies listed starting a lines 229, 255, 277).
6. EPA gives less weight to responses seen at doses above 1000 mg/kg/day in all rat studies, even though no dose exceeds 5% of feed. Considering that these findings are in studies with only 50-60 animals per group, that no study appears to have exceeded a maximum tolerated dose (as defined by the EPA and others), it is not clear why EPA does not accept these findings and then do an appropriate margin-of-exposure evaluation or linear extrapolation from these data to show a lack of risk in humans.
7. EPA's summary, which states that *"In 5 of the 9 rat studies conducted with glyphosate, no tumors were identified for detailed evaluation."* is misleading and fails to properly characterize the broad array of findings in these data (lines 291-322). In short, three of these studies were inadequate leaving 2 studies in Sprague-Dawley rats (1 positive) and four studies in Wistar rats (2 positive).
8. With only two studies in Sprague-Dawley rats, the strong positive response seen for thyroid c-cell carcinomas in female rats in one of these studies should be considered positive and due to exposure to glyphosate (lines 324-330)

9. I have individual comments on every mouse study evaluated by the EPA (lines 337-460).
10. EPA consistently dismisses significant findings in mouse studies because of a lack of a preneoplastic finding (studies listed starting a lines 342, 380). This presumes that all mechanisms by which chemicals induce tumors in animals will involve enough stages that there would be a histologically identifiable preneoplastic leasion from which final tumors are formed. This simply is not the case and this criteria is applied without any concern for its validity by the EPA.
11. EPA dismisses significant findings in ALL mouse studies because of a lack of a significant pairwise comparison even though there is a significant trend in violation of the GCRA (lines 337-460).
12. EPA gives less weight to responses seen at doses above 1000 mg/kg/day in all mouse studies, even though no dose exceeds 5% of feed. Considering that these findings are in studies with only 50-60 animals per group, that no study appears to have exceeded a maximum tolerated dose (as defined by the EPA and others), it is not clear why EPA does not accept these findings and then do an appropriate margin-of-exposure evaluation or linear extrapolation from these data to show a lack of risk in humans.
13. EPA uses an outside historical control dataset in one study (start line 380) to dismiss findings and fails to use an equally valid historical control data set identified by the IARC to assess the importance of renal tumors in another study (start line 342). A full evaluation of this second study using the historical control data identified by the IARC supports a strong positive finding in this study (lines 350-365).
14. EPA relies on two-sided p-values for trend tests when one-sided p-values would be more appropriate for identifying adverse effects (lines 367-370; 410-413; Tables 2,4,6)
15. EPA has serious errors in the use of a historical control population that uses data from animals that lived 24 months to compare to response in a study that only went 18 months (lines 388-408). When properly applied, the finding is significant compared to the historical control rate.
16. EPA excludes three positive findings in one study, identified by the European Food Safety Agency for which I sent them data prior to this current EPA review being released (lines 425-434)
17. EPA excludes positive results in a study in Swiss Albino mice because there is an infection in the animals that are not seen in any of the data evaluated by others and for which no documentation is provided (lines 445-460)
18. EPA summarizes the mouse data incorrectly (as they did with rats) when they state that *"No tumors were identified for detailed evaluation in 2 of the 6 mouse carcinogenicity studies."* One study had inadequate dosing and should have been excluded, and one study used Glyphosate trimesium salt rather than pure glyphosate. The remaining four mouse studies all had at least one positive finding (lines 464-475)
19. EPA did not analyze the consistency across mouse studies on the findings relating to renal tumors. I did (Tables 1-3, lines 498-532). Note, all studies were adjusted to an estimated 24 month response using the poly-3 adjustment (lines 487-517)

20. EPA did not analyze the consistency across mouse studies on the findings relating to malignant lymphomas. I did (Tables 1, 4-5, lines 536-541). Note, all studies were adjusted to an estimated 24 month response using the poly-3 adjustment (lines 487-517)
21. EPA did not analyze the consistency across mouse studies on the findings relating to heamngiosarcomas. I did (Tables 1, 6-7, lines 566-599). Note, all studies were adjusted to an estimated 24 month response using the poly-3 adjustment (lines 487-517)
22. Trends in male mice for malignant lymphomas and hemangiosarcomas remained even after doses above 1000 mg/kg/day were excluded (Tables 4-7, lines 536-599).
23. My conclusion is that the mouse data clearly indicates that glyphosate can induce malignant lymphomas and hemangiosarcomas in male CD-1 mice, even when doses above 1000 mg/kg/day are eliminated. There is also a suggestion that glyphosate can induce hemangiomas in female CD-1 mice. The mouse data also demonstrate that glyphosate can induce malignant lymphomas in male CD-1 mice and male Swiss Albino mice. Finally, the renal tumors seen in the CD-1 mice also appear in the Swiss Albino mice, supporting the role glyphosate plays in inducing these tumors. This is clearly sufficient evidence of the carcinogenicity of glyphosate in mice. (lines 573-600)

In summary, these data demonstrate an association in humans to NHL, evidence in rats for thyroid tumors, and very strong evidence in mice for renal tumors, hemangiosarcomas and malignant lymphomas. EPA's exclusion of doses above 1000 mg/kg/day is unscientific and their argument of a lack of significance above this dose is unsupported.

In every case where EPA could choose between a public health protective choice where slight weaknesses in a study or a lack of a very strong finding could raise concerns versus a choice where every study must be perfect and definitive otherwise it is not used, EPA has chosen to discard positive findings leaving them to finally conclude there is no concern. These data simply do not support a finding that glyphosate is *"not likely to be carcinogenic to humans"*.

EPA should declare glyphosate a probable human carcinogen and go on to do a risk assessment to determine if human exposure is sufficient to warrant concern. That resulting risk assessment should be reviewed by the Science Advisory Panel.

DETAILED TECHNICAL REVIEW

Human Evidence

The EPA's final conclusion on the evidence from human exposures to glyphosate and the risk of NHL is as follows:

Page 68: *"Based on the weight-of-evidence, the agency cannot exclude chance and/or bias as an explanation for observed associations in the database. Due to study limitations and contradictory results across studies of at least equal quality, a conclusion regarding the association between glyphosate exposure and risk of NHL cannot be determined based on the available data. The agency will continue to monitor the literature for studies and any updates to the AHS will be considered when available."*

The Agency provides many reasons for this finding. I would summarize them as follows:

1. *"All meta-analysis estimates reported were non-statistically significant except the meta-risk ratio reported by IARC (2015), which was borderline significant with the lower limit of the 95% CI at 1.03"*

Comment: In fact, there were three groups that did meta-analyses. Two were reported as significant (Schinasi and Leon, 2014 and IARC, 2015), although the IARC (2015) corrected an issue they saw with the Schinasi and Leon analysis. The IARC study showed a meta-RR of 1.3 with a confidence bound of (1.03-1.65). The other group (Chang and Delzel, 2016) provided four separate meta-analyses, all of which are reported as having a meta-RR of 1.3 with associated confidence bounds ranging from (1.0-1.6) to (1.0-1.8). Chang and Delzell presented only 1 significant digit for the lower confidence bounds and since their model 1 is exactly the same as the IARC model, they also had at least one significant finding. In fact they characterize their findings as *"we found marginally significant positive meta-RRs for the association between glyphosate use and risk of NHL"*. Thus, the data across all studies, when combined, point to a positive association between glyphosate and NHL in humans.

2. The exposure-response relationship seen in Eriksson et al. (2008) and McDuffie et al. (2001), even though significant, contradicted the exposure-response seen in the Agricultural Health Study (AHS).

Comment: There were 92 cases of NHL in the AHS, with 77.2% (71 cases) having some exposure, whereas the analysis of the tertiles to investigate exposure response relationships, used only 61 cases. Thus, 14% of the exposed cases were excluded. In comparison, both Eriksson et al. (a highly rated study by EPA) and McDuffie et al. were able to characterize all exposed individuals into their exposure groupings with zero loss. To characterize the exposure-response relationship in the AHS as superior to the other two studies is inappropriate.

3. Control for confounding varied across studies and there is a strong potential for

confounding by co-exposures to other pesticides.

Comment: This is correct with some studies doing better than others. However, the magnitude of the impact of this confounding differs by study as well. They cite the one case, Eriksson where the effect estimate went from 2.02 (1.10-3.71) unadjusted to 1.51 (0.77-2.94) adjusted. Others included in the meta-analysis are as follows: DeRoos et al. (2005), 1.2 (0.7-1.9) unadjusted, 1.1 (0.7-1.9) adjusted; DeRoos et al., 2003, 2.1 (1.1-4.0) unadjusted, 1.6 (0.9-2.8) adjusted; Hardell et al., 2002, 3.01 (1.08-8.52) unadjusted, 1.85 (0.55-6.20) adjusted). Orsi et al. (RR 1.0 (0.5-2.2)) and McDuffie et al. (RR 1.2 (0.83-1.74)) did not do analyses adjusting for other pesticides. EPA could remove these studies from the meta-analysis and redo it, but it is unlikely to dramatically change the overall results.

The EPA also expressed concern that what they see as a reduction when you correct for other pesticide exposures would carry over for other confounders. This is highly speculative since many of the NHL patients had no exposure to glyphosate and there are likely truck operators and mechanics (diesel exhaust fumes), factory workers (solvents) and other outdoor workers (UV radiation) in the cases and controls and the result of correcting for the confounders could go either way.

However, it is fair to say that confounding could not be ruled out in these studies.

4. Recall bias is a concern, especially in the case-control studies.

Comment: I agree.

5. The highest risk measures are coming from studies that would likely have lower exposures to glyphosate.

Comment: This is entirely speculative and is based upon an ecological assessment (glyphosate use has increased dramatically over time) and not upon actual data pertaining to the studies at hand. Nor does it fully account for the time since first exposure for the studies done with earlier cohorts.

6. The follow-up time in the DeRoos et al. (2005) study is sufficient that it should be given more weight than the other studies.

Comment: As noted by Portier et al., the median follow-up time in the AHS study was 6.7 years (not 7) and there is a question of whether this is long enough. EPA actually provides a solid argument for why there is concern. EPA gave three publications that they suggest puts the latency period for NHL between 1 and 25 years. Kato et al. (2005) in a high quality population-based, incidence case-control study looking at the relationships between organic solvent exposure and NHL in women found statistical significance only for women occupationally exposed prior to 1970 (cases and controls were recruited between 1995 and 1998) and cited two other studies with similar results (no reference given). They concluded this long latency was either due to higher exposures prior to 1970 or "*at least a 25 year*

latency period is required for NHL induction by these exposures". Weisenburger (1992), in discussing the problems with pathological identification of NHL and the known mechanisms in 1992 states that "The latency for NHL following an environmental exposure is largely unknown" then goes on to say that following chemotherapy for Hodgkin's disease, "the median latency is 5-6 years" based upon 44 case reports from two publications. I was unable to get a copy of one publication, but the publication by Jacquillat et al (1991) showed 24 patients, 17 of whom received radiation therapy along with chemotherapy, 5 radiation alone, three chemotherapy alone and one unknown. The latency ranged from 1 to 11 years in this paper (median 5.5 years) and up to 16 years in the other (abstract review only) These are rather extreme exposures relative to those from glyphosate and it would not be surprising for the glyphosate lag time to be longer than that from chemotherapy and radiation treatment, as suggested by Weisenberger et al. I was unable to obtain a copy of the third paper (Fontana et al., 1998) and the abstract provides no information on lag times.

The rest of the arguments are speculative dealing mostly with years of exposure prior to the beginning of the AHS. Without an analysis including this prior information on exposure with concurrent exposure, it is unclear that the resulting relative risks would go down or up.

Summary: The conclusion by the EPA that "*the association between glyphosate exposure and risk of NHL cannot be determined based on the available data*" fails to account for the overall strength of this evidence and the nature of that evidence. Using the Bradford-Hill criteria for causality described in the 2005 Guidelines for Carcinogenic Risk Assessment (GCRA), I would note that the observations are consistent (relative risks are positive, meta-analyses are positive), significant (in the meta-analysis), not specific (and as noted in the GCRA *"although the presence of specificity may support causality, its absence does not exclude it"*), temporally observed, shows a biological gradient, is coherent with the animal evidence (discussed later), has no experimental evidence from humans, and has no support from structure-activity relationships. So, is causality plausible here? Yes, absolutely. Is it demonstrated? No, clearly not. Are the findings possibly the result of chance, bias and or confounding? Yes, but more unlikely than likely.

The IARC Working Group concluded that there was "*limited evidence of carcinogenicity in humans*" from exposure to glyphosate where, as defined in the IARC Preamble, limited evidence means "*a positive association has been observed between exposure to the agent and cancer for which a causal interpretation is considered to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence.*" This is a more accurate description of these data than that used by the EPA. If chance, bias and confounding could be ruled out, the IARC Working Group would have classified this as a *"known human carcinogen"*, a much stronger finding. By arguing that "*the association between glyphosate exposure and risk of NHL cannot be determined based on the available data*", the EPA has given no weight to the human evidence in their final evaluation.

Animal Carcinogenicity Studies

According to the EPA, of the 9 available rat studies, 4 showed treatment related effects in various organs and of the 6 mouse studies they evaluated, 4 showed treatment effects in three tumors. In all cases, the EPA considers these findings to be not treatment related. I will first address the interpretations of individual studies, then discuss the entire package of studies.

Let's begin by repeating guidance from the GCRA :

"Other signs of treatment-related toxicity associated with an excessive high dose may include (a) significant reduction of body weight gain (e.g., greater than 10%), (b) significant increases in abnormal behavioral and clinical signs, (c) significant changes in hematology or clinical chemistry, (d) saturation of absorption and detoxification mechanisms, or (e) marked changes in organ weight, morphology, and histopathology. It should be noted that practical upper limits have been established to avoid the use of excessively high doses in long-term carcinogenicity studies of environmental chemicals (e.g., 5% of the test substance in the feed for dietary studies or 1 g/kg body weight for oral gavage studies [OECD, 1981])."

and

"A trend test such as the Cochran-Armitage test (Snedecor and Cochran, 1967) asks whether the results in all dose groups together increase as 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 that of 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."

and

"Generally speaking, statistically significant increases in tumors 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 somewhat lower than average."

and

"The most relevant historical data come from the same laboratory and the same supplier and are gathered within 2 or 3 years one way or the other of the study under review; other data should be used only with extreme caution."

These guidelines are critical in the discussions that follow. I will note that the EPA assessment cites OCSPP 870.4200 and OCSPP 870.4300 several times referring to an upper limit for evaluating the high dose in a carcinogenicity study. Theses guidelines give multiple guidance for how to select the appropriate dose. Here are the first two:

(i) For risk assessment purposes, at least three dose levels should be used, in

addition to the concurrent control group. Dose levels should be spaced to produce a gradation of effects. A rationale for the doses selected must be provided.

(ii) The highest-dose level should elicit signs of toxicity without substantially altering the normal life span due to effects other than tumors. The highest dose should be determined based on the findings from a 90–day study to ensure that the dose used is adequate to assess the carcinogenic potential of the test substance. Thus, the selection of the highest dose to be tested is dependent upon changes observed in several toxicological parameters in subchronic studies. The highest dose tested need not exceed 1,000 mg/kg/day.

Nowhere in this guidance does it state that the high dose **cannot exceed** 1,000 mg/kg/day; just that it **does not need to exceed** that number. The EPA notes this fact on Page 69 of the Report, but then later interprets it as a hard limit for excluding doses. Because other data are used to justify the high dose that have not been presented here, we must assume that the highest doses used in the Guideline studies were at or near the maximum-tolerated dose (MTD) and wholly appropriate for the overall evaluation. Thus, 1,000 mg/kg/day is not a threshold for determining where to cut off the data. The only document discussing excessive doses is the QCRA which uses >5% in feed for feeding studies and all doses used here are below that threshold.

Rat Studies

Burnett et al., 1979 (MRID 00105164): As noted by EPA, this study is inadequate due to insufficiently high dose. This study should not be considered negative.

Lankas, 1981 (MRID 00093879): This study in Sprague-Dawley rats was considered inadequate due to the highest dose being far below the MTD. However, the study did see an increase in testicular tumors. These tumors were dismissed because of a non-monotonic dose-response (0%, 6%, 2%, 12% in increasing dose), a lack of pre-neoplastic findings and a range of historical controls (mean 4.5%, range 3.4% to 6.7%) that was higher than seen in the controls, inflating the p-value (as noted in the GCRA, this argument is not an acceptable argument). Nonetheless, the finding in the high exposure group is clearly significant against concurrent controls and, had they presented all of the historical control evidence, might have been significant there as well. Since no data for this tumor is presented for any other study, it is hard to determine if this finding is unique among the studies.

Stout and Ruecker, 1990 (MRID 41643801): The Sprague-Dawley rats in this study were given doses considerably higher (max 1183 mg/kg/day) than those in the Lankas study and was considered adequate by the EPA for evaluation, although they warn that tumor doses in the highest group will be given less weight because it is so high. They found a statistically significant increase in adenomas of the liver and the pancreatic islet cells in males. For pancreatic tumors, EPA points to a lack of clear dose-response (2%, 18%, 10% 15%) and unusually low background response (historical controls provided were 5% mean, 2.9%, 8.5%, 5.8%, 1.8%, 8.3%, 5.0% and 5.1%, all in control groups larger than the concurrent control in this study; since 2% is near 1.8% and only 7 controls are given, this is not an unusually low

response). For liver adenomas (5%, 4%, 6%, 15%), EPA cites a lack of pairwise significance, a plateau of dose-response in the middle dose groups and no preneoplastic lesions as reasons to reject these findings. No historical control data is presented.

In female rats, thyroid C-cell adenomas and combined adenomas and carcinomas were significantly elevated by trend test but not by pairwise comparison. Because of this, they concluded *"although there may be an indication of a dose-response in females, the increases observed in the glyphosate treated groups were not considered to be different than those observed in the concurrent controls"* ignoring their Guidelines regarding *"Significance in either kind of test is sufficient to reject the hypothesis that chance accounts for the result."* Here, preneoplastic lesions were observed, but no monotonic dose-response so they were ignored. Thyroid tumors in male rats were marginally significant (p~0.08).

Atkinson et al., 1993a (MRID 496317023): No adverse effects reported in Sprague-Dawley rats given doses in the same range as the Stout and Ruecker study. No data provided.

Brammer, 2001 (MRID 49704601): This is a two-year study in Wistar rats which showed a statistically significant trend in liver adenomas in male rats (0%, 4%, 0%, 10%) with a maximum dose of 1498 mg/kg/day. EPA provides three reasons for dismissing these findings: non-monotonic dose-response, higher survival in the controls, and multiple comparisons p-value adjustment.

Pavkov and Wyand 1987 (MRIDs 40214007, 41209905, 41209907): This is again a study in Sprague-Dawley rats (substrain given for this study). This study showed no significant findings. The EPA did not comment on the dosing used, however, the maximum dose used in this study was 55.7 mg/kg/day, not much difference from the doses used in Burnett (30 mg/kg/day) and Lankas (34 mg/kg/day) and far lower than doses showing no toxicity in Sprague-Dawley rats. This study should be considered inadequate by the EPA.

Suresh, 1996 (MRID 49987401): This two-year study in Wistar rats using a maximum dose of 886 mg/kg/day saw no significant increases in any tumors. Again, no details are given on tumors appearing in other studies.

Enemoto, 1997 (MRID 50017103-50017105): Also conducted in Wistar rats, but with a maximum dose of 1247 mg/kg/day, demonstrated no increases in tumors. Again, no details are given on tumors appearing in other studies.

Wood et al., 2009a (MRID 49957404): In a last study performed in Wistar rats with a maximum dose of 1229.7 mg/kg/day, a significant increase in female rat mammary tumors (adenomas and carcinomas combined) was observed (4%, 6%, 2%, 16%). EPA dismissed these findings based upon multiple comparisons and no pre-neoplastic lesions.

Excel, 1997: Excluded by the EPA because they had insufficient information on the study and an industry-sponsored review of the literature (Greim et al., 2015) stated it was *"unreliable"*. Greim et al. had multiple errors and considerable missing data (pointed out to EPA in a previous mailing) making it an unreliable source for this decision. No information

is given on this study in any available documents I was able to find including the review by EFSA.

Summary and Comments on the Rat Studies

All told, there are 9 rat studies presented, four in Wistar rats and 5 in Sprague-Dawley rats.

EPA states that *"In 5 of the 9 rat studies conducted with glyphosate, no tumors were identified for detailed evaluation."*, but two of these studies have inadequate dosing to identify.

They also state *"Some of the tumor incidences at the highest dose tested (approaching or exceeding 1,000 mg/kg/day for almost all studies) were statistically significant from concurrent controls using raw (unadjusted) p-values; however, none of the pairwise comparisons were found to be statistically significant following adjustment for multiple comparisons, except the testicular tumors seen in a single study. Furthermore, these high-dose tumors were given less weight."* However, as noted below in my calculation of the limit of 5% of compound in diet, the dose can easily go over 2000 mg/kg/day before reaching this value. They have confused the maximum gavage dose with the maximum dietary dose. These findings should carry equal weight as all other doses.

Three of the Sprague-Dawley rat studies used doses so low that the statistical power to detect an effect was compromised. Even still, one of these studies saw an increase in testicular tumors that was not noted in any other study and could be disregarded (provided there really is no response for this tumor in the other studies). In the remaining two studies (Stout/Ruecker and Atkinson), the EPA argues the highest dose *"exceeds the highest dose recommended in the test guidelines on how to conduct carcinogenicity studies"*. According to Laaksonen et al. (, Lab. Anim. 47(4) 245-56, 2013), Sprague-Dawley rats eat, on average, about 600g/kg/week at study start and about half that at 2 years. Based on the guidelines, 5% in diet is acceptable and on a daily basis would be between 2.1 g/kg/day to 4.3 g/kg/day; thus the <2 g/kg/day used in these two studies should be acceptable. This argument is not supported for these studies. These two studies differed on their findings of cancer with Atkinson negative for all cancers and Stout/Ruecker positive for two cancers, one in females and one in males. The remaining reasons for dismissing these findings include a lack of preneoplastic findings and a non-monotonic dose-response.

The thyroid tumors in female rats Stout/Ruecker) should be considered a positive finding. The dose-response is clear and the marginal findings in males should increase the concern for this tumor. There is no reason to believe that adenomas and carcinomas MUST arise from preneoplastic lesions in thyroid C-cell tumors. The rates from the other study for these tumors are no presented, but even if they had been, how do you judge one positive study against one negative study? The public protective decision in this case should be to conclude these tumors arose as a function of exposure to glyphosate.

The remaining tumors can be debated; in all cases where a decision could go either way, EPA dismisses findings rather than accepts them.

Mouse Studies

Reyna and Gordon, 1973 (MRID 00061113): This study is new to the EPA assessment. This study used doses as high as 50 mg/kg/day, far below the maximum doses used in the other studies that were below the maximum tolerated dose. In essence, this study is inadequate and should not be used for making a decision.

Knezevich and Hogan, 1983 (MRID 00130406): This is the first 24 month study in CD-1 mice at dietary doses of up to 6069 mg/kg/day according to EPA. EPA does not show their conversion from ppm in feed to mg/kg/day so it is unknown why this number is different from the European Food Safety Authority (EFSA) which lists the highest dose as 5874 mg/kg/day. The actual high dose used, according to EFSA, was 30,000 ppm or 3% in feed, below the EPA threshold given in the GCRA. According to EPA, *"No effect on survival was observed"* suggesting this high dose did not exceed the MTD.

This study saw an increase in kidney tubular cell adenomas and carcinomas (2%, 0%, 0%, 6%), a very rare tumor in these mice. Four reasons were given for discounting this finding: *"1) renal tubular cell tumors are spontaneous lesions for which there is a paucity of historical control data for this mouse stock; 2) there was no statistical significance in a pairwise comparison of treated groups with the concurrent controls and there was no evidence of a statistically significant linear trend; 3) multiple renal tumors were not found in any animal; and 4) compound-related nephrotoxic lesions, including pre-neoplastic changes, were not present in male mice in this study"*. In fact, the one-sided p-value (alternative is an increased risk) for this study was 0.03. In 1986, the EPA did have an adequate historical control population for these tumors and found they were highly statistically significant. The IARC also identified an adequate historical control population (Chandra and Frith, 1994) who reported only 1 tumor in 725 CD-1 mice also supporting a highly significant finding. As noted earlier, the second reason violates the QCRA if the one-sided test is applied. The third argument is not supported with such a small number of affected animals and a very rare tumor and the fourth reason, while arguable, presumes there would be a preneoplastic lesion rather than a unique mutational event to begin the cancer process.

Note: The raw p-value presented in Table 4.12 is for a two-sided test, a one-sided test is more appropriate here and has a raw p-value of 0.034. If the true control rate is 0.0014 as noted by Chandra and Frith (1994), the probability of seeing a finding more extreme than the one noted here is 0.0017. Even if the background is as high as 1%, the p-value would be 0.026.

Atkinson, 1993b (MRID 49631702): This 24 month study in CD-1 mice showed an increase in hemangiosarcomas (0%, 0%, 0%, 9%) which was statistically significant (p=0.003) with a marginally significant comparison between control and high dose of 0.053. The only negative comment given by the EPA on this study was *"however, the incidence of hemangiosarcomas at the high-dose was not statistically significant when compared to the concurrent controls"*, thus excluding the finding from the trend test because of a non-significant pairwise test, in violation of the QCRA.

Wood et al., 2009b (MRID 49957402): This study, also in CD-1 mice, was for 80 weeks (approximately 18 months) with a high dose in males of 810 mg/kg/day (again, not exceeding the 5% dose in feed). There was no effect on survival suggesting the study did not exceed the MTD. There was a monotonic increase in lung adenocarcinomas (10%, 10%, 14%, 22%) and a monotonic increase in malignant lymphomas (0%, 2%, 4%, 10%). For the lung cancers, the EPA again argued a lack of significance for pairwise comparisons (in violation

of their QCRA) and there was no evidence of progression from adenomas to carcinomas.

For the malignant lymphomas, the EPA noted that *"For this strain of mouse, the mean incidence for untreated animals is approximately 4.5% (range: 1.5%-21.7%) based on historical control data from Charles River (59 studies performed from 1987-2000; Giknis and Clifford, 2005) and Huntingdon Laboratories (20 studies from 1990-2002; Son and Gopinath, 2004)."* These controls are not from the same laboratory at the same time, but EPA did paraphrase the QCRA noting that these data *"should be used with caution"* whereas the GCRA states *"other data should be used only with extreme caution"*. In this case they did neither. The paper by Son and Gopinath documents the numbers of tumors seen in animals that die prior to 80 weeks out of 1453 males in 20 control groups. They saw a total of 36 animals with lymphomas, for a raw rate of 2.4%; however this is a lower bound on the rate since they did not look at all animals at 80 weeks to get obtain the number that are alive and having a tumor. It is not clear how EPA interpreted these numbers in their presentation. The study by Giknis and Clifford (2005) had 52 studies (not 59) and only 26 of them were for 18 months; the rest were for 2 years and these last 26 would be inappropriate as a historical control. The numbers cited by the EPA *("4.5% (range: 1.5%-21.7%)")* are directly out of Giknis and Clifford for all 52 studies and the range fails to include the 11 studies with no tumors (lower end of range is 0). In the 26 studies ending at 18 months, Giknis and Clifford saw tumor incidence as follows (0/60, 0/50, 0/50, 0/50, 0/50, 0/50, 0/50, 0/50, 1/69, 1/50, 1/50, 1/50, 1/50, 1/50, 1/50, 1/50, 2/60, 2/59, 2/53, 2/50, 2/47, 2/46, 3/60, 3/59, 4/49, 7/50) thus ranging from 0% to 14% with a weighted mean of 2.5%.

NOTE: The p-value sited by EPA for the trend test is the two-sided p-value; a one sided p-value is more appropriate and the correct value is 0.0043. If you assume that 2.5% is the historical control rate, the probability of seeing a more significant finding that the one seen in this study is 0.0079.

<u>Sugimoto, 1997 (MRID 50017108 - 50017109):</u> In another study in CD-1 mice (with sub strain noted), mice were given, for 18 months, a maximum dose of 40,000 ppm of glyphosate which is 4% in the diet, again below the 5% in feed set by the QCRA. The second highest dose was 0.8% in diet. This study demonstrated a clear dose-response for hemangioma in female mice (0%, 0%, 4%, 10%) with a p-value for trend of p=0.002 by EPA's calculation. There were no treatment effects on survival suggesting this dose did not exceed the MTD. This tumor was not considered treatment related by the EPA because of no pairwise significance with the high dose versus control using a multiple comparisons analysis (the uncorrected p-value is 0.028 and the corrected p-value is 0.055).

What is not mentioned by the EPA but was evaluated by the EFSA, was the dose-response trend for hemangiosarcoma in male mice for which the one-sided p-value for trend is 0.008. Here the responses are 0%, 0%, 0% and 4%, a very low response rate. However, this is only an 18 month study, so low rates of tumors are to be expected.

What is also not mentioned are the malignant lymphomas and kidney tumors also found in males in this study (EFSA, 2015). The renal tumors had rates of 0%, 0%, 0%, 4% (the same as the hemangiosarcomas in males) with a p-value for trend of 0.008. The malignant lymphomas had rates of 4%, 4%, 0%, 12% with a p-value for trend of 0.008. I will compare these rates to those seen in the other studies later.

Pavkov and Turnier, 1987 (MRIDs 40214006, 41209907): This is a two-year chronic toxicity study in CD-1 mice with a maximum dose of 991 mg/kg/day. They list this study as completely negative for any cancer findings. However, this study evaluated Glyphosate trimesium salt (52.6% pure). No details on this study are provided by the EPA and I could find no other regulatory body that has reviewed this study nor is it listed in the Greim et al. (2015) manuscript. It is also the only carcinogenicity study with such a low percentage of pure glyphosate.

Kumar, 2001: This 18-month chronic carcinogenicity study in Swiss Albino mice with high-dose exposures of 10,000 ppm (1% diet) was excluded by the EPA *"due to the presence of a viral infection within the colony, which confounded the interpretation of the study findings"*. No information on this viral infection is given in the EPA Assessment. It is not possible to determine where this information on a viral infection came from. In the most recent draft classification document on glyphosate by the European Chemical Agency, they state that *"in the study report itself, there was no evidence of health deterioration due to suspected viral infection and, thus, the actual basis of EPA's decision is not known"* when referring to this study. The only reference I can find is from the paper by Greim et al. who down-rated the study *"based on speculation of a viral infection within the colony"*.

This study is important as they saw increases in kidney tumors (0%, 0%, 2%, 4%) and malignant lymphomas (20%, 30%, 32%, 38%) with one-sided p-values for trend of 0.04 and 0.05 respectively. While these are not strikingly strong p-values, they show a consistency in the male mouse data for these tumors.

Summary and Comments on the Mouse Studies

EPA concluded that *"No tumors were identified for detailed evaluation in 2 of the 6 mouse carcinogenicity studies."* One of these mouse studies should have been excluded because of the low doses used in the study. The other study has no details provided by the EPA or any other regulatory body and uses Glyphosate trimesium salt (52.6% pure).

EPA then concluded *"In the remaining 4 mouse studies, 3 observed a statistically significant trend in tumor incidences in the hemangiosarcomas, lung adenomas, malignant lymphomas or hemangiomas; however, the agency determined that none of the tumors observed in the mouse are treatment related."* In fact, there were 5 additional studies since they excluded the one study in Swiss Albino mice because of an infection in the study animals that appears to be speculative. Let's consider these 5 remaining studies. Since the hemangiomas only occurred in one study in female mice, I will not discuss it further.

Table 1 provides a summary of the findings in the 5 studies for which I could find sufficient data to make a comparison across the three main tumor findings in male mice: renal tumors, hemangiosarcoms and malignant lymphomas. A review of all of the studies in one simple picture illustrates the consistency of the findings across the various studies. Now, let's compare the actual tumor rates to see how they compare.

Table 1: Cancer findings in studies of glyphosate in male mice

1983[5]	Crl:CD-1	24	4,841	+[3]		
1993[5]	?:CD-1	24	1,000		+	+/-[4]
1997	CrJ:CD-1	18	4,843	+[5]	+[5]	+[5]
2001	Swiss	18	1,460	+[5]	**No Data**	+[5]
2009	Crl:CD-1	18	810			+

Cancer increases in risk generally as a power of length of exposure (1). This relationship was used to develop a means to adjust the length of time an animal is on a study, enabling a scientist to determine risk at the end of two-years, the typical time used for animal bioassays (2, 3). This is called the Poly-3 adjustment. The US National Toxicology Program uses the Poly-3 test to evaluate significance in their animal bioassays. Now you will note that three of the mouse studies were only conducted for 18 months. (Comparing 18 month studies with 24 month studies without making an adjustment for the differences in length of exposure is like comparing cancer rates in 40 year-olds exposed for 20 years to cancer rates in 65 year-olds exposed for 45 years and concluding they are not consistent with each other; the conclusion is meaningless because the correct evaluation was not done.) Thus, in order to compare all 5 studies, we must use the Poly-3 adjustment to extrapolate the 18 month studies to estimate what we think the cancer risk would have looked like at 24 months. The adjustment decreases the number of animals without tumors in all groups in the 18 month studies by $(18/24)^{3}$. The one-sided p-values for both the unadjusted trend test and the poly-3 adjusted trend test are given in Table 2 for male mouse renal tumors.

---

[1] months
[2] mg/kg/day
[3] indicates p-value for trend <0.05
[4] p=0.08
[5] not evaluated by the EPA

Table 2: Analysis of Male Mouse Renal Tumors From the Individual Studies

83	Crl:CD-1	24	157, 814, 4841	1/50, 0/49, 1/50, 3/50	0.03 (0.03)
93	?:CD-1	24	100, 300, 1000	2/50, 2/50, 0/50, 0/50	0.94 (0.94)
97	CrJ:CD-1	18	165, 838, 4348	0/50, 0/50, 0/50, 2/50	0.008 (0.009)
01	SW	18	15, 151, 1460	0/49, 0/49, 1/50, 2/50	0.04 (0.04)
09	Crl:CD-1	18	71, 234, 810	0/51, 0/51, 0/51, 0/51	-

As an example of how the Poly-3 adjustments work, consider a comparison of the high-dose renal tumor response in the 1983 study (3/50=6%) to the high-dose response in the 1997 study (2/50=4%). In the 1997 study, 48 animals had no tumors at 18 months; the poly-3 adjustment reduces this to 20.25 leading to an incidence estimate of 2/22.25=9%. Because the Poly3 test effectively reduces the number of animals on study, even though the incidence estimate goes up, the p-value for the trend test goes down. Numerous evaluations of the validity of the poly-3 adjustment have been published in the peer-reviewed literature and it seems to work very well.

Now that the lengths of the studies have been adjusted, the next question to ask is whether this dose-response is consistent across all of the studies or whether there are anomalies. Combining all of the studies into one analysis can help us to evaluate this question; if the pooled data are no longer significant or less significant, the studies are not consistent and do not complement each other. Combining all of the studies into one pooled analysis and performing a trend analysis on the pooled data yields highly significant findings (Table 3, Line 1). Excluding the Swiss Albino mouse study and only using the CD-1 mice also yields a significant trend (Table 3, Line 2). Repeating these analyses with the Poly-3 adjusted data does not alter the significant findings. Since EPA is concerned about doses above 1000 mg/kg/day, I excluded doses above this dose and re-analyzed the data. The results of the restricted analysis are shown in Table 3, Lines 3-4. Without the doses above 1000 mg/kg/day, the effect disappears.

Table 3: Pooled Analysis of Male Mouse Renal Tumors

All Combined	CD-1 and Swiss	0.0004 (0.001)
CD-1 Combined	CD-1	0.001 (0.001)
All Combined, doses>1000 dropped	CD-1 and Swiss	0.80 (0.84)
CD-1 Combined, doses>1000 dropped	CD-1	0.85 (0.86)

Tables 4 and 5 repeat these analyses for malignant lymphomas. Because of the different backgrounds between the Swiss mice and the CD-1 mice, when they are all combined, the joint analysis is not significant (Table 5, line 1). Removing the Swiss mouse study and only evaluating the CD-1 mice leads to highly significant trends in all analyses (Table 5, lines 2). A significant trend remains in CD-1 mice even after removing the doses>1000 mg/kg/day (Table 5, line 4) suggesting this is not a high-dose only effect.

Table 4: Analysis of Male Mouse Malignant Lymphoma From the Individual Studies

1983	Crl:CD-1	24	157, 814, 4841	2/50, 5/49, 4/50, 2/50	0.51 (0.51)
1993	?:CD-1	24	100, 300, 1000	4/50, 2/50, 1/50, 6/50	0.08 (0.08)
1997	CrJ:CD-1	18	165, 838, 4348	2/50, 2/50, 0/50, 6/50	0.008 (0.012)
2001	SW	18	15, 151, 1460	10/49, 15/49, 16/49, 19/49	0.05 (0.09)
2009	Crl:CD-1	18	71, 234, 810	0/51, 1/51, 2/51, 5/51	0.004 (0.005)

Table 5: Pooled Analysis of Male Mouse Malignant Lymphoma

All Combined	CD-1 and Swiss	0.17 (0.19)
CD-1 Combined	CD-1	0.02 (0.01)
All Combined, doses>1000 dropped	CD-1 and Swiss	0.86 (0.93)
CD-1 Combined, doses>1000 dropped	CD-1	0.03 (0.05)

Tables 6 and 7 repeat these analyses for hemangiosarcomas. The findings in the Swiss mouse were unavailable so Tables 6 and 7 only contain analyses of the CD-1 mouse data. All pooled analyses are highly significant (Table 7) and they remain significant if doses>1000 are excluded (Table 7, line 2). So again, this is not a high dose-only effect.

Table 6: Analysis of Male Mouse Hemangiosarcomas From the Individual Studies

1983	Crl:CD-1	24	157, 814, 4841	0/50, 0/49, 1/50, 0/50	0.63 (0.63)
1993	?:CD-1	24	100, 300, 1000	0/50, 0/50, 0/50, 4/50	0.0004 (0.0004)
1997	CrJ:CD-1	18	165, 838, 4348	0/50, 0/50, 0/50, 2/50	0.008 (0.009)
2001	SW	18	15, 151, 1460	No Data	-
2009	Crl:CD-1	18	71, 234, 810	0/51, 0/51, 0/51, 0/51	-

Table 7: Pooled Analysis of Male Mouse Hemangiosarcomas

CD-1 Combined	CD-1	0.02 (0.03)
CD-1 Combined and Doses Pooled[1]	CD-1	0.02 (0.02)
CD-1 Combined, doses>1000 dropped	CD-1	<0.0001 (<0.0001)
CD-1 Combined, doses>1000 dropped and Doses Pooled[2]	CD-1	0.0003 (0.0003)

In summary, the results seen for renal tumors, malignant lymphomas and hemangiosarcomas in male mice in the 4 CD-1 studies for which the data were available are consistent and have a much stronger trend when all of the data are combined. The trend tests for malignant lymphomas and hemangiosarcomas in these studies remain significant when doses above 1000 mg/kg/day are eliminated.

EPA's approach has been to eliminate each study separately, generally by arguing the dose is too high (even though no signs of exceeding the MTD are apparent and their guidelines do not support the cut-off they are using), that there are no precursor lesions (suggesting cancer cannot arise without precursor lesions which is not a scientific necessity), and that the pairwise comparisons are not significant so the trend test should be ignored (in violation of their own guidelines). In addition, EPA has failed to present all of the positive tumor sites seen in these mouse studies, they have incorrectly used (probably inappropriate) historical controls and when these are used correctly a significant finding remains, they have included studies that should have been dismissed due to power issues, have included a study for which there is almost no available information other than the one paragraph they have presented, and have not evaluated the data across the studies to look for consistency in the response for tumors that appear in multiple studies. In essence, this is a very weak scientific evaluation of the available mouse carcinogenicity data.

My conclusion is that the mouse data clearly indicates that glyphosate can induce malignant lymphomas and hemangiosarcomas in male CD-1 mice, even when doses above 1000 mg/kg/day are eliminated. There is also a suggestion that glyphosate can induce hemangiomas in female CD-1 mice. The mouse data also demonstrate that glyphosate can induce malignant lymphomas in male CD-1 mice and male Swiss Albino mice. Finally, the renal tumors seen in the CD-1 mice also appear in the Swiss Albino mice supporting the role glyphosate plays in inducing these tumors. This is clearly sufficient evidence of the carcinogenicity of glyphosate in mice.

1. Portier CJ, Hedges JC, Hoel DG. Age-specific models of mortality and tumor onset for historical control animals in the National Toxicology Program's carcinogenicity experiments. *Cancer Res* (1986) **46**(9):4372-8. PubMed PMID: 3731095.
2. Bailer AJ, Portier CJ. Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples. *Biometrics* (1988) **44**(2):417-31. PubMed PMID: 3390507.
3. Portier CJ, Bailer AJ. Testing for increased carcinogenicity using a survival-adjusted quantal response test. *Fundam Appl Toxicol* (1989) **12**(4):731-7. PubMed PMID: 2744275.

www.agri-pulse.com/articles/7584-oh-brother-croplife-questions-makeup-of-glyphosate-panel

Oh, brother: CropLife questions makeup of glyphosate panel

Steve Davies (/authors/2-steve-davies)
October 12, 2016

WASHINGTON, Oct. 12, 2016 - When the Scientific Advisory Panel begins four days of meetings next week to examine the carcinogenic potential of glyphosate, Christopher Portier won't be there.

But his brother Kenneth, whom EPA picked to serve as one of the scientists reviewing the world's most widely used herbicide, will be in attendance.

And that has the crop-protection industry concerned.

Christopher Portier, former director of the U.S. National Center for Environmental Health at the Centers for Disease Control and Prevention, served as an "invited specialist" on the International Agency for Research on Cancer panel that determined glyphosate is "probably carcinogenic to humans."

He has been a lightning rod for criticism since the IARC assessment came out in March 2015 and sparked a worldwide pushback from principal manufacturer Monsanto and the ag chemical industry in general to prove that it is not.

The European Food Safety Authority
released
(https://www.efsa.europa.eu/en/efsajournal/pub/4302)
its own analysis, which concluded glyphosate is not likely to be carcinogenic, after which Portier led a group of 95 scientists who
wrote a letter to EFSA
(https://www.efsa.europa.eu/en/press/news/160113)
and
published a study
(http://jech.bmj.com/content/early/2016/03/03/jech-2015-207005.full)
disputing that characterization and taking issue with EFSA's methodology.

EXHIBIT
15-21
9/5/17

Christopher's brother, Kenneth Portier, is also a scientist: He's vice president of the Statistics & Evaluation Center at the American Cancer Society, and has served on more than 60 Scientific Advisory Panels for EPA.

Despite those qualifications, his family ties have CropLife America raising questions about the SAP.

"It's hard to know if it's a fair panel," said Janet Collins, CLA's senior vice president of science and regulatory affairs, when asked about the SAP members last week. (Technically, the appointed scientists are members of the Science Review Board. The SAP is a permanent body that contributes a few of its members to panels that examine issues arising under the Federal Insecticide, Fungicide, and Rodenticide Act.)

When questioned specifically about Kenneth Portier, Collins noted that Christopher Portier has been "leading the charge" in defending IARC's monograph and speculated that the brothers have "probably talked about" this area of common interest.

"I'm not intending to call this particular gentleman out," she said, referring to Kenneth Portier. "I'm not in any way disparaging him or his credentials." But she added, "There appears to be a conflict of interest."

"Somebody who wanted to could argue that he shouldn't be on the panel," she said.

Agri-Pulse
could not reach Kenneth Portier for comment, but Christopher defended his brother's credentials. In a phone interview, he said that Kenneth is "perfectly well qualified" to be on the glyphosate SAP.

"I don't believe anybody in their right mind would say he's not perfectly qualified. If they're trying to say he's biased, good luck, how are you going to prove it?"

"My brother chaired EPA's SAP for seven years," he continued. "Nobody has ever questioned his integrity or his scientific acumen – in fact, his statistical acumen. They're welcome to question it now, but they're just going to get a lot of negative feedback on it because he's done an excellent job for EPA over the years."

Portier also said that he and his brother don't always see eye to eye. "He doesn't agree with everything I believe," Christopher said.

Asked whether he had spoken with his brother about glyphosate, Christopher Portier said, "In broad terms. I told him it's a battle between hazard and risk, and that he does understand."

"It's always an interesting tactic, I find, to change the message," Christopher said. "The message is not what I'm saying is wrong, the message becomes, 'But his brother's on the panel.' I'd prefer if they'd come straight at me and come after my message."

Portier also addressed the criticism that because he works part time for the Environmental Defense Fund – and did at the time of the IARC review – that his involvement with the IARC monograph is somehow tainted.

"I work for them two days a week, mostly on air pollution and air pollution modeling, and on climate change and climate change modeling." The work, he said, has "nothing to do with pesticides."

"Nobody has paid me a cent to do what I'm doing with glyphosate," he said. "I have no conflict of interest whatsoever."

Christopher Portier has submitted
comments
(http://www.agri-pulse.com/uploaded/glyphosate-portier-sap-comments.pdf)
to EPA in advance of the SAP meeting, at which the scientists will consider the studies and methods used by EPA
to conclude in a recent paper
(http://www.agri-pulse.com/uploaded/glyphosate-issue-paper-epa.pdf)
that glyphosate is not likely carcinogenic.

In his comments, Portier said EPA incorrectly downplayed the significance of rat and mouse studies because of the size of the doses, even though the doses did not exceed 5 percent of the animals' body weight.

The data he examined, he said, "demonstrate an association in humans to (non-Hodgkin lymphoma), evidence in rats for thyroid tumors, and very strong evidence in mice for renal tumors, hemangiosarcomas and malignant lymphomas. EPA's exclusion of doses above 1,000 mg/kg/day is unscientific and their argument of a lack of significance above this dose is unsupported."

"In every case where EPA could choose between a public health protective choice where slight weaknesses in a study or a lack of a very strong finding could raise concerns, versus a choice where every study must be perfect and definitive, otherwise it is not used, EPA has chosen to discard positive findings, leaving them to finally conclude there is no concern. These data simply do not support a finding that glyphosate is 'not likely to be carcinogenic to humans.'"
Also commenting
(https://www.regulations.gov/document?D=EPA-HQ-OPP-2016-0385-0221)
were some of the scientists who recently prepared a paper that found glyphosate is not likely be carcinogenic. Intertek Scientific and Regulatory Consultancy Services put together the panel for that paper on commission from Monsanto.

That paper said that "even without data IARC did not include, there is no support for IARC's conclusion that glyphosate is 'probably carcinogenic to humans.'"

In their comments, five of the 15 scientists on the Intertek panel praised EPA for "an excellent and thorough review of glyphosate."

"We agree with the agency that it is important to give more weight to studies evaluating endpoints that measured gene mutations and chromosomal aberrations (i.e. permanent DNA damage) than to endpoints reflecting DNA events that may be transient or reversible such as primary DNA damage (e.g., comet assays)," they said.

The SAP meeting will be held Oct. 18-21 in Arlington, Virginia. The online docket for the meeting is here
(https://www.regulations.gov/docket?D=EPA-HQ-OPP-2016-0385)
.

To: Jones, Jim[redacted]
From: Chris Portier
Sent: Wed 5/4/2016 11:38:18 AM
Subject: Fwd: glyphosate: POLITICO on EPA report

;;;;Jim,
FYI.

C.

Subject: glyphosate: POLITICO on EPA report

GLYPHOSATE STORM'S A-BREWIN': The U.S. Environmental Protection Agency has made a preliminary finding that glyphosate is unlikely to cause cancer in humans — but the agency isn't ready to go public yet. The EPA briefly posted online an October 2015 final report from its Cancer Assessment Review Committee, which concluded glyphosate is "not likely to be carcinogenic to humans." It then pulled it from its website. The committee said evidence from existing epidemiological studies and tests of lab animals doesn't meet the bar for classifying the herbicide as a carcinogen. An agency spokesperson told POLITICO the report was removed because assessment was ongoing. "Our assessment will be peer reviewed and completed by end of 2016," said the spokesperson.

— Why this matters for the EU: A political scrum over what to do about glyphosate is underway in the EU. Parliament voted to extend the chemical's authorization for seven years, the Commission is pushing for 10, but the real decision comes in a Plant, Animal, Food and Feed Committee meeting on May 18-19. Advocates for banning glyphosate altogether cite a March 2015 study by International Agency for Research on Cancer, which said it caused cancer. Glyphosate's political supporters cite a November study with the opposite conclusions. This latter group might now have another study in their arsenal — and from a reputable U.S. government agency. "In line with the 90,000 pages, and 3,300 studies already published in support of the reapproval of glyphosate, the EPA report casts yet more doubt on the conclusions of IARC," a spokesperson for the European Crop Protection Association told Morning Agri. Greenpeace EU, which opposes using glyphosate as long as there is no scientific consensus, told Morning Agri it had not yet read the study and so couldn't comment. More: http://reut.rs/23mbxYf.

EXHIBIT
15-22
9/13/17

From: "Lowit, Anna" <[redacted]>
Subject: FW: Sorry
Date: June 24, 2016 at 8:18:40 PM GMT+2
To: "[redacted]" <[redacted]>

Hi Chris

Jim Jones forwarded me some files from you. thanks for sending them. I have a quick Q for you.

in this PPT file, what is the citation(s) for the metaanalysis of the animal tumor data?

Thanks
Anna

Sent from my Windows Phone

From: Jones, Jim
Sent: 6/24/2016 7:43 AM
To: Housenger, Jack; Lowit, Anna
Subject: FW: Sorry

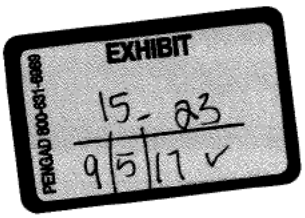
As per my conversation with Jack. Jim

-----Original Message-----
From: Chris Portier [redacted]
Sent: Thursday, June 23, 2016 2:17 PM
To: Jones, Jim <[redacted]>
Subject: Sorry

Jim,

I had an error in one Table that I had to correct. New version attached.

C,



From: "Lowit, Anna" <[redacted]>
Subject: RE: Sorry
Date: June 24, 2016 at 9:11:58 PM GMT+2
To: Chris Portier <[redacted]>

Ditto on the terse emails, it's too easy on the phone to be quick and even rude! I do it too.

Thanks for the quick response.

Would you mind sharing the code? I'm interested in the analysis, it's a different approach to the data compared to all the others "floating around".

Sent from my Windows Phone

From: Chris Portier
Sent: 6/24/2016 2:22 PM
To: Lowit, Anna
Subject: Re: Sorry

Anna,

Oh, and I wanted to say Hi Anna. Sometimes my emails are a bit short.

If you need any background from me, I'll be happy to help you out. I am also in DC if you want to meet and discuss this.

C.

On Jun 24, 2016, at 2:18 PM, Lowit, Anna <[redacted]> wrote:
Hi Chris

Jim Jones forwarded me some files from you. thanks for sending them. I have a quick Q for you.

in this PPT file, what is the citation(s) for the metaanalysis of the animal tumor data?

Thanks
Anna

Sent from my Windows Phone

From: Jones, Jim
Sent: 6/24/2016 7:43 AM
To: Housenger, Jack; Lowit, Anna
Subject: FW: Sorry

As per my conversation with Jack. Jim

-----Original Message-----
From: Chris Portier [redacted]
Sent: Thursday, June 23, 2016 2:17 PM
To: Jones, Jim <[redacted]
Subject: Sorry

Jim,

I had an error in one Table that I had to correct. New version attached.

C,

<FiguresandTablesEPA.pptx>

From: "Lowit, Anna" <[redacted]>
Subject: RE: Sorry
Date: June 27, 2016 at 2:17:34 AM GMT+2
To: Chris Portier <[redacted]>

Thanks!

Sent from my Windows Phone

From: Chris Portier
Sent: 6/26/2016 6:42 PM
To: Lowit, Anna
Subject: Re: Sorry

Anna,

Per your request. I believe these are all of the files you will need. Let me know if these do not work for you. I had some minor errors in the tables again because I was not taking direction into account for the trend test (up or down). That is now fixed.

C.

From: "Lowit, Anna" <[redacted]>
Subject: RE: Sorry
Date: June 24, 2016 at 9:23:09 PM GMT+2
To: Chris Portier <[redacted]>

That would be great ☺

Sent from my Windows Phone

From: Chris Portier
Sent: 6/24/2016 3:14 PM
To: Lowit, Anna
Subject: Re: Sorry

No problem. Shall I clean it up a bit first. I can get it to you by monday.

On Jun 24, 2016, at 3:11 PM, Lowit, Anna <[redacted]> wrote:
Ditto on the terse emails, it's too easy on the phone to be quick and even rude! I do it too.

Thanks for the quick response.

Would you mind sharing the code? I'm interested in the analysis, it's a different approach to the data compared to all the others "floating around".

Sent from my Windows Phone

From: Chris Portier
Sent: 6/24/2016 2:22 PM
To: Lowit, Anna
Subject: Re: Sorry

Anna,

Oh, and I wanted to say Hi Anna. Sometimes my emails are a bit short.

If you need any background from me, I'll be happy to help you out. I am also in DC if you want to meet and discuss this.

C.

On Jun 24, 2016, at 2:18 PM, Lowit, Anna <[redacted]> wrote:
Hi Chris

Jim Jones forwarded me some files from you. thanks for sending them. I have a quick Q for you.

in this PPT file, what is the citation(s) for the metaanalysis of the animal tumor data?

Thanks
Anna

Sent from my Windows Phone

From: Jones, Jim
Sent: 6/24/2016 7:43 AM
To: Housenger, Jack; Lowit, Anna
Subject: FW: Sorry

As per my conversation with Jack. Jim

-----Original Message-----
From: Chris Portier [redacted]
Sent: Thursday, June 23, 2016 2:17 PM
To: Jones, Jim <[redacted]
Subject: Sorry

Jim,

I had an error in one Table that I had to correct. New version attached.

C,

<FiguresandTablesEPA.pptx>

White, Dylan

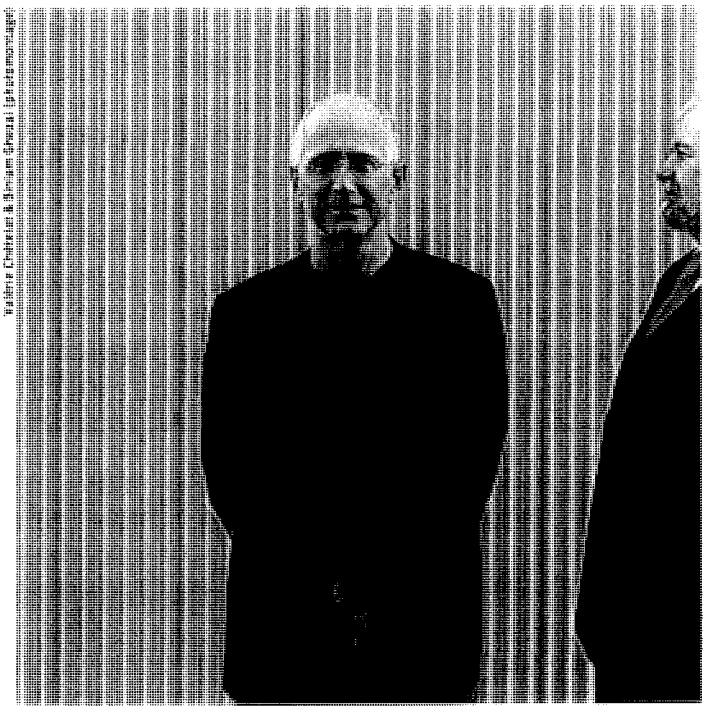
From:	Chris Portier [redacted] on behalf of Chris Portier
Sent:	Monday, March 07, 2016 6:17 AM
To:	Dr. Christopher Portier
Subject:	Keeping you all informed...
Attachments:	Horizons_Glyphosate.pdf; ATT00001.htm; voting-calculator-results.pdf; ATT00002.htm
Follow Up Flag:	Follow up
Flag Status:	Flagged

I did a Pro-Con piece for the Swiss science magazine Horizons a month or so ago on glyphosate and cancer and it just came out. The head of the EFSA pesticides program did the Con. The words he chose were very carefully nuanced so the reader does not see the debate about trend analysis but it is written in the piece. Very interesting. (Attached)

EXHIBIT
15-24
9/5/17

Is glyphosate carcinogenic?

Glyphosate is the world's most widely used herbicide. While it's important in controlling weeds, its possible effects on humans are hotly debated among scientists.



Walter F. Chatterjee & Miriam Ghezzal | Spektrum.ch/Image

Yes

says the environmental health researcher Christopher Portier

How strong is the evidence that humans are more likely to get cancer when widely exposed to glyphosate? This question centres on three issues: finding evidence in humans, finding evidence in laboratory animals and finding evidence of a molecular mechanism by which glyphosate might cause cancer.

Some 26 cancer studies have been carried out on humans who have been exposed to glyphosate formulations. Most of them found no connection. Nine of these studies examined non-Hodgkin lymphoma. Four case-control studies, when pooled, showed there to be an association between this cancer and glyphosate, as did two other case-control studies. The studies of higher quality adjusted for multiple exposures to other pesticides but still demonstrated an association, with the length of exposure increasing the strength of the association. However, these studies had certain limitations that made it impossible to rule out bias or other confounding factors. The conclusion we must draw is that glyphosate formulations are associated with non-Hodgkin lymphoma in humans, but there is only limited evidence of causality.

Five laboratory studies were carried out on mice, and nine on rats. All five of the mouse studies displayed increased tumour growth in at least one site. Three studies showed growth in kidney tumours, which rarely occur in mice; two studies showed an increase in hemangiosarcomas (a cancer arising in the blood vessels); and two studies also showed growth in malignant lymphomas. With the exception of growth in a few non-malignant tumours, none of the rat studies showed any effect. The conclusion is that glyphosate causes various tumours in laboratory mice.

> "There is evidence of a mechanism by which glyphosate causes cancer".

As to the molecular mechanism, publicly available data demonstrates that glyphosate and glyphosate formulations cause DNA damage in human and animal cells as well as in laboratory animals, but so far not in bacterial cells. In two studies, glyphosate formulations also induced DNA damage in the blood cells of exposed humans. In human and other cells, glyphosate and glyphosate formulations have been shown to induce free oxygen radicals that are capable of damaging DNA. The conclusion is that there is indeed evidence of a mechanism by which glyphosate causes cancer.

From all this information, it is reasonable to conclude that, at sufficient levels of exposure, glyphosate and glyphosate formulations are probably carcinogenic to humans.

Christopher Portier is the former Director of the US National Institute of Environmental Health. He lives in Switzerland and wrote the 'Open letter: Review of the Carcinogenicity of Glyphosate by EFSA and BfR' to the European Commission that was signed by 95 scientists from around the world.

No

says Jose Tarazona from the European Food Safety Authority

The European Food Safety Authority (EFSA) has recently reviewed the toxicological profile of glyphosate and proposed new toxicological reference values for risk assessment. The EFSA did not confirm the recent classification of glyphosate as probably carcinogenic by the International Agency for Research on Cancer (IARC).

The IARC considered there to be "limited evidence in humans" for an association between glyphosate and non-Hodgkin lymphoma, while for the EFSA the evidence was insufficient to support such a classification. As the evidence from studies in humans alone had been insufficient for concluding that glyphosate is carcinogenic, the assessment of evidence in laboratory animals was key, and led to the different conclusions of the two bodies.

Significant trends in reports on industry-sponsored studies have been observed by the IARC. The EFSA searched the recent, large database of animal carcinogenicity studies in its entirety, but found no significant differences between control and treatment groups in the studies that were valid. Reviewing the biological relevance of the incidences observed, the EFSA noted that the statistical trends were the result of bias, driven by secondary toxicity at excessively high doses, or chance results not related to glyphosate treatment.

> "The lab results don't show a dose-response".
>
> Jose Tarazona

It is well known that excessive toxicity can affect the carcinogenic responses in bioassays. Such toxicity can cause effects such as cell death with associated regenerative cell proliferation. This can lead to tumour development as a secondary consequence, and is unrelated to the intrinsic potential of the substance to cause tumours at lower, less toxic doses.

The observed incidences were within the historical range observed in untreated animals. The laboratory results did not show a dose-response, and remain unconfirmed by equivalent studies at similar or higher doses. Therefore, besides the absence of statistically significant differences with the concurrent controls, the observed tumour incidences also lacked biological relevance.

The EFSA also concluded that glyphosate is unlikely to cause DNA damage, as has been confirmed by a large number of studies showing no effect. However, effects were reported for glyphosate formulations containing other ingredients, and the EFSA's assessment of a surfactant frequently used in these formulations revealed some concerns. This led the EFSA to recommend carrying out further assessments regarding the possibility of DNA damage being caused by formulated products.

Jose Tarazona is the Head of the Pesticides Unit at the European Food Safety Authority (EFSA) and vice-chairman of the EU Scientific Committee on Health and Environmental Risks.

Arch Toxicol
DOI 10.1007/s00204-017-2009-7

REPLY

Re: Tarazona et al. (2017): Glyphosate toxicity and carcinogenicity: a review of the scientific basis of the European Union assessment and its differences with IARC. doi: 10.1007/s00204-017-1962-5

Christopher J. Portier[1] · Peter Clausing[2]

Received: 18 May 2017 / Accepted: 1 June 2017
© Springer-Verlag GmbH Germany 2017

Recently Tarazona et al. (2017) explained the methodology used by the European Food Safety Authority (EFSA 2015) for the scientific assessment of carcinogenicity as applied to glyphosate. We noted a number of inaccuracies in this paper which may have affected the outcome of the assessment, i.e., glyphosate is considered as non-carcinogenic. These problems relate to the evaluation of individual studies, the objective evaluation of the combined evidence and the weight of evidence approach.

The authors point out that one mouse carcinogenicity study (study N) was "found unreliable after *detailed assessment* due to the occurrence of a viral infection *in all groups including controls*" (EFSA 2015, Table 4, footnote). This decisive statement is in contradiction to the contents of the draft report by the European Chemical Agency (BAuA 2016) concerning the same study. There it is stated that "No information is available on possible abundance of oncogenic viruses in the mouse colonies from which the animals used in the glyphosate studies were obtained" (BAuA 2016, p. 72). According to the draft BAuA report, the "detailed assessment" seems to be based on a remark from the U.S. EPA observer during a teleconference, but as stated in the draft BAuA report "in the study report itself, there was no evidence of health deterioration due to suspected viral infection, and thus, the actual basis of EPA's decision is not known" (BAuA 2016, p. 72). In addition, the incidence of malignant lymphomas in the control group of this study was 20% as compared to an average incidence of 18.4% in the historical control database from five earlier studies (BAuA 2016, p. 67). This negligible difference does not suggest that oncogenic viruses increased the incidence of malignant lymphoma in this particular study.

In comparison, mouse study B, was used in the EFSA evaluation (Tarazona et al. 2017) despite serious concerns regarding the quality of the data with regard to malignant lymphomas. According to regulatory documents (Germany 2015; BAuA 2016), the assessment of malignant lymphomas in this study was "based on histological examination of lymph nodes with macroscopic changes", a wholly unacceptable pathological assessment by OECD guidelines (OECD 2009). However, even this description is wrong since the individual animal data from the study show animals with lymphoreticular neoplasia in the thymus without any lymph node macroscopic changes and animals with lymph node macroscopic changes not examined histopathologically. Finally, it was only "assumed" (BAuA 2016, p. 71) that the "lymphoreticular neoplasia" identified in study A were equivalent to malignant lymphomas without histopathological confirmation. Including study B but excluding study N shows a clear bias by Tarazona et al. against positive findings.

EFSA (2015) also approached the statistical analysis of the data from these studies using a "balancing" between the statistical significance found in trend tests versus the lack of statistical significance in pair-wise comparisons. EFSA (2015) used two-sided tests in the statistical analysis of the tumor incidences while conformity with good scientific practice and OECD guidance 116 (OECD 2012) would have required one-sided tests. If the authorities had followed these recommendations, the significance of the findings would have tilted to "significance" for both pair-wise testing and trend tests. For the only two studies in CD-1 mice exposed for 18 months, the one-sided p values for incidence of malignant lymphomas in control animals

✉ Peter Clausing
pcl@jpberlin.de

[1] Thun, Switzerland

[2] PAN Germany, Nernstweg 32, 22765 Hamburg, Germany

Published online: 15 June 2017

EXHIBIT
15-25
9/5/17

Table 1 Additional tumors with significant ($p < 0.05$) trends (exact Cochran-Armitage linear trend test in proportions) in the carcinogenicity studies not cited by Tarazona et al. (2017)

Study species	Tumor type Sex; incidences	p value (one-sided)
C Mouse	Hemangioma Females; 0/50, 0/50, 2/50, 5/50*	0.002
D Mouse	Lung adenocarcinomas Males; 5/51, 5/51, 7/51, 11/51	0.028
G Rat	Thyroid follicular cell adenomas and carcinomas Males; 0/50, 0/50, 0/50, 2/50, 2/49	0.034
E Rat	Thyroid c-cell carcinomas Females; 1/47, 0/49, 2/50, 6/47	0.003
J Rat	Kidney adenoma Male; 0/50, 0/50, 0/50, 4/50	0.004
K Rat	Hepatocellular adenoma Males; 0/52, 2/52, 0/52, 5/52*	0.008
L Rat	Skin keratoacanthoma Males; 2/51, 3/51, 0/51, 6/51	0.030
L Rat	Mammary gland adenomas and adenocarcinomas Males; 2/51, 3/51, 1/51, 8/51*	0.007

* These groups have a significantly increased ($p<0.05$) incidence of tumors relative to the controls by the Fisher Exact Test in addition to a significantly positive trend test finding

versus high dose animals are 0.134 and 0.028 for studies C and D (Fisher's exact test), respectively, while one-sided exact p values for the Cochran–Armitage trend tests are 0.02 and 0.008 for studies C and D, respectively.

As can be derived from Table 31 of the draft CLH report, studies C, D and N show dose–response relationships for malignant lymphomas (BAuA 2016).

Tarazona et al. (2017) give five reasons for dismissing all of the positive findings in all of the studies. First, they claim to have "balanced" the positive trend test findings against lack of significance in pairwise tests. The guidelines they cite clearly state that "Significance in either kind of test is sufficient to reject the hypothesis that chance accounts for the result" (OECD 2012). While appropriate to require the determination of the statistical methods in the study plan (OECD 2009), this should not prevent regulatory authorities from applying a more comprehensive statistical analysis. Thus, this "balancing" appears to violate their guidance. Second, they cite a lack of consistency in multiple animal studies. However, to arrive at this conclusion, they group together studies with different strains, different durations, different pathology (see comment about Study B above) done at different times in different labs. Table 5 is an excellent example of an incorrect assessment comparing across mouse studies with different durations, with different substrains and markedly different pathology. Third, they argue that they are seeing effects only at or above the MTD. No "excessive toxicity" was found in any of the mouse or rat carcinogenicity studies (Germany 2015). The reduced body weight in high dose animals seen in some of these studies was associated with an even higher reduction in food consumption—not surprising at dietary concentrations of 30,000 ppm of glyphosate or higher. No evidence for excessive toxicity in any of the studies was provided by Tarazona et al. (2017). In addition, this "limit dose" of 1000 mg/kg does not apply to some studies with positive findings. For example, the high dose in study D was 810 mg/kg. Study D had significant increases in malignant lymphomas and lung adenocarcinomas in males. Fourth, they claim there is a lack of preneoplastic lesions, yet, for example, there was a significant increase in bilateral chronic interstitial nephritis ($p = 0.008$, exact trend test) in Study A which also showed kidney tumors. In addition, it is not clear what preneoplastic lesions they would be looking for when dealing with malignant lymphomas or hemangiosarcomas. Finally, they exclude positive findings falling within the range of the historical controls. Despite formal statistical methods for using historical control data appropriately in an evaluation (Fung et al. 1996; Greim et al. 2003; Haseman 1984; Peddada et al. 2007), EFSA (Germany 2015), EChA (BUaU 2016) and many other regulatory agencies (e.g., EPA, 2016) continue to use this inappropriate rule when evaluating these types of data. Formal statistical analysis of the mouse kidney, malignant lymphoma and hemangiosarcoma findings using historical controls results in two marginal findings becoming significant and did not reverse any of the positive findings relative to the concurrent controls.

Finally, a considerable number of tumors were missed in the evaluation by both EFSA and EChA (Table 1). Our calculations were based on data in the supplemental information provided by Greim et al. (2015). These significant increases were not mentioned by either EFSA (Germany

2015) or ECHA (BAuA 2016). In addition, Table 6 of Tarazona et al. (2017) failed to mention thyroid c-cell carcinoma in female rats (Study F), also seen in Study E, and their Table 4 included a non-chronic study (I).

In conclusion, the weight of evidence discussed by Tarazona et al. (2017) needs to be re-assessed taking into account the five key issues raised above. They clearly need to be more precise in their evaluation of the evidence and more cautious in their indiscriminate exclusion of positive findings.

References

BAuA (2016) Proposal for harmonised classification and labelling. Based on Regulation (EC) No 1272/2008 (CLP Regulation), Annex VI, Part 2. Substance Name: N-(phosphonomethyl)glycine; Glyphosate (ISO). European Chemicals Agency http://echa.europa.eu/documents/10162/13626/clh_report_glyphosate_en.pdf

European Food Safety Authority (2015) Conclusion on the peer review of the pesticide risk assessment of the active substance glyphosate. EFSA J 13:4302–4302

Fung KY, Krewski D, Smythe RT (1996) A comparison of tests for trend with historical controls in carcinogen bioassay. Can J Stat 24:431–454

Germany (2015) Final addendum to the renewal assessment report on glyphosate, compiled by EFSA

Greim H, Gelbke H-P, Reuter U, Thielmann HW, Edler L (2003) Evaluation of historical control data in carcinogenicity studies. Hum Exp Toxicol 22:541–549

Greim H, Saltmiras D, Mostert V, Strupp C (2015) Evaluation of carcinogenic potential of the herbicide glyphosate, drawing on tumor incidence data from fourteen chronic/carcinogenicity rodent studies. Crit Rev Toxicol 45:185–208

Haseman JK (1984) Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Environ Health Perspect 58:385–392

OECD (2009) Carcinogenicity studies. OECD Guideline No. 451. http://www.oecd-ilibrary.org/environment/test-no-451-carcinogenicity-studies_9789264071186-en;jsessionid=17ikkmub37voi.x-oecd-live-02

OECD (2012) Guidance document 116 on the conduct and design of chronic toxicity and carcinogenicity studies, supporting test guidelines 451, 452 and 453, 2nd edn. Series on testing and assessment No 116. Env/jm/mono(2011)47

Peddada SD, Dinse GE, Kissling GE (2007) Incorporating historical control data when comparing tumor incidence rates. J Am Stat Assoc 102:1212–1220

Tarazona JV, Court-Marques D, Tiramani M, Reich H, Pfeil R, Istace F, Crivellente F (2017) Glyphosate toxicity and carcinogenicity: a review of the scientific basis of the European Union assessment and its differences with IARC. Toxicol, Arch. doi:10.1007/s00204-017-1962-5

The glyphosate saga: an example of influence of unsound science and interest groups in public health decision making

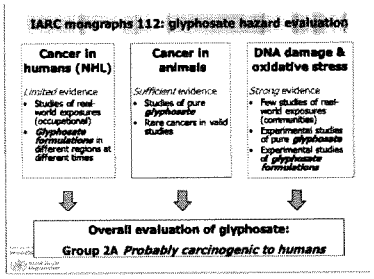
Xaver Baur, Daniele Mandrioli, Lygia Therese Budnik, Ellen Silbergeld, Philip J. Landrigan, Christopher J. Portier

Xaver Baur, Charité University Medicine, Berlin, Germany; EOM Society, Berlin, Germany, ***Ramazzini Fellow***
Daniele Mandrioli, Cesare Maltoni Cancer Research Center of the Ramazzini Institute, Bologna, Italy, ***Ramazzini Fellow***
Lygia Therese Budnik, University of Hamburg, Hamburg, Germany, EOM Society, Berlin, Germany, ***Ramazzini Fellow***
Ellen Silbergeld, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA, ***Ramazzini Fellow***
Philip J. Landrigan, Mount Sinai School of Medicine, New York City, USA, ***Ramazzini Fellow***
Christopher J. Portier, Maastricht University, the Netherlands & Environmental Health Consultant, Switzerland

Background

The International Agency for Research on Cancer (IARC) evaluated the carcinogenic hazard of the herbicide glyphosate. It concluded that the data for glyphosate meet the criteria for classification as a probable human carcinogen (group 2A). On the other hand, the European Food Safety Authority (EFSA) and German Federal Institute for Risk Assessment (BfR) came to the conclusion that glyphosate carcinogenicity is unlikely(1). FAO/WHO did not challenge the IARC findings, but concluded that cancer risk was unlikely at typical human ingestion levels. Ongoing evaluations of glyphosate carcinogenicity by the US Environmental Protection Agency (US EPA) and the European Chemical Agency (ECHA) have both been published as drafts still under review. We summarized the methods used by different public health decision makers and their limits in terms of science, transparency and conflicts of interests.

Figure 1: Evaluation of glyphosate by IARC: evidence of its carcinogenicity in humans, animals and mechanistic *models (Fig: courtesy: Kate Z. Guyton).*



Public Health Decision Making Processes

IARC: IARC applied a transparent and rigorous scientific process by 17 publicly identified independent experts, and a highly professional standardized evaluation of all publically available studies(2). The IARC Working Group found an association between NHL and glyphosate based on the available human evidence and significant carcinogenic effects in laboratory animals for rare kidney tumours and hemangiosarcoma in two mouse studies and benign tumours in two rat studies. The IARC WG summarized that there was strong evidence of genotoxicity and oxidative stress for glyphosate, entirely from publicly available research, including findings of DNA damage in the peripheral blood of exposed humans. The IARC WG concluded that there is sufficient evidence of carcinogenicity in animals, limited evidence of carcinogenicity in humans and strong evidence for two carcinogenic mechanisms.

EFSA/BfR: The EFSA/BfR evaluation of glyphosate carcinogenicity, based on the Renewal Assessment Report (RAR) for glyphosate prepared by the Rapporteur Member State (BfR), concludes that: "glyphosate is unlikely to pose a carcinogenic hazard to humans"(3). No authors or contributors are listed for either document by BfR and EFSA(4). The use of confidential data submitted to the BfR makes it impossible for any scientist not associated with BfR or EFSA to review this conclusion. Three known experts from the chemical industry are members of the Pesticide Committee of the BfR(5).

EFSA classified the human evidence as 'very limited' and then dismissed any association of glyphosate with cancer without clear explanation or justification. Ignoring established guidelines by OECD and ECHA, cited in their report, EFSA dismissed evidence of renal tumours in three mouse studies, hemangiosarcoma in two mouse studies and malignant lymphoma in two mouse studies. Thus, EFSA incorrectly discarded all findings of glyphosate-induced cancer in animals as chance occurrences. EFSA ignored important laboratory and human mechanistic evidence.

FAO/WHO: The joint FAO/WHO meeting of pesticide residues concluded that "glyphosate is unlikely to pose a carcinogenic risk to humans from exposure through the diet". This is both a risk and a hazard statement; no independent hazard statement was prepared. A transparent disclosure of the criteria used for assessing carcinogenic risk is missing. No reference was provided in the document. No list of authors and disclosure of conflicts of interests was provided in the document. Nevertheless, different members, including the chairman, of the joint FAO/WHO meeting of pesticide residues had been reported to have financial conflicts of interests with industry(6).

ECHA: Glyphosate classification is also currently under review by ECHA. A draft of the report by ECHA has been published for public comments(7). It concludes that no hazard classification for its carcinogenicity is warranted. But males in all five mouse carcinogenicity studies considered by ECHA to be of acceptable quality show a statistically significant increase in the incidence of one or several tumor types. Importantly, the finding of an increased incidence of malignant lymphoma in animals is further supported by the results of epidemiological studies indicating an association between glyphosate exposure and Non-Hodgkin lymphoma. This clearly exceeds the criteria for classification as a carcinogen as given in CLP Regulation, documented on page 95 of the Dossier(8).

US EPA: Glyphosate is currently undergoing a registration review by US EPA. The recently proposed classification by US EPA is that glyphosate is not likely to be carcinogenic at doses relevant for human health risk assessment(9). This is again a risk and hazard statement with no independent hazard statement provided. However, this is based on some speculation, i.e. that the more positive epidemiological studies should have had lower relative risks than other studies; further, it is assumed that there were previous exposures in the greatly weighted Agricultural Health Study which has a rather short follow-up time. US EPA's interpretation that "the association between glyphosate exposure and the risk of NHL cannot be determined based on the available data" does not correctly characterize the human data presented. Findings on multiple myeloma that were included in the IARC evaluation were not considered adequately.

The evaluation of the animal carcinogenicity data and mechanistic data for glyphosate missed important findings, but basically follows along the same lines as the EFSA/BFR review, suggesting a lack of independent review.

Conclusion

The glyphosate issue is just one example of inappropriate corporate influences on public health regulations by use of unsound scientific reviews. These economically motivated activities leave a resulting health burden on society. We call for increased sensitivity, full transparency and the implementation of effective rules governing decision making bodies(10).
We urge the Collegium Ramazzini to again support an IARC evaluation of carcinogenicity(11), since the most appropriate and scientifically based evaluation of the cancers reported in humans and laboratory animals as well as supportive mechanistic data is that glyphosate is a probable human carcinogen(1). We suggest that common commercial formulations of GBHs should be prioritized for inclusion in government-led toxicology testing programs such as the U.S. National Toxicology Program, as well as for biomonitoring as conducted by the U.S. Centers for Disease Control and Prevention(12).

1 - http://jech.bmj.com/content/early/2016/03/03/jech-2015-207005.full
2 - http://monographs.iarc.fr/ENG/Monographs/vol112/mono112-09.pdf
3 - European Food Safety Authority. Conclusion on the peer review of the pesticide risk assessment of the active substance glyphosate. EFSA J 2015;13:4302)
4 - http://registerofquestions.efsa.europa.eu/roqFrontend/outputLoader?output=ON-4302
5 - http://corporateeurope.org/food-and-agriculture/2015/04/glyphosate-saga-independent-scientific-advice-according-germany-uk
6 - https://www.theguardian.com/environment/2016/may/17/ unwho-panel-in-conflict-of-interest-row-over-glyphosates-cancer-risk
7 - https://echa.europa.eu/documents/10162/13626/clh_report_glyphosate_en.pdf
8 - http://www.pan-germany.org/download/ PAN_Germany_Comment_on_CLH-Report_regarding_Carcinogenicity_1607.pdf
9 - https://www.epa.gov/pesticides/scientific-advisory-panel-meet-cancer-potential-glyphosate
10 - http://ehp.niehs.nih.gov/1509880/
11 - http://www.collegiumramazzini.org/download/12-SeventeenthCRStatement(2015).pdf

EXHIBIT
PENGAD 800-631-6989
15-26
9/5/17

CURRICULUM VITAE
Christopher J. Portier, Ph.D.

Personal Data: Birth Date [redacted]
Birthplace [redacted]

Address: [redacted]

Education:

1981	Ph.D.	(Biostatistics), University of North Carolina, Chapel Hill
1979	M.S.	(Biostatistics), University of North Carolina, Chapel Hill
1977	B.S.	(Mathematics), summa cum laude, Nicholls State University

Employment:

2013-present	**Consultant** to various governmental agencies (multiple countries)
2013-2014	**Senior Visiting Scientist**, International Agency for Research on Cancer, Lyon, France
2013-present	**Senior Contributing Scientist**, Environmental Defense Fund, New York City, NY
2010-2013	**Director**, National Center for Environment Health, Centers for Disease Control and Prevention, Atlanta, GA
2010-2013	**Director**, Agency for Toxic Substances and Disease Registry, Atlanta, GA
2009 – 2010	**Senior Advisor to the Director**, National Institute of Environmental Health Sciences and National Toxicology Program, Research Triangle Park, North Carolina.
2009 – 2010	**Visiting Scientist**, National Research Centre for Environmental Toxicology (EnTox), Queensland, Australia
2006 - 2009	**Associate Director**, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina.
2006 - 2009	**Director, Office of Risk Assessment Research**, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina.
1993 – 2010	**Head, Environmental Systems Biology** (originally Stochastic Modeling), Laboratory of Molecular Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina.
2000 - 2006	**Associate Director, National Toxicology Program**, National Institute of Environmental Health Sciences, Division of Intramural Research, Research Triangle Park, North Carolina.
2000 - 2006	**Director, Environmental Toxicology Program**, National Institute of Environmental Health Sciences, Division of Intramural Research, Research Triangle Park, North Carolina.
2006-2007	**Scientific Advisor to the Director**, Public Health and the Environment Department, World Health Organization, Geneva, Switzerland (detail from NIEHS – four months)
1993 - 2005	**Chief, Laboratory of Computational Biology and Risk Analysis** (originally the Laboratory of Quantitative and Computational Biology), National Institute of Environmental Health Sciences, Division of Intramural Research, Research Triangle Park, North Carolina.

EXHIBIT
15-27
9/5/17 v
PENGAD 800-631-6989

1996 - 2000	**Associate Director for Risk Assessment**, Environmental Toxicology Program National Institute of Environmental Health Sciences, Division of Intramural Research, Research Triangle Park, North Carolina.
1990 - 1993	**Head, Risk Methodology Section**, National Institute of Environmental Health Sciences, Division of Biometry and Risk Assessment, Research Triangle Park, North Carolina.
1987, 1992, 1990	**Guest Scientist**, German Cancer Research Center, Heidelberg, Germany.
1978 - 1990	**Mathematical Statistician**, National Institute of Environmental Health Sciences, Division of Biometry and Risk Assessment, Research Triangle Park, North Carolina.
1977	**Mathematician**, Computer Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
1976	**Undergraduate Research Trainee**, Neutron Physics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

University Affiliations:

2014 – present	Visiting Professor, Department of Toxicogenomics, Maastricht University, The Netherlands
2013 – present	Honorary Professor, National Research Centre for Environmental Toxicology, University of Queensland, Brisbane, Australia
2011 – present	Adjunct Professor, Department of Environmental Health, Emory University, Atlanta, GA, USA
2009 – 2010	Visiting Professor, University of Queensland, Brisbane, Australia
1986 - 2007	Adjunct Professor of Biostatistics, University of North Carolina, School of Public Health, Chapel Hill, North Carolina.
1990-1992	Adjunct Professor of Statistics, University of Waterloo, Waterloo, Ontario, Canada

Honors & Awards:

- 2013 President's Dream Green Team Award for "A Human Health Perspective on Climate Change"
- Fellow, World Innovation Foundation, 2006
- Society of Toxicology, Risk Assessment Specialty Section, Paper of the Year, 2006
- Society of Toxicology, Risk Assessment Specialty Section, Paper of the Year, 2005
- Outstanding Risk Practitioner Award, International Society for Risk Analysis, 2000.
- Elected Fellow, International Statistical Institute, 2000.
- Outstanding Performance Award, National Institute of Environmental Health Sciences, numerous dates.
- Commendation for Sustained High Quality Work Performance, National Institute of Environmental Health Sciences, numerous dates.
- Merit Award, National Institute of Health, 1998.
- Board of Publications, Best Paper Award, Society of Toxicology, 1995.
- Distinguished Achievement Award, Section on Statistics and the Environment, American Statistical Association, 1995.
- Spiegelman Award presented by the American Public Health Association to the most outstanding public health statistician under the age of 40, 1995.
- Best-applied statistics paper, Centers for Disease Control, 1993.
- Elected Fellow, American Statistical Association, 1992.

- Elected Foreign Correspondent, Russian National Academy of Natural Sciences, 1992.
- First recipient of the James E. Grizzle Distinguished Alumnus Award, The Department of Biostatistics, The University of North Carolina, 1991.

Professional Societies Membership:

Society of Toxicology, American Public Health Association, International Statistics Institute, Society for Risk Analysis, EcoHealth Society, American Association for Cancer Research, American Association for the Advancement of Science

Editorial Activities:

- Editor in Chief - The Open Environmental Journal (2008 to 2010)
- Associate Editor – Frontiers in Predictive Toxicity (2010 to present)
- Associate Editor - Environmental Health Perspectives (1987-2006)
- Associate Editor - Risk Analysis: An International Journal (1989-2003)
- Editorial Board – Environmental and Ecological Statistics (2004-2007)
- Associate Editor – Statistics in Medicine (1998-2002)
- Associate Editor - Biometrics (1997-99)
- Editorial Board Member/Reviewer (different dates): Biometrika, Cancer Research, Communications in Statistics, Fundamental and Applied Toxicology, Journal of Applied Toxicology, Journal of the American Statistical Association, Journal of Toxicology and Environmental Health, Science, Mathematical Biosciences, Journal of Mathematical Biology, Carcinogenesis, Science, PNAS, Toxicological Sciences

Advisory & Review Committees:

2015 – 2016	Member, Committee to Review the Draft Interagency Report on the Impacts of Climate Change on Human Health in the United States, National Research Council, National Academy of Sciences, USA
2010 – present	Member, Science Advisory Group on Electromagnetic Fields and Health, Netherlands Organisation for Health Research and Development
2009 – 2010	Coordinating Lead Author, Interagency Working Group on Climate Change and Health
2009 – present	Member, Institute of Medicine Roundtable on Environmental Health Sciences Research and Medicine
2009 – 2012	Member, National Academies of Science Roundtable on Science and Technology for Sustainability
2009	Member, WHO Advisory group on the health implications of the use of DDT to reduce risks of malaria.
2005 – 2010	Chair, Subcommittee on Toxics and Risk, President's National Council on Science and Technology
1997 - 2012	Advisor, *World Health Organization*, International Program on Chemical Safety, EMF Project.
2008 – 2010	Member, Environmental Protection Agency, Science Advisory Board
2007 – 2010	Member, International Life Sciences Institute, Health and Environmental Sciences Institute, Subcommittee on Susceptible Populations
2008	Center Review Committee, Canadian National Science and Engineering Research Council Chair in Risk Assessment

2008	Chair, International Agency for Research on Cancer Monographs Advisory Group, Lyon, France
2008	Advisory Group, Center for Environmental Oncology, University of Pittsburgh Cancer Institute
2007.	Chair, WHO Workshop on Low Cost Options for Reducing Exposures to ELF-EMF, Geneva
2007.	Invited Participant, International Program on Chemical Safety Workshop on Aggregate and Cumulative Risk Assessment, Washington, DC.
2006	Rapporteur, International Agency for Research on Cancer, Scientific Advisory Group to Plan Volume 100 of the IARC Monograph Series
2005	Chair, International Agency for Research on Cancer, Scientific Advisory Board on the Preamble to the Cancer Monograph Series
2005	Chair, World Health Organization Expert Panel on Health Criteria Document for Extremely Low Frequency Electric and Magnetic Fields
2003 – 2005	Co-Chair, Subcommittee on Health and Environment, President's National Council on Science and Technology
2003	Ad-Hoc member, EPA Science Advisory Board, Review of Children's Cancer Risk Assessment Supplement to Cancer Guidelines
2002 – 2006	Co-Chair, Subcommittee on Mercury, President's National Council on Science and Technology
2000 – 2007	Member, Finish Academy of Sciences Centers of Excellence Program Science Advisory Committee
2000	Reviewer, *Congressional Research Service, Library of Congress*; Research needs relevant to children's environmental health risks.
1998 - 2004	Member and Chair, *Environmental Protection Agency*, FIFRA Science Advisory Panel.
1997 - 2006	Member, National Occupational Research Agenda Team, *National Institute of Occupational Safety and Health.*
1995 - 2000	Advisor, *Australian Health Council,* Risk Assessment Methodology, Member *NHMRC* Steering Committee on Cancer Risk Assessment Guidelines.
1992 - 2000	Member, *EPA* Dioxin Reassessment Working Group.
1985 - 2007	Thesis director for graduate students, Department of Biostatistics, *University of North Carolina - Chapel Hill, North Carolina.*
1997	Advisor, *Netherlands National Health Council,* Risk Assessment Methodology.
1997	Reviewer, *Air Force Office of Scientific Research.*
1996 - 1997	Temporary Advisor, *World Health Organization,* Expert Committee on Food Additives.
1996	Advisor, *Environmental Protection Agency;* Evaluation of the benchmark dose methodology.
1996	Advisor, *Environmental Protection Agency;* Evaluation of risks from exposure to PCBs.
1996	Expert Review Committee, *Environmental Protection Agency;* Cancer dose-response for PCB's.
1995 - 1996	Member, *California Environmental Protection Agency,* Risk Assessment Advisory Committee.

1994 - 1997	Science Advisory Panel, *Public Broadcasting System Production* "Poisons in the Womb".
1991 - 1995	Ad-Hoc Member, *Environmental Protection Agency*, Science Advisory Panel.

Legislative Hearings:

- Glyphosate Carcinogenicity, European Parliament, Brussels, December 2015
- Glyphosate Carcinogenicity, German Parliament, Berlin, July 2015
- Lead and Children's Health, Senate Committee on Environment and Public Works, July, 2012
- Asthma and Children's Health, Senate Committee on Environment and Public Works, May, 2012
- Contaminated Drywall, Senate Committee on Commerce, Science and Transportation, December, 2012.
- Camp Lejeune Contaminated Drinking Water, House Committee on Science and Technology, September, 2010.
- Autism and Vaccines, House Committee on Government Reform, December, 2002.

US Government Service Activities:

- Member, President's Task Force on Environmental Justice 2010-2013
- Member, President's Task Force on Children's Environmental Health 2009-2013
- Member, National Toxicology Program Executive Committee 2010-2013
- Financial Support and International Press Conference for research on "The Health Benefits of Tackling Climate Change" appearing as a series in Lancet, November 25, 2009
- Organizing Committee, White House Stakeholder briefing on Climate Change and Human Health, Old Executive Office Building, November 2009.
- Member, US Delegation, World Climate Congress, Geneva (September 2009)
- Member, US Delegation, Global Risk Communication Dialogue (2008-2009)
- Member, NIEHS Corrective Action Plan Management Committee (2008-2009)
- Primary focus, all interagency activities on hazards and risk (2006 to present)
- Co-Organizer, NIEHS/EPA Workshop on Children's Environmental Health, RTP, NC, January, (2007)
- Co-Organizer, NIEHS/NTP Workshop on the Identification of Targets for the HTS Roadmap Project (2007)
- Coordinator, NIEHS/EPA Review of the Children's Environmental Health Centers Program (2006-2007)
- Organizing Committee, Global Environmental Health Initiative, NIEHS (2006 to 2009)
- NIEHS Leadership Council (2005 to 2009)
- Organizer, formal collaborative agreements between NTP and Ramazzini Foundation (2001 to 2006)
- Organizer, formal collaborative agreements between NTP and Korean NTP (2002 to 2006)
- NIEHS Title 42 Review Committee (2003 to 2004)
- NIEHS Executive Committee and Operations Update Committee (2000 to 2005)
- NIEHS Leadership Retreats, DERT Retreats, DIR Retreats (all years since 1997)
- Presenter, NIEHS-sponsored National Academy of Sciences Committee on Emerging Issues in Environmental Health, November, 2001
- Organizer and presenter, National Toxicology Program Executive Committee Meetings (multiple dates since 2000)
- Organizer and presenter, National Toxicology Program Board of Scientific Counselors (multiple dates since 1998)

- Organizer, Joint NIEHS/US Geological Survey Interagency Program on Exposure Assessment, April 2001 to present)
- Organizer, US-Vietnam Scientific Conference on the Health and Environmental Effects of Agent Orange/Dioxin in Vietnam, March, 2002
- Organizing Committee, National Toxicology Program/EPA/FDA Scientific Conference on the Allergenicity of Genetically Modified Food, November, 2001
- NIEHS Town Hall Meeting, Los Angeles California, November, 2001
- NTP Research Directions, NAEHSC, Research Triangle Park, NC. May, 2001.
- NCI Study Section Center Presite Meeting, Seattle, Washington, January, 2001.
- Program committee member, *NIEHS/Colorado State University* conference on the Application of Technology to Chemical Mixture Research, 2001.
- Coordinating Core Committee, National Center for Toxicogenomics, NIEHS, 2000 to present
- Organizer, Joint US-Vietnam Consultation on Research on Agent Orange Health Effects in Vietnam. Singapore, 2000
- *ICCVAM/NICEATM*, Up-and-Down Procedure Peer Review Meeting, 2000.
- Chairman, *NIEHS* Risk Assessment Research Committee, 1995-present.
- Discussant, *NIEHS/PNNL* Workshop on Human Biology Models for Environmental Health Effects, 2000.
- Risk Assessment Coordinator, *NIEHS* US *RAPID* Program for the Evaluation of Health Risks from Exposure to Electric and Magnetic Fields, 1996-99.
- Organizer and Chair, Four Public Comment Sessions on the report of the *NIEHS/DOE* Working Group on the Health Effects of Exposure to Electric and Magnetic Fields, 1998.
- Organizer and Co-Chair, *NIEHS/DOE* Working Group on the Health Effects of Exposure to Electric and Magnetic Fields, 1998.
- Scientific Organizing Committee, *NIEHS* Workshop on Risk Assessment Issues Associated with Endocrine Disrupting Chemicals, 1998.
- Organizer, *NIEHS/DOE* Science Research Symposium on the Health Effects of Exposure to Electric and Magnetic Fields I: Biophysical Mechanisms and *In Vitro* Experimentation, 1998.
- Organizer, *NIEHS/DOE* Science Research Symposium on the Health Effects of Exposure to Electric and Magnetic Fields II: Epidemiological Findings, 1998.
- Organizer, *NIEHS/DOE* Science Research Symposium on the Health Effects of Exposure to Electric and Magnetic Fields III: *In Vitro* and Clinical Research Findings, 1998.
- Head, Toxicokinetics Faculty, *NIEHS*, 1994-97.
- Coordinator/Director, *NIEHS/ATSDR* Interagency Course on Mechanistic Modeling in Environmental Risk Assessment, 1996.
- Organizer, *NIEHS/EPA* Workshop on Research Priorities for New Risk Assessment Guidelines, 1996.
- Co-Organizer, *National Institute of Statistical Sciences*, *NIEHS/EPA* Workshop on Mechanistic Modeling in Risk Assessment, 1995.
- Scientific Coordinator and Mission Director, *NIEHS* "Mission to Vietnam" to assess the potential for scientific collaboration on the impact of Agent Orange on the Vietnamese Population, 1995.
- Chairman, *NIEHS* Computer Science Focus Group, 1995.
- Discussant, National Toxicology Program Workshop on Mechanistic Modeling in Toxicology, *NIEHS*, 1995.
- Discussant, National Toxicology Program Workshop on Mechanisms of Carcinogenesis, *NIEHS*, 1995.
- Co-Organizer, International Conference on The Role of Cell Proliferation in Carcinogenesis, co-

sponsored by *NIEHS*, *The Chemical Industry Institute of Toxicology*, *The International Life Sciences Institute* and *The American Industrial Health Council*, 1992.

- Organizer and Director, Scientific Basis of Animal Carcinogenicity Testing, Moscow, Russia, co-sponsored by the *International Agency for Research on Cancer*, *NIEHS*, *Health and Welfare Canada* and *The All-Union Cancer Research Center*, 1991.
- Chairman, Computer Technology Advisory Forum, *NIEHS*, 1989.
- Organizer and Director, Design and Analysis of Long-Term Animal Carcinogenicity Experiments, Lyon, France, co-sponsored by the *International Agency for Research on Cancer* and the *NIEHS*, 1988.

Extramural Activities:

- Member, NRC Committee to review the Draft Interagency Report on the Impacts of Climate Change on Human Health in the United States, Washington, DC, 2015
- Expert Scientist, International Agency for Research on Cancer Monograph Meeting on Some Organophosphate Pesticides and Herbicides, Lyon, France, March, 2015
- Overall Chair, International Agency for Research on Cancer Monograph Meeting on Diesel and Gasoline Engine Exhausts and related compounds, Lyon, France, June, 2012
- Advisor to Wellcome Trust at "International Research Futures Symposium on Global Change, Economic Sustainability, and Human Health", London, England, March, 2012.
- Expert Panel Member for review of Hollings Marine Laboratory, National Oceanographic and Atmospheric Agency, Charleston, USA, February, 2012.
- Chair, Mechanism Subgroup, International Agency for Research on Cancer Monograph Meeting on Radiofrequency Electric and Magnetic Fields, Lyon, France, May, 2011
- Advisor, Greek Ministry Health, Working group on hexavalent chromium in the environment, January, 2011
- Member, WHO Consultation on Human Health Risks from DDT, Geneva, Switzerland, November, 2010
- Associate Editor, Frontiers in Predictive Toxicity, 2010 – 2011
- Scientific Advisor, Health Investigation Levels Workshop, Canberra, Australia, January, 2010
- Chair, IARC Working Group, IARC Monograph 100-G, Lyon, France, October, 2009
- Scientific Organizing Committee, VII World Congress on Alternatives and Animal Use in Life Sciences, Rome, Italy, September, 2009
- Chair, Research Directions Working Group, World Health Organization Consultation on Global Research on Climate Change and Health, October, 2008.
- Editor-in-Chief, The Open Environment Journal, May 2008-August, 2010
- Member, EPA Science Advisory Board, July, 2008-present
- Working Group Member, IARC Monograph 98 - Fire-fighting, Painting and Shift-work, Lyon, France, November, 2007
- Chair, WHO Extremely Low Frequency Magnetic and Electric Fields Workshop on Intervention Strategies, June, 2007
- Special Advisor to the Director, Program on Public Health and the Environment, WHO, Geneva, May-July, 2007
- Member, International Life Sciences Institute Working Group on Susceptible Populations, March, 2007 – present
- Special Advisor to the Director, Program on Public Health and the Environment, WHO, Geneva, November, 2006-January, 2007

- Breakout Group Chair, International Workshop on Uncertainty and Variability in PBPK Modeling, RTP, NC USA, October, 2006
- Member, Health Effects Sciences Institute Committee on Sensitive Subpopulations and Groups, Washington, DC, 2006 to present
- Rapporteur, Steering Committee for developing the 100th Monograph of the International Agency for Research on Cancer, Lyon, France, September, 2006
- Co-Organizer, parallel workshops on the advancement of PBPK modeling in risk assessment, Research Triangle Park, November, 2006, Corfu, Greece, April, 2007.
- Organizer, Alternative Models in Developmental Neurotoxicity, Alexandria, Virginia, March, 2006.
- Organizer, NTP High Throughput Screening Workshop, Washington, DC, December, 2005
- Organizer, ISRTP Meeting on Alternative Methods in Toxicology, Baltimore, Maryland, November, 2005
- Organizer, NTP 25th Anniversary Meeting, Washington, DC, May, 2005
- Organizer, IPCS/WHO Workgroup on Dose-Response Modeling, Geneva, Switzerland, September, 2004
- Organizer, Consultation on harmonization of toxicological research between the NTP, Ramazzini Foundation and the European Union, European Congress of Toxicology, Florence, Italy, September, 2003.
- Member, WHO Workgroup on the epidemiology of cellular phone toxicity, Tskuuba, Japan, September, 2003.
- Program Committee, 12th International Conference on Global Warming, Boston, Massachusetts, May 2003
- Program Committee, International Conference on Cancer Risk Assessment, Athens, Greece, August, 2003
- Chair, WHO Public Consultation on Risk Communication, Luxembourg, February, 2003.
- Chair, WHO Committee on Establishing a Plan for Implementation of the Precautionary Principle in Risk Management. Luxembourg, February, 2003.
- Presenter (on behalf of US Government), National Academy of Sciences Panel on the Use of Third Party Toxicity Research with Human Research Participants, December, 2002
- Member, US Science Delegation, United Nations Environmental Program Consultation on Organic Mercury, September, 2002
- Science Panel Member, IARC Carcinogenicity Review of ELF-EMF, Lyon, France, June, 2001.
- Reviewer, Finish Ministry of Health Centers of Excellence Program, Helsinki, April, 2001.
- *EPA* dioxin reassessment peer review workshop and public comment session, Washington, DC, 2000.
- Organizer: Dioxin Dose-Response Working Group Meeting, Fort Collins, Colorado, February, 2000.
- Chair, Spiegelman Award Committee, *American Public Health Association*, 1998.
- Chair, *Bioelectromagnetics Society* Symposium on the use of Transgenic Animals in Evaluating Health Risks from Exposure to Cellular Phones, St. Petersburg, Florida, 1998.
- Member, *World Health Organization* International Program on Chemical Safety, Workshop on Issues in Cancer Risk Assessment, 1998.
- Advisor, *Joint Committee on Food Additives, World Health Organization/Food and Agriculture Organization*. Evaluation of certain food additives and contaminants
- Member, US Government Methylene Chloride Risk Characterization Science Committee, 1996-1998.
- Scientific Organizing Committee, *Colorado State University* Workshop on Biomedical Advances on Chemical Mixtures, 1997.
- *National Academy of Sciences*, Institute of Medicine, Committee on Funding Future Agent Orange

Research in Vietnam, 1996.

- Discussant, Workshop on the role of Endocrine Disruptors in Human Health, 1995.
- Advisor to *Australian Health Council* on Risk Assessment Methodology, Member *NHMRC* Steering Committee on Cancer Risk Assessment Guidelines
- Participant, International Program on Chemical Safety of the *World Health Organization* Workshop on Chemical Risk Assessment, London, England, 1995.
- Participant, *IARC* Workshop on Receptor-Mediated Carcinogenesis, Lyon, France, 1994.
- Co-Organizer, Symposium on Quantitative Risk Assessment, *German Cancer Research Center*, Heidelberg, Germany, 1993.
- Participant, *IARC* Monograph on Risk Assessment Methodology, *International Agency for Research on Cancer*, Lyon, France, 1993.
- Thesis advisor for graduate student, *University of Waterloo*, Waterloo, Ontario, Canada. 1991-93.
- Co-Organizer, *Russian Academy of Sciences* Informatics and Cybernetics Research Award, 1992.
- Official Observer, IARC Monograph on the Biological Effects of Ultraviolet Radiation, *International Agency for Research on Cancer*, Lyon, France, 1992.
- Member, *International Life Sciences Institute*, Dose-Response Working Group, 1991.
- Participant in Banbury Conference on Human Health Risks from Exposures to Dioxins, Banbury Conference Center, Cold Spring Harbor, New York, 1990.
- Co-Chairman, Session on Biostatistical Developments in Cancer Research, *15th International Cancer Congress*, Hamburg, Germany, 1990.
- Participant in *Environmental Protection Agency* Workshop on Risk Assessment Guidelines, Virginia Beach, Virginia, 1989.

Invited Presentations (present-1999)

"Glyphosate Carcinogenicity", Swiss Society of Toxicology, Basel, November, 2016

"Glyphosate Carcinogenicity", Concerned Scientists of Switzerland Annual Meeting, Zurich, December, 2015

"Should the precautionary principle be invoked for RF-EMF", BIOEM 2015, Asilomar, CA, USA, June 2015

"IARC Monograph Review Process and Glyphosate", Deutscher Bundestag, Berlin, Germany, June 2015

"The Exposome: Why does it matter for public health?", Association of Public Health Laboratories, Dr. Katherine Kelly Distinguished Lecture, Indianapolis, USA May 2015

"A bioinformatics/biostatistics approach to systems toxicology", Maastricht University, Maastricht, Netherlands, March, 2015

"Mechanistic Data, Cellular Pathways, and Cancer Classification", 39th Annual Toxicology Forum Meeting, Washington D.C., USA, January, 2015

"Current Issues in Environmental Public Health", Nicholls State University, Thibodaux, LA, April, 2014

"Review of Approaches to the Quantification of Risks", International Agency for Research on Cancer, Lyon, France, November, 2013

"The Gene-Environment-Disease Interactome", Maastricht University, September, 2013

"Toxicogenomics and Electromagnetic Fields", BEMS Annual Meeting, Thesaloniki, Greece, June 2013

"Toxicogenomics and Risk Assessment", International Symposium on Toxicogenomics and Human Health, Paris, France, May, 2013

"Extreme Weather, Climate and Health: Putting Science into Practice", Climate and Health Conference, Washington, DC, February, 2013

"Biofuels and Human Health: CDC's activities", Our Energy Future and Health, National Academy of Sciences, Washington, DC, USA, January, 2013

"Global Environmental Health", Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, January, 2013
"Addressing Superfund Sites", SBRP Annual Meeting, Washington, DC, October, 2012
"The future of cancer risk assessment", 6th NCC International Symposium- Management of Carcinogenic Hazard: Recent Progress and Future Perspectives, Seoul, Korea, June, 2012
"ATSDR Reorganization and New Directions", National Toxicology Program Executive Committee, Washington, DC, May, 2012
"Hydraulic Fracturing, National Concerns", Institute of Medicine Roundtable on Environmental Health, Washington, DC, April, 2012.
"Using Health Impact Assessment to Guide National Policy", National Health Impact Assessment Meeting, Washington, DC, April, 2012
"The Great Debate", Society of Toxicology Meetings, San Francisco, USA, March, 2012.
"CDC and Environmental Health", Society of Toxicology Meetings, San Francisco, USA, March, 2012.
"Keynote Address", National Climate Assessment Initial Workshop, Charleston, SC, February, 2012.
"Integration of Climate, Weather and Health", American Meteorological Society, New Orleans, USA, January, 2012.
"Biomonitoring and Environmental Health", CINVESTAV, Mexico Department of Health, Mexico City, Mexico, January, 2012.
"What is health risk assessment", International Conference on EMF and Health, Brussels, Belgium, November, 2012.
"Risks from multiple chemicals in polluted communities", International Conference on Chemical Mixtures, Wasington, USA, October, 2011.
"Healthy Homes: The Scientific Support", National Healthy Homes Conference, Denver, CO, June, 2011
"Climate Change and Human Health", Keynote Address, Yale School of Public Health Annual Alumni Day, New Haven, Ct, June, 2011
"Environmental Public Health", Keynote Address, International Conference on Sustainable Remediation, Amherst, MA, June, 2011
"The gene-environment disease interactome for humans", Society of Toxicology Annual Meeting, Washington, DC, March, 2011
"Environmental Exposure Science in the 21st Century", NAS Committee on Human and Environmental Exposure Science in the 21st Century, Washington, DC, February, 2011
"Human Health Effects and the Impact of Climate Change on Our Oceans", Annual National Conference on Science, Policy, and the Environment, Washington, Dc, January, 2011
"A Vision for the National Center for Environmental Health and Agency for Toxic Substances and Disease Registry", NAS Committee on Emerging Issues in Environmental Health, Washington, DC, December, 2010
"Predicting Health Risks Using Gene-Expression Data", EPA NextGEN Workshop, Research Triangle Park, NC, October, 2010
"Building Sustainable Environments", NAS Workshop: Pathways to Sustainable Development, Atlanta, GA, September, 2010
"Future Directions in Environmental Public Health", State Environmental Health Directors Annual Meeting, Portland, OR, September 2010
"The gene-environment disease interactome for humans", AACR Workshop on The Future of Molecular Epidemiology, Miami, Fl, June, 2010
"Emerging methods for determining chemical hazards", Keynote Address, Human Health Hazard Indicators Workshop, Sacramento, March, 2010.
"Dose-response and risk assessment considerations of melamine in infants", Society of Toxicology Annual Meetings, Sal Lake City, March, 2010
"The gene-environment disease interactome for humans", University of Queensland Centre for Clinical Research, Brisbane, Australia, February, 2010

"Research Needs for Climate Change and Human Health", Queensland Health and ENTOX, February, 2010

"Using systems biology to develop the gene-environment disease interactome for humans", ENTOX, Brisbane, Australia, December, 2009

"Quantifying the health risks of dioxins and dioxin-like compounds", ENTOX, January, 2010

"Epigenetics and its use in toxicology and risk assessment", Keynote Adress, Australian College of Toxicology and Risk Assessment Annual Meeting, Canberra, December, 2009.

"The changing shape of regulatory toxicology in the United States", Office of Chemical Safety & Environmental Health, Canberra, Australia, December, 2009

"The changing shape of regulatory toxicology in the United States", NICNAS, Sydney, Australia, December, 2009

"Development and Use of the Gene-Environment-Disease Interactome for Humans", 60th Anniversary of the Department of Biostatistics, University of North Carolina, Chapel Hill, NC, October, 2009

"Using systems biology to develop the gene-environment disease interactome for humans", Seventh World Congress on Alternatives and Animal Use in Life Sciences, Rome, Italy, September, 2009

"Use of 'Omics' Technologies to Enhance Risk Assessment", International Workshop on Genomics in Cancer Risk Assessment, Venice, Italy, August, 2009.

"Proteomics, Genomics and the Human Gene-Environment-Disease Interactome", First International Radiation Proteomics Workshop, Munich, Germany, May, 2009.

"Molecular Biology and Regualtory Toxicology", Society of Toxicology, Annual Meeting, Baltimore, Maryland, March, 2009.

"Using Modern Research Tools to Enhance Regulatory Toxicology", SOT Satelite Meeting on Development Of Toxicological And Environmental Public Health Infrastructures In Africa: Understanding The Premise And Mapping The Approach, Baltimore, MD, March 2009.

"Health Implications of Climate Change", Navy and Marine Public Health Conference, Norfolk, Va., March, 2009.

"Climate Change and Minority Health", Second Annual National Conference on Health Disparities, University of the Virgin Islands, St. Croix, December, 2008.

"Climate Change and Human Health", the Environmental Mutagen Society Annual Meetings, Puerto Rico, October, 2008.

"Climate Change and Human Health", The Puerto Rico Chamber of Commerce, October, 2008.

"Using Stem Cells in Environmental Health Research", California Stem Cells and Predictive Toxicology Initiative Workshop, Berkeley, CA, July, 2008.

"Developing the Disease-Environment Interactome", IMBA-FGF Workshop on Developing Omics Technologies to Assess Unclear Risks", Berlin, Germany, May, 2008

"Environmental Systems Biology", 40th Annual Interface Symposium on the interface between computer science and statistics, Durham, NC, May, 2008

"Environmental causes of childhood leukemia and modern environmental health research", Summary Address, ICNIRP/WHO/BS Symposium on Research Advances in Childhood Leukemia, Berlin, Germany, May, 2008.

"Using systems biology as a tool to understand environmental health", Atlantic Coast Symposium on the Mathematical Sciences in Biology and Medicine, Raleigh, NC, April, 2008.

"Biological networks and high-throughput screening", Society of Toxicology, Seattle, Washington, March, 2008.

"Environment and cancer: strategies for identifying new hazards". Workshop on the Environment and Cancer, American Cancer Society, Atlanta, Georgia, USA. January, 2008.

"A systems approach to human health risk assessment", Environmental Health Sciences Decision Making: Risk Management, Evidence and Ethics. US National Academy of Sciences, Washington, DC, USA. January, 2008.

"Finding targets for high-throughput screening linking genetics, genomics, pathways and human disease". NTP Biological Screening Program Seminar Series. January, 2008.

"The forest for the trees: A systems approach to environmental health research". Keynote Address, KTL-DEH Center of Excellence Program, Lamalo, Finland. December, 2007.
"Toxicology for the 21st Century", ECNIS Research Colloquium, Engelheim, Germany, October, 2007.
"Evidence-Based Decision Making in Public Health", First International Workshop on Evidence-Based Toxicology, Cuomo, Italy. October, 2007.
"Genes, Pathways and Diseases", World Health Organization Headquarters, Geneva, Switzerland. September, 2007
"Pharmacokinetics and dynamics of ketamine in horses", University of Bern, June, 2007
"Uncertainty and Variability in Risk Assessment", NAS Workshop on Characterizing Uncertainty: Subgroup on Uncertainty in Estimating Low-Dose Risk from High-Dose Data, June, 2007
"The utility and interpretation of high-throughput screening data in risk assessment with focus on the NTP programs", International Workshop on "Evaluating upstream endpoints for improved decision making and risk assessment", Berkeley, Ca. May, 2007
"Future directions of risk assessment at the World Health Organization", International Program for Chemical Safety Harmonization Steering Committee Meeting, Berlin, Germany, May, 2007
"Chipping away at environmental health risk assessment", GEMS annual meeting, Research Triangle park, NC, April, 2007
"The forest for the trees: A systems approach to environmental health", National Conference on Science, Policy and the Environment, Washington, DC, February, 2007
"Emerging technologies and their application in risk assessment", National Research Council Standing Committee on Risk Analysis Issues and Reviews, Washington, DC, December, 206
"Stochastic systems biology modeling", International Congress on Systems Biology, Tokyo, Japan, October, 2006
"Acceptable risk: environmental health research and public health", National Symposium on Acceptable Risk in Clinical Medicine, RTP, NC, September, 2006
"Stochastic Systems Biology", Society for Industrial and Applied Mathematics Annual Meeting, Raleigh, NC, August, 2006.
"Identifying and Quantifying Gene Interaction Networks", Academia Sinica, Tapei, Taiwan, April, 2006.
"Estimating Health Risks from Environmental Exposures", Medical School, Cheng Kung National University, Taiwan, April, 2006.
"Systems Biology and Environmental Health Research", National Institute of Environmental Studies, Tsukuba, Japan, April, 2006.
"Cancer Research and Risk Assessment: Looking to the Future", National Cancer Institute, Japan, April, 2006
"Cancer Research and Risk Assessment: Looking to the Future", Keynote Address, Cancer and the Environment Symposium, Duke University, March, 2006
"Environmental Systems Biology", Mount St Mary's College, Maryland, March, 2006.
"Bioinformatics in High Throughput Screening: A proposal", ILSI Annual Meeting, Puerto Rico, January, 2006.
"Gene Regulatory Networks in Cancer Risk Assessment, German Cancer Research Center, Heidelberg, Germany, December, 2005.
"Alternative Methods in Toxicology; Problems and Solutions", Keynote Address, Workshop on Alternative Methods in Toxicology, Baltimore, Maryland, November, 2005
"Future Directions of the National Toxicology Program", American Public Health Association Annual Meeting, New Orleans, Louisiana, November, 2005.
"Risk Assessment", A Mini-Course, University of Finland, Kuopio, Finland, October, 2005.
"Mechanism-Based Modeling as an Alternative to Animals in Toxicology", 5th World Conference on Alternatives To Animals Used in the Life Sciences, Berlin, Germany, August, 2005.
"Environmental Systems Biology", Toxicology Forum, Aspen, Colorado, July, 2005.
"Mechanistic Implications of Modifiers of Chemical Toxicity", Keynote Lecture, Workshop on Modifiers of Chemical Toxicity: Implications for Human Health Risk Assessment, Poros, Greece, June, 2005.

"Future Directions in Screening for Toxicants", Committee on the Future of Toxicology, National Academy of Sciences, Washington, DC, May, 2005.
"Toxicogenomics and the Future of Cancer Risk Assessment", International Agency for Research on Cancer, Lyon France, May, 2005.
"Statistical Contributions to Federal Advisory Committees", Eastern North American Region of the Biometrics Society Annual Meeting, Austin, Texas, March, 2005.
"The Future of Long-Term Animal Carcinogenicity Studies", Invited Debate, Society of Toxicology Annual Meetings, New Orleans, Louisiana, March, 2005.
"Future Directions of the National Toxicology Program", Society of Toxicology Annual Meetings, New Orleans, Louisiana, March, 2005.
"Role of the National Toxicology Program in Risk Assessment", Federal-State Oncology and Risk Analysis Committee, Madison, Wisconsin, October, 2004.
"Identifying and Quantifying Cancer Risks using Mechanistic Data", Cancer and The Environment: NCI Science Writers Seminar, National Institutes of Health, Bethesda, Maryland, October, 2004.
"Toxicogenomics and the Future of Risk Assessment", Office of Science and Technology Policy, Washington, DC, August, 2004
"A Vision for the National Toxicology Program", Toxicology Forum, Aspen, Colorado, June, 2004
"Health and Environment", Joint Program on Climate Variability and Human Health, Atlanta, Georgia, March, 2004
"Dose-Response Analysis of Toxicogenomic Data", NIEHS Toxicogenomics Faculty, Research Triangle Park, NC, December, 2003.
"Dose-Response Analysis of Toxicogenomic Data", DERT Retreat, Southern Pines, NC, December, 2003.
"Toxicogenomics and the Research Directions of the National Toxicology Program", Keynote Address, Conference on Toxicology in the 21st Century, Korean Food and Drug Administration, Seoul, Korea, November, 2003.
"Dioxin and Agent Orange: What is known and what is suspected", Keynote Address, NCST/NIEHS Conference of Environmental Mediation for Dioxin Soil Contamination, Hanoi, Vietnam, November, 2003.
"Dose-Response Analysis of Toxicogenomic Data", NCSU/NIEHS Workshop on Bioinformatics and Risk Assessment, Research Triangle Park, NC, October, 2003.
"Health Effects of Electric and Magnetic Fields: Current Research Directions", Joint WHO/Japanese Ministry of Health Public Meeting on Cellular Radiation, Tokyo, Japan, September, 2003.
"Future Directions in Toxicology and the National Toxicology Program", NIEHS Public Liaison Meeting, New York City, September, 2003.
"Chipping Away at Risk Assessment: Genomics, Proteomics, Metabonomics and Cancer Risk Assessment", Society of Toxicologic Pathology, Savannah, Ga. June, 2003.
"Cancer Modeling; An Overview", Statistical Methods in Cancer Research, Radiation Effects Research Foundation, Kyoto, Japan, March, 2003 (presented by H. Toyoshiba due to war in Iraq and US Govt. responsibilities)
"Statistical methods for evaluating population exposures using CDC's Environmental Report Card", CDC, Atlanta, Ga. March, 2003.
"Bystander Effects in Carcinogenesis", Society of Toxicology Meetings, Salt Lake City, Utah, March, 2003
"Validation and use of genetically-modified mouse models as alternatives in carcinogenicity testing", ILSI Workshop on Genetically Modified Mouse Models, Alexandria, Va., February, 2003.
"The Future of Genomics in Toxicology", Mississippi State University, Oxford, Mississippi, February, 2003
"NIEHS priorities in ecological research", USGS Director's Retreat, Washington, DC. January, 2003.
"The US National Toxicology Program", Keynote Lecture at Opening Ceremonies for the Korean National Toxicology Program, Seoul, Korea, November, 2002

"The NIEHS/NTP Initiatives on Alternative Models in Toxicology", Science Advisory Board Meeting of the Center for the Advancement of Alternatives in Toxicology, Johns Hopkins University, Baltimore, Maryland, November, 2002

"The Analysis and Interpretation of Two-Stage Liver Bioassays in the Rat", Workshop on Hepatic Preneoplasia: Quantitative Evaluation in Carcinogenesis Bioassays and Relevance for Human Hepatocarcinogenesis, Heidelberg, Germany, June, 2002

"Chipping Away at Risk Assessment: Genomics, Proteomics, Metabonomics and Cancer Risk Assessment", Office of Environmental Health Hazard Assessment, California Department of Health, Sacramento, California, June, 2002

"Mechanistic Modeling of Stochastic Endpoints", NIEHS Conference on Future Directions in Biostatistics, Research Triangle Park, North Carolina, May, 2002

"Mechanistic Models of Skin Carcinogenesis", North Carolina State University, Raleigh, North Carolina, May, 2002

"Endocrine Dismodulation and Cancer", Workshop on Light, Endocrine Systems and Cancer, University of Cologne, Cologne, Germany, May, 2002

"Mechanism-Based Quantitative Analysis of Carcinogenesis Data", New York Academy of Sciences, April, 2002.

"Emerging Issues in Cancer Risk Assessment", Sunrise Expert Seminar, American Association for Cancer Research Annual Meetings, San Francisco, California, April, 2002

"Quantitative Evaluation of Health Risks from Dioxin", Joint US-Vietnam Scientific Conference on the Health and Environmental Effects of Agent Orange/Dioxin in Vietnam, Hanoi, Vietnam, March, 2002

"Cancer Risk Assessment", Hanoi Medical School/Vietnam Environmental Protection Agency/Vietnam Ministry of Health Joint Seminar, Hanoi Vietnam, January, 2002

"Mechanism-Based Modeling of Genomics Data", University of Berne, Switzerland, October, 2001.

"Multistage Models of Carcinogenesis", International Biometrics Conference Satellite Meeting, Fukuoka, Japan, September, 2001. (given by H. Toyoshiba due to conflict in scheduling)

"Discussion of CDC's Report on Levels of Environmental Contaminants in Human Tissues", National Public Radio Talk of the Nation Science Friday, March, 2001.

"Mixtures and Models in Environmental Health Risk Assessment", *NIEHS/Colorado State University* conference on the Application of Technology to Chemical Mixture Research, January, 2001

"Biological and biophysical research at extremely low- and radio-frequencies", Forschungsgemeinschaft Funk, Bad Münstereifel, Germany, December 2000.

"Harmonization of study evaluation and health risk assessment", 2nd International EMF Seminar, Chinese Ministry of Health, WHO and ICNIRP, Xian, China, October 2000.

"QRA: extrapolations (animals to humans; high dose low dose)", First International Course on Scientific Basis of Carcinogen Assessment – Quo Vadis?, NIVA, Naantali, Finland, September 2000.

"Statistical and biological models and toxicological issues in risk assessment", Center for Environmental Health Risk Assessment (CERA) workshop on Future Needs in Risk Assessment, Stockholm, Sweden, September 2000.

"Innovative use of mechanistic considerations in forming quantitative risk assessment models", Workshop on Future Research For Improving Risk Assessment Methods, National Institute for Occupational Safety and Health, Aspen, Colorado, August 2000.

"Latest scientific findings on dioxin", Dioxin Summit, University of California at Berkeley, California, August 2000.

"Stochastic modeling in carcinogenesis and development", Virtual Human Workshop, Research Triangle Park, NC. June 2000.

"Evaluating health risks; methods and examples", Federal Office for Public Health, Bern, Switzerland. May 2000.

"Decisions about environmental health risks: What are the key questions and how does this apply to melatonin", Low frequency EMF, visible light, melatonin and cancer", International Symposium, University of Cologne and the German Research Society, Cologne, Germany. May 2000.

"Children's environmental health-risk assessment issues and challenges", California Environmental Protection Agency, Office of Environmental Health Assessment, Oakland, CA. April 2000.

"Children's environmental health risks", WHO/NIEHS Pacific Rim Meeting on Children's Health, Manila, April, 2000.

"Toxicology and environmental health seminar series 2000", University of Florida, Gainesville, Florida. April 2000.

"Biologically based population health risk assessment models", University of Ottawa, Department of Epidemiology and Community Medicine and the Institute of Population Health, Ottawa, Canada. March 2000.

"Molecular epidemiology: a new tool in cancer prevention", Applying science to regulatory policy and public health, Keystone Symposium, Taos, New Mexico. February 2000.

"Statistical methods used for evaluating chemical safety in the environment", Mathematical Research Institute, Oberwolfach, Germany. February 2000.

"Mechanism-based mathematical modeling in health risk assessment" NC State University, January, 2000.

"Methods for analyzing and quantifying EMF health risks", WHO/ICNIRP scientific meeting, Erice, Sicily, Italy. November 1999.

"Harmonization of cancer and non-cancer risk assessment", The Society of Toxicology, Washington, DC. November 1999.

"Risky business: evaluating the safety of environmental exposures", Illinois Environmental Health Association, 1999 Annual Education Conference, Peoria, IL. October 1999.

"Challenges in the use of experimental and epidemiological data in health risk assessment", Symposium on "Statistical methods in epidemiology and demography", Department of Biostatistics at the University of North Carolina in Chapel Hill, Chapel Hill, North Carolina. October 1999.

"Linking toxicokinetics and toxicodynamics in biologically-based dose response models", Dioxin 99, 19th International Symposium on Halogenated Environmental Organic Pollutants and Persistent Organic Pollutants (POPs), Venice, Italy. September 1999.

"How did WHO decide on a TDI for dioxin", Faculty of Environmental Studies in Nagasaki University and the Department of Medical Informatics in Kyushu Medical School, Nagasaki, Japan. September 1999.

"PBPK Models for Estrogen", Workshop To Evaluate Research Priorities For Endocrine Active Compound Risk Assessment Methods, CMA, EPA, NIEHS, Research Triangle Park, North Carolina. August 1999.

"Risk assessment activities in the National Toxicology Program; risk assessment methodology, and mechanistic modeling of endocrine disruptors and dioxins", Harvard Symposium on Persistent Organic Pollutants (POPS): A Public Health Perspective, Boston, Massachusetts. June 1999.

"Do extremely low frequency electric and magnetic fields pose a health risk?", 21st Annual Meeting of the Bioelectromagnetics Society, Long Beach, California. June 1999.

"Probabilistic assessment of cancer dose response data", Probabilistic Risk Assessment Workshop, Sarasota, Florida. February 1999.

"Linking toxicology and epidemiology: the role of mechanistic modeling", CDC and ATSDR Symposium on Statistical Methods, Emerging Statistical Issues In Public Health For The 21st Century, Atlanta, Georgia. January 1999.

"Evaluation Of Health Risks From Exposure To Electric And Magnetic Fields; Completion Of A 2-Year Review Process", Toxicology Round Table, Virginia-Maryland Regional College of Veterinary Medicine, Blacksburg, Virginia. January 1999.

Interagency Agreements and Intramural Research Grants

US Environmental Protection Agency (1993-1994). Development and implementation of a mechanistic model for 2,3,7,8-TCDD. $25K.

Agency for Toxic Substances and Disease Registry (1994-1996). Study of the relationship between chemical structure/activity and dose-response shape and magnitude for carcinogens and development of physiologically based pharmacokinetic models for risk assessment. $796K.

NIEHS, Intramural Research Grant (1995-1998). Development of a mechanistic model of the impact of electromagnetic fields on melatonin. $180K.

NIEHS, Interagency Coordinator for Risk Assessment, EMFRAPID Program (1996-1998). Evaluation of potential risks from exposure to electric and magnetic fields. $2 million.

US Environmental Protection Agency (1999). Update to Dose-Response Chapter, Dioxin Reassessment. $25K.

NIH, Office of Research on Minority Health (1999-2002) GIS/Resampling method for evaluating data on environmental justice. $70K.

Direction of Ph.D. Theses:

A Bailer. *The effects of treatment lethality on tests of carcinogenicity*. Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, 1986.

P Williams. *Estimating tumor incidence rates using the method of moments and maximum likelihood estimation combined*. Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, 1989.

G Carr. *The analysis of data on adverse reactions to chemicals in developmental toxicology*. Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, 1989.

S Liu. *Estimating parameters in a two-stage model of carcinogenesis using information on enzyme-altered foci from initiation-promotion experiments*. Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, 1993.

CD Sherman. *Multipath/multistage models of carcinogenesis*. Department of Statistics and Actuarial Sciences, University of Waterloo, Waterloo, Ontario, Canada, 1994.

C Lyles. *Cell labeling data: Models and parameter estimation*. Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, 1995.

F Ye. *The equal slopes test for benchmark doses*. Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, 2001

S Whitaker. *Development of a biologically-based mathematical model of fetal development*. Department of Mathematics, North Carolina State University, Raleigh, North Carolina, 2000.

R Helms. *Homeostatic feedback control of growth on multistage cancer models*. Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, 2001.

Journal Articles: (peer reviewed)

1. Cote I, Andersen ME, Ankley GT, Barone S, Birnbaum LS, Boekelheide K, et al. The Next Generation of Risk Assessment Multi-Year Study-Highlights of Findings, Applications to Risk Assessment, and Future Directions. *Environ Health Perspect* (2016) **124**(11):1671-82. doi: 10.1289/EHP233. PubMed PMID: 27091369; PubMed Central PMCID: PMCPMC5089888.

2. Parham F, Portier CJ, Chang X, Mevissen M. The Use of Signal-Transduction and Metabolic Pathways to Predict Human Disease Targets from Electric and Magnetic Fields Using in vitro Data in Human Cell Lines. *Frontiers in public health* (2016) **4**:193. doi: 10.3389/fpubh.2016.00193. PubMed PMID: 27656641; PubMed Central PMCID: PMCPMC5013261.

3. Portier CJ, Armstrong BK, Baguley BC, Baur X, Belyaev I, Belle R, et al. Differences in the carcinogenic evaluation of glyphosate between the International Agency for Research on Cancer (IARC) and the European Food Safety Authority (EFSA). *Journal of epidemiology and community health* (2016) **70**(8):741-5. doi: 10.1136/jech-2015-207005. PubMed PMID: 26941213; PubMed Central PMCID: PMCPMC4975799.

4. Sand S, Parham F, Portier CJ, Tice RR, Krewski D. Comparison of Points of Departure for Health Risk Assessment Based on High-Throughput Screening Data. *Environ Health Perspect* (2016). doi: 10.1289/EHP408. PubMed PMID: 27384688.

5. Scinicariello F, Portier C. A simple procedure for estimating pseudo risk ratios from exposure to non-carcinogenic chemical mixtures. *Archives of toxicology* (2016) **90**(3):513-23. doi: 10.1007/s00204-015-1467-z. PubMed PMID: 25667015.

6. Smith MT, Guyton KZ, Gibbons CF, Fritz JM, Portier CJ, Rusyn I, et al. Key Characteristics of Carcinogens as a Basis for Organizing Data on Mechanisms of Carcinogenesis. *Environ Health Perspect* (2016) **124**(6):713-21. doi: 10.1289/ehp.1509912. PubMed PMID: 26600562; PubMed Central PMCID: PMCPMC4892922.

7. McPartland, J., Dantzker, H.C., Portier, C. J. Building a robust 21st century chemical testing program at the U.S. Environmental Protection Agency: recommendations for strengthening scientific engagement, Environ Health Perspect 2015. 123 (1): p. 1-5.

8. Smith, M.T., Gibbons, C.F., Fritz, J.M., Rusyn, I., Lambert, P., Kavlock, R., Hecht, S.S., Bucher, J., Caldwell, J.C., Demarini, D., Cogliano, V., Portier, C., Paan, R., Straif, K., Guyton, K.Z., Key Characteristics of Carcinogens and an Approach to using Mechanistic Data in their Classification, Environ Health Perspect 2015 (in press)

9. Thomas, R., Thomas, R.S., Auerbach, S. S., Portier, C. J., Biological networks for predicting chemical hepatocarcinogenicity using gene expression data from treated mice and relevance across human and rat species. PLoS One, 2013. **8**(5): p. e63308.

10. Scinicariello, F., Buser, M.C., Mevissen, M., Portier, C.J., Blood lead level association with lower body weight in NHANES 1999-2006. Toxicol Appl Pharmacol, 2013. 273(3): p. 516-23.

11. Thomas R, Portier CJ., Gene Expression Networks, Methods Mol Biol. 2013;930:165-78.

12. Aylward LL, Kirman CR, Schoeny R, Portier CJ, Hays SM., Evaluation of Biomonitoring Data from the CDC National Exposure Report in a Risk Assessment Context: Perspectives across Chemicals. Environ Health Perspect. 2012 Dec 11. [Epub ahead of print] PMID: 23232556

13. Sand, S., Portier, C.J., Krewski, D. A Signal-to-noise crossover dose as the point of departure for risk assessment. Environmental Health Perspectives. 119(12):1766-74, 2011

14. Gohlke, J.M., Thomas, R., Woodward, A., Campbell-Lundrum, D., Pruss-Ustun, A., Hales, S., Portier, C.J. Estimating the global public health implications of electricity and coal consumption. Environmental Health Perspectives 2011 119 (6): 821-6

15. McHale CM, Zhang L, Lan Q, Vermeulen R, Li G, Hubbard AE, Porter KE, Thomas R, Portier CJ, Shen M, Rappaport SM, Yin S, Smith MT, Rothman N. Global gene expression profiling of a population exposed to a range of benzene levels. Environ Health Perspect. 2011 May;119(5):628-34.

16. Prause AS, Guionaud CT, Stoffel MH, Portier CJ, Mevissen M. Expression and function of 5-hydroxytryptamine 4 receptors in smooth muscle preparations from the duodenum, ileum, and pelvic flexure of horses without gastrointestinal tract disease. Am J Vet Res. 2010 Dec;71(12):1432-42.

17. Luke, N.S., DeVito, M.J., Portier, C.J., El-Masri, H.A., Employing a mechanistic model for the MAPK pathway to examine the impact of cellular all-or-none behavior on overall tissue response, Dose-Response 2010 8(3): 347-67.

18. Crump, KS, Chen, C., Chiu, W.A., Louis, T.A., Portier, C. J., Subramaniam, R.P., Wgite, P.D., What role for biologically-based Dose-Response Models in Estimating Low-Dose Risk. Env. Health Persp. 2010 118(5):585-8

19. Parham F, Austin C, Southall N, Huang R, Tice R, Portier C. Dose-Response modeling of High-Throughput Screening Data. J Biomol Screen. 2009 **14**(10), 1216-27

20. Hines RN, Sargent D, Autrup H, Birnbaum LS, Brent RL, Doerrer NG, Cohen Hubal EA, Juberg DR, Laurent C, Luebke R., Olejniczak K, Portier CJ, Slikker W. Approaches for assessing risks to sensitive populations: lessons learned from evaluating risks in the pediatric population. Tox. Sci. 2010 **113** (4), 4-26.

21. Portier, C. Toxicological decision making on hazards and risks – status quo and the way forward: current concepts and schemes of science-driven decision making – an overview. *Human and Experimental Toxicology* 2009 **28**(2-3), 123-125

22. Prause, A.S., Stoffel, M.H., Portier, C.J., Mevissen, M., Expression and function of 5-HT7 receptors in smooth muscle preparation from equine duodenum, ileum, and pelvic flexure, *Research in Veterinary Science* 2009 **87**(2), 292-299

23. Boyd, W.A., Smith, M. V., Kissling, G. E., Rice, J., R., Snyder, D. W., Portier, C. J., Freedman, J. H. Application of a Mathematical Model to Describe the Effects of Chlorpyrifos on Caenorhabditis elegans Development, *PLoS ONE* 2009 **4**(9): e7024. doi:10.1371/journal.pone.0007024

24. Smith MV, Boyd WA, Kissling GE, Rice JR, Snyder DW, et al. A Discrete Time Model for the Analysis of Medium-Throughput C. elegans Growth Data. *PLoS ONE* 2009 **4**(9): e7018. doi:10.1371/journal.pone.0007018

25. Gohlke, J. M., Stockton, P.S., Sieber, S., Foley, J., Portier, C. J. AhR-mediated gene expression in the developing mouse telencephalon. *Reproductive Toxicology* 2009 (in press) doi:10.1016/j.reprotox.2009.05.067

26. Thomas, R., Gohlke, J., Parham, F., Smith, M., Portier, C. (2009) Choosing the right path: enhancement of biologically-relevant sets of genes or proteins using pathway structure. *Genome Biology* 2009 **10**(4), R44.

27. Julia M Gohlke, Reuben Thomas, Yonqing Zhang, Michael C Rosenstein, Allan P Davis, Cynthia Murphy, Carolyn J Mattingly, Kevin G Becker, Christopher J Portier, Genetic and Environmental Pathways to Complex Disease. *BMC Systems Biology* 2009 May 5, 3:46.

28. Schmitz, A., Portier, C. J., Thurmann, W., Theurillat, R., Mevissen, M. Stereoselective biotransformation of ketamine in equine liver and lung microsomes. *J. Vet. Pharm. And Therapeutics* 2008 **31** (5): 446-455

29. Xia, M; Huang, R; Witt, KL; Southall, N; Fostel, J; Cho, MH; Jadhav, A; Smith, CS; Inglese, J; Portier, CJ; Tice, RR; Austin, CP Compound cytotoxicity profiling using quantitative high-throughput screening. *Env. Health Perspectives* 2008 **116** (3): 284-291

30. Gohlke, J. M., Armant, O., Parham, F., M., Smith, M., V., Zimmer, C., Castro, D., S., Nguyen, L., Parker, J., S., Gradwohl, G., Guillemot, F., Portier, C. J. Characterization of proneural gene regulatory network during mouse telencephalon development., *BMC Biology* 2008 **6** (15)

31. Subramaniam, R. P., Chen, C., Crump, K. S., Devoney, D., Fox, J., F., Portier, C. J., Schlosser, P. M., Thompson, C. M., White, P. Uncertainties in Biologically-based modeling of formaldehyde-induced respiratory cancer risk: Identification of key issues. *Risk Analysis* 2008 28(4): 907-923

32. Buehler, M., Steiner, A., Meylan, M., Portier, C. J., and Mevissen, M., In vitro effects of bethanechol on smooth muscle preparations obtained from abomasal fundus, corpus and antrum of dairy cows. Research in Vet. Sci. **2008 84** (3), 444-451

33. Barton, H.A., W.A. Chiu, R.W. Setzer, M.E. Andersen, A.J. Bailer, F.Y. Bois, R.S. Dewoskin, S. Hays, G. Johanson, N. Jones, G. Loizou, R.C. Macphail, C.J. Portier, M. Spendiff, and Y.M. Tan, Characterizing Uncertainty and Variability in Physiologically-based Pharmacokinetic (PBPK) Models: State of the Science and Needs for Research and Implementation. *Toxicol Sciences* 2007 **99** (2) 395-402.

34. Pfeiffer, J.B., M. Mevissen, A. Steiner, C.J. Portier, and M. Meylan, In vitro effects of bethanechol on specimens of intestinal smooth muscle obtained from the duodenum and jejunum of healthy dairy cows. Am J Vet Res, 2007. **68**(3): p. 313-22.

35. Smith, M., Miller, C., Kohn, M., Walker, N.J., Portier, C. J., Absolute estimation of initial concentrations of amplicon in a real-time PT-PCR process, *BMC Bioinformatics* 2007 **8**(1), 409

36. Toyoshiba, H., Sone, H., Yamanaka, T., Parham, F., Irwin, R., Boorman, G., and Portier, C. Gene network analysis suggests differences between high and low doses of acetaminophen. Toxicology and Applied Pharmacology *2006* 215 (3), 306-316

37. Knobloch M, Portier CJ, Levionnois OL, Theurillat R, Thormann W, Spadavecchia C, Mevissen M. Antinociceptive effects, metabolism and disposition of ketamine in ponies under target-controlled drug infusion. Toxicology And Applied Pharmacology 2005 216 (3): 373-386

38. Portier, C.J., H. Toyoshiba, H. Sone, F. Parham, R.D. Irwin, and G.A. Boorman, Comparative analysis of gene networks at multiple doses and time points in livers of rats exposed to acetaminophen. Altex 2005 **23 Suppl**: p. 380-4.

39. Resnik, D.B. and C. Portier, Pesticide testing on human subjects: weighing benefits and risks. Environ Health Perspect. 2005 **113**(7): p. 813-7.

40. Walker, N. J., Crockett, P. W., Nyska, A., Brix, A. E., Jokinen, M. P., Sells, D. M., Hailey, J. R., Easterling, M., Haseman, J. K., Yin, M., Wyde, M. E., Bucher, J. R., and Portier, C. J. Dose-additive carcinogenicity of a defined mixture of "dioxin-like compounds". *Environ Health Perspect* 2005 113, 43-48.[1]

41. Boorman, G. A., Irwin, R. D., Vallant, M. K., Gerken, D. K., Lobenhofer, E. K., Hejtmancik, M. R., Hurban, P., Brys, A. M., Travlos, G. S., Parker, J. S., and Portier, C. J. Variation in the hepatic gene expression in individual male Fischer rats. *Toxicol Pathol* 2005 33, 102-110.

42. Toyoshiba, H., Walker, N.J., Bailer, A.J. and Portier, C.J. Evaluation of toxic equivalency factors for induction of cytochromes P450 CYP1A1 and CYP1A2 enzyme activity by dioxin-like compounds. *Toxicol Appl Pharmacol 2004 194*, 156-68.

43. Trachsel, D., Tschudi, P., Portier, C.J., Kuhn, M., Thormann, W., Scholtysik, G. and Mevissen, M. Pharmacokinetics and pharmacodynamic effects of amiodarone in plasma of ponies after single intravenous administration. *Toxicology and Applied Pharmacology 2004, 195,* 113-125.

44. Bucher, J. R., and Portier, C. Human carcinogenic risk evaluation, Part V: The national toxicology program vision for assessing the human carcinogenic hazard of chemicals. *Toxicol Sci* 2004 82, 363-366.

45. Germolec, D. R., Kashon, M., Nyska, A., Kuper, C. F., Portier, C., Kommineni, C., Johnson, K. A., and Luster, M. I. The accuracy of extended histopathology to detect immunotoxic chemicals. *Toxicol Sci* 2004 82, 504-514.

46. Germolec, D. R., Nyska, A., Kashon, M., Kuper, C. F., Portier, C., Kommineni, C., Johnson, K. A., and Luster, M. I.. Extended histopathology in immunotoxicity testing: interlaboratory validation studies. *Toxicol Sci* 2004 78, 107-115.

47. Shelby, M., Portier, C., Goldman, L., Moore, J., Iannucci, A., Jahnke, G., and Donkin, S.. NTP-CERHR expert panel report on the reproductive and developmental toxicity of methanol. *Reprod Toxicol* 2004 18, 303-390.

48. Smith, M. V., Nyska, A., and Portier, C. Application of a statistical dynamic model investigating the short-term cellular kinetics induced by riddelliine, a hepatic endothelial carcinogen. *Toxicol Sci* 2004 80, 258-267.

49. Toyoshiba, H., Yamanaka, T., Sone, H., Parham, F. M., Walker, N. J., Martinez, J., and Portier, C. J. Gene interaction network suggests dioxin induces a significant linkage between aryl hydrocarbon receptor and retinoic acid receptor beta. *Environ Health Perspect* 2004 112, 1217-1224.[2]

---

[1] Awarded outstanding published paper in 2005 by the Risk Assessment Specialty Section of the Society of Toxicology

[2] Awarded outstanding published paper in 2004 by the Risk Assessment Specialty Section of the Society of Toxicology

50. Yamanaka, T., Toyoshiba, H., Sone, H., Parham, F. M., and Portier, C. J. The TAO-Gen algorithm for identifying gene interaction networks with application to SOS repair in E. coli. *Environ Health Perspect* 2004 112, 1614-1621

51. Schecter, A., Lucier, G. W., Cunningham, M. L., Abdo, K. M., Blumenthal, G., Silver, A. G., Melnick, R., Portier, C., Barr, D. B., Barr, J. R., Stanfill, S. B., Patterson, D. G., Jr., Needham, L. L., Stopford, W., Masten, S., Mignogna, J., and Tung, K. C. Human consumption of methyleugenol and its elimination from serum. *Environ Health Perspect* 2004 112, 678-680.

52. Bucher, J and Portier, C. Human carcinogenic risk evaluation, Part V: The national toxicology program vision for assessing the human carcinogenic hazard of chemicals. Toxicol Sci. 2004 Dec;82(2):363-6

53. Chhabra, R.S., Bucher, J.R., Wolfe, M. and Portier, C. Toxicity characterization of environmental chemicals by the US National Toxicology Program: an overview. *Int J Hyg Environ Health, 2003 206*, 437-45.

54. Germolec, D.R., Nyska, A., Kashon, M., Kuper, C.F., Portier, C., Kommineni, C., Johnson, K.A. and Luster, M.I. Extended Histopathology in Immunotoxicity Testing: Interlaboratory Validation Studies. *Toxicol Sci.* 2003 78(1) 107-115

55. Koken, P.J., Piver, W.T., Ye, F., Elixhauser, A., Olsen, L.M. and Portier, C.J. Temperature, air pollution, and hospitalization for cardiovascular diseases among elderly people in Denver. *Environ Health Perspect 2003 111*, 1312-7.

56. Whitaker, S.Y., Tran, H.T. and Portier, C.J. (2003) Development of a biologically-based controlled growth and differentiation model for developmental toxicology. *J Math Biol, 46*, 1-16.

57. Hauser, D.S., Mevissen, M., Weiss, R., Portier, C.J., Scholtysik, G., Studer, U.E. and Danuser, H. Effects of ketanserin and DOI on spontaneous and 5-HT-evoked peristalsis of the pig ureter in vivo. *Br J Pharmacol 2002 135*, 1026-32.

58. Kim, A.H., Kohn, M.C., Portier, C.J. and Walker, N.J. Impact of physiologically based pharmacokinetic modeling on benchmark dose calculations for TCDD-induced biochemical responses. *Regul Toxicol Pharmacol 2002 36*, 287-96.

59. Kohn, M.C., Melnick, R.L., Ye, F. and Portier, C.J. Pharmacokinetics of sodium nitrite-induced methemoglobinemia in the rat. *Drug Metab Dispos* 2002, *30*, 676-83.

60. Koo, J.W., Parham, F., Kohn, M.C., Masten, S.A., Brock, J.W., Needham, L.L. and Portier, C.J. The association between biomarker-based exposure estimates for phthalates and demographic factors in a human reference population. *Environ Health Perspect* 2002, *110*, 405-10.

61. Parham, F.M., Matthews, H.B. and Portier, C.J. A physiologically based pharmacokinetic model of p,p'-dichlorodiphenylsulfone. *Toxicol Appl Pharmacol* 2002, *181*, 153-63.

62. Portier, C.J. Endocrine dismodulation and cancer. *Neuroendocrinol Lett* 2002, *23 Suppl 2*, 43-7.

63. Simmons, P.T. and Portier, C.J. Toxicogenomics: the new frontier in risk analysis. *Carcinogenesis* 2002, *23*, 903-5.

64. Weiss, R., Mevissen, M., Hauser, D.S., Scholtysik, G., Portier, C.J., Walter, B., Studer, U.E. and Danuser, H. Inhibition of human and pig ureter motility in vitro and in vivo by the K(+) channel openers PKF 217-744b and nicorandil. *J Pharmacol Exp Ther* 2002, *302*, 651-8.

65. Whitaker, S.Y. and Portier, C. A controlled growth and differentiation model for non-monotonic dose-response. *Human and Ecological Risk Assessment* 2002, *8*, 1739-1755.

66. Kohn, M.C., Walker, N.J., Kim, A.H. and Portier, C.J. Physiological modeling of a proposed mechanism of enzyme induction by TCDD. *Toxicology* 2001, *162*, 193-208.

67. Mevissen, M., Denac, H., Schaad, A., Portier, C.J. and Scholtysik, G. Identification of a cardiac sodium channel insensitive to synthetic modulators. *J Cardiovasc Pharmacol Ther* 2001, *6*, 201-12.

68. Portier, C.J. Linking toxicology and epidemiology: the role of mechanistic modeling. *Stat Med* 2001, *20*, 1387-93.

69. Willems, B.A., Melnick, R.L., Kohn, M.C. and Portier, C.J. A physiologically based pharmacokinetic model for inhalation and intravenous administration of naphthalene in rats and mice. *Toxicol Appl Pharmacol* 2001, *176*, 81-91.

70. Ye, F., Piver, W.T., Ando, M. and Portier, C.J. Effects of temperature and air pollutants on cardiovascular and respiratory diseases for males and females older than 65 years of age in Tokyo, July and August 1980-1995. *Environ Health Perspect* 2001, *109*, 355-9.

71. Halmes, N.C., Roberts, S.M., Tolson, J.K. and Portier, C.J. Reevaluating cancer risk estimates for short-term exposure scenarios. *Toxicol Sci* 2000, *58*, 32-42.

72. Kohn, M.C., Parham, F., Masten, S.A., Portier, C.J., Shelby, M.D., Brock, J.W. and Needham, L.L. Human exposure estimates for phthalates. *Environ Health Perspect* 2000, *108*, A440-2.

73. Portier, C.J. and Parham, F.M. Comments on a biochemical model of cyclophosphamide hematotoxicity. *J Toxicol Environ Health A* 2000, *61*, 525-8.

74. Portier, C., Sherman, C.D. and Kopp-Schneider, A. Multistage, stochastic models of the cancer process: a general theory for calculating tumor incidence. *Stochastic Environmental Research and Risk Assessment* 2000, *14*, 173-179.

75. Sherman, C.D. and Portier, C.J. Calculation of the cumulative distribution function of the time to a small observable tumor. *Bull Math Biol* 2000, *62*, 229-40.

76. Smith, M.V. and Portier, C.J. Incorporating observability thresholds of tumors into the two-stage carcinogenesis model. *Math Biosci* 2000, *163*, 75-89.

77. Tritscher, A.M., Mahler, J., Portier, C.J., Lucier, G.W. and Walker, N.J. Induction of lung lesions in female rats following chronic exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Toxicol Pathol* 2000, *28*, 761-9.

78. Walker, N.J., Tritscher, A.M., Sills, R.C., Lucier, G.W. and Portier, C.J. Hepatocarcinogenesis in female Sprague-Dawley rats following discontinuous treatment with 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Toxicol Sci* 2000, *54*, 330-7.

79. Andersen, M.E., Conolly, R.B., Faustman, E.M., Kavlock, R.J., Portier, C.J., Sheehan, D.M., Wier, P.J. and Ziese, L. Quantitative mechanistically based dose-response modeling with endocrine-active compounds. *Environ Health Perspect* 1999, *107 Suppl 4*, 631-8.

80. el Masri, H. and Portier, C. Replication potential of cells via the protein kinase C-MAPK pathway: application of a mathematical model. *Bulletin of Mathematical Biology* 1999, *61*, 379-398.

81. El-Masri, H.A., Bell, D.A. and Portier, C.J. Effects of glutathione transferase theta polymorphism on the risk estimates of dichloromethane to humans. *Toxicol Appl Pharmacol* 1999, *158*, 221-30.

82. Melnick, R.L., Sills, R.C., Portier, C.J., Roycroft, J.H., Chou, B.J., Grumbein, S.L. and Miller, R.A. Multiple organ carcinogenicity of inhaled chloroprene (2-chloro-1,3-butadiene) in F344/N rats and B6C3F1 mice and comparison of dose-response with 1,3-butadiene in mice. *Carcinogenesis* 1999, *20*, 867-78.

83. Piver, W.T., Ando, M., Ye, F. and Portier, C.J. Temperature and air pollution as risk factors for heat stroke in Tokyo, July and August 1980-1995. *Environ Health Perspect* 1999, *107*, 911-6.

84. Walker, N.J., Portier, C.J., Lax, S.F., Crofts, F.G., Li, Y., Lucier, G.W. and Sutter, T.R. Characterization of the dose-response of CYP1B1, CYP1A1, and CYP1A2 in the liver of female Sprague-Dawley rats following chronic exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Toxicol Appl Pharmacol* 1999, *154*, 279-86.

85. El-Masri, H.A. and Portier, C.J. Physiologically based pharmacokinetics model of primidone and its metabolites phenobarbital and phenylethylmalonamide in humans, rats, and mice. *Drug Metab Dispos* 1998, *26*, 585-94.

86. Jung, D., Berg, P.A., Edler, L., Ehrenthal, W., Fenner, D., Flesch-Janys, D., Huber, C., Klein, R., Koitka, C., Lucier, G., Manz, A., Muttray, A., Needham, L., Papke, O., Pietsch, M., Portier, C., Patterson, D., Prellwitz, W., Rose, D.M., Thews, A. and Konietzko, J. Immunologic findings in workers formerly exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin and its congeners. *Environ Health Perspect* 1998, *106 Suppl 2*, 689-95.

87. Kopp-Schneider, A., Portier, C. and Bannasch, P. A model for hepatocarcinogenesis treating phenotypical changes in focal hepatocellular lesions as epigenetic events. *Math Biosci* 1998, *148*, 181-204.

88. Murrell, J.A., Portier, C.J. and Morris, R.W. Characterizing dose-response: I: Critical assessment of the benchmark dose concept. *Risk Anal* 1998, *18*, 13-26.

89. Parham, F.M. and Portier, C.J. Using structural information to create physiologically based pharmacokinetic models for all polychlorinated biphenyls. II. Rates of metabolism. *Toxicol Appl Pharmacol* 1998, *151*, 110-6.

90. Portier, C.J. and Bell, D.A. Genetic susceptibility: significance in risk assessment. *Toxicol Lett* 1998, *102-103*, 185-9.

91. Portier, C.J. and Ye, F. U-shaped dose-response curves for carcinogens. *Hum Exp Toxicol* 1998, *17*, 705-7; discussion 716-8.

92. Sherman, C.D. and Portier, C.J. Eyes closed: simple, intuitive, statistically sound, and efficient methods for estimating parameters of clonal growth cancer models. *Risk Anal* 1998, *18*, 529-34.

93. Blumenthal, G.M., Kohn, M.C. and Portier, C.J. A mathematical model of production, distribution, and metabolism of melatonin in mammalian systems. *Toxicol Appl Pharmacol* 1997, *147*, 83-92.

94. Gaido, K.W., Leonard, L.S., Lovell, S., Gould, J.C., Babai, D., Portier, C.J. and McDonnell, D.P. Evaluation of chemicals with endocrine modulating activity in a yeast-based steroid hormone receptor gene transcription assay. *Toxicol Appl Pharmacol* 1997, *143*, 205-12.

95. Parham, F.M., Kohn, M.C., Matthews, H.B., DeRosa, C. and Portier, C.J. Using structural information to create physiologically based pharmacokinetic models for all polychlorinated biphenyls. *Toxicol Appl Pharmacol* 1997, *144*, 340-7.

96. Portier, C. and el Masri, H. Statistical research needs in mechanistic modeling for carcinogenic risk assessment. *Stat Methods Med Res* 1997, *6*, 305-15.

97. Sherman, C.D. and Portier, C.J. The two-stage model of carcinogenesis: overcoming the nonidentifiability dilemma. *Risk Anal* 1997, *17*, 367-74.

98. Buchanan, J.R. and Portier, C.J. The use of data on biologically reactive intermediates in risk assessment. *Adv Exp Med Biol* 1996, *387*, 429-37.

99. Bucher, J.R., Portier, C.J., Goodman, J.I., Faustman, E.M. and Lucier, G.W. Workshop overview. National Toxicology Program Studies: principles of dose selection and applications to mechanistic based risk assessment. *Fundam Appl Toxicol* 1996, *31*, 1-8.

100. Kohn, M.C., Sewall, C.H., Lucier, G.W. and Portier, C.J. A mechanistic model of effects of dioxin on thyroid hormones in the rat. *Toxicol Appl Pharmacol* 1996, *136*, 29-48.

101. Luster, M.I., Portier, C., Simeonova, P. and Munson, A.E. Relationship between chemical-induced immunotoxicity and carcinogenesis. *Drung Information Journal* 1996, *30*, 281-286.

102. Melnick, R.L., Kohn, M.C. and Portier, C.J. Implications for risk assessment of suggested nongenotoxic mechanisms of chemical carcinogenesis. *Environ Health Perspect* 1996, *104 Suppl 1*, 123-34.

103. Portier, C.J. and Lyles, C.M. Practicing safe modeling: GLP for biologically based mechanistic models. *Environ Health Perspect* 1996, *104*, 806.

104. Portier, C.J., Kopp-Schneider, A. and Sherman, C.D. Calculating tumor incidence rates in stochastic models of carcinogenesis. *Math Biosci* 1996, *135*, 129-46.

105. Portier, C.J., Sherman, C.D., Kohn, M., Edler, L., Kopp-Schneider, A., Maronpot, R.M. and Lucier, G. Modeling the number and size of hepatic focal lesions following exposure to 2,3,7,8-TCDD. *Toxicol Appl Pharmacol* 1996, *138*, 20-30.

106. Sherman, C.D. and Portier, C.J. Stochastic simulation of a multistage model of carcinogenesis. *Math Biosci* 1996, *134*, 35-50.

107. Kopp-Schneider, A. and Portier, C.J. Carcinoma formation in NMRI mouse skin painting studies is a process suggesting greater than two stages. *Carcinogenesis* 1995, *16*, 53-9.

108. Sherman, C.D. and Portier, C.J. Quantitative analysis of multiple phenotype enzyme-altered foci in rat hepatocarcinogenesis experiments: the multipath/multistage model. *Carcinogenesis* 1995, *16*, 2499-506.

109. Hoel, D.G. and Portier, C.J. Nonlinearity of dose-response functions for carcinogenicity. *Environ Health Perspect* 1994, *102 Suppl 1*, 109-13.

110. Kohn, M.C., Lucier, G.W. and Portier, C.J. The importance of biological realism in dioxin risk assessment models. *Risk Anal* 1994, *14*, 993-1000.

111. Kohn, M.C. and Portier, C.J. A model of effects of TCDD on expression of rat liver proteins. *Prog Clin Biol Res* 1994, *387*, 211-22.

112. Kopp-Schneider, A., Portier, C.J. and Sherman, C.D. The exact formula for tumor incidence in the two-stage model. *Risk Anal* 1994, *14*, 1079-80.

113. Kopp-Schneider, A. and Portier, C.J. A stem cell model for carcinogenesis. *Math Biosci* 1994, *120*, 211-32.

114. Luster, M.I., Portier, C., Pait, D.G. and Germolec, D.R. Use of animal studies in risk assessment for immunotoxicology. *Toxicology* 1994, *92*, 229-43.[3]

115. Portier, C.J., Lucier, G.W. and Edler, L. Risk from low-dose exposures. *Science* 1994, *266*, 1141-2.

116. Portier, C.J. Biostatistical issues in the design and analysis of animal carcinogenicity experiments. *Environ Health Perspect* 1994, *102 Suppl 1*, 5-8.

117. Sherman, C.D., Portier, C.J. and Kopp-Schneider, A. Multistage models of carcinogenesis: an approximation for the size and number distribution of late-stage clones. *Risk Anal* 1994, *14*, 1039-48.

118. Bailer, A.J. and Portier, C.J. An index of tumorigenic potency. *Biometrics* 1993, *49*, 357-65.

119. Carr, G.J. and Portier, C.J. An evaluation of some methods for fitting dose-response models to quantal-response developmental toxicology data. *Biometrics* 1993, *49*, 779-91.

120. Fenger, J., Forslund, J., Grandjean, P., Gron, P., Jensen, F., Keiding, L., Kjaergaard, S., Larsen, J., Lewtas, J., Lynge, E., Moller, H., Moller, L., Nielsen, P., Ostenfeldt, N., Pilsgaard, H., Portier, C., Poulsen, E., Rastogi, S., Skov, T., Thomsen, A., Ulbak, K. and Osterlind, A. Report: The Working Group On Cancer And The Non-Occupational Environment. *Pharmacology and Toxicology* 1993, *72*, 167-171.

121. Kohn, M.C. and Portier, C.J. Effects of the mechanism of receptor-mediated gene expression on the shape of the dose-response curve. *Risk Anal* 1993, *13*, 565-72.

122. Kohn, M.C., Lucier, G.W., Clark, G.C., Sewall, C., Tritscher, A.M. and Portier, C.J. A mechanistic model of effects of dioxin on gene expression in the rat liver. *Toxicol Appl Pharmacol* 1993, *120*, 138-54.

---

[3] Board of Publications, Best Paper Award, Society of Toxicology, 1995

123. Lucier, G.W., Portier, C.J. and Gallo, M.A. Receptor mechanisms and dose-response models for the effects of dioxins. *Environ Health Perspect* 1993, *101*, 36-44.

124. Luster, M.I., Portier, C., Pait, D.G., Rosenthal, G.J., Germolec, D.R., Corsini, E., Blaylock, B.L., Pollock, P., Kouchi, Y., Craig, W. and et al. Risk assessment in immunotoxicology. II. Relationships between immune and host resistance tests. *Fundam Appl Toxicol* 1993, *21*, 71-82.

125. Maronpot, R.R., Foley, J.F., Takahashi, K., Goldsworthy, T., Clark, G., Tritscher, A., Portier, C. and Lucier, G. Dose response for TCDD promotion of hepatocarcinogenesis in rats initiated with DEN: histologic, biochemical, and cell proliferation endpoints. *Environ Health Perspect* 1993, *101*, 634-42.

126. Meier, K.L., Bailer, A.J. and Portier, C.J. A measure of tumorigenic potency incorporating dose-response shape. *Biometrics* 1993, *49*, 917-26.

127. Melnick, R.L., Huff, J., Barrett, J.C., Maronpot, R.R., Lucier, G. and Portier, C.J. Cell proliferation and chemical carcinogenesis: symposium overview. *Environ Health Perspect* 1993, *101 Suppl 5*, 3-7.

128. Melnick, R.L., Huff, J., Barrett, J.C., Maronpot, R.R., Lucier, G. and Portier, C.J. Cell proliferation and chemical carcinogenesis: a symposium overview. *Mol Carcinog* 1993, *7*, 135-8.

129. Portier, C.J., Kopp-Schneider, A. and Sherman, C.D. Using cell replication data in mathematical modeling in carcinogenesis. *Environ Health Perspect* 1993, *101 Suppl 5*, 79-86.

130. Portier, C., Tritscher, A., Kohn, M., Sewall, C., Clark, G., Edler, L., Hoel, D. and Lucier, G. Ligand/receptor binding for 2,3,7,8-TCDD: implications for risk assessment. *Fundam Appl Toxicol* 1993, *20*, 48-56.[4]

131. Kopp-Schneider, A. and Portier, C.J. Birth and death/differentiation rates of papillomas in mouse skin. *Carcinogenesis* 1992, *13*, 973-8.

132. Luster, M.I., Pait, D.G., Portier, C., Rosenthal, G.J., Germolec, D.R., Comment, C.E., Munson, A.E., White, K. and Pollock, P. Qualitative and quantitative experimental models to aid in risk assessment for immunotoxicology. *Toxicol Lett* 1992, *64-65 Spec No*, 71-8.

133. Luster, M.I., Portier, C., Pait, D.G., White, K.L., Jr., Gennings, C., Munson, A.E. and Rosenthal, G.J. Risk assessment in immunotoxicology. I. Sensitivity and predictability of immune tests. *Fundam Appl Toxicol* 1992, *18*, 200-10.

134. Piegorsch, W.W., Carr, G.J., Portier, C.J. and Hoel, D.G. Concordance of carcinogenic response between rodent species: potency dependence and potential underestimation. *Risk Anal* 1992, *12*, 115-21.

135. Tritscher, A.M., Goldstein, J.A., Portier, C.J., McCoy, Z., Clark, G.C. and Lucier, G.W. Dose-response relationships for chronic exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin in a rat tumor promotion model: quantification and immunolocalization of CYP1A1 and CYP1A2 in the liver. *Cancer Res* 1992, *52*, 3436-42.

---

[4] Best-applied statistics paper, Centers for Disease Control, 1993.

136. Williams, P. and Portier, C. Analytic expressions for maximum likelihood estimators in a non-parametric model of tumor incidence and death. *Communications in Statistics - Theory and Methods* 1992, *21*, 711-732.

137. Williams, P. and Portier, C. Explicit solutions for constrained maximum likelihood estimators in survival/sacrifice experiments. *Biometrika* 1992, *74*, 717-729.

138. Carr, G.J. and Portier, C.J. An evaluation of the Rai and Van Ryzin dose-response model in teratology. *Risk Anal* 1991, *11*, 111-20.

139. Kopp-Schneider, A. and Portier, C.J. Distinguishing between models of carcinogenesis: the role of clonal expansion. *Fundam Appl Toxicol* 1991, *17*, 601-13.

140. Kopp-Schneider, A., Portier, C.J. and Rippmann, F. The application of a multistage model that incorporates DNA damage and repair to the analysis of initiation/promotion experiments. *Math Biosci* 1991, *105*, 139-66.

141. Portier, C.J. and Kopp-Schneider, A. A multistage model of carcinogenesis incorporating DNA damage and repair. *Risk Anal* 1991, *11*, 535-43.

142. Rogan, W.J., Blanton, P.J., Portier, C.J. and Stallard, E. Should the presence of carcinogens in breast milk discourage breast feeding? *Regul Toxicol Pharmacol* 1991, *13*, 228-40.

143. Bailer, A.J. and Portier, C.J. A note on fitting one-compartment models: non-linear least squares versus linear least squares using transformed data. *J Appl Toxicol* 1990, *10*, 303-6.

144. Portier, C.J. and Edler, L. Two-stage models of carcinogenesis, classification of agents, and design of experiments. *Fundam Appl Toxicol* 1990, *14*, 444-60.

145. Kopp, A. and Portier, C.J. A note on approximating the cumulative distribution function of the time to tumor onset in multistage models. *Biometrics* 1989, *45*, 1259-63.

146. Portier, C.J. and Kaplan, N.L. Variability of safe dose estimates when using complicated models of the carcinogenic process. A case study: methylene chloride. *Fundam Appl Toxicol* 1989, *13*, 533-44.

147. Portier, C.J. and Bailer, A.J. Testing for increased carcinogenicity using a survival-adjusted quantal response test. *Fundam Appl Toxicol* 1989, *12*, 731-7.

148. Portier, C.J. and Bailer, A.J. Two-stage models of tumor incidence for historical control animals in the National Toxicology Program's carcinogenicity experiments. *J Toxicol Environ Health* 1989, *27*, 21-45.

149. Bailer, A.J. and Portier, C.J. Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples. *Biometrics* 1988, *44*, 417-31.

150. Bailer, A.J. and Portier, C.J. An illustration of dangers of ignoring survival differences in carcinogenic data. *J Appl Toxicol* 1988, *8*, 185-9.

151. Portier, C.J. Species correlation of chemical carcinogens. *Risk Anal* 1988, *8*, 551-3.

152. Portier, C. Life table analysis of carcinogenicity experiments. *Journal of the American College of Toxicology* 1988, *7*, 575-582.

153. Portier, C.J. Statistical properties of a two-stage model of carcinogenesis. *Environ Health Perspect* 1987, *76*, 125-31.

154. Portier, C.J. and Hoel, D.G. Issues concerning the estimation of the TD50. *Risk Anal* 1987, *7*, 437-47.

155. Portier, C.J. and Bailer, A.J. Simulating failure times when the event of interest is unobservable with emphasis on animal carcinogenicity studies. *Comput Biomed Res* 1987, *20*, 458-66.

156. Portier, C.J. and Dinse, G.E. Semiparametric analysis of tumor incidence rates in survival/sacrifice experiments. *Biometrics* 1987, *43*, 107-14.

157. Kupper, L.L., Portier, C., Hogan, M.D. and Yamamoto, E. The impact of litter effects on dose-response modeling in teratology. *Biometrics* 1986, *42*, 85-98.

158. Portier, C.J., Hedges, J.C. and Hoel, D.G. Age-specific models of mortality and tumor onset for historical control animals in the National Toxicology Program's carcinogenicity experiments. *Cancer Res* 1986, *46*, 4372-8.

159. Portier, C.J. Type 1 error and power of the linear trend test in proportions under the National Toxicology Program's modified pathology protocol. *Fundam Appl Toxicol* 1986, *6*, 515-9.

160. Portier, C. Estimating the tumor onset distribution in animal carcinogenesis experiments. *Biometrika* 1986, *3*, 371-378.

161. Portier, C.J. and Hoel, D.G. Design of animal carcinogenicity studies for goodness-of-fit of multistage models. *Fundam Appl Toxicol* 1984, *4*, 949-59.

162. Krewski, D., Crump, K.S., Farmer, J., Gaylor, D.W., Howe, R., Portier, C., Salsburg, D., Sielken, R.L. and Van Ryzin, J. A comparison of statistical methods for low dose extrapolation utilizing time-to-tumor data. *Fundam Appl Toxicol* 1983, *3*, 140-60.

163. Portier, C. and Hoel, D. Low-dose-rate extrapolation using the multistage model. *Biometrics* 1983, *39*, 897-906.

164. Portier, C. and Hoel, D. Optimal design of the chronic animal bioassay. *J Toxicol Environ Health* 1983, *12*, 1-19.

Journal Articles: (reviews and other)

165. McPartland, J., Dantzker, H.C., Portier, C.J., Building a robust 21st century chemical testing program at the U.S. Environmental Protection Agency: recommendations for strengthening scientific engagement, Environ Health Perspect, 2015. 123(1): 1-5.

166. Portier, C.J., L.R. Goldman, and B.D. Goldstein, Inconclusive findings: now you see them, now you don't! Environ Health Perspect, 2014. 122(2): p. A36.

167. Portier, C., ATSDR in the 21st Century, Journal of Environmental Health, 74(7):30-1, 2011

168. Portier, C. Comprehensive Public Health. Public Health Rep. 2011 126: Supp 1:3-6

169. Glass, R., Rosenthal, J., Jessup, C., Birnbaum, L., Portier, C. Tackling the research challenges of health and climate change. Env. Health Perspectives 2009 117 (12), A534-A535

170. Woodruff TJ, Zeise L, Axelrad DA, Guyton KZ, Janssen S, Miller M, et al. 2008. Meeting report: moving upstream-evaluating adverse upstream end points for improved risk assessment and decision-making. Environmental health perspectives 116(11): 1568-1575.

171. Portier, C. Risk Factors for Childhood Leukemia: Discussion and Summary, *Radiation Protection Dosimetry* 2009 **132**(2), 273-274.

172. Gohlke, J. M., Hrynkow, S. H., Portier, C. J., Health, Economy and Environment: Sustainable Energy Choices for a Nation, *Environmental Health Perspectives 2008* **116** (6), 236-237

173. Kissling, G. E., C. J. Portier and J. Huff. MtBE and cancer in animals: Statistical issues with poly-3 survival adjustments for lifetime studies. *Regul Toxicol Pharmacol.* 2008 **50** (3) 428-429

174. Barton, H. A., W. A. Chiu, R. Woodrow Setzer, M. E. Andersen, A. J. Bailer, F. Y. Bois, R. S. Dewoskin, S. Hays, G. Johanson, N. Jones, G. Loizou, R. C. Macphail, C. J. Portier, M. Spendiff and Y. M. Tan. Characterizing uncertainty and variability in physiologically based pharmacokinetic models: state of the science and needs for research and implementation. Toxicol Sci 2007 **99**(2): 395-402.

175. Gohlke, J. M. and C. J. Portier (2007). The forest for the trees: a systems approach to human health research. Environ Health Perspect 2007**115**(9): 1261-3.

176. Portier, C. J., Suk, W. A., Schwartz, D. A., Filling the translation-policy gap. *Env. Health Perspectives* 2007 **115** (3) A125

177. Becker, R. A., C. J. Borgert, S. Webb, J. Ansell, S. Amundson, C. J. Portier, A. Goldberg, L. H. Bruner, A. Rowan, R. D. Curren and W. T. Stott (2006). "Report of an ISRTP workshop: progress and barriers to incorporating alternative toxicological methods in the U.S." Regul Toxicol Pharmacol 2006 **46**(1): 18-22.

178. TCDD in female Sprague-Dawley rats. *Organohalogen Compounds* 1996, *29*, 222-227.

179. Portier, C.J. and D.A. Schwartz, The NIEHS and the National Toxicology Program: an integrated scientific vision. Environ Health Perspect, 2005. **113**(7): p. A440.

180. Portier, C. J., and Schwartz, D. A.. The NIEHS and the National Toxicology Program: an integrated scientific vision. *Environ Health Perspect* 2005, **113**, A440.

181. Portier, C.J., Suk, W. A. and D.A. Schwartz, Filling the translation-policy gap. Environ Health Perspect, 2005. **115**(3): p. A125.

182. Portier, C.J. Comments on the International Symposium on Light, Endocrine Systems and Cancer. *Neuroendocrinol Lett* 2002, *23 Suppl 2*, 79-81.

183. Portier, C. Chipping away at environmental health risk assessment. *Risk Policy Report* 2001, *8*, 37-38.

184. Portier, C. Decisions about environmental health risks: What are the key questions and how does this apply to melatonin? *Zentralblatt fur Arbeitsmedizin* 2000, *50*, 312-314.

185. Portier, C. Risk ranges for various endpoints following exposure to 2,3,7,8-TCDD. *Food Addit Contam* 2000, *17*, 335-46.

186. Brooks, E., Kohn, M., Van Birgelen, A., Lucier, G. and Portier, C. Stochastic models for papilloma formation following exposure to TCDD. *Organohalogen Compounds* 1999, *41*, 522-524.

187. Portier, C., Edler, L., Jung, D., Needham, L., Masten, S.A., Parham, F. and Lucier, G. Half-lives and body burdens for dioxin and dioxin-like compounds in humans estimated from an occupational cohort in Germany. *Organohalogen Compounds* 1999, *42*, 129-138.

188. Tritscher, A., Mahler, J., Portier, C., Lucier, G. and Walker, N.J. TCDD-induced lesions in rat lung following oral exposure. *Organohalogen Compounds* 1999, *42*, 285-288.

189. Lucier, G. and Portier, C. New Cancer Risk Guidelines. *Issues in Science and Technology* 1996, *13*, 10.

190. Portier, C. and Kohn, M. A biologically-based model for the carcinogenic effects of 2378-

191. Portier, C., Lucier, G. and Damstra, T. Dioxin research in Vietnam. *Science* 1995, *270*, 901.

192. Portier, C. and Kohn, M. Receptor Mediated Carcinogenesis and Dioxin. *Organohalogen Compounds* 1995, *26*, 267-272.

193. Edler, L. and Portier, C. Pharmacokinetic modeling of receptor/enzyme pathways for the action of dioxins. *Chemosphere* 1994, *25*, 239-242.

194. Edler, L., Portier, C. and Kopp-Schneider, A. Zur existenz von schwellenwerten: Wissenschaftliche methode oder statistisches artefakt in der risikoabschatzung. *Zbl Arbeitsmed* 1994, *44*, 16-21.

195. Kohn, M., Lucier, G., Sewall, C. and Portier, C. Physiological modeling of the effects of TCDD on thyroid hormones in the rat. *Organohalogen Compounds* 1994, *21*, 223-226.

196. Portier, C., Kohn, M., Sherman, C.D. and Lucier, G. Modeling the number and size of hepatic focal lesions following exposure to 2378-TCDD. *Organohalogen Compounds* 1994, *21*, 393-397.

197. Sherman, C.D. and Portier, C. The multipath/multistage model. *Informatik und Epidemiologie in Medizin und Biologie* 1994, *25*, 250-254.

198. Sherman, C.D. and Portier, C. The multipath/multistage model of carcinogenesis. *Organohalogen Compounds* 1994, *21*, 451-455.

199. Portier, C.J. Mechanistic modeling and risk assessment. *Pharmacol Toxicol* 1993, *72 Suppl 1*, 28-32.

200. Portier, C. Quantitative Risk Assessment. *ChemTech* 1990, *20*, 484-487.

Books/Book Chapters:

201. International Agency for Research on Cancer. Shiftwork, painting and firefighting (C. Portier member). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 98*. 2008 In press.

202. Trong, L., Portier, C.J. (eds) Proceedings of the Viet Nam – United States Scientific Conferences on Human Health and Environmental Effects of Agent Orange/Dioxin, Part 1: Environmental Effects - 3-6 March, 2002. Ha Noi, Vietnam, 2004.

203. Trong, L., Portier, C.J. (eds) Proceedings of the Viet Nam – United States Scientific Conferences on Human Health and Environmental Effects of Agent Orange/Dioxin, Part 2: Human Health Effects - 3-6 March, 2002. Ha Noi, Vietnam, 2004.

204. DeVito, M., Kim, A.H., Walker, N., Parham, F. and Portier, C. Dose-response modeling for 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Dioxins and Health*; 2 ed.; John Wiley and Sons: Hoboken, NJ, 2003.

205. World Health Organization. Establishing a dialogue on Risks from Electromagnetic Fields. Sahl, J., Bonner, P., Kemp, R., Kheifets, L., Repacholi, M., van Deventer, E., Vogel, E. and Portier, C. WHO Press, 2002

206. International Agency for Research on Cancer. Non-ionizing Radiation, Part I: Static and Extremely Low Frequency Electric and Magnetic Fields (C. Portier member). I *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume* 80. 19-26 2002

207. Wolfe, M., Boorman, G., Olden, K. and Portier, C. Power-Line Frequency Electric and Magnetic Fields: Effects on Human Health. *Human Health Risks from Exposure to ELF-EMF*; ICNIRP: Maastricht, The Netherlands, 2000.

208. Joint Committee on Food Additives (including C. Portier), World Health Organization/Food and Agriculture Organization. (1999) Evaluation of certain food additives and contaminants: Forty-ninth report of the Joint FAO/WHO Expert Committee on Food Additives. JCFA. WHO Technical Series #884. 1999.

209. Piver, W., Ando, M., Ye, F. and Portier, C. Temperature and air pollution as risk factors for cerebral vascular diseases in Tokyo, July-August 1980-1995. Proceedings of the 10-th Global Warming International Conference 1999

210. National Institute of Environmental Health Sciences. NIEHS Report on Health Effects from Exposure to Power-Line Frequency Electric and Magnetic Fields. Portier, C.J., Wolfe, M.S., Boorman, G.A., Bernheim, N.J., Galvin, M.J., Newton, S.A., Parham, F.M. and Olden, K.O. NIEHS 1999

211. Willems, B., Portier, C. and Lucier, G. Mechanism-based carcinogenic risk assessment of estrogens and estrogen-like compounds. *The Handbook of Environmental Chemistry*; Springer-Verlag: Heidelberg, 1999.

212. Portier, C.J. and Germolec, D. Immunotoxicology. *Encyclopedia of Biostatistics*; Wiley and Sons: Sussex, England, 1998 pp 1998-2000.

213. Portier, C.J. and Wolfe, M.S. Risk Communication: The focus in the NIEHS RAPID Program's review of EMF health Hazards. *Risk Perception, Risk Communication and Its Application to EMF Exposure*; ICNIRP: Oberschleisheim, Germany, 1998; pp 295-302.

214. Portier, C.J. Risk ranges for various endpoints following exposure to 2,3,7,8-TCDD. *WHO-ECEH/IPCS Assessment of health risk of dioxins*; International Program for Chemical Safety: Geneva, 1998; pp 23.

215. Portier, C.J. and Wolfe, M.S. (Eds.) EMF Science Review Symposium Breakout Group Report for Epidemiology Research Findings; National Institute of Environmental Health Sciences: Research Triangle Park, North Carolina, 1998.

216. Portier, C.J. and Wolfe, M.S. (Eds.) EMF Science Review Symposium Breakout Group Report for Clinical and *In Vivo* Laboratory Findings. In *NIH Publication Number 98-4400*; National Institute of Environmental Health Sciences: Research Triangle Park, North Carolina, 1998.

217. Portier, C.J. and Wolfe, M.S. (Eds.) Assessment of Health Effects from Exposure to Power-Line Frequency Electric and Magnetic Fields. In *NIH Publication Number 98-3981*; National Institute of Environmental Health Sciences: Research Triangle Park, North Carolina, 1998; pp 508.

218. Sherman, C.D. and Portier, C.J. Multistage carcinogenesis Models. *Encyclopedia of Biostatistics*; Wiley and Sons: Sussex, England, 1998; pp 2808-2814.

219. Portier, C.J. and Wolfe, M.S. Linking science to decisions: A strategy for electric and magnetic fields. *Biological Effects of ELF Electric and Magnetic Fields*; ICNIRP: Oberschleisheim, Germany, 1997; pp 211-218.

220. Portier, C.J. and Wolfe, M.S. (Eds.) EMF Science Review Symposium Breakout Group Report for Theoretical Mechanisms and *In Vitro* Research Findings; National Institute of Environmental Health Sciences: Research Triangle Park, North Carolina, 1997.

221. Buchanan, R. and Portier, C.J. The use of data on biologically reactive intermediates in risk assessment. *Biological reactive intermediates V : basic mechanistic research in toxicology and human risk assessment*; 387 ed.; Plenum Press: New York, 1995.

222. Luster, M., Portier, C., Pait, D.G., Rosenthal, G. and Germolec, D. Immunotoxicology and Risk Assessment. *Methods in immunotoxicology*; Wiley-Liss: New York, 1995; pp 51-68.

223. Portier, C.J. Quantitative models for cancer dose-response relationships: parameter estimation. *Low-Dose Extrapolation of Cancer Risks*; ILSI Press: Washington, 1995; pp 123-134.

224. Bailer, A.J. and Portier, C.J. Modeling risks from water contaminants: the application of concentration-response models. *Water contamination and health : integration of exposure assessment, toxicology, and risk assessment*; Dekker: New York, 1994; pp 447-466.

225. Kohn, M.C. and Portier, C.P. A model of the effects of TCDD on expression of rat liver proteins. *Receptor Mediated Biological Processes: Implications for Evaluating Carcinogens*; Wiley-Liss: New York, 1994; pp 211-222.

226. Kohn, M.C., Lucier, G.W. and Portier, C.J. Receptor mechanisms and dioxin risk assessment. *Hazardous Waste and Public Health: International Congress on the Health Effects of Hazardous Waste*; Princeton Press: Princeton, NJ, 1994; pp 421-432.

227. Portier, C.J. and Sherman, C.D. The potential effects of chemical mixtures on the carcinogenic process within the context of the mathematical multistage model. *Risk Assessment of Chemical Mixtures: Biological and Toxicological Issues*; Academic Press: New York, 1994; pp 665-686.

228. Portier, C.J., Hoel, D.G., Kaplan, N.L. and Kopp, A. Biologically based models for risk assessment. *IARC Sci Publ* 1990, 20-8.

229. Portier, C. Quantitative risk assessment. *Carcinogenicity and Pesticides*; American Chemical Society: New York, 1989; pp 164-174.

230. Portier, C. Design of long-term animal carcinogenicity experiments: dose allocation, animal allocation and sacrifice times. *Statistical Methods in Toxicological Research*; Gordan and Breech: New York, 1989; pp 455-467.

231. Portier, C. Utilizing biologically-based models to estimate carcinogenic risk. *Scientific Issues in Quantitative Cancer Risk Assessment*; Birkhauser: New York, 1989; pp 252-266.

232. Kaplan, N., Hogan, M., Portier, C. and Hoel, D. An evaluation of the safety factor approach in risk assessment. *Developmental Toxicology: Mechanisms and Risk*; Cold Spring Harbor Laboratory: New York, 1988.

233. Portier, C. (1985) Optimal dose/animal allocation for terminal sacrifice studies. Proceedings of the ASA Conference on Long-Term Animal Carcinogenicity Studies, 42-50.

234. Portier, C.J., Hoel, D. and VanRyzin, J. Statistical analysis of the carcinogenesis bioassay data relating to the risks from exposure to 2,3,7,8-tertachlorodibenzo-p-dioxin. *Public Health Risks of the Dioxins*; W. Kaufmann: Los Altos, 1984; pp 99-120.

Reports and Agency Publications:

235. A Human Health Perspective on Climate Change, Report of the Interagency Working Group on Climate Change and Health (C. Portier, chair), NIEHS, RTP, NC 2009

236. IARC Report of the Advisory Working Group for Priorities for future IARC Monographs, International Agency for Research on Cancer, Lyon, France 2008

237. NIEHSCenters for Children's Environmental Health and Disease Prevention Research Program, Review Panel Report, NIEHS, RTP, NC 2007

238. WHO Environmental Health Criteria Monograph 238; Extremely Low Frequency Electric and Magnetic Fields, (C. Portier, Chair), pp 543, World Health Organization Press, Geneva, Switzerland – 2007 ISBN 978 92 4 157238 5

239. IARC Report of the Advisory Group to Plan Monograph Volume 100: A Review of Human Carcinogens, International Agency for Research on Cancer, Lyon, France 2006

240. NIEHS Report of the Expert Panel to the NIEHS on Thimerosal Exposure in Pediatric Vaccines: Feasibility of Studies Using the Vaccine Safety Datalink, NIEHS, RTP, NC 2006

241. IARC Report of the Advisory Group to Revise the IARC Monographs Preamble International Agency for Research on Cancer, Lyon, France 2005

242. NTP. A National Toxicology Program for the 21st Century: A roadmap to achieve the NTP vision, pp. 4. National Toxicology Program/National Institute of Environmental Health Sciences, Research Triangle Park, NC. 2004

243. NTP. NTP Current directions and evolving strategies, pp. 4. National Toxicology Program/National Institute of Environmental Health Sciences, Research Triangle Park, NC. 2004

244. NTP. NTP Vision: Toxicology in the 21st Century: The role of the National Toxicology Program, pp. 4. National Toxicology Program/National Institute of Environmental Health Sciences, Research Triangle Park, NC 2003

245. USEPA Science Advisory Panel (C. Portier, Chair). Review of Characterization of Epidemiology Data Relating to Prostate Cancer and Exposure to Atrazine. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2003-02.

246. USEPA Science Advisory Panel (C. Portier, Member). Review of Proposed Science Policy: PPAR-alpha Agonist-Mediated Hepatocarcinogenesis in Rodents and Relevance to Human Health Risk Assessments. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2003-04.

247. USEPA Science Advisory Panel (C. Portier, Chair). Review of Physiologically-Based Pharmacokinetic/Pharmacodynamic Modeling: Preliminary Evaluation and Case Study for the N-Methyl Carbamate Pesticides: A Consultation. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2003-05.

248. USEPA Science Advisory Panel (C. Portier, Member). Comments on Determination of the Appropriate FQPA Safety Factor(s) in the Organophosphorous Pesticide Cumulative Risk Assessment: Susceptibility and Sensitivity to the Common Mechanism, Acetylcholinesterase Inhibition. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2002-03.

249. USEPA Science Advisory Panel (C. Portier, Member). Review Of Draft Termite Bait Product Performance Testing Guideline. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2002-04.

250. USEPA Science Advisory Panel (C. Portier, Chair). Review of Corn Rootworm Plant-Incorporated Protectant Insect Resistance Management and Non-Target Insect Issues. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2002-05.

251. USEPA Science Advisory Panel (C. Portier, Member). Review of Stochastic Human Exposure and Dose Simulation Model (SHEDS). FIFRA-SAP. US EPA Science Advisory Panel Report No. 2002-06.

252. IARC Monographs Volume 80 (C. Portier member) Non-Ionizing Radiation, Part 1: Static and Extremely Low-Frequency (ELF) Electric and Magnetic Fields, 429 pages; 2002

253. USEPA Science Advisory Panel (C. Portier, Member). Comments on Methods to Conduct a Preliminary Cumulative Risk Assessment for Organophosphorous Chemicals. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2002-01.

254. USEPA Science Advisory Panel (C. Portier, Chair). BT Plant-Pesticides Risk and Benefit Assessment. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2001-07.

255. USEPA Science Advisory Panel (C. Portier, Chair). Review of dietary exposure evaluation model (DEEM) and MaxLIP pesticide residue decompositing procedures and software. FIFRA-SAP. Science Advisory Panel Report No. 2001-01B.

256. USEPA Science Advisory Panel (C. Portier, Chair). Review of dietary exposure evaluation model (DEEM). FIFRA-SAP. EPA Science Advisory Panel Report No. 2001-01C.

257. USEPA Science Advisory Panel (C. Portier, Chair). Consultation on Development and Use of Distributions of Pesticide Concentrations in Drinking Water for FQPA Assessments. FIFRA-SAP. EPA Science Advisory Panel Report No. 2001-01D.

258. USEPA Science Advisory Panel (C. Portier, Member). Review of Probabilistic models and methodologies: Advancing the ecological risk assessment process in the EPA Office of Pesticides Programs. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2001-06.

259. USEPA Science Advisory Panel (C. Portier, Chair). Review of Common Mechanism of Action of Thiocarbamates and Dithiocarbamates. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2001-12.

260. USEPA Science Advisory Panel (C. Portier, Member). Review of Preliminary Cumulative Hazard and Dose-Response Assessment for Organophosphorus Pesticides: Determination of Relative Potency and Points of Departure for Cholinesterase Inhibition. FIFRA-SAP. US EPA Science Advisory Panel Report No. 2001-11.

261. USEPA Science Advisory Panel (C. Portier, Chair). Review of Characterization and Non-Target Organism Data Requirements for Protein Plant-Pesticides and- Cumulative Risk Assessment Methodology Issues of Pesticide Substances that Have a Common Mechanism of Toxicity. FIFRA-SAP. EPA-SAP-99-06. February 4, 2000.

262. USEPA Science Advisory Panel (C Portier, Member). Review of food allergenicity of CRY9C Endotoxin and other non-digestible proteins. FIFRA-SAP. USEPA Science Advisory Panel Report 2000-01A.

263. USEPA Science Advisory Panel (C Portier, Member). Review of atrazine: hazard and dose-response assessment and characterization. FIFRA-SAP. USEPA Science Advisory Panel Report No. 2000-05.

264. USEPA Science Advisory Panel (C. Portier, Chairman). Review of Statistical Methods for Use of Composite Data in Acute Dietary Exposure Assessment. FIFRA-SAP. EPA-SAP 99-03B. May 26, 1999.

265. USEPA Science Advisory Panel (C. Portier, Chairman). Review of the Use of Watershed-derived Percent Crop Areas as a Refinement Tool in FQPA Drinking Water Exposure Assessments for Tolerance Reassurance. FIFRA-SAP. EPA-SAP 99-03C. May 26, 1999.

266. USEPA Science Advisory Panel (C. Portier, Member). *Review of Burkholderia cepacia*: Risk Assessment of a Biopesticide with Affinities to a Human Opportunistic Pathogen. FIFRA-SAP. EPA-SAP-99-04. September 30, 1999.

267. USEPA Science Advisory Panel (C. Portier, Member). A Consultation on Protocol Design to Assess Acute Neurotoxicity Studies following Oral Administration of Pesticides. FIFRA-SAP. EPA-SAP-99-04B. September 30, 1999.

268. USEPA Science Advisory Panel (C. Portier, Member). Review of Higher Tier Ecological Risk Assessment for Chlorfenapyr. FIFRA-SAP. EPA-SAP-99-04C. September 30, 1999.

269. USEPA Science Advisory Panel (C. Portier, Chair). Review of -Spray Drift: Review of Proposed Pesticide Deposition Curves to Adjacent Areas. FIFRA-SAP. EPA-SAP-99-04C. September 30, 1999.

270. USEPA Science Advisory Panel (C. Portier, Member). Review of Guidance Document for Small Scale Prospective Ground Water Monitoring Studies. FIFRA-SAP. SAP98-01. November 19, 1998.

271. USEPA Science Advisory Panel (C. Portier, Member). Review of Common Mechanism of Action of Organophosphates. FIFRA-SAP. March, 1998.

272. USEPA Science Advisory Panel (C. Portier, Member). Review of Suggested Probabilistic Risk Assessment Methodology for Evaluating Pesticides That Exhibit a Common Mechanism of Action. FIFRA-SAP. March, 1998.

273. USEPA Science Advisory Panel (C. Portier, Member). Review of the Use of FQPA 10x Safety Factor to Address Special Sensitivity of Infants and Children to Pesticides. FIFRA-SAP. March, 1998.

274. USEPA Science Advisory Panel (C. Portier, Member). Review of Post Application Exposure Guidelines: Series 875-Group B. FIFRA-SAP. March, 1998.

275. USEPA Science Advisory Panel (C. Portier, Member). Review of Proposed Methods for Basin-scale Estimation of Pesticide Concentrations in Flowing Water and Reservoirs for Tolerance Reassessment. FIFRA-SAP. July, 1998.

276. USEPA Science Advisory Panel (C. Portier, Member). Review of Linear Low Dose Extrapolation for Cancer Risk Decisions: Sources of Uncertainty and How They Affect the Precision of Risk Estimates. FIFRA-SAP. July, 1998.

277. USEPA Science Advisory Panel (C. Portier, Member). Review of DDVP (Dichlorvos) Risk Issues. FIFRA-SAP. July, 1998.

278. USEPA Science Advisory Panel (C. Portier, Member). Review of FQPA 10x Safety Factor: Status Report. FIFRA-SAP. July, 1998.

279. USEPA Science Advisory Panel (C. Portier, Member). Review of Chlorothalonil: Mechanism for the Formation of Renal and Forestomach Tumors. FIFRA-SAP. July, 1998.

280. Environmental Protection Agency. Dose Response Modeling for TCDD. Lucier, G., Gallo, M., Portier, C.J., Bayard, S., Cooper, K., Georgopolous, P., McGrath, L., Andersen, M., DeVito, M., White, P., Kedderis, L., Mills, J. and Silbergeld, E. EPA/600/BP-92/001b 1994

281. Portier, C.J. Optimal bioassay design under the Armitage-Doll multistage model of carcinogenesis. PhD, University of North Carolina. 1981.

IPCS

INTERNATIONAL PROGRAMME ON CHEMICAL SAFETY

WHO ILO UNEP

Environmental Health Criteria 239

Principles for Modelling Dose–Response for the Risk Assessment of Chemicals

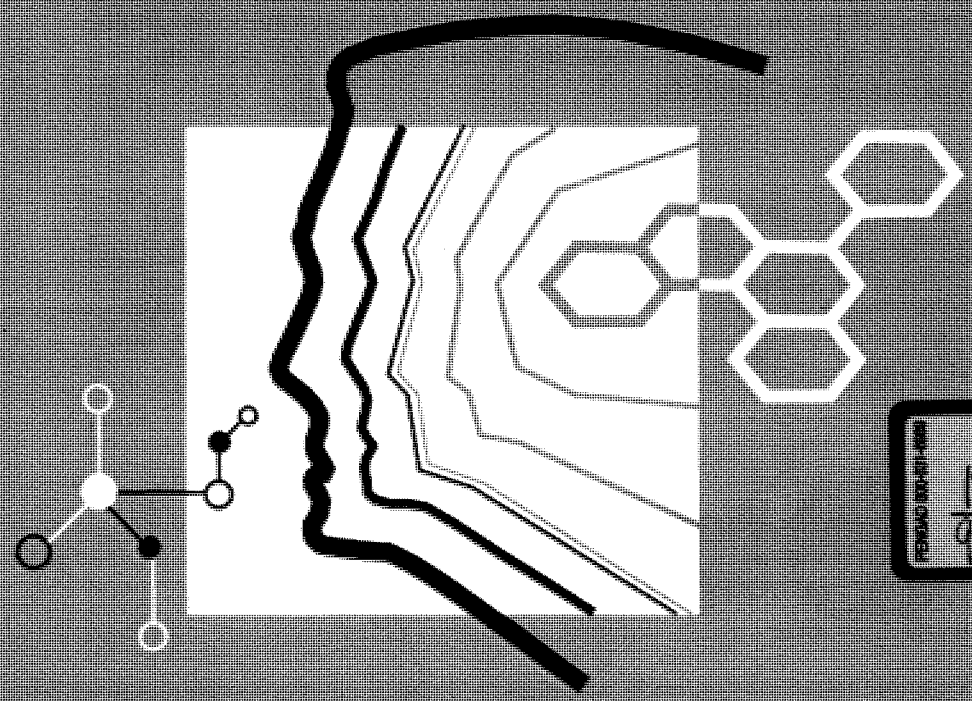


EXHIBIT
15-28
9/7/17

IOMC INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS
A cooperative agreement among UNEP, ILO, FAO, WHO, UNIDO, UNITAR and OECD

World Health Organization

This report contains the collective views of an international group of experts and does not necessarily represent the decisions or the stated policy of the United Nations Environment Programme, the International Labour Organization or the World Health Organization.

Environmental Health Criteria 239

PRINCIPLES FOR MODELLING DOSE–RESPONSE FOR THE RISK ASSESSMENT OF CHEMICALS

First draft prepared by the WHO Task Group on Environmental Health Criteria on Principles for Modelling Dose–Response for the Risk Assessment of Chemicals

Published under the joint sponsorship of the United Nations Environment Programme, the International Labour Organization and the World Health Organization, and produced within the framework of the Inter-Organization Programme for the Sound Management of Chemicals.

World Health Organization

The **International Programme on Chemical Safety (IPCS)**, established in 1980, is a joint venture of the United Nations Environment Programme (UNEP), the International Labour Organization (ILO) and the World Health Organization (WHO). The overall objectives of the IPCS are to establish the scientific basis for assessment of the risk to human health and the environment from exposure to chemicals, through international peer review processes, as a prerequisite for the promotion of chemical safety, and to provide technical assistance in strengthening national capacities for the sound management of chemicals.

The **Inter-Organization Programme for the Sound Management of Chemicals (IOMC)** was established in 1995 by UNEP, ILO, the Food and Agriculture Organization of the United Nations, WHO, the United Nations Industrial Development Organization, the United Nations Institute for Training and Research and the Organisation for Economic Co-operation and Development (Participating Organizations), following recommendations made by the 1992 UN Conference on Environment and Development to strengthen cooperation and increase coordination in the field of chemical safety. The purpose of the IOMC is to promote coordination of the policies and activities pursued by the Participating Organizations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

WHO Library Cataloguing-in-Publication Data

Principles for modelling dose-response for the risk assessment of chemicals.

(Environmental health criteria ; 239)

1.Chemicals. 2.Dose-response relationship, Drug. 3.Dose-response relationship, Radiation. 4.Risk assessment. 5.Environmental exposure. I.World Health Organization. II.Inter-Organization Programme for the Sound Management of Chemicals. III.Series.

ISBN 978 92 4 157239 2 (NLM classification: QV 38)
ISSN 0250-863X

© World Health Organization 2009

All rights reserved. Publications of the World Health Organization can be obtained from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 3264; fax: +41 22 791 4857; e-mail: bookorders@who.int). Requests for permission to reproduce or translate WHO publications – whether for sale or for noncommercial distribution – should be addressed to WHO Press, at the above address (fax: +41 22 791 4806; e-mail: permissions@who.int).

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use.

The named authors alone are responsible for the views expressed in this publication.

This document was technically and linguistically edited by Marla Sheffer, Ottawa, Canada, and printed by Wissenchaftliche Verlagsgesellschaft mbH, Stuttgart, Germany.

CONTENTS

ENVIRONMENTAL HEALTH CRITERIA ON PRINCIPLES FOR MODELLING DOSE–RESPONSE FOR THE RISK ASSESSMENT OF CHEMICALS

PREAMBLE ix

PREFACE xvii

ACRONYMS AND ABBREVIATIONS xix

1. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS 1

1.1 Summary 1
1.2 Conclusions 5
1.3 Recommendations 6

2. INTRODUCTION 9

2.1 Background 10
2.2 Scope 10

3. RISK ANALYSIS 13

3.1 Decision paradigms 13
3.2 Risk analysis paradigms 13
3.3 Motivations and considerations for producing a formal risk assessment 15
3.3.1 Transparency and justification 15
3.3.2 Public health and individual health 16
3.3.3 Quantification and computation 17
3.3.4 Cost of assessment 17
3.4 Risk assessment 17
3.4.1 Problem formulation 18
3.4.1.1 Defining the question 19
3.4.1.2 Prior knowledge 19
3.4.1.3 Desired outcomes 19
3.4.2 Risk assessment outcomes 20

4. DOSE–RESPONSE MODELLING: BASIC CONCEPTS 22

4.1 Introduction 22
4.2 What is dose? 23
4.3 What is response? 24
4.4 What is a model? 25
4.5 What is dose–response modelling? 27
4.6 Risk versus safety in dose–response modelling 31
4.7 Summary 33

5. DOSE–RESPONSE MODELLING: WHY AND WHEN TO USE IT 34

5.1 Historical perspectives 34
5.1.1 The no-observed-adverse-effect level approach to acceptable/tolerable daily intake 35
5.1.2 The benchmark dose approach to acceptable/tolerable daily intake 38
5.2 Points of consideration 40
5.2.1 General aspects of definition 41
5.2.2 Estimation procedure 42
5.2.3 Uncertainty 43
5.2.4 Study design 44
5.2.5 Biological information 46
5.2.6 Comparison of experimental results 46
5.2.7 Risk management perspectives 47
5.3 Implementation issues 47
5.4 Summary 48

6. PRINCIPLES OF DOSE–RESPONSE MODELLING 49

6.1 Data 49
6.1.1 Selection of data 49
6.1.2 Data types 49
6.2 Models and distributions 51
6.2.1 Dose–response models 51
6.2.1.1 Continuous dose–response models 51
6.2.1.2 Quantal dose–response models 53
6.2.1.3 Thresholds 53
6.2.1.4 Severity (degree of effect) 55
6.2.1.5 Modelling with covariates 57

6.2.1.6 Biologically based dose–response models 57
6.2.2 Statistical distributions 59
6.2.2.1 Continuous distributions 59
6.2.2.2 Discrete distributions 60
6.3 Model fitting and estimation of parameters 61
6.3.1 Criterion function 62
6.3.2 Search algorithms 62
6.4 Model comparison 63
6.5 Representing uncertainty 65
6.5.1 Sampling error 65
6.5.2 Study error 66
6.5.3 Model error 66
6.6 Benchmark dose and benchmark response selection 73
6.7 Summary 76

7. COMMUNICATING THE RESULTS OF DOSE–RESPONSE MODELLING 78

7.1 Introduction 78
7.2 Incorporation of the outputs of dose–response modelling into risk assessments 80
7.3 Derivation of health-based guidance values 81
7.4 Estimation of the margin of exposure 82
7.5 Quantitative estimations of the magnitude of the risk at levels of human exposure 83
7.6 Presentation of results 84
7.6.1 Tables 84
7.6.2 Graphs 84
7.7 Risk assessment context and questions 89
7.8 Synopsis of approach to modelling 89
7.8.1 Data sets 90
7.8.2 Uncertainty 90
7.9 Explaining/interpreting the output of the dose–response analysis 91
7.9.1 Outputs in the observable biological range 91
7.9.1.1 Health-based guidance values 92
7.9.1.2 Margin of exposure 92
7.9.2 Outputs outside the observable biological range 94
7.9.2.1 Prediction of risks at specified exposure levels 95

7.9.2.2 Prediction of exposure levels producing specified risk levels 96
7.9.2.3 Uncertainty analyses 96
7.10 Issues for risk managers 97
7.10.1 Risk assessment issues 97
7.10.1.1 Population versus individual effects 97
7.10.1.2 Risk characterization 98
7.10.2 Risk management issues 98
7.10.2.1 Risk management options 98
7.10.2.2 Cost–benefit and risk–benefit analyses 99
7.10.2.3 Acceptable level of risk 99

8. CONCLUSIONS AND RECOMMENDATIONS 101

8.1 Conclusions 101
8.2 Recommendations 102

REFERENCES 103

ANNEX 1: TERMINOLOGY 111

RESUME, CONCLUSIONS ET RECOMMANDATIONS 123

RESUMEN, CONCLUSIONES Y RECOMENDACIONES 130

NOTE TO READERS OF THE CRITERIA MONOGRAPHS

Every effort has been made to present information in the criteria monographs as accurately as possible without unduly delaying their publication. In the interest of all users of the Environmental Health Criteria monographs, readers are requested to communicate any errors that may have occurred to the Director of the International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland, in order that they may be included in corrigenda.

Environmental Health Criteria

PREAMBLE

Objectives

In 1973, the WHO Environmental Health Criteria Programme was initiated with the following objectives:

(i) to assess information on the relationship between exposure to environmental pollutants and human health, and to provide guidelines for setting exposure limits;
(ii) to identify new or potential pollutants;
(iii) to identify gaps in knowledge concerning the health effects of pollutants;
(iv) to promote the harmonization of toxicological and epidemiological methods in order to have internationally comparable results.

The first Environmental Health Criteria (EHC) monograph, on mercury, was published in 1976, and since that time an ever-increasing number of assessments of chemicals and of physical effects have been produced. In addition, many EHC monographs have been devoted to evaluating toxicological methodology, e.g. for genetic, neurotoxic, teratogenic, and nephrotoxic effects. Other publications have been concerned with epidemiological guidelines, evaluation of short-term tests for carcinogens, biomarkers, effects on the elderly, and so forth.

Since its inauguration, the EHC Programme has widened its scope, and the importance of environmental effects, in addition to health effects, has been increasingly emphasized in the total evaluation of chemicals.

The original impetus for the Programme came from World Health Assembly resolutions and the recommendations of the 1972 UN Conference on the Human Environment. Subsequently, the work became an integral part of the International Programme on Chemical Safety (IPCS), a cooperative programme of WHO, ILO, and UNEP. In this manner, with the strong support of the new

partners, the importance of occupational health and environmental effects was fully recognized. The EHC monographs have become widely established, used, and recognized throughout the world.

The recommendations of the 1992 UN Conference on Environment and Development and the subsequent establishment of the Intergovernmental Forum on Chemical Safety with the priorities for action in the six programme areas of Chapter 19, Agenda 21, all lend further weight to the need for EHC assessments of the risks of chemicals.

Scope

Two different types of EHC documents are available: 1) on specific chemicals or groups of related chemicals; and 2) on risk assessment methodologies. The criteria monographs are intended to provide critical reviews on the effect on human health and the environment of chemicals and of combinations of chemicals and physical and biological agents and risk assessment methodologies. As such, they include and review studies that are of direct relevance for evaluations. However, they do not describe *every* study carried out. Worldwide data are used and are quoted from original studies, not from abstracts or reviews. Both published and unpublished reports are considered, and it is incumbent on the authors to assess all the articles cited in the references. Preference is always given to published data. Unpublished data are used only when relevant published data are absent or when they are pivotal to the risk assessment. A detailed policy statement is available that describes the procedures used for unpublished proprietary data so that this information can be used in the evaluation without compromising its confidential nature (WHO (1990) Revised Guidelines for the Preparation of Environmental Health Criteria Monographs. PCS/90.69, Geneva, World Health Organization).

In the evaluation of human health risks, sound human data, whenever available, are preferred to animal data. Animal and in vitro studies provide support and are used mainly to supply evidence missing from human studies. It is mandatory that research on human subjects is conducted in full accord with ethical principles, including the provisions of the Helsinki Declaration.

The EHC monographs are intended to assist national and international authorities in making risk assessments and subsequent risk management decisions. They represent a thorough evaluation of risks and are not, in any sense, recommendations for regulation or standard setting. These latter are the exclusive purview of national and regional governments.

Procedures

The following procedures were followed in the development and publication of this EHC. A designated IPCS Staff Member (Dr Sam Page and subsequently Dr A. Tritscher), responsible for the scientific content of the document, served as the Responsible Officer (RO). The IPCS editor was responsible for layout and language.

The WHO Planning Group for the IPCS Harmonization Project on Dose–Response Modelling met on 10 October 2002 in Geneva to develop an outline and proposed time frame for the project. A first draft working paper, including contributions from several additional authors, was prepared by Drs C. Carrington and M. Bolger and distributed to the Task Group prior to the Task Group meeting, which was held from 13 to 17 September 2004. The first draft working paper was revised during the Task Group meeting and during a subsequent internal Task Group Internet forum. This revised draft was available on the IPCS web site for external review and comment. Comments received are available on request from the WHO Secretariat. They were reviewed by the Task Group, and necessary additions and revisions to the document were made.

The Task Group members serve as individual scientists, not as representatives of any organization, government, or industry. All individuals who as authors, consultants, or advisers participate in the preparation of the EHC monograph must, in addition to serving in their personal capacity as scientists, inform the WHO Secretariat if at any time a conflict of interest, whether actual or potential, could be perceived in their work. They are required to sign a declaration of interest statement. The Chairpersons of Task Groups are briefed on their role and responsibility in ensuring that these rules are followed. Such a procedure ensures the transparency and probity of the process. Their function is to evaluate the accuracy, significance, and relevance of the information in the document. A

summary and recommendations for further research and improved safety aspects are also required. The composition of the Task Group is dictated by the range of expertise required for the subject of the meeting and, where possible, by the need for a balanced geographical distribution.

WHO PLANNING GROUP FOR THE IPCS HARMONIZATION PROJECT ON DOSE–RESPONSE MODELLING

Members

Dr P.M. Bolger, Food and Drug Administration, College Park, MD, United States of America (USA)

Professor E. Dybing, Norwegian Institute of Public Health, Oslo, Norway

Dr L. Edler, German Cancer Research Center, Heidelberg, Germany

Dr M. Hartley, National Industrial Chemicals Notification and Assessment Scheme (NICNAS), Sydney, Australia

Dr A. Knaap, National Institute of Public Health and the Environment (RIVM), Bilthoven, Netherlands

Dr C. Portier, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

Professor A. Renwick, University of Southampton, Southampton, United Kingdom

Secretariat

Dr T. Damstra, International Programme on Chemical Safety, World Health Organization, Research Triangle Park, NC, USA

Dr S. Page, International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland

Ms C. Sonich-Mullin, International Programme on Chemical Safety, World Health Organization, Cincinnati, OH, USA

Ms C. Vickers, International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland

WHO TASK GROUP ON ENVIRONMENTAL HEALTH CRITERIA ON PRINCIPLES FOR MODELLING DOSE–RESPONSE FOR THE RISK ASSESSMENT OF CHEMICALS

Dr S. Page and Dr A. Tritscher, IPCS, served as the Responsible Officer (RO) and were responsible for the preparation of the final document and for its overall scientific content. Marla Sheffer, Ottawa, Canada, was the IPCS editor responsible for layout and language.

* * *

Risk assessment activities of IPCS are supported financially by the Department of Health and Department for Environment, Food & Rural Affairs, Food Standards Agency, United Kingdom; Environmental Protection Agency, Food and Drug Administration, and National Institute of Environmental Health Sciences, USA; European Commission; German Federal Ministry of Environment, Nature Conservation and Nuclear Safety; Health Canada; Japanese Ministry of Health, Labour and Welfare; and Swiss Agency for Environment, Forests and Landscape.

* * *

Task Group members

Dr P.M. Bolger, Food and Drug Administration, College Park, MD, USA

Professor A. Boobis, Imperial College London, London, United Kingdom

Dr C. Carrington, Food and Drug Administration, College Park, MD, USA

Dr V. Cogliano, International Agency for Research on Cancer, Lyon, France

Professor E. Dybing, Norwegian Institute of Public Health, Oslo, Norway

Dr L. Edler, German Cancer Research Center, Heidelberg, Germany

Professor E. Faustman, University of Washington, Seattle, WA, USA

Dr M. Healy, Food Standards Australia New Zealand, Canberra, Australia

Mr J. Howlett, Food and Agriculture Organization of the United Nations, Rome, Italy

Dr J. Kleinberg, European Food Safety Authority, Brussels, Belgium

Dr A. Knaap, National Institute of Public Health and the Environment (RIVM), Bilthoven, Netherlands

Dr J. Larsen, Danish Institute of Food and Veterinary Research, Solberg, Denmark

Dr D. Lovell, University of Surrey, Surrey, United Kingdom

Dr C. Portier, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

Professor A. Renwick, University of Southampton, Southampton, United Kingdom

Professor T. Sanner, Institute for Cancer Research, Oslo, Norway

Dr J. Schlatter, Swiss Federal Office of Public Health, Zurich, Switzerland

Dr R. Setzer, Jr, Environmental Protection Agency, Research Triangle Park, NC, USA

Professor W. Slob, National Institute of Public Health and the Environment (RIVM), Bilthoven, Netherlands

Dr A. Wadge, Food Standards Agency, London, United Kingdom

Professor G. Williams, New York Medical College, Valhalla, NY, USA

Secretariat

Dr S. Page, International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland

Dr A. Tritscher, International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland

Other contributors

Professor E. Calabrese, University of Massachusetts, Amherst, MA, USA

Dr I. Dewhurst, Pesticides Safety Directorate, York, United Kingdom

Dr D. Gaylor, Gaylor and Associates, Eureka Springs, AR, USA

Professor P. Grandjean, Environmental Medicine, University of Southern Denmark, Odense, Denmark

Dr I. Mangelsdorf, Fraunhofer Institute of Toxicology and Experimental Medicine, Hanover, Germany

Dr S. Sand, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

Dr K. Victorin, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

PREFACE

The International Programme on Chemical Safety (IPCS) was initiated in 1980 as a collaborative programme of the United Nations Environment Programme (UNEP), the International Labour Organization (ILO), and the World Health Organization (WHO). One of the major objectives of IPCS is to improve scientific methodologies for assessing the effects of chemicals on human health and the environment. As part of this effort, IPCS publishes a series of monographs, called Environmental Health Criteria (EHC) documents, that evaluate the scientific principles underlying methodologies and strategies to assess risks from exposure to chemicals.

This EHC is part of the ongoing review of the underlying scientific bases for decision-making in chemical risk assessment by IPCS. It involves specific consideration of the area of dose–response assessment in the evaluation of information from toxicological studies in animals and from human clinical and epidemiological studies. It covers toxicants with threshold effects and those for which there may be no practical threshold, such as substances that are genotoxic and carcinogenic. The discussions are concerned with that subset of cause–effect relationships commonly referred to as dose–response models, which are typically used to characterize the biological effects of intentional (e.g. drugs and nutrients) and unintentional (e.g. contaminants) exposure to chemicals.

This EHC is intended primarily to provide descriptive guidance for risk assessors in using dose–response modelling in hazard characterization. It will also provide mathematical modellers with an appreciation of issues to be considered when modelling in the context of the risk assessment process. Risk managers will be able to obtain a general understanding of the applications and limitations of dose–response modelling. For both risk assessors and risk managers, some considerations for communicating the results of risk assessments that use dose–response modelling are presented.

The efforts of all who helped in the preparation, review, and finalization of the monograph are gratefully acknowledged. Special thanks are due to Health Canada, the Ministry of Health of Japan, the United Kingdom Food Standards Agency and the United States

National Institute of Environmental Health Sciences for their financial support of the project.

ACRONYMS AND ABBREVIATIONS

ADI	acceptable daily intake
AIC	Akaike's information criterion
ALT	alanine aminotransferase
BMD	benchmark dose
BMD_{10}	benchmark dose at 10% risk
BMDL	lower confidence limit on the benchmark dose
BMDS	Benchmark Dose Software (United States Environmental Protection Agency)
BMR	benchmark response
CDF	cumulative distribution function
CSAF	chemical-specific adjustment factor
DDT	dichlorodiphenyltrichloroethane
DNA	deoxyribonucleic acid
DRM	dose–response modelling
EHC	Environmental Health Criteria
ED_{10}	effective dose for a 10% risk
EPI	exposure potency index
F	Frequency
FAO	Food and Agriculture Organization of the United Nations
f_x	dose–response function
f(x)	dose–response function
IPCS	International Programme on Chemical Safety
JECFA	Joint FAO/WHO Expert Committee on Food Additives
LD_{50}	median lethal dose
LED_{10}	lower bound on the effective dose resulting in a 10% increase in risk (ED_{10})
LL	log-likelihood
LOAEL	lowest-observed-adverse-effect level
MOE	margin of exposure
NOAEL	no-observed-adverse-effect level
NOEL	no-observed-effect level
RfD	reference dose
SD	standard deviation

T_{25}	chronic daily dose that gives 25% of the animals tumours above background at a specific tissue site
TDI	tolerable daily intake
UF	uncertainty factor
WHO	World Health Organization

1. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

1.1 Summary

Dose–response modelling (DRM), for use in quantitative risk assessment and ultimately for informing public health decisions about chemical exposures, can be described as a six-step process. The first four steps—data selection, model selection, statistical linkage, and parameter estimation—constitute dose–response analysis. These steps relate to the process through which a mathematical description of the data is obtained in order to evaluate predicted responses for known doses or to obtain dose estimates when a chosen response is of interest. The fifth step involves the integration of the results of the dose–response analysis with estimates of exposure for the purposes of guiding public health decisions. The final step, which can optionally be applied earlier in DRM, involves an assessment of the quality of the dose–response analysis and the sensitivity of model predictions to the assumptions used in the analysis.

The characterization of dose–response relationships in animal and human studies has been a major component of hazard characterization and has been used in the extrapolation of incidences of adverse effects in the range of human exposure levels. Over the years, a variety of methods have been developed to accommodate such relationships, for improving extrapolation to low doses and deriving health-based guidance values, such as acceptable daily intakes (ADIs), tolerable daily intakes (TDIs), and reference doses (RfDs). DRM may prove useful in risk assessments for making better use of available data and for providing tools to evaluate the quality of data and the ensuing uncertainties in dose–response estimates.

In general, DRM estimates are based on data from the entire dose–response curve for the critical effect. The standard no-observed-adverse-effect level (NOAEL) approach can be regarded as a special, simplified case of dose–response analysis, as it identifies a single dose that is assumed to be without an appreciable

adverse effect. The dose–response model reflects the characteristics of the dose–response curve, particularly in providing estimates of the slope. In the case of a regression framework, it provides standard error and confidence intervals for the model parameters. A disadvantage of using the NOAEL approach is that it is not possible to quantify the degree of variability and uncertainty that may be present, whereas other dose–response models can facilitate the analysis of sensitivity and uncertainty. Consideration of a dose–response model can optimize study design and clarify the need for additional studies. The NOAEL approach incorporates biological information through the application of "expert" but subjective judgement. Full DRM has the potential for a more "science-rich" analysis through the more formal quantitative inclusion of, for example, factors and covariates into the models. Estimates derived from DRM enhance the ability to compare quantitatively different experiments, effects, and compounds within a common framework. DRM can enhance risk and safety assessments as well as provide opportunities to consider the likelihood of effects outside the observable range.

The choice of the models to be used depends upon the type of data. The models should include a model for dose–response and a model for the variability of the data. Once models are fit to a data set, the degree to which they individually describe the data can be evaluated using goodness-of-fit measures. In addition, their ability to describe the data with respect to each other may be compared. Uncertainties about the inferences that result from such models fall into four main categories: statistical uncertainty of inferences due to variability among the responses of experimental subjects, experimental errors (e.g. imperfect randomization, dosing errors, unfavourable dose location), variability among experiments due to unavoidable differences in experimental execution, and uncertainty due to the fact that the "true model" for the data is unknown. Dose–response analysis needs to address all four sources of variability and uncertainty whenever possible.

One particularly important application of DRM is the calculation of benchmark doses (BMDs). These are doses at which it is inferred that a particular level of response would occur. When appropriate data are available, BMDs are an alternative to the NOAEL approach in the calculation of health-based guidance values. When extrapolation is necessary, the uncertainty associated

with a prediction should be represented. Here it is especially important to include model uncertainty.

Full DRM offers the potential to provide additional information for the risk manager. The output of DRM should be directed towards addressing specific questions about the likelihood of adverse health effects. It can be presented in three principal ways. Firstly, it can be used for the establishment of a health-based guidance value, such as an ADI, TDI, or RfD, analogous to current procedures based on a NOAEL or lowest-observed-adverse-effect level (LOAEL). DRM can be a more scientifically robust method for determining health-based guidance values. Secondly, the output from DRM can be used in risk management to estimate a margin of exposure (MOE), by calculation of the ratio of the dose corresponding to a given limit of response to a human exposure level. Thirdly, based on the modelled dose–response relationship, the output can be a quantitative estimation of the magnitude of the risk/health effect at the level of human exposure, with the generally accepted assumption that the uncertainty factors used cover the uncertainties about differences in sensitivity between individuals and species. DRM can provide better information on the likelihood of effects at low doses that are below the levels observed in biological systems and can also provide better estimates of the statistical uncertainties surrounding estimates of likely effects.

Two factors that can impact the type of outputs from DRM exercises and that may be of importance to the risk manager are multiple data sets and uncertainties. DRM can be used with exposure data to identify subpopulations at risk. DRM can also be used to assist risk managers in determining priorities and evaluating the consequences of proposed interventions aimed at reducing the risk. For risk communication, the use of DRM techniques offers opportunities and challenges. DRM evaluations can produce information in several formats, including dose–response functions that allow, along with estimates of exposure, the prediction of risks at specified exposure levels and functions that allow the estimation of exposure levels resulting in specified risks. This includes estimates of the possible risk at intakes above a health-based guidance value, such as an ADI. DRM evaluations also offer approaches to compare competing risks or benefits and provide a focus on uncertainties that can influence the predicted risk.

However, unless the situation of risk is viewed at the population level, there is the risk communication problem that while explaining the level of risk in those circumstances where there is no safe level of exposure, some percentage of the population will be predicted to experience some effect deemed to be adverse. It must be recognized that the use of DRM requires a certain quantity and quality of data, as well as specific expertise.

The potential "ongoing" use of the estimates from DRM can, from a risk management perspective, give an improved characterization for decision-making by:

- providing information about what happens above the health-based guidance value (magnitude and types of health impacts);
- showing benefits from different regulatory actions;
- giving the decision-maker a "more-than-one-point" appreciation of the data;
- promoting consistency in decisions, if appropriate adjustments are made for differences in effect, effect level, species, and study design; and
- allowing for an iterative interaction between the risk assessor and risk manager on a continuous and ongoing basis.

The use of DRM and probabilistic assessment techniques to quantitatively describe variability and uncertainty brings new challenges in risk communication. Some of these challenges are:

- explaining that some percentage of the population is predicted to exceed the safety level and/or experience an adverse effect;
- explaining the level of risk in those circumstances where there is assumed to be no safe level of exposure;
- comparing competing risks or benefits;
- providing a focus on uncertainties that influence the predicted risk; and
- explaining that a risk estimate pertains to what may occur at a population level, rather than the individual level, and noting that this is also the case for the ADI/TDI approach.

1.2 Conclusions

- Full DRM can be considered a more sophisticated or robust alternative to the NOAEL approach in all cases where suitable dose–response data are available (e.g. several dose groups with different response levels).

- For quantal dose–response data, the interest is often in low response (incidence) levels. This may call for low-dose extrapolation by several orders of magnitude (e.g. for tumour incidences). However, equally plausible dose–response risk models may result in highly divergent low estimates. A currently applied approach is to estimate a BMD_{10} (dose at 10% risk) and linearly extrapolate from that point downwards, as a conservative approach. Another option, currently under development, is to apply a Bayesian approach that considers the various models all together.

- For continuous dose–response data, two approaches of DRM exist. One is to transform the continuous data into quantal data. The other is to consider continuous dose–response data as information on the severity of the effect and therefore as a function of dose. In the latter approach, measurable changes of effect are often close to response levels considered as adverse (e.g. 10% inhibition of cholinesterase), and the low-dose extrapolation problem is minor or non-existent.

- For the purpose of deriving an ADI, TDI, or RfD, DRM may be used for deriving a BMD, to be used as a point of departure in the same way as the NOAEL is used (i.e. the same uncertainty factors would be applied to the BMD as to the NOAEL).

- DRM may also be used for estimating risks at a given (human) exposure level. For risks in terms of incidences (quantal data), this may involve low-dose extrapolation.

- DRM exercises can provide information on uncertainties associated with the data and identify factors contributing to uncertainties in risk estimates.

- Application of DRM for all end-points can be cost prohibitive, so it is efficient to pre-select the apparently more sensitive end-points. In some cases, however, it is not easy to identify the most sensitive end-points by visual inspection, so all of the end-points may need to be modelled.

- The BMD and the lower confidence limit of the BMD (BMDL) should always be reported, so that the quality of the data and the model fit are clear and potencies can be compared on the basis of the BMD.

- The output of the different models used in DRM should be presented.

1.3 Recommendations

- Toxicity testing protocols (e.g. Organisation for Economic Co-operation and Development guidelines) should be reviewed for optimization for BMD and other DRM approaches, including optimal designs for the number of animals and number of doses for different dose–response curves. Additional research is needed for the development of optimal study designs. Guidance should be developed for combining existing studies with a view to DRM.

- Better guidance needs to be developed for combined analysis of different data sets for more precisely estimating BMDs.

- Better understanding of when and how to use the benchmark response (BMR) needs to be developed.

- Better understanding of the shape of the dose–response curve at low doses needs to be developed. Additional research is needed to determine the biological basis for extrapolation (e.g. by using biomarkers, tumour precursors, genetically modified animals, and toxicokinetics for target dose estimation).

- Improved guidance needs to be developed for risk communication based on the results of DRM and probabilistic assessment techniques. This should include communication of

the types of uncertainty and the relation to statistical variability, imprecision, and the use of confidence intervals.

- The use of DRM should be reviewed and additional general principles for its use developed when more experience becomes available.

2. INTRODUCTION

The International Programme on Chemical Safety (IPCS) and other public health organizations have recognized the importance of the harmonization of procedures to enhance the quality of risk assessments, to improve the transparency of the risk assessment process, and to facilitate risk communication.

Public health decisions on the plausible risks of chemical exposures can include several possible outcomes. The ultimate goal is to implement a risk management action that will produce the desired reduction of risk. Among the first objectives of a risk assessment is the determination of the presence or absence of a cause–effect relationship. If there is sufficient plausibility for the presence of such a relationship, then dose–response information is needed and will be subject to an analysis of a dose–response relationship.

Extrapolation is a fundamental problem in the quantitative health risk assessment of exposures to chemicals that produce toxicity in experimental systems. Adverse health effects of chemicals are, in the absence of human data, typically evaluated in laboratory animals at significantly higher doses than the levels to which humans may be exposed. Also, for certain substances for which the exposure can be controlled, such as food additives and residues of pesticides and veterinary drugs, the quantification of the risk above the level of exposure that has been assessed to be safe (e.g. the acceptable daily intake [ADI]) can be difficult. This is particularly true in cases of temporary excursions above an ADI.

The use of mathematical and statistical approaches in hazard characterization is increasing. Although dose–response models have been available for some time, their use has been somewhat limited because of a lack of either appropriate scientific information or agreed-upon approaches and methods for how to obtain and use available dose–response information appropriately. Dose–response modelling (DRM) involves a number of choices based upon scientific experience, data availability, and mathematical tractability and can take on many different forms and be used in many different

ways. A recent review of the available quantitative approaches for hazard characterization noted that mathematical modelling of the dose–response relationship could improve the risk assessment process (Edler et al., 2002).

Dose–response models may improve and generate more reliable predictions, but they can never be proved to be completely correct. Therefore, it is necessary to rely on scientific judgement to determine the utility of risk predictions from DRM in making public health decisions. It is important to remember that risk numbers derived from DRM can be misleading for a variety of reasons; like any other tool used in science, DRM needs to be utilized in a broader context of all of the available scientific knowledge. Although mathematical and statistical rigor are important factors in risk assessment, the final standard that prevails remains biological plausibility. It is this inherent uncertainty and its communication for which modelling and quantitative risk assessment can be particularly valuable.

2.1 Background

This Environmental Health Criteria report (EHC) is intended primarily to provide descriptive guidance for risk assessors in using DRM in hazard characterization. It will also provide mathematical modellers an appreciation of the issues to be considered when modelling in the context of the risk assessment process. Risk managers will be able to obtain a general understanding of the applications and limitations of DRM. For both risk assessors and risk managers, some considerations for communicating the results of risk assessments that use DRM are presented.

2.2 Scope

This EHC is part of the ongoing review of the underlying scientific bases for decision-making in chemical risk assessment by IPCS. It involves specific consideration of the area of dose–response assessment in the evaluation of information from toxicological studies in animals and from human clinical and epidemiological studies; it does not include consideration of other aspects of quantitative risk assessment, such as physiologically based modelling. It covers toxicants with threshold effects and those for which there may be no practical threshold, such as substances

that are genotoxic and carcinogenic. The discussions are concerned with that subset of cause–effect relationships commonly referred to as dose–response models, which are typically used to characterize the biological effects of intentional (e.g. drugs and nutrients) and unintentional (e.g. contaminants) exposure to chemicals. Dose–response models are also commonly used in microbiological risk assessments (e.g. WHO, 2004a).

This document focuses primarily on experimental animal studies. In DRM of human epidemiological data, several important issues should be considered:

- *Impact of imprecision of the dose estimate*. This issue differs substantially from the situation with experimental animal studies. In observational studies, where the dose is not a matter of design, this imprecision is likely to be substantial.

- *Absence of a true control group*. In many observational studies, there may not be any subjects who are completely free from exposure. The response at zero exposure cannot be observed and has to be estimated.

- *Shape of the dose–response curve at low doses*. The shape may depend on both the outcome parameter and the toxicant. For most contaminants, insufficient information is available, and the impact on uncertainties must therefore be considered.

- *Confounder adjustment*. In epidemiological studies, confounder adjustment must be included. Decisions therefore need to be made as to which confounders to include in the DRM.

- *Meta-analysis*. If more than one study is available, a meta-analysis can provide improved information on the dose–response models.

Many of the considerations in this EHC are also relevant to ecotoxicological studies.

This EHC is intended to provide guidance in a number of areas relevant to DRM. Initially, there is a discussion of the risk analysis paradigm (chapter 3) and the basic concepts of DRM (chapter 4). In

chapter 5, the use of DRM is described, including comparing the no-observed-adverse-effect level (NOAEL) approach with the benchmark dose (BMD) modelling method. Chapter 6 provides the principles of DRM, including data considerations, model descriptions, model fitting and parameter estimation, model comparisons, and uncertainty. This chapter also includes discussion of BMD approaches. Chapter 7 discusses the provision of scientific advice by risk assessors to the risk managers. This chapter includes an explanation of the output of the dose–response analysis and the strengths and weaknesses of DRM. The final chapter, chapter 8, summarizes the conclusions of the EHC and provides recommendations for future research.

There is only limited treatment of the mathematical and statistical considerations for DRM. References and links are provided for more in-depth treatments, modelling tools, and examples.

3. RISK ANALYSIS

3.1 Decision paradigms

A risk analysis decision paradigm is a formal representation of a process that distinguishes the scientific bases from the risk management objectives and generally contains a component in which the probability of harm is estimated. This component of the decision paradigm is referred to as the risk assessment. As a probability calculation, a risk assessment will include both a statement of the objective under consideration (i.e. the harm) and the basis for the assertion that the harm may occur (i.e. the probability).

3.2 Risk analysis paradigms

The first risk analysis paradigm for public health was proposed by the National Academy of Sciences of the United States of America (NRC, 1983) and focused on assessing the risk of cancer from exposure to chemicals in food. The decision process was divided into three major steps: research, risk assessment, and risk management. The risk assessment process was further divided into hazard identification, dose–response assessment, exposure assessment, and risk characterization. Risk management is the decision-making process involving the consideration of political, social, economic, and technical factors with relevant risk assessment information relating to a hazard so as to develop, analyse, select, and implement appropriate risk mitigation options. Risk management comprises three elements: risk evaluation, emission and exposure control, and risk monitoring.

In the National Academy of Sciences paradigm, the principal steps were considered to be sequential, with the decision process commencing with research and concluding with the decision. A drawback of this sequential concept is an absence of the recognition of the influence that the risk analysis might have on data collection or of the impact that political, social, and economic objectives may have on the need to identify the hazard.

More recent examinations of risk assessment/analysis methodology have paid much closer attention to the influence of risk management on the risk assessment process (NRC, 1994, 1996; Presidential Commission, 1997; Renwick et al., 2003). Rather than insist that management be insulated from the risk assessment process for the sake of preserving scientific objectivity, it is acknowledged that risk management should interact with risk assessment for the scope of the analysis, particularly in problem formulation. The focus on this interaction leads to the notion that the relationship between risk assessment and risk management is an interactive, often iterative, and circular process (see Figure 1).

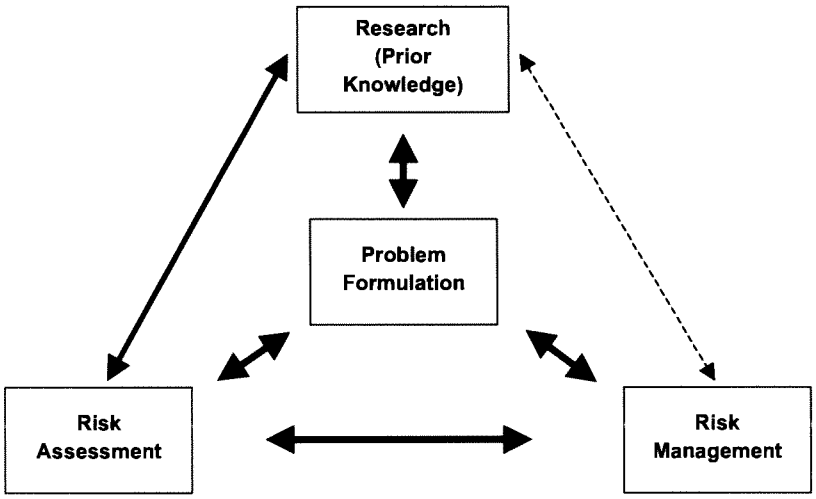


Fig. 1. Interactions of risk assessment with risk management.

As a general rule, formal risk assessments are preceded by preliminary risk assessments. These are usually subjective and informal and may be initiated from inside or outside the risk assessment and scientific communities. A key consideration of these preliminary risk assessments is whether or not a formal risk assessment is necessary. The transition process from preliminary assessments to formal risk assessments has been described as problem formulation (Renwick et al., 2003). It is an iterative process that facilitates the critical interface between risk assessment and risk management. Risk communication, with stakeholder involvement, is particularly essential during the problem formulation.

As the risk analysis paradigm evolved, the need for risk communication as an integral part was recognized (see Figure 2). Risk communication not only is the interactive exchange of information and opinions among risk assessors and risk managers, but necessarily includes all interested parties. The issues of risk, risk-related factors, and risk perceptions should involve interactive exchange throughout the risk assessment process. The communication of the results of the risk assessment as the basis of the risk management decisions demands transparency and appreciation for the uncertainties involved.

Fig. 2. The risk analysis paradigm.

3.3 Motivations and considerations for producing a formal risk assessment

There are several different reasons for preparing a formal risk assessment. The relative importance of these different motivations may influence the scope or the methodology used.

3.3.1 Transparency and justification

A major function of formal risk assessment is to serve as a transparent justification of a public health decision, whereby each step and assumption are clearly described. A key reason for undertaking such an assessment is to separate clearly scientific knowledge from values. Formal risk assessments are almost always

performed for notable public health issues where there is a wider interest in the political, social, and economic consequences of such assessments. Identifying the public health objectives before the technical analysis will allow participation in the debate of the other issues involved without necessarily requiring involvement in the scientific discussion. There may be areas of a risk assessment that can be obscure to someone not privy to their development. As a result, transparency is often audience dependent, relative to the level of comprehension and involvement. Since less can be taken for granted, the extent of the explanation required will increase as the audience broadens and its level of interest increases and sophistication decreases. Producing records of an assessment with varying degrees of technical detail may be a useful objective.

The World Trade Organization, under the Agreement on the Application of Sanitary and Phytosanitary Measures, has recognized the importance of harmonized science-based risk assessments. The World Trade Organization has specifically cited the standards, guidelines, and recommendations of the Codex Alimentarius Commission as reflecting international consensus regarding the requirements to protect human health from foodborne hazards. The Codex Alimentarius Commission has formally adopted the risk analysis paradigm in its decision-making (Codex Alimentarius Commission, 2003). Other organizations have also adopted this paradigm (European Commission, 2000).

3.3.2 Public health and individual health

A public health risk assessment is concerned with a population. The behavioural, environmental, or biological characteristics will vary among individuals in the population of concern. This variation is considered in probabilistic approaches and determines the statistical nature of health risk measures and conclusions made on populations. A risk assessment may need to describe or model these individual characteristics to produce a prediction of what might be expected to happen in the population. Specifying the population with which the risk assessment is concerned may be an important part of the problem formulation. In a public policy setting, the population will generally be defined by the risk managers, often in view of social, economic, and other considerations.

3.3.3 Quantification and computation

Public health issues often involve matters of degree, particularly in regard to level of exposure and risk, and may be defined by measures of quantity or statistical rates. If an uncertainty analysis is conducted, knowledge may be quantified as a matter of degree. Although judging matters of degree does not require the use of numbers, communication of degree does. Quantitative risk assessment approaches, including DRM, can be valuable in providing information to address these issues.

Formal risk assessments often involve the interaction of multiple quantitative measures that may lead to extensive and complicated calculations. Particularly in DRM, mathematical and statistical considerations are often complex. Although computers can carry out these calculations more accurately and quickly, knowledge of the scientific basis and experience with the applications of DRM are essential in order to avoid misinterpreting and incorrectly communicating the outcomes.

3.3.4 Cost of assessment

Risk assessments take time and effort to develop. The time and effort required will increase with the complexity of the problem and often with the degree of transparency that is required. The level of scientific detail addressed by the models and the level of documentation needed may vary with the nature and magnitude of the motivations for producing the risk assessment in the first place. In order to tailor the risk assessment to the decision problem, it may be desirable to develop the risk assessment by an iterative process that commences with the simplest possible statement of the problem and becomes more complicated as the risk assessment is developed.

3.4 Risk assessment

The risk assessment paradigm, incorporating problem formulation, is illustrated in Figure 3 (based on Renwick et al., 2003).

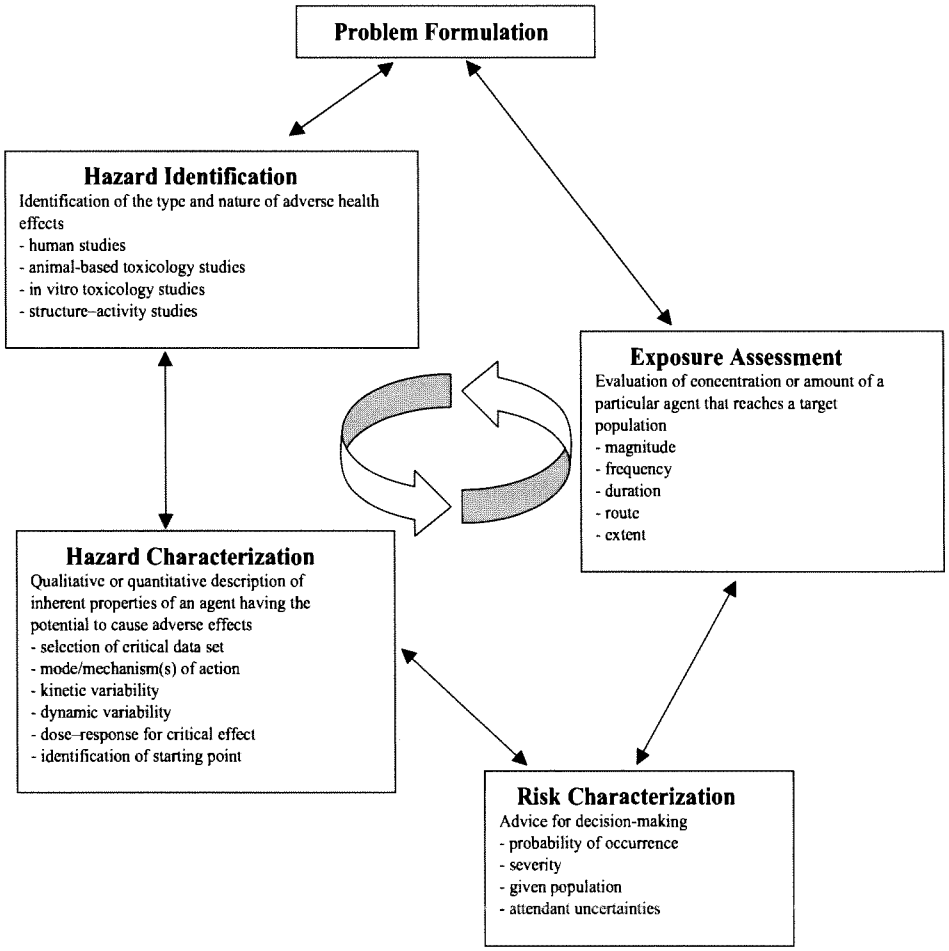


Fig. 3. Risk assessment (adapted from Renwick et al., 2003).

3.4.1 Problem formulation

Problem formulation is the initial phase in a risk assessment that determines if a detailed risk assessment is necessary and, if so, whether it is possible. Further, it serves as the transition from an informal risk assessment to a formal risk assessment. Problem formulation requires at least some preliminary consideration of the hazard identification, hazard characterization, and exposure assessment and usually proceeds in iterative stages. The output is a plan for the risk assessment process, which can be changed as the risk assessment progresses.

3.4.1.1 *Defining the question*

Among additional considerations are those that address who should be involved in the risk assessment and risk management processes. The transparency of a risk assessment will depend on how well these are described. It is not necessary to establish beyond all doubt that there is a cause–effect relationship in order to conduct a risk assessment. The suspicion that there may be such a relationship is sufficient. The consideration of the evidence for or against the supposition is often an integral part of the analysis. Identifying the problem may be politically controversial. That is, it may constitute a risk management issue that must be resolved before the risk assessment may be used as the justification for a decision. Non-scientific controversy may be diverted from the risk assessment by separating the valuation of the effect from the risk assessment per se (i.e. the risk assessment may be used as part of a cost–benefit analysis, but the cost–benefit analysis is not part of the risk assessment). Predicting the occurrence of an event is not part of an expression of the level of public health concern. However, suggesting that the problem is big enough to merit a formal risk assessment does imply that the risk may be of some significance.

3.4.1.2 *Prior knowledge*

Organizing information regarding public health issues that may involve many details and complex cause–effect relationships may benefit from the methodical collection and evaluation of prior knowledge of the agent, exposure to the agent, and possible biological effect(s) resulting from exposure to the agent. This is essential for determining the feasibility of a detailed assessment. Prior knowledge is also important for prioritizing and directing the risk assessment. Organization of information may also instigate and support specialization; different experts may produce or oversee different parts of the risk assessment. This information may in turn influence the conception of the problem (where management specifies the objective of the analysis) and also may influence additional research that may be needed.

3.4.1.3 *Desired outcomes*

The desired outcomes of the problem formulation are:

- explicit questions to be answered in the risk characterization to meet the needs of the risk manager;
- determination of the resources that are needed and available; and
- time frame for completing the assessment.

3.4.2 Risk assessment outcomes

The advice to risk managers that is formulated in the risk characterization may be qualitative or quantitative. Quantitative advice includes:

- health-based guidance values;
- estimates of the risks at different levels of exposure;
- exposure-based estimates used with low levels of exposure (e.g. threshold of toxicological concern); and
- risks at minimum and maximum intakes (e.g. nutrients).

Qualitative advice includes:

- statements/evidence that the agent is of no toxicological concern owing to the absence of toxicity even at high exposure levels (e.g. ADI "not specified");
- statements/evidence that the agent is "safe" in the context of a specified use; and
- recommendations for avoidance, minimization, or reduction of exposure.

Risk characterization should include all key assumptions and a clear explanation of the uncertainties in the risk assessment. It should also include information on susceptible subpopulations, including those with greater potential exposure and/or specific physiological conditions or genetic factors. At present, this is limited, and generic approaches have to be used (e.g. 10 × 10 uncertainty factors for interspecies differences and human variability). The advice to risk managers can be in the form of a comparison of the relative risks resulting from choosing different risk management options.

The risk assessment that is produced is followed by either a risk management decision or a request for further analysis, which may

influence the further research that is conducted. In one sense, the risk assessment process may never end. However, from a risk management standpoint, there is usually some imperative and timeline that conclude the process. Therefore, in another sense, the risk assessment ends when the risk management decision is made. The record produced by a risk assessment stands as a justification for a decision at the time the decision is made. However, with additional information, such as that which can reduce the uncertainties identified in the risk assessment, the risk assessment/analysis may be reopened. It is also possible that additional information can increase uncertainty.

4. DOSE–RESPONSE MODELLING: BASIC CONCEPTS

4.1 Introduction

Toxicology is the science of identifying and quantifying harmful or adverse effects of chemical and physical agents in the human environment. This can be accomplished by observations in humans (i.e. epidemiological and clinical studies), experimental studies using animal models (i.e. in vivo bioassays), or cellular and molecular studies. All these approaches have firmly established the principle of dose–response. Accordingly, dose–response toxicities of chemicals can be and have been expressed quantitatively (e.g. the median lethal dose, or LD_{50}).

However, scientific data alone are not sufficient to allow a decision to be made regarding the potential toxicity of chemicals and other agents that humans encounter; it is the analysis and interpretation of these data that lead to a scientifically supported decision regarding potential health effects. Many analytical processes have been developed to address the evaluation of the toxicities of chemicals, ranging from very simple approaches based solely upon the identification of the possibility of a hazard (NTP, 2002; Cogliano et al., 2004; USEPA, 2005) to much more complicated approaches incorporating biological mechanisms, complicated mathematical models, bioavailability in humans, and direct predictions of chemically induced changes in disease incidence in the affected human population (Portier & Kohn, 1996; Kim et al., 2002). All of these methods have two basic steps in common: analysis of the dose–response information and implementation of the results of that analysis to formulate a conclusion. The combined two-step approach will be referred to as DRM.

This chapter describes the elements that embody DRM. Most of the information presented is found in more extensive detail in other chapters of this guidance document. This chapter sets the stage for discussion of dose/exposure–response modelling by briefly

answering the questions: What is dose? What is response? What is a model? It then goes on to introduce the reader to the types of data and information that may have an impact on the development of dose–response models.

4.2 What is dose?

It is critical when performing dose–response analyses to have a clear concept of what is meant by "dose" and how it applies to the response. There are three basic types of "dose" that arise from scientific investigations: the administered or external dose, the internal (absorbed) dose, and the target or tissue dose. External dose denotes the amount of a chemical or other agent administered to an experimental animal or human in a controlled experimental setting by some specific route at some specific frequency. In the terminology used by the Joint FAO/WHO Expert Committee on Food Additives (JECFA), intake (dietary exposure) refers to external dose. Internal dose is the amount determined by toxicokinetics to be systemically available. It is a consequence of absorption, distribution, metabolism, and excretion of the chemical. The tissue dose is the amount that is distributed to and present in a specific tissue of interest. The three are, of course, related, and each can be used to express dose–response.

Two other parameters are important: the dose frequency and duration of dosing. Dosing can be acute, subchronic, or chronic. For simplicity, the term dose in DRM will be used as an inclusive term referring to all three forms of dose described above. In general, units of dose should reflect the magnitude, frequency, and duration over which it applies. Dose can be expressed in a multitude of metrics. Some of these metrics include daily intake (e.g. ng/kg body weight per day), total body burden (e.g. ng/kg body weight), body burden averaged over a given period of time, or tissue concentration (ng/kg).

For humans, where dosing of xenobiotics is not intentional, the term exposure is used for the external dose. In epidemiological studies, exposure is rarely known, and best estimates are made using several assumptions and/or biomonitoring of tissue (usually blood) concentrations at very few time points, often many years

after what is believed to be the period of first/highest exposure. Sometimes, when laboratory animals are used for DRM, the dose used in the animal study is transformed to an equivalent human exposure prior to modelling. Exposure assessment is the qualitative and/or quantitative evaluation of the likely intake of chemical agents via food, as well as exposure from other sources, if relevant (WHO, 1997). In this situation, models of exposure linked to response data may be used to develop a dose–response model. However, limited knowledge of the events controlling absorption and tissue distribution (especially in humans at low levels of exposure), metabolism, and excretion and the other molecular and biochemical processes that ultimately lead to particular responses contribute to the uncertainty in these analyses.

4.3 What is response?

Response, in this context, generally relates to an observation or effect seen in a laboratory cell culture, an animal, or a human following exposure. These end-points cover a broad range of observations, from early responses, such as biochemical alterations, to more complicated responses, such as cancer and developmental defects. Responses can be either adaptive or adverse (e.g. Williams & Iatropoulos, 2002). The latter are defined as a change in the morphology, physiology, growth, development, reproduction, or lifespan of an organism or subsystem (e.g. subpopulation of cells) 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. These are critical responses that are likely to underlie an adverse health effect in humans. The responses are sometimes species and/or tissue specific and have different degrees of variation across individuals. Nevertheless, there is some commonality across species, and there are known linkages between some responses (e.g. DNA damage is a precursor for mutations). DRM can address each response, provide insight into their quantitative similarity across species and tissues, and link responses in a mechanistically reasonable manner.

Response is generally considered to vary across experimental units (animals, humans, cell cultures) in the same dose group in a random fashion. This random variation is usually assumed to follow some statistical distribution describing the frequency of any given

response for a population. In general, statistical distributions are characterized by their central tendency (usually the mean or average value) and their effective range (usually based on the standard deviation). Most responses of interest in the context of dose–response assessment fall into one of four basic categories:

- *Quantal responses.* Quantal responses generally relate to the number of experimental units responding in a given period of time (e.g. the proportion of animals with a tumour in a cancer bioassay).

- *Counts.* Count data generally relate to a discrete number of items measured in a single experimental unit (e.g. number of papillomas on the skin).

- *Continuous measures.* Continuous measures generally take on any value in a defined range (e.g. body weight).

- *Ordered categorical measures.* Ordinal categorical measures generally take on one value from a small set of ordered values (e.g. tumour severity grades).

Sometimes it is useful to convert continuous data into proportions (e.g. number of animals outside a clinically relevant range for an immune system marker) or categories (e.g. measured degree of liver necrosis converted to minimal, moderate, or extensive).

For each of these different data types, there will be some differences in how they will be handled for DRM; as a general rule, however, the goal of DRM is to describe the mean and variance of the response as a function of exposure and/or time.

4.4 What is a model?

Dose–response models are mathematical models used to characterize the relationship between dose and response for a given set of scientific data. Mathematical models consist of three basic components: assumptions used to derive the model, a functional form for the model, and parameters that are components of the

functional form. For example, the simplest dose–response model is a linear model to describe a continuous response (see Figure 4).

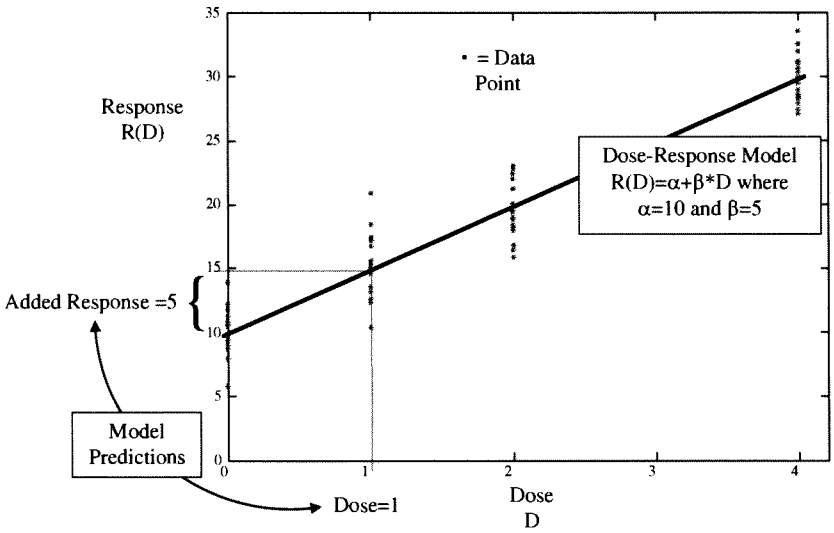


Fig. 4. Dose–response illustration displaying a linear model fit to continuous data for which prediction of the dose associated with an added response of 5 units (not designated) is a dose of 1 unit (not designated).

For this model, the key components are:

- *Assumptions*: Mean added response is proportional to dose.
- *Functional form*: $R(D) = \alpha + \beta \cdot D$, where R(D) is the mean response as a function of dose, denoted D.
- *Parameters*: α is a parameter describing the mean response in the control (unexposed) group, and β is a parameter describing the mean change in response per unit dose.

Dose–response models range from very simple models, such as the linear model described above, to extremely complicated models for which the eventual functional form cannot easily be expressed as a single equation (e.g. biologically based dose–response models). Models can also be linked, meaning that one model could describe part of the dose–response process while another describes the remainder of the process. For example, in most cases for chemical carcinogenesis, cancer risk is more closely linked to tissue

concentration than to administered dose. Given data on dose, tissue concentration, and tumour response, one can use a toxicokinetic model to relate dose to tissue concentration and use a multistage cancer model to relate tissue concentration to response. The two models combined are needed to describe the dose–response relationship.

Dose–response models may incorporate other information into the model form. Age and time-on-study are commonly used in DRM, but other factors, such as species/strain/human ethnicity, sex, body weight, etc., have also been used to expand the utility of dose–response models.

4.5 What is dose–response modelling?

DRM can be described by six basic steps, with a variety of options at each step (Table 1). The first four steps, which will be referred to as dose–response analysis, are aimed at the analysis of the data available for DRM. Dose–response analysis provides the linkage of a model to dose–response data for the purposes of predicting response to a given dose or predicting dose from a given response. The last two steps deal with implementation and evaluation of the analysis results.

Table 1. Basic steps in dose–response modelling

Step	Description	Options	Section links for chapter 6
1. Data selection	Determine the response to be modelled, and select appropriate data	End-point, data quality, sample size, data utility, data availability	6.1
2. Model selection	Choose the type of model to be applied to the data	End-point, data availability, purpose	6.2.1
3. Statistical linkage	Assumes that statistical distributions describe the response	End-point, data type, model choice, software availability	6.2.2
4. Parameter estimation	Combine the first three steps in an appropriate computer program to obtain estimates of the model parameters	Linkage function, software availability, variance	6.3
5. Implementation	Use the estimated model parameters and the model	Outputs, target selection, model predictions, BMD,	6.3

Step	Description	Options	Section links for chapter 6
	formula to predict response/dose as needed	direct extrapolation	
6. Evaluation	Examine the sensitivity of the resulting predictions to the assumptions used in the analysis	Model comparison, uncertainty	6.4, 6.5, 6.6

Step 1 involves selection of the appropriate data for DRM. The type of data available can have a marked impact on the complexity of the model that can be used. For example, whereas two points can be used to identify the slope of a line, it takes at least three points to identify the shape of a more complex dose–response relationship (e.g. straight line versus two connected lines). The issue of whether there are enough data to support a given model is quite complex (Portier, 1994) and is discussed in greater detail in section 6.1. In general, the data can restrict the type of model that can be used.

The second step is then to choose an appropriate model. Many choices exist for modelling dose–response data, and examples of some of the possible choices are presented in chapter 6. These models have been generally divided into two categories: empirical and biologically based models. Empirical models generally refer to functional forms for which there is limited mechanistic justification (e.g. the linear model above). Most of the DRM that has been done to date has focused on the use of empirical models. Biologically based models generally have functional forms that are derived from some basic principles about the onset and progression of disease in a biological system. These models are generally functionally complicated and require that experience in mathematics, statistics, and computer science be linked to experience with biological mechanisms. Mechanistic models also generally have greater data needs than do empirical models.

The third step requires the choice of a statistical linkage between the data and the model. The most common linkage method is to assume a statistical distribution for the response and use that distribution to derive a mathematical function describing the quality of the fit of the model to the data. However, a considerable amount of DRM has been done by simpler linkage functions, such as drawing a straight line through the data points. The advantage of

choosing a formal statistical linkage is the ability to test hypotheses and derive confidence intervals for model predictions.

In DRM, fitting the model to the data is the fourth step. Since the primary components of a model are the parameters that define the model, curve fitting simply involves choosing values for the parameters in the model. If a formal statistical linkage has been developed for linking the data to the model, then the parameters are chosen such that they "optimize" the value of the linkage function. For example, a common choice is to link the data to the model using the squared distance, denoted $[R(d_i) \bullet o_{ij}]^2$, between the predicted value from the model, denoted $R(d_i)$, and the observed value, denoted o_{ij}. These squared differences can be summed across all data points, and model parameters are chosen to minimize this sum; this is the common least-squares algorithm. Simpler methods can also be used to estimate model parameters. For example, by drawing a line through the data points, the parameters in the linear model can be estimated directly, since the value of can be estimated as the point where the line crosses the y-axis (zero dose) and the value of can be estimated by calculating the slope of the drawn line. Formal optimization is a better choice for modelling than ad hoc procedures, which often do not meet the criterion of transparency.

The fifth step in DRM is to make the inferences necessary to develop measures to protect public health. In its simplest form, a dose–response model allows the prediction of the response if the dose is known and the calculation of the dose if the aim is to target a specific level of response. In addition, implementation of the dose–response analysis (steps 1–4) also encompasses the extrapolation of results from the specific responses seen for the experiment being modelled to other exposure scenarios and other doses. This step can also involve an extrapolation from a laboratory species to humans. Usually, when making a prediction, the emphasis is on the change in response seen in the treated animals compared with the response seen in the controls. The different types of data (quantal, count, continuous, categorical) require different methods for predicting changes in response beyond the normal response. In general, the targets used for additional response fall into the categories of added response (simply subtract control

response), relative response (fold change relative to control response), and extra response (added response scaled to range from zero to the maximum possible response). Each of these choices can impact the final decision, so care should be taken to understand why a specific choice is made. Figure 4 illustrates some of the basic components of DRM for the simple linear model case and added response.

Measures used by public health agencies to prevent excess exposure to a hazardous agent generally fall into the categories of direct banning or limiting exposure. DRM could inform both choices, although its major impact is in the area of limiting exposure. Several methods on how to use DRM in this context have been proposed. The simplest is to use the predicted model to find the dose associated with a negligible (e.g. one in a million) or zero response over control. In general, this results in extrapolation far beyond the range of the data, which creates a great deal of uncertainty. A second approach is to use the dose–response model to identify a dose with a known response at or slightly below the observable range (the limit of scientific certainty) and use other models to get into a range where the response is assumed to be virtually unchanged relative to the control response. In this approach, a functional model structure can be used, such as a straight line, or something simpler, such as uncertainty factors (UFs), to identify a safe level of exposure. All of these options are discussed in chapter 6.

The basic steps in DRM shown in Table 1 can be repeated to consider other options in the process in order to understand the impact of choices on the predictions from DRM. This final step (step 6) in DRM is aimed at understanding the sensitivity of the analysis to specific choices and judging the overall quality of the final predictions. The simplest way to evaluate sensitivity is by considering several choices and determining if the results dramatically change. Depending on the degree of difference between choices, there could be value in performing a formal analysis of the quality of the fit of the model to the data. Other methods can also be used to assess the impact of choices used in the modelling on the eventual outcome, such as uncertainty analysis and Bayesian mixing. In some cases, step 6 is performed before step 5, with a focus on the assumptions used in the dose–response

analysis, and/or after step 5, with a focus on the assumptions used for implementation. These steps are further described in chapter 6.

4.6 Risk versus safety in dose–response modelling

Risk as used in this discussion is the direct estimation of the likelihood or degree of an event or its prevalence in a human population as a function of exposure. Given sufficient data in humans in the range of exposure where there is concern, it is possible to obtain scientifically supported estimates of risk. In most cases, the data used to develop dose–response models are not from studies in humans in the range of exposures that humans generally encounter. The most common type of data used for DRM comes from experiments in laboratory animals, generally at administered doses significantly exceeding the exposures that humans encounter. Even when human data are available and suitable for dose–response analysis, they are generally from selected populations, such as workers in occupational settings, whose exposures differ from those of the general population.

Thus, in many cases, dose–response analyses need to be extrapolated from an observable region where scientific support is available to a region where scientific support is weaker or non-existent. For dose–response analyses based on human studies, the extrapolation is generally a downward extrapolation to different exposure levels, but extrapolations can also be to different life stages (e.g. fetus, child) or to different populations with different environmental factors that might affect exposure (e.g. dietary differences). For dose–response analyses based upon laboratory data using animals, there is the additional problem of extrapolating from animals to humans.

Most of the methods used to implement the results of a dose–response analysis (step 5) address these extrapolation issues. The methods that have been used for extrapolation are diverse and sometimes contentious, with different countries, and even different agencies within a given country, using different approaches. The strategies used for extrapolation basically fall into two categories: those aimed at using estimates of risk for exposures outside of the

range of the data used in the dose–response analysis, and those aimed at establishing safety without using an estimate of risk.

Estimates of risk and the dose associated with that risk generally require extrapolation from the data on responses and doses to a lower dose range. These extrapolations can be done using the model (step 2) that was fit (step 4) to the data (direct estimation) or a different model, usually a line, extending from the lowest dose to a point of zero risk. The latter approach is generally envisioned to be conservative, assuming that the true risk is less than would be estimated by this second model at all doses below the dose for which scientific support is clear. In contrast, methods used to establish safety for a given dose without presenting an estimate of risk rely upon the concept that a dose that is sufficiently distant from the lowest dose associated with the observable range will be safe. This is generally done using uncertainty factors that have been developed over years of experience. In some cases where the general human exposure is estimated, however, the difference between the estimated exposure and the dose at the lowest edge of scientific support is used (margin of exposure, or MOE).

Regardless of how dose–response analysis is performed, additional methods are employed to extrapolate to humans. These methods are also varied, ranging from the use of additional uncertainty factors to more complicated modelling schemes based upon differences in toxicokinetics and toxicodynamics between humans and animals.

The term "risk assessment" is generally used to describe the entire process of making a public health decision regarding a specific chemical or agent. However, risk assessment can be defined further to differentiate between analyses aimed at establishing safety (as defined above) and analyses aimed at estimating risks. In this case, "safety assessment" would refer to the decision process aimed at establishing safety, whereas "risk assessment" would refer to assessments aimed at estimating risks that are part of a larger decision process. Safety assessments are more often used in cases where exposure can be controlled, such as for food additives and residues of pesticides and veterinary drugs in foods.

4.7 Summary

DRM, as used for informing public health decisions about chemical exposures, is a six-step process. The first four steps constitute dose–response analysis and relate to the process through which a mathematical description of the data is obtained in order to evaluate predicted responses for known doses or to obtain dose estimates when a chosen response is of interest. The fifth step involves the implementation of the results of the dose–response analysis for the purposes of guiding a public health decision. The final step, which can optionally be applied earlier in DRM, involves an assessment of the quality of the dose–response analysis and the sensitivity of model predictions to the assumptions used in the analysis. DRM, because it involves a large number of choices based upon scientific experience, can take on many different forms and be used in many different ways. The remaining chapters of this report focus on the range of choices available for each step in the process and some guidance to be used in making these choices.

5. DOSE–RESPONSE MODELLING: WHY AND WHEN TO USE IT

Dose–response analysis is a major part of the hazard characterization within the risk assessment paradigm and has been used in the past for both the characterization of dose–response relationships observed in animal bioassays as well as the low-dose extrapolation of incidences of adverse effects to the range of human exposure levels. Dose–response analysis includes the use of the NOAEL (pairwise testing) for deriving health-based guidance values such as the ADI and the use of DRM (fitting functions).

5.1 Historical perspectives

It has always been a challenge to extrapolate from effects observed in experimental animal bioassays to potential effects in humans in order to protect humans from potentially harmful chemical exposures. A variety of approaches have been developed.

The prototype chemical safety assessment uses the ADI methodology, which was introduced by Lehman & Fitzhugh (1954) and has come to be widely employed for the derivation of health-based guidance values (IPCS, 1987). The ADI was originally devised as a procedure for the regulatory approval of food additives. Since food additives are deliberately added, the process often defines what the regulatory agency is willing to accept as a legal standard of safety. The same methodology is used to derive health-based guidance values for chemical contaminants. However, because "acceptable" was deemed to be an inappropriate term for chemical contaminants, the term "tolerable" was used instead (i.e. tolerable daily intake, or TDI). Comparable terms that have been used are provisional maximum tolerable daily intake (IPCS, 1987) and reference dose (RfD) (Barnes & Dourson, 1988). Other similar methods exist for different types of exposures, such as for compounds with accumulating properties—for example, provisional maximum tolerable weekly intake or provisional maximum tolerable monthly intake (IPCS, 1987; WHO, 2002).

5.1.1 The no-observed-adverse-effect level approach to acceptable/tolerable daily intake

Calculation of the ADI based on the NOAEL approach for the case of quantal data is summarized in Table 2.

Table 2. NOAEL-derived ADI for the case of quantal data

Step	NOAEL-derived ADI
1. Data selection	Sufficient sample sizes, at least one dose with "no" effect and one dose with effect. Relevant end-points in a relevant species are important for any approach.
2. Model selection	Statistical method $R(D) = \begin{cases} 0 & \text{if response at dose D is not significantly different from control response} \\ 1 & \text{if response at dose D is significantly different from control response} \end{cases}$
3. Statistical linkage	Pairwise statistical tests between dose groups and control group.
4. Parameter estimation	Assessment of point of departure $NOAEL = D_{NOAEL}$ where $R(D) = 0$ for all $D \leq D_{NOAEL}$ and $R(D) = 1$ for all $D > D_{NOAEL}$ This procedure presupposes that all doses below the NOAEL are non-significant and all doses above the LOAEL are significant. This is often not the case.
5. Implementation	$ADI = \frac{NOAEL}{UFs}$ where UF is uncertainty factor.
6. Evaluation	Statistical power analysis should be performed to check if the test was sensitive enough to detect relevant effects.

Selecting the data needed to calculate the ADI based on the NOAEL approach (step 1) is similar to choosing the data to be used

for more complicated modelling; the better data sets have an appropriate number of relevant doses, sufficient sample sizes, and relevant end-points in a relevant species. The next step in calculating an ADI is to determine the NOAEL, which is the highest concentration or dose of a chemical, found by experiment or observation, that causes no detectable adverse effect, as defined above. This includes a statistical method (step 2), statistical linkage (step 3), and a method of assessment of a point of departure (step 4) that describes the identification of the NOAEL. Consider a response procedure, R(D), of the form:

$$R(D) = \begin{cases} 0 \text{ if response at dose D is not significantly different from control response} \\ 1 \text{ if response at dose D is significantly different from control response} \end{cases}$$

The statistical linkage (step 3) between this procedure and the data is represented by the statistical test used to determine if a response at any given dose is different from the control response. When the response is non-significant, we simply *act as if* the effect were in fact zero. Obviously, we cannot conclude that the effect actually *is* zero. When the NOAEL approach is chosen, the statistical test is used to decide upon the existence of a statistically significant increase (e.g. at the 5% level) over background (e.g. the control group) for each dose level separately. The selection of the NOAEL (step 4) is then achieved by choosing the largest dose, D_{NOAEL}, for which all smaller doses have R(D) = 0 and all larger doses have R(D) = 1. Mathematically, this assessment can be written as:

$$\text{NOAEL} = D_{NOAEL} \text{ where } R(D) = 0 \text{ for all } D \leq D_{NOAEL} \text{ and } R(D) = 1 \text{ for all } D > D_{NOAEL}$$

This procedure presupposes that all doses below the NOAEL are non-significant and all doses above the LOAEL are significant. This is not always the case.

The ADI methodology specifies that an acceptable dose of a chemical may be calculated by dividing the NOAEL by appropriate uncertainty factors (also called safety factors). Uncertainty factors are default factors used to account for both uncertainty and variability.

Historically, an uncertainty factor of 100-fold has been used to convert the NOAEL from an animal study into a health-based guidance value (Lehman & Fitzhugh, 1954; Dourson & Stara, 1983; IPCS, 1987). Additional uncertainty factors may be used to allow for database deficiencies, such as the absence of a chronic study (IPCS, 1994). The default 100-fold uncertainty factor may be seen to represent the product of two separate 10-fold factors that allow for interspecies differences and human variability (IPCS, 1987; Renwick & Lazarus, 1998). The recognition that the original 100-fold uncertainty factor could be considered to represent two 10-fold factors allowed some flexibility, because different factors could be applied to the NOAEL from a study in humans and from a study in animals. The concept of chemical-specific adjustment factors (CSAFs) (IPCS, 1994, 2005) was introduced to allow appropriate data on species differences and/or human variability in either toxicokinetics (fate of the chemical in the body) or toxicodynamics (actions of the chemical on the body) to modify the relevant default 10-fold uncertainty factor. The strategy used by WHO/IPCS in the NOAEL/ADI approach involves replacing the original 100-fold uncertainty factor with CSAFs where there are adequate data (IPCS, 1994, 2005).

Regardless of the quantities chosen for the uncertainty factor, the prediction (step 5) of the ADI from NOAEL-based DRM is given by the equation:

$$\text{ADI} = \frac{\text{NOAEL}}{\text{UFs}}$$

Step 6 can be extended to the evaluation of the sensitivity of the ADI to the assumed values of the uncertainty factors.

Some scientists have raised concerns regarding the use of the NOAEL to determine an ADI. The greatest concern is that the NOAEL tends to yield lower ADIs for chemicals for which there are more or better data. Therefore, stakeholders using usually more costly, better data are "punished" (Crump, 1984; Dourson et al., 1985; Kimmel & Gaylor, 1988; Barnes et al., 1995; Slob & Pieters, 1998).

5.1.2 *The benchmark dose approach to acceptable/tolerable daily intake*

The BMD concept was introduced as an alternative to the NOAEL approach (Crump, 1984; Kimmel & Gaylor, 1988). The BMD method has a number of advantages, including the possibility to extrapolate outside the experimental dose range and respond appropriately to sample size and the associated uncertainty.

Calculation of the ADI based on the BMD approach is summarized in Table 3 for the case of quantal data. A generic form of the BMD and benchmark dose lower confidence limit (BMDL) is presented in this table. In this document, a variety of response levels, such as 1%, 5%, and 10%, will be discussed.

Table 3. BMD-derived ADI (Weibull model) for the case of quantal data

Step	BMD-derived ADI
1. Data selection	Sufficient number of doses with different response levels and a sufficient number of *total* subjects.
2. Model selection	Fit dose–response model (e.g. Weibull model).
3. Statistical linkage	Predicted fractions are linked to observed fractions, and their "distance" is minimized by optimizing some fit criteria function (e.g. likelihood function based on assumed distribution).
4. Parameter estimation	Choose an appropriate response, p, in the range of experimental response. Estimate $BMDL_p$, the 95% lower confidence bound on the BMD_p, where $\frac{R(BMD_p) - R(0)}{1 - R(0)} = p$
5. Implementation	$ADI = \frac{BMDL_p}{UFs}$
6. Evaluation	Sensitivity of BMD to model choice can be checked by fitting various models.

In choosing the data (step 1) for BMD modelling, the same basic considerations apply as for the NOAEL method. In addition, studies showing a graded monotonic response with a significant dose-related trend work best. This is generally true for all DRM analyses.

Choosing a model (step 2) for the BMD method is dependent upon the types of data available and the characteristics of the response being modelled. Complicated models will require a larger number of dose groups than simpler models. Several models have been proposed for each type of data. In the United States Environmental Protection Agency's Benchmark Dose Software (BMDS) program, a number of routinely used models are cited (http://www.epa.gov/ncea/bmds/). As an example, assuming the availability of data that represent the proportion of animals responding to a given exposure with an adverse effect (e.g. cancer) from each dose group, one model choice could be the Weibull model, which has the form:

$$R(D) = \alpha + (1-\alpha)(1-e^{-(\beta * D)^{\gamma}})$$

where α is the proportion responding in the unexposed group, β describes the increase in probability of adverse effect per unit dose, and γ describes the shape of the dose–response curve (e.g. $\gamma >> 1$ implies threshold-like behaviour; $\gamma = 1$ implies log-linear behaviour).

The statistical linkage (step 3) between the data and the model can assume a number of different forms, as described previously (section 4.5) and in section 6.2. For quantal data, it is appropriate to assume that the data are binomially distributed for each dose group. Estimating model parameters (step 4) for the BMD method can also be based upon a variety of different methods. For the Weibull example, one routinely used approach would be to choose the parameters that maximize the binomial-based log-likelihood.

The concept of the BMD comes from the idea that it is desirable to use a dose–response model to capture the general pattern of response for all dose groups in the experimental data set, but there was some dose, the BMD, below which predictions would be tenuous. This BMD can be selected in a number of ways (e.g. Barnes et al., 1995; Murrell et al., 1998), but the most common way is to choose an excess response, the benchmark response, or BMR (p), below which there was insufficient support from the data. A common choice for BMR is p = 10%. Once the BMR (p) is selected, the BMD, specifically denoted BMD_p, is calculated according to the following equation, if the extra risk formula is used:

$$\frac{R(BMD_p) - R(0)}{1 - R(0)} = p$$

Empirical investigations showed for a large and representative set of compounds that the 95% statistical lower bound on the estimated BMD may be regarded as an analogue to a NOAEL, and substituting one with the other would result in similar ADIs (Crump, 1984; Barnes et al., 1995). As with all aspects of modelling, many choices exist for calculating confidence bounds, and these are discussed further in chapter 6.

Having chosen a method for estimating a 95% statistical lower bound on BMD_p, which can be called $BMDL_p$, the ADI can be calculated as follows:

$$ADI = \frac{BMDL_p}{UFs}$$

In this calculation, the values of the uncertainty factors could be the same as those used for the NOAEL or adjusted to account for a slightly different interpretation for the $BMDL_p$ relative to the NOAEL (Renwick et al., 2003).

The BMD method includes the determination of the response at a given dose, the dose at a given response, and their confidence limits. Using extrapolation of the dose–response model below the biologically observable dose range, the response at specified (lower) dose levels can be estimated as well as the dose corresponding to a specific response level.

5.2 Points of consideration

The use of DRM in general for hazard characterization is possible when a sufficient amount of dose–response information is available, either from an experimental animal bioassay or from a human study (epidemiological study or clinical trial). As shown in the previous section, the BMD can be considered as an alternative point of departure for deriving an ADI in those situations where a NOAEL would have been used as a point of departure in current procedures. In addition, DRM may be helpful in those situations

where there is a need for low-dose extrapolation (e.g. substances that are genotoxic and carcinogenic). It should be noted, however, that extrapolation from a single model that fits the data in the observed range cannot be justified, since other models fitting the data equally well may result in substantially different estimates of low-dose risk. Bayesian approaches are currently under development, which take into account both statistical uncertainties in the data and model uncertainty (see section 6.5). In practice, linear extrapolation from a BMD_{10} (or ED_{10}, effective dose for a 10% risk, approximately equal to BMD_{10}) is often applied as a simple method for low-dose extrapolation. This is considered a conservative approach. As another application, DRM may be used to estimate risks at any given (human) exposure level. Since human exposure levels are usually lower than the doses in the observed range in animal studies, methods for low-dose extrapolation may also be needed in this application.

5.2.1 General aspects of definition

The NOAEL is a parameter derived directly from the observed dose–response data and is defined as the highest administered dose at which the effect is still not significantly different from that at dose 0 (see section 5.1). The NOAEL is based on a multiple test procedure performed along the applied dose series. It lacks further detailed statistical properties compared with a parameter of a dose–response model, for which the precision of the estimate can be quantified.

The dependence of the NOAEL on the statistical significance test, however, tends to penalize chemicals for which there are more or better data by giving a higher estimate for those chemicals with less precise data. This problem does not occur in DRM. In fact, the opposite relationship holds: it penalizes studies with few or poor data.

The NOAEL approach can be formally considered a dichotomous procedure, where no effect is assumed to be present below the NOAEL and where an expression of the critical effect is present above the NOAEL (see section 5.1 and Table 2). Given the typical animal studies used in toxicology, the effect size that can be detected by a statistical test may be larger than 10% (additional

risk). Therefore, the NOAEL may be expected to be a dose at which the effect is in reality somewhere between 0% and 10% or more. In contrast, the BMD is a dose for which the size of the effect has been predefined, and thus it is under the control of the risk assessor. Furthermore, while the dichotomy of the NOAEL approach does not provide quantitative information about risk above the ADI, such information might be derived from fitted dose–response models, where such dichotomy does not exist.

In general, DRM estimates are based on data from the entire dose–response curve for the critical effect. The standard NOAEL approach can be regarded as a special, simplified case of dose–response analysis, as it identifies a single dose that is assumed to be without an appreciable adverse effect. The dose–response model thus reflects the characteristics of the dose–response curve, particularly in providing estimates of slope. In the case of a regression framework, it provides standard error and confidence intervals for the model parameters.

5.2.2 Estimation procedure

NOAELs are restricted by the set of doses used in the specific studies. An important consequence is that the NOAEL may be either below or above the threshold it aims to approximate, assuming one exists. When the true threshold is higher than the NOAEL, the distance between the two can be expected to be limited (related to the dose spacing used). However, when the true threshold is lower than the NOAEL, the distance between the two is unlimited: the true threshold could be anywhere between zero and the NOAEL.

The actual value of the NOAEL depends strongly on the following characteristics of the study design:

- *Group size.* The power to detect a NOAEL at some dose level is directly dependent on the sample sizes chosen at those dose levels (Gaylor, 1989). The larger the group size, the smaller the potential true effect size at the NOAEL.

- *Dose location.* Since the NOAEL is an applied dose that did not show significant effects, while the next higher dose applied did show significant effects, the NOAEL can only be one of the

doses actually applied in the study. A particularly disturbing disadvantage of the NOAEL approach is that in some cases no NOAEL can be assessed because the lowest applied dose showed effects.

- *Experimental variation*. Larger experimental variation between subjects will result in lower statistical power and, hence, higher NOAELs. In quantal data, this phenomenon is somewhat hidden, but in continuous data, it is directly visible: it is reflected by the scatter in the data per dose group. This experimental variation comprises various things: biological (e.g. genetic) variation between subjects, variation in experimental conditions (e.g. time of feeding, location in experimental room, time of section or interim measurements), and measurement errors.

DRM-derived estimates are based on interpolation, and these estimates are not restricted to the actually applied doses. DRM can also be used on a study where no NOAEL (only a LOAEL) can be defined, so in this situation another study may be unnecessary. A comparison of different models can be useful. When multiple models are fit to the same data and produce widely varying BMD estimates, caution should be used in interpreting the results, as this could indicate insufficient data for modelling (see chapter 6).

It should be noted that in comparison with the NOAEL approach, implementation of the full DRM approach may lead to differences between the NOAEL and the BMD in individual data sets. However, on average, BMDs that represent the lower confidence interval on the dosage giving a BMR of 5% or 10% tend to be quite similar to the NOAELs (Allen et al., 1994). Therefore, in data sets where DRM cannot be applied, the NOAEL may serve as a reasonable surrogate of the BMD.

5.2.3 Uncertainty

A modelling approach facilitates both sensitivity and uncertainty analyses. Uncertainty (see section 6.5) can be expressed numerically when the doses and responses are linked by a model. Such numerical analyses can also be subject to sensitivity analyses, to test the contribution of different aspects of the database or of

model characteristics to the overall uncertainty. The uncertainty in the risk estimates that arises from aspects of study design, such as dose spacing, sample size, and biological variability, can be assessed in a dose–response model. While uncertainty factors are amenable to uncertainty analysis (Slob & Pieters, 1998), the threshold procedure of the NOAEL is not readily amenable to quantitative estimation of uncertainty or to a sensitivity analysis.

A disadvantage of using the threshold procedure of the NOAEL for the estimation of a point of departure ("starting point") for formulating advice to risk managers is that it is not possible to quantify the degree of variability and uncertainty that may be present. The NOAEL is assumed to be a dose without biologically significant effects. This assumption is more likely to be valid in toxicological studies with larger sample sizes.

5.2.4 Study design

A design optimal for the NOAEL approach could limit the use of DRM, and vice versa. While the NOAEL approach requires sufficient sample sizes within dose groups (to warrant statistical power), the DRM approach requires a sufficient number of dose groups (to warrant a description of the whole dose–response). Given the restrictions on the total number of animals used in a single study, these two requirements may not be compatible.

An important point to bear in mind is that DRM can be used on studies carried out in the past and based on the traditional designs (with three dose groups and a control). Some have argued that optimal designs for dose–response models may have the advantage for animal welfare that fewer animals could be used (Slob et al., 2005).

While DRM provides uncertain estimates when the number of dose groups is too small, the determination of both the BMD/BMDL and the NOAEL may prove inadequate at different points when the number of animals per dose group is too small. For example, when the critical effect is seen in a larger experimental animal, such as the dog, with few animals per dose group, the NOAEL may be high owing to the insensitivity of the test. The BMD/BMDL approach, however, can be used to evaluate sparse dose–response data and quantify the inherent uncertainty. However,

even here, where an apparent dose–response relationship in the data remains, the BMD/BMDL may also provide very uncertain estimates. Therefore, a typical four-dose study with a few animals per dose may in practice be unreliable whatever method, NOAEL or BMD, is applied. However, the advantage of DRM is that this uncertainty is made visible, whereas in the NOAEL approach it remains hidden.

DRM reduces the need for more experiments when a small degree of extrapolation is needed (e.g. when the doses used are near the human exposure level). In contrast, the NOAEL approach may require further experiments where no clear NOAEL (or LOAEL) can be identified. This can be illustrated by the study of Allen et al. (1996) on developmental risk assessment in rats exposed to boric acid in their diet. This study failed to establish a NOAEL; however, the BMD approach could have been applied, thereby avoiding the need for repeat studies (see also section 5.2.2 above). Distributing the total number of animals over more dose groups does not result in poorer performance, despite the smaller number of animals per dose group, as shown by Slob et al. (2005). The above example of Allen et al. (1996) suggests that the BMD approach provides a reasonable basis for appropriately comparing and combining studies, as opposed to ad hoc combinations of study results.

A major advantage of DRM is the ability to estimate risks within the observable range of effects. In animal studies, it is possible to estimate risk over the full range of doses used. Estimation of risks outside the observable range will be more and more unreliable when risks get smaller and smaller (Murrell et al., 1998). Some studies (e.g. Sand et al., 2002) have investigated the effect of model dependence at different response levels.

Some experts have argued that extrapolation to risk levels outside the observable range might be warranted when there are indications that the same toxicological mechanism is active in both the extrapolation region and the experimental region of the model fit. However, mathematical models that adequately describe the full complexity of the mechanisms involved are very rare; even then, it needs to be additionally assumed that the parameters estimated from the data (i.e. the observable range) are adequate for the low-dose

range and have sufficient precision to make the prediction (see also section 6.5.3).

5.2.5 Biological information

The NOAEL approach incorporates biological information through the application of "expert" but subjective judgement. Full DRM has the potential for a more "science-rich" analysis through the more formal quantitative inclusion of factors/covariates into the models, in the case of both human epidemiological and animal data.

Such an approach can lead to more certain estimates, centred on a toxicologically based concept of estimating the dose–response relationship on the basis of all available biological knowledge, using empirical data and applying statistical inference. More complicated models can be developed on the basis of toxicokinetics and toxicodynamics.

5.2.6 Comparison of experimental results

NOAELs derive from an algorithmic analysis of the results of a single experiment. Meta-analysis on data such as NOAELs across a range of studies on a specific chemical is possible, such as when data are insufficient to build a dose–response model, but may be limited by the statistical properties of the NOAEL estimates.

Estimates derived from full DRM, however, enhance the ability to compare different experiments, effects, and compounds using a common framework. The estimates obtained may provide a test of consistency among different studies that may use different dose levels. DRM methodology can be used to describe dose–response relationships in different studies (e.g. rat and mouse, chronic and subchronic exposure, healthy and diseased animals) if suitable data sets exist.

Rules for combining studies, however, need to be developed. Descriptions of the dose–response on the same end-points in different studies may be integrated to provide a cohesive picture of the chemical's toxicity. The values obtained using DRM may result in estimates for each end-point on the basis of biological and functional relevance.

5.2.7 Risk management perspectives

The potential use of the estimates from DRM can, from a risk management perspective, give an improved characterization for decision-making by:

- giving the decision-maker a "more-than-one-point" appreciation of the data;
- providing information about what happens above the safety level (magnitude and types of health impacts);
- quantifying benefits in risk reduction from different regulatory actions;
- promoting consistency in decisions, if appropriate adjustments are made for differences in effect, effect level, species, and study design; and
- allowing for an iterative interaction between the risk assessor and risk manager on a continuous and ongoing basis.

5.3 Implementation issues

In the case of the BMD, there are a number of decisions to be made in applying the method and determining a BMDL: for example, which mathematical model to use; what degree of confidence to use in calculating confidence limits; what response level to predetermine as the BMR (e.g. BMR = 1%, 5%, or 10% incidence of an effect, or a 5% or 10% change in a continuous endpoint, such as body weight or red blood cell counts). It is often not clear what response level (BMR) can be considered as non-adverse. For example, should a 5% decrease in red blood cell counts be considered as adverse, or should a smaller (or larger) change be chosen? Should up to a 5% increased incidence in hepatocellular hypertrophy be considered as acceptable in an animal study, or is a maximum of 10% increase adequate? These and other choices need additional discussion among toxicologists and clinicians. Although an explicit statement on the BMR is an improvement compared with the generally unknown response level associated with a NOAEL, choices of a BMR need consensus building.

5.4 Summary

The characterization of dose–response relationships in animal and human studies has been a major component of hazard characterization. Over the years, a variety of methods have been developed to accommodate such relationships. DRM may be regarded as the most adequate approach for analysing dose–response data, provided that a suitable data set of animal or human dose–response data is available.

The standard NOAEL approach identifies a single dose that is assumed to be without appreciable effect, whereas the BMD is based on data from the entire dose–response curve, estimated for the critical effect. Although the effect at the NOAEL is assumed to be zero, it will be non-zero in many cases, although to what extent remains unknown. The size of the effect at the BMD is made explicit and, as far as possible, is based on toxicological knowledge. While the uncertainty in a NOAEL cannot be quantified, the uncertainty in a BMD can be quantified by a confidence interval. The use of DRM may call for different guidelines for optimal study designs, as the number of dose groups should be sufficiently large. Distributing the total number of animals over more dose groups may be done without loss of precision. DRM can more effectively compare different experiments, effects, and compounds. While risks above the ADI based on a NOAEL cannot be quantified, such may be possible for exposures exceeding the ADI based on DRM. For estimating risks below the observable range, extrapolation based on a single fitted model is unwarranted. Here, linear extrapolation may be considered as a conservative approach. Currently, more advanced methods are being developed (e.g. Bayesian approaches) for low-dose extrapolation based on DRM.

6. PRINCIPLES OF DOSE–RESPONSE MODELLING

6.1 Data

6.1.1 Selection of data

When considering which data to use from a set of available toxicity studies on a particular compound, it may not be effective to do a dose–response analysis for each observed end-point in each study. As a first step, one may omit studies that have obviously larger NOAELs compared with the other studies. In this way, one may, for example, select for a given type of toxic response (e.g. chronic, developmental) for the most sensitive species. For a given study, many end-points may have been measured. End-points not showing a clear dose–response on visual inspection can be omitted. Then, based on the toxicological impact together with the apparent magnitude of the response, a selection of end-points can be made as candidates for modelling. It would be very helpful if submitted studies included an annex with plots (in addition to tables) of observed data points for each end-point, possibly with fitted curves to the plots, to enhance the process of selecting end-points. At a minimum, these should be included for end-points showing evident effects.

After selecting the potentially relevant end-points, one must decide whether each dose–response data set is actually amenable for a dose–response analysis. Generally, it is desirable to have at least three or four different doses (including controls). In addition, the associated effect levels need to be different from each other; it is preferable to have at least three different response levels.

6.1.2 Data types

There are various types of response data, and these can be categorized in various ways. The main distinction relevant for effects is that between quantal and continuous data. Quantal data relate to an effect that is observed or not in each individual subject (laboratory animal or human). Hence, for each dose, the number of subjects responding out of the number of subjects available is

reported. In continuous data, a quantitative measurement is associated with each individual subject. As an intermediate type of data, ordinal data reflect (ordered) severity categories—that is, they are qualitative data but with a rank order (e.g. histopathological severity data). When the categories are non-ordered, they are called categorical data, but these are rare for response data. Finally, count data form another class of data (i.e. discrete data), but in practice they can often be treated as continuous data (see also section 4.3).

Although the type of data is important for statistical reasons (see section 6.2.2 on distributions), the distinction between quantal and continuous data also has a crucial impact on interpretation of results and their ensuing use in risk assessment. In the case of quantal dose–response data, information on the change of incidence with dose is available at one particular degree of effect. For example, the incidence of cleft palate may increase as dose increases, but under the categories "no cleft palate" and "cleft palate", there is no information about the degree of the effect. In ordinal and continuous data, in contrast, information on both the degree of effect and the incidence is available as a function of dose. So, for example, cleft palate might be categorized into an ordinal variable with levels "no clefting", "mild clefting", "moderate clefting", and "severe clefting", or it might be quantified in a continuous variable as, for example, the fraction of closure. The relationship between the average response and dose gives information about how exposure changes the degree of effect. For instance, a plot of (average) red blood cell count may show the decrease in mean red blood cell count (i.e. the degree of the effect) as a function of dose. By also considering the individual data points, information on the incidence can be derived as well. For example, an estimate of the fraction of individuals with red blood cell counts less than some critical value can be derived.

When using animals as a model for human response, the observed dose–response information is assumed to mimic the dose–response in humans to some approximation. It might be argued that this assumption is more plausible for degree of effect than for incidence. The problem is that the observed dose–incidence relationship for animals largely reflects the variation in the animals used, which is highly controlled in a laboratory experiment. Hence, it may not mimic the human variation.

6.2 Models and distributions

6.2.1 Dose–response models

6.2.1.1 Continuous dose–response models

The models listed in Table 4 are some of the forms that may be used to describe the relationship between dose and the magnitude of a response on a continuous scale in an individual. When combined with a statistical distribution (e.g. normal or lognormal), these equations can also be used to describe the relationship between dose and a continuous response in a population, where the continuous model corresponds to the central estimate.

Dose–response data are often adjusted by subtracting the (mean) control value from each individual observation. However, this procedure does not account for the fact that the background response level in the controls is, like the response level in the experimental groups, subject to sampling error. A better approach is to account for the background response in the model with a parameter that needs to be estimated from the data. Among the many ways in which this can be done, the following are three of the simplest:

1. $y = a + f_x(D)$
2. $y = a \times f_x(D)$
3. $y = f_x(a + D)$

where D is dose, a is the background term, and f_x may be any dose–response function. For some assessments, there may be mechanistic information that makes one form preferable to another. For example, the first form is preferable for modelling an influence that produces the effect independently, the second corresponds to the idea of normalizing the response as a fraction of the background response, and the third reflects a contribution from another agent acting by the same mechanism.

Table 4. Continuous dose–response models

Name(s)	Notes	Equation for response	Parameter explanations
Michaelis-Menten law of mass action	A theoretical account of enzyme- or receptor-based activity where the rate of action is a function of the rate of association (ka) and the rate of dissociation (kd).	$= \text{RMax}\frac{[S]}{K_M + [S]}$	RMax is the maximum rate of the reaction, [S] is the substrate concentration, and K_M is the Michaelis-Menten constant, which is equal to ka/kd.
Hill equation log-logistic	A modification of the Michaelis-Menten equation that supposes that the occupation of multiple sites or receptors is required for the production of an effect.	$= \text{RMax}\frac{D^n}{K_D{}^n + D^n}$	RMax is the maximum response, D is the dose, K_D is the reaction constant for the drug–receptor interaction, and n is the number of (hypothetical) binding sites.
First-order exponential	If the interaction between a chemical and a target site is irreversible, then the rate of the reaction is determined by the rate of association (ka) only.	$= \text{RMax}\left(1 - e^{-rD}\right)$	RMax is the maximum response, D is the dose, and r is the exponential rate constant.
Power	Simple exponential model.	$= \beta D^{\alpha}$	D is the dose, α is the shape parameter, and β is the scale parameter.
Linear	Although there is usually no biological theory to suggest it, linear models are often justified by their simplicity; linear models have but a single parameter.	$= mD$	D is the dose and m is the slope.

6.2.1.2 Quantal dose–response models

Quantal dose–response functions describe the relationship between dose and the frequency of a particular outcome in a population (see Table 5). For a group of homogeneous or nearly identical individuals, the relationship between dose and frequency can be described with a step function where all subjects either respond or fail to respond at any given dose. However, because variability is ubiquitous in living organisms, quantal dose–response data typically show gradually increasing incidence with dose. One interpretation of this is that individual subjects differ in tolerance to the agent, which can be described by a statistical tolerance distribution. Hence, any cumulative distribution function may be used as a quantal dose–response function. Other models have been derived from statistical assumptions about how the agent might exert its effect in an organism, such as the gamma multi-hit model.

Background response rates should, just as in the case of continuous data, be accounted for by incorporating an additional parameter in the dose–response model. The two simplest ways of doing this are:

1. $y = a + (1 - a)\, f(x)$
2. $y = f(x + a)$

where f(x) is any dose–response function (varying from 0 to 1). As with continuous data, correcting the data for background response prior to the dose–response analysis is statistically unsound. The background response level should be estimated simultaneously with the dose–response model and be treated in the same way as the observed responses in the other dose groups.

6.2.1.3 Thresholds

The term "threshold" can be used in three different senses. First, it is used in a scientific sense to indicate a level of exposure at which no effect occurs (e.g. there is a physical stimulus, but there is no response). Second, a threshold may be thought of as a level at which there may or may not be an effect, but it is too small to be observed (e.g. a NOAEL). In this case, it is the perceptual limitation of an observer or analyst, rather than the actual subject of the experiment, that is being described. As a third meaning, a "practical

Table 5. Quantal dose–response models

Name(s)	Theoretical basis	Equation for frequency (F)	Parameter explanations
Step function	No variability.	If $D < T$, $F = 0$ If $D \geq T$, $F = 1$	D is the dose and T is the threshold parameter.
One-hit (single-hit)	Hit theory models employ the use of a rate to describe the interaction between a group of causal agents (e.g. molecules) and a group of targets (e.g. a human population).	$= 1 - e^{-(\alpha+\beta D)}$	D is the dose, e is Euler's constant, α is a location parameter, and β is the slope parameter.
Gamma multi-hit	An expansion of the one-hit model, which is based on the notion that multiple hits or events are required to produce a particular effect.	$= \Gamma(\text{gamma}*D, k)$	Γ() is the incomplete gamma CDF, D is the dose, gamma is a rate parameter, and k is the number of hits required to produce the effect.
Probit normal	A descriptive model based on a normal or Gaussian distribution.	$= \Phi(\alpha + D*\beta)$	Φ() is the normal CDF, D is the dose, α is a location parameter, and β is the slope parameter.
Logistic	The statistical logistic model is also a descriptive tool with no theoretical basis.	$= \frac{1}{1 + e^{-\alpha - D \times \beta}}$	D is the dose, α is a location parameter, and β is the slope parameter.
Weibull	A flexible descriptive model originally developed to describe survival data in demography.	$= e^{-(\alpha + (\beta \times D)^{\gamma})}$	D is the dose, α is the background parameter, β is the slope parameter, and γ is an exponent.

CDF, cumulative distribution function

threshold" is a response where the consequences are determined to be trivial and not worth further consideration.

A threshold in the first sense may be incorporated into a model. The introduction of a threshold parameter truncates the dose–response relation at a threshold dose:

- Below threshold, the effective dose is zero.
- Above threshold, the effective dose is the dose minus threshold.

Threshold terms generally are difficult to estimate accurately and have large confidence limits.

6.2.1.4 Severity (degree of effect)

The severity of toxic responses is rarely used in DRM other than in a qualitative manner (e.g. tumour formation vs reduced fertility). However, one may also consider severity or degree of response in a quantitative way at the level of a single end-point. As noted above, the dose–response of continuous end-points may be directly interpreted as a dose-related change in degree of effect—for example, a per cent decrease in haematocrit (Woutersen et al., 2001) or a per cent change in body weight (see Figure 5, which represents the dose–response relationship between body weight and exposure to the mycotoxin deoxynivalenol) (Pieters et al., 2004). Here, a certain degree of effect (5% reduction in body weight) is chosen for deriving the BMD. The BMDL is then defined as the dose associated with a particular (e.g. 5%) change in degree of effect for that end-point.

An important advantage of defining a BMR in terms of degree of effect based on continuous response data is that values for the BMR that may be considered non-adverse are within or close to the range of observations. Therefore, low-dose extrapolation may not be needed or needed only to a small extent when continuous end-points are considered.

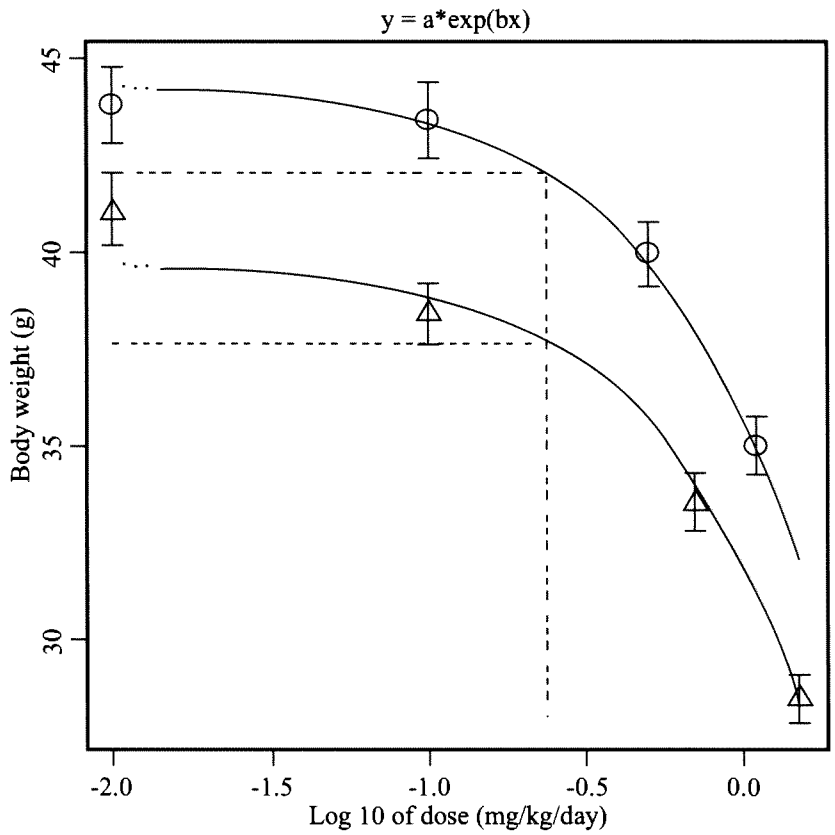


Fig. 5. Dose–response model fitted to male (circles) and female (triangles) body weights plotted against log dose (exposure to the mycotoxin deoxynivalenol). The plotted marks represent the (geometric) means of about 40 mice, with 90% confidence intervals. The BMD associated with a BMR of 5% is estimated at 0.24 mg/kg body weight (log equivalent = −0.62), with a lower confidence bound of 0.22 mg/kg body weight (log equivalent = −0.66). The latter value can be considered as a BMDL for this end-point (adapted from Pieters et al., 2004).

In the case of a histopathological end-point resulting in ordinal data, a dose–response function may be fit using categorical regression, and the BMDL associated with a particular degree of effect (e.g. minimal or mild) may be estimated (e.g. Piersma et al., 2000; Woutersen et al., 2001).

Categorical regression may also be applied at a higher level—that is, in an analysis of multiple studies (Hertzberg & Miller, 1985; Hertzberg, 1991; Hertzberg & Wymer, 1991). In this application of categorical regression, severity categories are defined covering disparate end-points. Most of these applications focus on estimating

the likelihood that a given category of severity may occur at a given dose level.

6.2.1.5 Modelling with covariates

In some circumstances, it is desirable to include variables in addition to an exposure variable in dose–response models. For example, in epidemiological studies, it is common to model disease risk in terms of not only exposure, but also age, sex, socioeconomic status, smoking status, and other measurements that may be relevant to the disease state. These other factors may be correlated with exposure status because of the way in which the sample was taken. Then, unless the proper covariates are included in a model for the relationship between exposure and the health end-point, the effect of exposure will be incorrectly estimated. In bioassay studies, in which animals are randomized to treatment groups, this sort of confounding cannot, in principle, occur, but it may be useful to include a covariate such as sex to account for some of the variability in a related measure (see Figure 6).

6.2.1.6 Biologically based dose–response models

While biological considerations may motivate the choice of one or several empirical models, the level of biological detail in such models is minimal. Thus, their credibility for interpolating and extrapolating a data set derives mainly from their fit to the data, as evaluated statistically. Another class of model, the biologically based dose–response model, is much more complicated and is explicitly designed to model the biological details that lead from initial exposure to a toxicant to the ultimate pathological outcome. Typically, such a model includes a physiologically based toxicokinetic model to describe the distribution and metabolism of the parent compound and toxic metabolites and other mechanistic, or toxicodynamic, models that link target tissue concentration to the ultimate response. The toxicodynamic part of the models may be relatively simple (e.g. when the outcome is inhibition of acetylcholinesterase in the model for chlorpyrifos; Timchalk et al., 2002) or as complicated as a fully elaborated stochastic model for carcinogenesis (Sherman & Portier, 1998). Such a model is really a quantitative expression of a set of biological hypotheses and, when rigorously tested against critical experiments, becomes a credible tool for extrapolating from experimental results into exposure

realms that are difficult or expensive to reproduce in controlled experiments. Such models are quite expensive to construct both in resources and in time and thus would be expected to be developed fully only for exposures and toxicities of the highest concern.

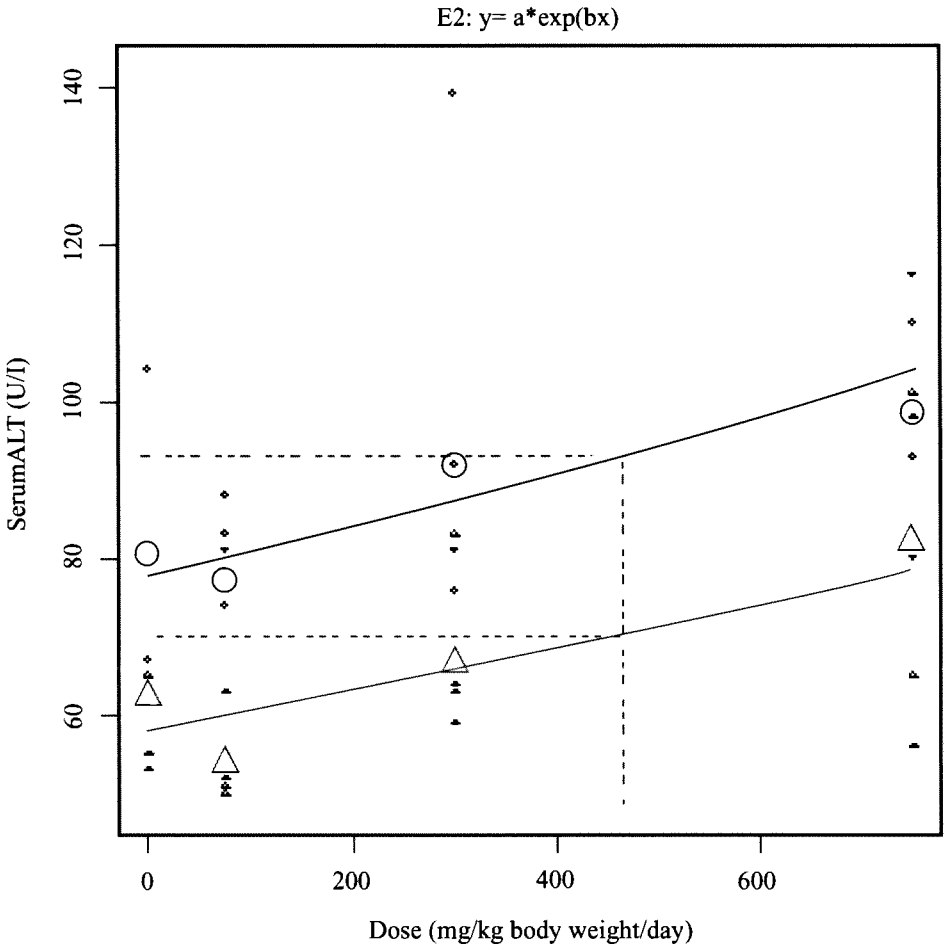


Fig. 6. Dose–response model fit to serum alanine aminotransferase (ALT) levels observed in males (circles) and females (triangles), where sex is treated as a covariate. In this case, the parameter a (background response level) differs between sexes, whereas parameter b and the residual variance (var) for the log(data) do not differ between sexes.

6.2.2 Statistical distributions

6.2.2.1 Continuous distributions

The normal or Gaussian distribution is symmetrical and defined from minus to plus infinity. It has two parameters: the mean and standard deviation, which control the location and scale of the distribution, respectively. Because sums of large numbers of small effects tend to be approximately normally distributed, this distribution is often used to describe variability and the variation of measurement error.

The lognormal distribution has two parameters: the geometric mean and the geometric standard deviation. It can be considered as a derivative of the normal distribution where the logarithms of the observed or predicted values are assumed to be normally distributed. This produces a skewed distribution on the original scale. Another consequence of using a lognormal distribution is that it will not generate negative values, which makes it more suitable for describing positive-only data sets and unsuited for values with negative values. Since many distributions are skewed and contain only positive numbers, the lognormal distribution often provides a good description. In addition, products of a large number of small effects tend to be approximately lognormally distributed. Since effects in biological measures tend to be multiplicative (proportional) rather than additive, the lognormal distribution is generally more suitable for biological measures.

The Weibull distribution is most commonly used to represent the survival or "lifetime" distribution of physical systems/products or biological systems, depending upon the context. In many applications, there is no explicit theoretical reasoning indicating that a Weibull distribution is appropriate or should be used, although the distribution does have some theoretical underpinning within the class of extreme value distributions. From a curve-fitting standpoint, the functional form of the distribution is simply a power transformation of the exponential model, which gives the model more flexibility for describing data. The multi-hit model is a special case of the Weibull model.

A more complete list of continuous distributions is given in Evans et al. (1993).

6.2.2.2 Discrete distributions

Discrete distributions describe responses on a finite or infinite scale, preferably count data; a special case is a response with a dichotomous quantal outcome of 0 or 1.

A Bernoulli distribution has an outcome of 1 or 0, corresponding to the occurrence or absence of an event that occurs with frequency f over an infinite sequence of trials. The Bernoulli distribution is then simply "1" with frequency f and "0" with frequency 1 − f. The Bernoulli trial is the basis of the binomial distribution, the definition of which subsumes the former.

The binomial distribution is defined as the distribution of a sum of a given number of Bernoulli trials with outcome of 1 or 0, denoting the occurrence or absence of a specified event, respectively. In toxicological applications, the number of trials is fixed by the experimental design, and the proportion of subjects in which the specified event occurs is the response to be estimated. As a result, the binomial distribution is the distribution typically used to estimate quantal response model parameters.

The Poisson distribution is a one-parameter distribution for a positive and discrete valued response. The domain of the response variable is any positive integer. The distribution was originally derived as a distribution of rare events: specifically, the number (n) of events occurring in a sequence of Bernoulli trials where the number of trials is large and the probability (P) of events per trial is small. Consequently, the Poisson distribution can be used as an approximation of the binomial distribution when n is large and P is small. The Poisson distribution is commonly used in analyses of epidemiological data when the study design involves prospectively following a cohort of subjects over a time period for which the expected incidence of adverse events is small relative to the cohort size.

A more complete list of discrete distributions can be found in Evans et al. (1993).

6.3 Model fitting and estimation of parameters

The general principles of parameter estimation and model fitting have been discussed in chapter 4. Two basic methodologies are available for model fitting: conventional, in which parameters are selected to minimize or maximize an objective function, and Bayesian, in which information in a data set is combined with prior information about model parameters, resulting in a posterior distribution for those parameters that reflects the degree of uncertainty about those parameters. For historical and computational reasons, "user-friendly" software designed for carrying out dose–response analysis and non-linear modelling in general has been restricted to using conventional methodologies, whereas Bayesian methods are implemented in packages that require more extensive programming and substantially greater understanding of the statistical details (for further details on Bayesian approaches, see Hasselblad & Jarabek, 1995; Gelman et al., 2004). While such software requires substantial statistical understanding for successful use of Bayesian methods and is thus beyond the scope of this document, even conventional methods require an understanding of some basic principles before outcomes from applying the software can be properly interpreted. Some general remarks are given below.

6.3.1 Criterion function

The general approach of fitting a model is to find parameter values for the model that optimize the fit of the model to the data. To that end, a criterion function is defined, reflecting the fit of the model. The goal is to find the parameter values that optimize the value of the criterion. For many models typically used, this can be achieved only by an iterative "trial and error" approach (see below).

In many applications, the logarithm of the likelihood function is used as the criterion. The likelihood directly derives from the distribution assumed for the scatter in the data. For quantal data, the binomial likelihood is typically used. For continuous data, the normal likelihood is often used, be it for the observed responses themselves or for the log-transformed responses. Note that maximizing the likelihood function for data that are assumed to be

normally distributed is in fact equivalent to minimizing the sum of squares.

6.3.2 Search algorithms

Computer software employs algorithms to find parameter values that optimize the fit of the model to the data, and the user does not need to worry about the exact nature of the calculations. However, some basic understanding of the search process is required in order to interpret the outcomes.

An iterative search algorithm tries to find "better" parameter values in a process by evaluating whether the fit can be improved by changing the parameter values through a trial and error process. More advanced algorithms operate by evaluating the slope of the likelihood at which the fit is improved for one or more parameter value changes (basically using the slope to "climb the likelihood function" as quickly as possible to find the top value). The algorithm can start searching only when the parameters have values to start with. Although the software often gives a reasonable first guess for the starting values, the user may have to change these. It is not unusual (in particular when the information in the data is hardly sufficient to estimate the intended parameters) that the end result depends on the starting values chosen, and the user should be aware of that.

The algorithm keeps on varying the parameter values until criteria for stopping are satisfied. There are two major reasons for the algorithm to stop the searching process:

1. The algorithm has converged (e.g. it has found a clear maximum in the log-likelihood function). In this case, the associated parameter values can be considered as the "best" estimates—e.g. the maximum likelihood estimate—if the likelihood was maximized. However, it can happen that the log-likelihood function has not one but more (local) maxima. This means that one may get other results when running the algorithm again, but with other start values. This can be understood by remembering that the algorithm can only "feel" the slope locally, so that it usually finds the optimum that is closest to the starting point.

2. The algorithm has not converged (i.e. the algorithm was not able to find a clear optimum in the likelihood function, but it stops because the maximum number of iterations [trials] is exceeded). This may occur when the starting values were poorly chosen, such that the associated model would be too far away from the data. Another reason could be that the information in the data is poor relative to the number of parameters to be estimated. For example, a dose–response model with five unknown parameters cannot be estimated with a study with four dose groups. As another example, the variation between the observations within dose groups may be large compared with the overall change in the dose–response. In these cases, the likelihood function may be very flat, and the algorithm cannot find a point where the function changes between increasing and decreasing. The user may recognize such situations by high correlations between parameter estimates (i.e. changing the value of one parameter may be compensated by another), leaving the model prediction practically unchanged.

6.4 Model comparison

The fundamental criterion for judging a model is that the selected model should describe the data, especially in regions of the dose–response where inferences are needed. Most fitting methods provide a global goodness-of-fit measure, usually providing a p-value. These measures quantify the degree to which the model predictions correspond to the data. Small p-values indicate a poor fit to the data. Since it is particularly important that the data be adequately described, it is recommended that a p-value of 0.1 be used to compute the critical value for goodness of fit, instead of the more conventional values of 0.05 or 0.01.

Another way to detect the form of these deviations from fit is with graphical displays. Plots should always supplement goodness-of-fit testing. For continuous data, it would be extremely helpful for plots that include data points to also include a measure of dispersion of those data points. In certain cases, the typical models used in DRM cannot fit the observed data, such as when the data are not monotonic or when the response rises abruptly after some lower doses that give only the background response. In these cases,

adjustments to the data (e.g. a transformation of dose) or the model (e.g. adjustments for unrelated deaths) may be helpful.

When fitting many different models to the same data, they generally will not all result in the same fit, and some care must be taken in choosing which model or models will be considered. In applying a statistical theory to this problem, one of four possible situations may arise:

1. The models form a nested series of models in the same family, in the sense that there is a "full" model, and other "restricted" models are derived from that full model by setting successively more parameters to a fixed value or, conversely, successively incorporating more parameters into the model. Likelihood ratio tests can be used to evaluate whether the improvement in fit afforded by estimating additional parameters is justified. The general form of the test is to calculate $2 \times (LL_{full} - LL_{restricted})$, where LL is log-likelihood, and compare this with a critical value from the chi-squared distribution with $P_{full} - P_{restricted}$ degrees of freedom (where P_x is the number of parameters estimated in model x).

2. The models are from the same family, but do not form a nested series. Some statistics, notably Akaike's information criterion (AIC is $-2LL + 2P$, where LL is the log-likelihood at the maximum likelihood estimates for the parameters, and P is the number of model degrees of freedom) can be used to compare models (Akaike, 1973; Burnham & Anderson, 2002). In this case, the model with the smallest AIC value is selected, although models with similar AIC values (differing by no more than about 4) are probably equivalent (Burnham & Anderson, 2002).

3. The models are not from the same family, but are fit using the same assumptions about the underlying probability distributions (e.g. all using a lognormal likelihood or all using a normal likelihood). In this case, Burnham & Anderson (2002) argue that AIC can still be used to identify the best model, but this appears to be a controversial point. Sand et al. (2002) have shown that it may be difficult to discriminate between the commonly used quantal dose–response models based on the AIC, which may be due to the fact that these models are quite

similar in their structure and include a similar number of parameters. In general, this case is still the subject of statistical research. At present, it will probably be adequate to use AIC to select a model as in the previous case, recognizing that this guidance may change.

4. Models do not use the same probability distribution. In this case, little formal statistical guidance is available. The plausibility of assumptions about the distribution of data needs to be examined by looking at the distribution of individual data. However, continuous data are often aggregated and reported as means and standard deviations, which eliminates the possibility of examining distributional assumptions. In these situations, the best that can be done is to rely on past experience with the end-points being modelled and select a reasonable probability distribution.

6.5 Representing uncertainty

Any parameters or predictions estimated from a given model are only point estimates and, to a larger or smaller extent, uncertain. This uncertainty arises from at least three sources:

1. *Sampling error*—the sampling error arising from inferences about a larger population from a single experiment;
2. *Study error*—the reality that dose–response estimates often differ among experiments with different experimental design, protocol, or uncontrolled circumstances; and
3. *Model error*—the fact that the "true" model is not known, which results in additional uncertainty when interpolating between doses, but even more so when extrapolating outside the dose range containing observations.

These three sources of uncertainty are briefly discussed below.

6.5.1 Sampling error

Uncertainty arising from sampling error with a single experiment is perhaps the easiest to evaluate and report. It may typically be quantified by a standard error or, preferably, by a

confidence interval. Confidence intervals may be calculated in several ways:

- plus or minus twice the parameter's standard error (provided by most dose–response software), which is estimated by the second derivative of the likelihood function (Hessian or information matrix);
- based on the profile of the log-likelihood function, using the chi-square approximation of the log-likelihood;
- bootstrap methods (see, for example, Efron, 1987; Efron & Tibshirani, 1993); and
- Bayesian methods, in particular if one has some preliminary knowledge of the plausible range of the parameter.

Various studies have compared the first three methods and concluded that the first may result in inaccurate intervals, whereas the second and third methods give similar results (see, for example, Moerbeek et al., 2004).

6.5.2 Study error

Uncertainty about the true value of a parameter that stems from variability among experiments can often be handled by treating the experiments as comprising an additional level of hierarchy, when the experiments are very similar in design and intent (e.g. same agent on the same end-point in the same strain and species). To characterize uncertainty in a statistical framework, it can be assumed that there is a population of experiments from which the ones at hand were selected (e.g. Davidian & Giltinan, 1995). As a result, the prediction or parameter of interest varies around a mean value among the members of that population of experiments, and an estimate of the mean and the degree of confidence can be derived. It should be noted that, even if data from only one experiment are available for analysis, this source of uncertainty still exists—it may be possible to quantify this uncertainty by analogy.

6.5.3 Model error

The third area of uncertainty, model uncertainty, is reflected by the question: to what extent do the data, possibly along with other knowledge about dose–response shape, constrain the set of possible dose–response shapes? A statistical model completely hinges on the

dose–response data, and the quality of the data is in fact the crucial aspect. In the fitting process, a model tries to hit the response at the observed doses. However, when a model is used to make inferences, interpolation between observed doses and extrapolation beyond the non-control doses are possible approaches. Thus, the model must also predict the response in the non-observed dose range. In other words, there are two aspects in evaluating the fitted model: one should assess not only if the model succeeded in describing the observed responses, but also if the model can be trusted to describe the non-observed responses where it is desirable to make inferences. The former aspect focuses on the quality of the model, the latter on the quality of the data. The following discussion elaborates on how to deal with the second of these two aspects (the first was addressed in section 6.4, Model comparison).

There are two ways to evaluate whether the data provide sufficient information to constrain the model and allow inference in some defined range outside of the range of the data. The fitted dose–response model should be visually inspected, to check if the data provide sufficient information to confine the model. Here, the question should be asked: if a curve is drawn through the data points by hand, could that be done in disparate ways? For instance, in the top panel of Figure 7, three curves have been drawn through the data points, each of which might be close to the true dose–response curve in the range between 2 and 5. In the bottom panel, however, it is very difficult to imagine that the true dose–response relationship would be very different from the (single) curve drawn here in the same range.

Another way to deal with this question is by comparing the outcomes from different fitted models. If the data do contain sufficient information to confine the shape of the dose–response relationship, different models fitting the data (nearly) equally well will result in similar fits and similar inferences. As an illustration, Figure 8 shows two different models fit to the same (continuous) data. Owing to the good quality of the data, they result in very similar estimated dose–response relationships. Inferences from dose–response models bear an additional level of uncertainty in proportion to the degree with which those inferences depend on the model used.

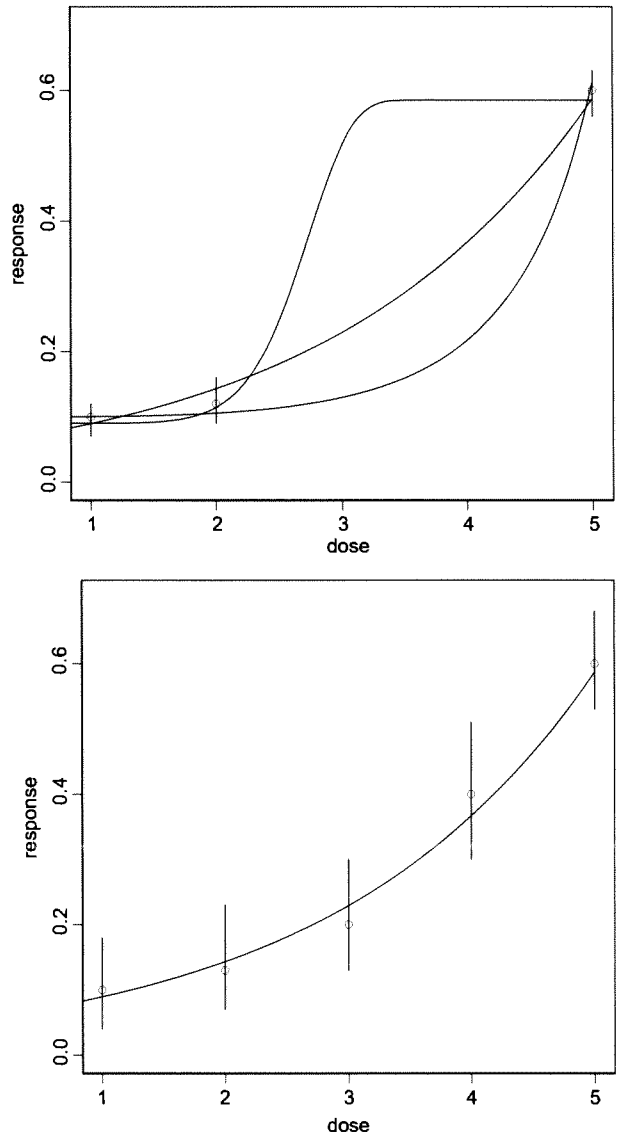


Fig. 7. Two data sets illustrating the idea of model uncertainty. In the top panel, the data (either quantal or continuous) do not contain sufficient information to confine the dose–response relationship in the range between 2 and 5: one may imagine various disparate curves that are all in agreement with the data, and hence they all might represent the true dose–response relationship. In the bottom panel, the data points prohibit the possibility of drawing disparate curves between 2 and 5.

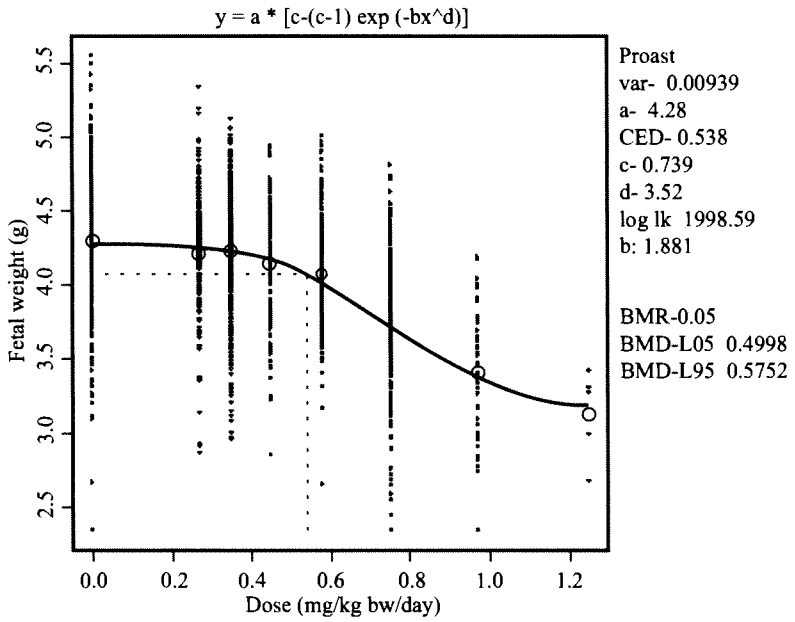
In current practice, there is a tendency to focus only on goodness of fit, and passing a formal goodness-of-fit test is often

regarded as sufficient evidence that the model is acceptable. This is unfortunate, since a goodness-of-fit test tends to be more easily passed for data with few dose groups or when few dose-related responses are noted and therefore non-observed responses are important or dominate. In addition, a goodness-of-fit test assumes that the experiment was carried out perfectly (i.e. perfectly random with respect to all potentially relevant experimental factors and actions). Clearly, this assumption is not realistic.

It is re-emphasized that a dose–response model, as long as it is not based on the mechanism of action of the particular chemical, serves only to smooth the observed dose–response relationship and to provide for a tool to assess confidence intervals. A statistical regression model itself has little, if any, biological meaning, and the choice of the model is to some extent arbitrary. It is the data, much more than the model, that should determine the dose–response relationship and any inferences derived from it. When different models (with similar goodness of fit and equal number of parameters) result in different estimates, this reflects a component of uncertainty that needs to be quantified and communicated with the estimate.

Dose–response models that are based on the mechanism of action of a particular chemical stand in opposition to statistical models as described here. Such mechanistic models contain information gleaned from biological theory and typically multiple experiments and therefore are less sensitive to data gaps (between dose groups). However, they do contain unknown parameters that need to be estimated from the data and thus require the resulting uncertainties to be quantified. Since such models are typically complex and idiosyncratic, little further general advice can be given, and it is suggested that professional statistical advice be sought in such cases.

Model uncertainty is particularly relevant to the issue of low-dose extrapolation. Here, the problem is that there may well be several models that are consistent with the data, as shown in the top half of Figure 9, and so give similar predictions in the range of the data, but whose predictions diverge at the low end of the dose range, as depicted in the lower half of Figure 9. One way to collect and represent model uncertainty in a risk assessment is through the



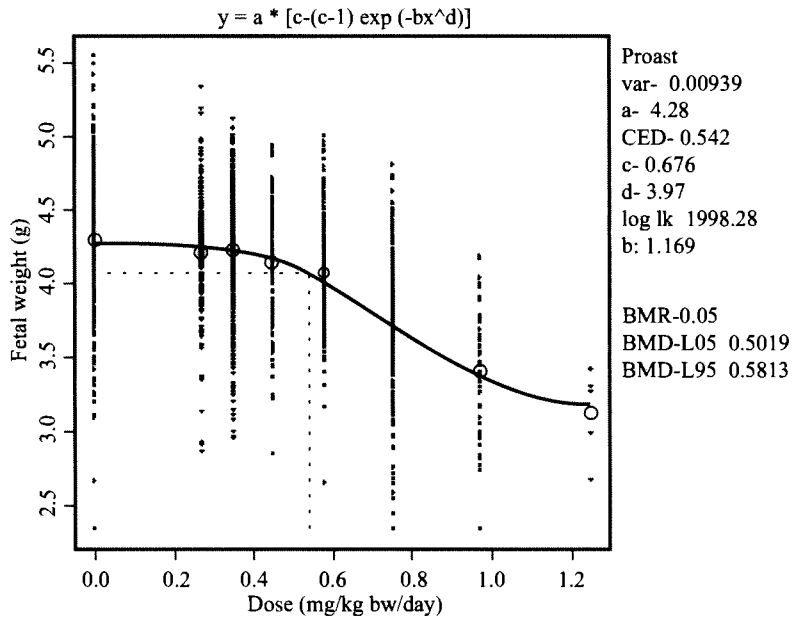
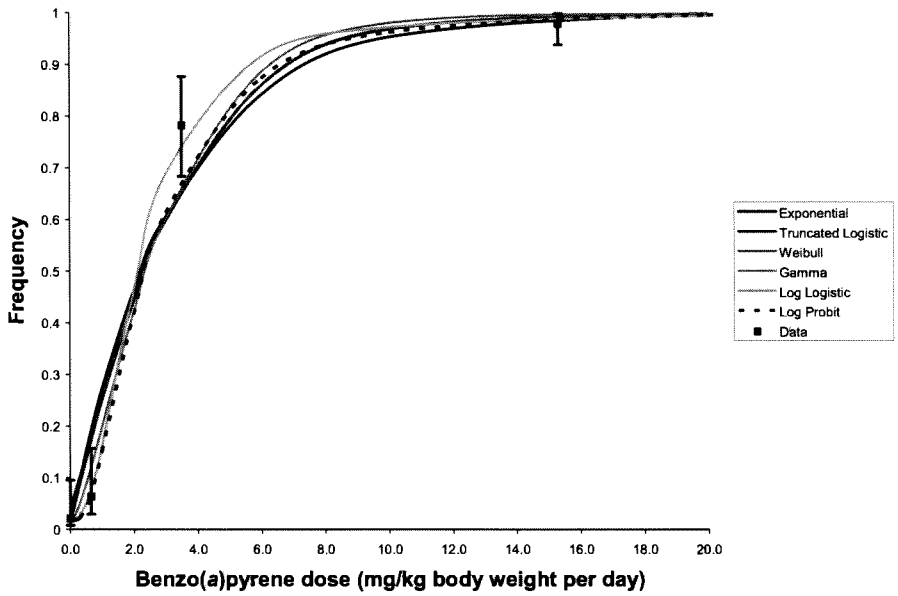


Fig. 8. Two different models (both with four parameters) fitted to the same data set resulting in similar dose–response relationships and similar BMD(L)s. Small circles indicate individual observations, large circles (geometric) group means.

use of probability trees (Rescher, 1969; Hacking, 1976). A probability tree is a logical construct that may be used to represent a set of mutually exclusive propositions. For example, if the three models depicted in the top panel of Figure 8 were equally well supported, then each model would have a probability of 0.33. If one model had a weight that was 6 times greater than the others, then it would have a probability of 0.75, whereas the others would have a probability of 0.125. Note that the probability of a model does not depend only on the strength of evidential support; it also depends on what other models are being considered. A model with little support may have a high probability if all the alternatives under consideration have even less support. Quantitative measures of model preference may be combined to produce an overall rank or to provide a formal measure of the weight of the evidence.

To some extent, all quantitative methods for assigning model probabilities rely on untestable assumptions or elements of judgement. Therefore, the simplest and most straightforward method for assigning probabilities to models is to simply give them all the same weight. This approach is implicit when the predictions from different models are simply listed (e.g. Ghani at al., 2000). Another relatively simple approach is to ask the experts to identify plausible theories and then apply probabilities to them (Evans et al., 1994; IPCS, 2000). These probabilities can then be updated to incorporate additional information in the data by using Bayesian methods. However, there are many formal techniques for assigning weights or probabilities to models (Bozdogan, 1987; Raftery et al., 1997). A semiformal approach may be used in which the same criteria discussed in the section for selecting models (section 6.2.1) may also be used to weight and assign probabilities to each alternative model considered (e.g. Carrington & Bolger, 2000). Model uncertainty may also be integrated with sampling error by using bootstrapping techniques. This involves repeatedly drawing random samples from the data set and refitting each data set with a set of models. The best models from each bootstrap are then retained in a probability tree to represent both parameter and model uncertainty.



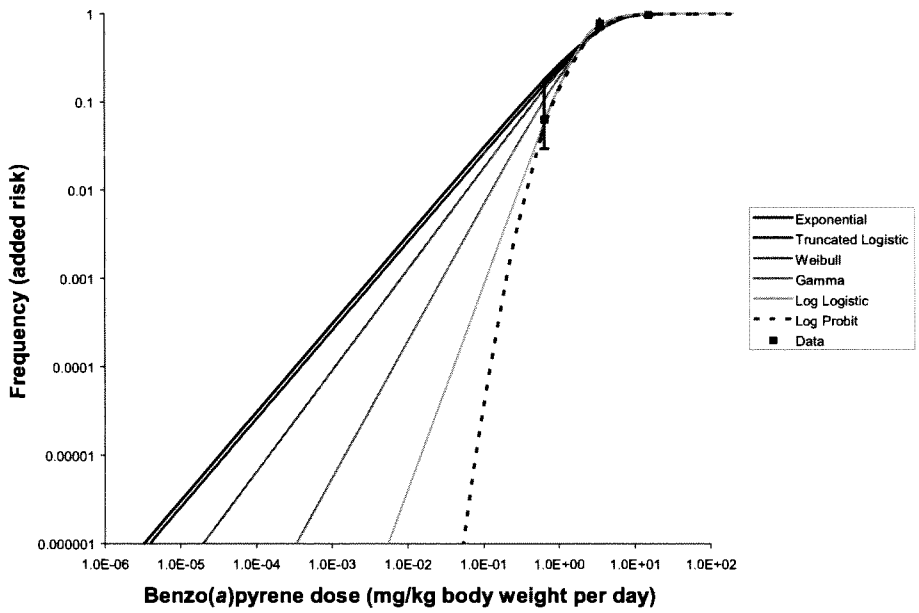


Fig. 9. Model uncertainty in low-dose extrapolation. Different models may all fit the data reasonably well (top), but yield highly divergent response estimates at low doses (bottom). The data and models are taken from Fitzgerald et al. (2004).

Alternatively, some people have addressed this uncertainty by choosing a subset of the models that appear to fit the data well. From these models, those with adequate fits are summarized with a range and associated variance. When choosing a final value for the BMD, these values can be aggregated by taking a mean or geometric mean to provide a central point estimate (National Health and Medical Research Council, 1999) or a value simply chosen through expert judgement (WHO, 2006).

6.6 Benchmark dose and benchmark response selection

One important use of DRM is the calculation of BMDs. A BMD is the dose at which it is inferred that a particular, prespecified level of response would occur. The methodology was introduced in Crump (1984) as an alternative to the use of NOAELs and LOAELs in dose–response assessment for determining quantities such as ADIs. The main advantages of the use of the BMD over NOAELs and LOAELs stems from the more complete use of dose–response data by BMD methods and from the fact that uncertainties about the value of a BMD can be quantified using statistical methodology. The uncertainty of a BMD may be expressed as a confidence interval, in which case the lower end of a one-sided 95% confidence interval is termed the BMDL, or as a full Bayesian posterior distribution.

The BMR is the response for which the BMD is to be calculated. There are both technical and policy aspects associated with selecting the BMR. The technical aspects have to do with just how the BMR is expressed; different types of end-points, such as quantal and continuous, require different treatments. Also, in somewhat more complicated situations, such as when covariates have been used in the modelling, the BMD depends on the BMR and possibly on the values of the covariates. Policy issues have to do with just how high or low down the dose–response curve the BMR should be. This section discusses the technical issues surrounding the choice of BMR and some of the consequences that need to be considered in making the policy decision about where to set the BMR, but it does not directly address the choice of its particular value.

The way in which the BMR is expressed depends upon the kind of response variable being modelled. For end-points with two states (affected/not affected), the BMR is usually expressed in a way that adjusts for background. Two equations are common. One is that of added risk (AR):

$$BMR_{AR} = f(BMD) - f(0)$$

where f_x represents the dose–response function evaluated at dose x. The other, which is probably most widely used, is extra risk (ER):

$$BMR_{ER} = \frac{f(BMD) - f(0)}{1 - f(0)}$$

where added risk is divided by the non-affected fraction of the non-exposed population. The response at the BMD_{ER} is always smaller than the response at the BMD_{AR} for the same numerical value of BMR when there is a background incidence. However, for small to moderate background response, the difference is small.

A third equation, common in epidemiological analyses, but applicable to animal studies as well, is relative risk (RR):

$$BMR_{RR} = f(BMD)/f(0)$$

BMRs for continuous end-points can be expressed directly in terms of changes in the mean response level or indirectly in terms of the fraction of experimental animals that exceed (or drop below) some critical level. For example, the BMD for mean adult body weight might be selected to be the dose at which the mean body weight drops below 90% of the body weight in controls or at which brain acetylcholinesterase activity is inhibited by 10% relative to control levels (this is often termed the critical effect size). One might also specify a fixed value or fixed drop in the mean, selecting, for example, the dose at which the mean nerve conduction velocity drops below a fixed rate or a fixed difference from that in unexposed individuals. For end-points that demonstrate a sigmoidal response, as does enzyme induction, it has been suggested (Murrell et al., 1998; see Gaylor & Aylward, 2004, for a contrary argument) that a formulation similar to extra risk be used: for these end-points, the authors suggest that the BMD is best characterized as the dose at

which the response is a specified fraction of the total dynamic range (e.g. the difference between background and maximum possible induction) of the response. The Gaylor & Aylward (2004) approach considers a certain setting within the definition of the response (i.e. a 1% change) and compares the uncertainties in the resulting BMD with the uncertainties in BMDs estimated using the specific setting in the "hybrid" approach. Thus, their conclusion may not hold in general terms (e.g. considering a 5% or 10% change in response relative to the total dynamic range).

Indirect or "hybrid" approaches have been advocated by Crump (2002) and Gaylor and his co-authors (Gaylor & Slikker, 1994; Kodell et al., 1995). In indirect approaches, the relationship between the mean of a continuous variable and dose is modelled, in the same manner as in the direct approaches. Next, a critical value for the continuous variable is determined that is to be considered as adverse, and an extra (or additional) risk BMR is selected for which to calculate a BMD. It is preferable that the critical value be based upon biological considerations, but it may otherwise be a value in the tail of the distribution of values in the control group. As the mean response increases, so will the fraction of subjects that exceed the previously determined critical value. The BMD is the dose at which the fraction exceeding the critical value corresponds to the fraction of affected animals associated with the BMR as defined for quantal data (e.g. BMR_{ER}).

It is possible to approximate the BMD as calculated in the previous paragraph (Crump, 1995) for a critical value corresponding to a "small" (e.g. 0.1–2%) risk in the control group and extra risk in the vicinity of 10%. This BMD corresponds approximately to the dose at which the mean of the response variable differs from the control mean by an amount equal to the standard deviation of the control group. This gives another way to specify a BMR for continuous variables, based on the variability of the animals used in the bioassays.

Both hybrid methods based on variability discussed above require that the variability be true interindividual variability, and not be due to large assay errors. They depend critically on the idea that extreme quantiles of an unexposed population may be thought of as affected in the same sense as an individual with the same value from

an exposed population. Sand et al. (2003) examined how the hybrid approach depends on the estimate of variance. Gaylor & Slikker (2004) discussed how different sources of variability may be separated.

In some cases, the dose is not the only independent variable in a dose–response model. For example, in epidemiological studies, often many covariates that help characterize an individual and that might influence the response variable and be incidentally associated with the exposure variable are included in analyses in an attempt to reduce bias in the estimates of the effects of exposure (see section 6.2.1.4). In developmental bioassays, characteristics of the dam or the litter as a whole (e.g. number of implantation sites) may be used as a covariate in the modelling to help explain some of the additional variation among litters usually seen in such studies. Even adult-only rodent bioassays are usually segregated by sex. Typically, then, the assessor needs to decide for which values of the covariates BMDs need to be calculated. When there are few, discrete covariates, it may make sense to calculate a separate BMD for each set of values (e.g. a BMD for both males and females). When covariates are continuous (or treated as such, as in number of implantation sites), in an animal bioassay, it is usual to pick a typical value in the control group. However, if BMD changes with the value of the continuous variable, a detailed analysis of the dependency should be undertaken (e.g. modelling the BMD as a function of that covariate). If the variable makes sense for extrapolation to the human situation, it might be informative to calculate the BMDs for several values of the covariate, to evaluate the sensitivity of the BMD to the range of covariate values for humans.

6.7 Summary

Data sets for DRM generally need to be selected to reflect the more sensitive end-points available, just to reduce potential workload. Models used depend upon the type of data (continuous, ordered categorical, quantal, or counts) and include a model for dose–response and a model for the variability of the data. Once models are fit to a data set, the degree to which they individually describe the data is evaluated using goodness-of-fit measures; in addition, their ability to describe the data with respect to each other may be compared using measures such as the AIC.

Uncertainty about the inferences that result from such models fall into three main categories: statistical uncertainty of inferences due to variability among responses in experimental subjects, variability among experiments due to unavoidable differences in experimental execution, and uncertainty due to the fact that different models yield different approximations of the true dose–response relationship. Dose–response analysis needs to address all three sources of uncertainty whenever possible.

One particularly important application of DRM is the calculation of BMDs, doses at which it is inferred that a particular level of response would occur. When data are available, BMDs are a better alternative than NOAELs or LOAELs in the calculation of guidance values such as ADIs or TDIs. When extrapolation is necessary, the uncertainty associated with any predictions made should be represented. It is often especially important to include model uncertainty.

7. COMMUNICATING THE RESULTS OF DOSE–RESPONSE MODELLING

7.1 Introduction

Risk communication has been defined as the "interactive exchange of information about (health or environmental) risks among risk assessors, managers, news media, interested groups and the general public" (IPCS, 2004). Risk communication has evolved with the rest of the risk analysis paradigm to embrace the "interactive" nature of the processes. The transition from monologue to reflexive dialogue in risk communication has necessitated awareness that risk perception issues are extremely important. The scientific, political, and social perspectives of bench scientists, risk assessors, risk managers, media, and the public can result in considerable misunderstandings and misinterpretations (Garvin, 2001). The preconception that scientific and technical knowledge and their application in risk analysis are value free and objective has often resulted in the marginalization of insights from other sources.

General public perception, resulting from health-based guidance approaches and terminology such as "ADI", "TDI", and "threshold", is that there is a bright line between "safe" and "unsafe". These approaches are not designed to incorporate risk and benefit dynamics and may not require or even allow an outside audience to become engaged in the decision process. For many considerations of chemical exposures, these dynamics do not have to be dealt with because the outcome of the safety/risk assessment provides a perfectly useful and acceptable answer to the risk manager. However, there are instances where these dynamics will need to be considered and evaluated.

The use of DRM and other probabilistic assessment techniques to quantitatively describe variability and uncertainty brings new challenges in risk communication. Some of these challenges are:

- explaining that a certain percentage of the population is predicted to experience some effect;

- explaining the level of risk in those circumstances where there is no safe level of exposure;
- comparing competing risks or benefits;
- providing a focus on uncertainties that are attendant to the predicted risk; and
- explaining that the risk generally is described at the population level, rather than the individual level, noting that this is also the case for the ADI/TDI approach.

In addition, one of the limitations of the current health-based guidance approach is that it gives no information about risk when the ADI/TDI is exceeded. For example, some subpopulations may exceed the health-based guidance value for dioxins, and the DRM approach may provide additional information that is useful for the risk manager and communicator.

An appreciation of the variability in most populations clearly impacts risk communication. This is particularly true for genotoxic carcinogens and other substances, such as lead, that are unavoidable contaminants and may be toxic at low levels. Using a point estimate to depict an entire population in the context of risk communication can be misleading, because it can suggest that the risks are larger for the entire population than they really are if upper percentile point estimates are used, and it ignores the fact that some portion of the population does have a somewhat higher level of risk. Becoming involved in a public decision requires a transformation from concern for an individual to concern for a population and thinking about variability as an inherent part of the problem rather than just a source of uncertainty.

In risk communication, uncertainties can facilitate dialogue. Uncertainty analysis can inform all the parties of what is known, what is not known, and the weight of evidence for what is only partially understood. However, there are currently no general criteria for the application of weight-of-evidence approaches. An appreciation of uncertainty, including uncertainty about variability, can lead to better consideration of the options for seeking better information, using a value-of-information approach (Thompson, 2002). However, in risk communication, "uncertainty" can be a double-edged sword. When the results of a probabilistic risk assessment are presented, uncertainty is specifically described rather

than managed by the use of a default factor. Since the responsibility for managing the uncertainty is left to the discretion of the management process, communicating the uncertainty to the participants in that process is very important.

The application of DRM and other probabilistic risk assessment techniques has the potential for improving risk analysis and public risk perception. There must be an acknowledgement of the limitations and weaknesses of the technical knowledge in addition to its strengths. There should also be the realization that there may be difficulties with risk comparisons and that social perceptions can drive precautionary considerations. There may not be agreement on how to interpret new information or on the appropriate criteria for making or reversing risk decisions. The critical contribution of probabilistic approaches is that they can improve the processes of risk assessment and risk management and thereby facilitate communication. As a result, participation in the decision process will be broadened.

7.2 Incorporation of the outputs of dose–response modelling into risk assessment

The output of dose–response analysis can be used in various ways, depending on problem formulation and the nature of the effect modelled. An output may be presented in three principal ways as the basis for advice on the possible health implications of human exposure:

1. establishment of a health-based guidance value, such as an ADI or TDI, which is a daily intake over a lifetime that is considered to be without appreciable health risk (this would be analogous to current procedures based on a NOAEL or LOAEL);
2. estimation of the MOE as the ratio between the dose–response output and the estimate of human exposure; and
3. quantitative estimation of the magnitude of the risk at the level of human exposure, derived from the modelled dose–response relationship.

The discussion below assumes that the dose used in the dose–response model was the external dose expressed in milligrams per kilogram body weight. The use of internal or target organ dose estimated by a physiologically based toxicokinetic model would

reduce the uncertainties of interspecies extrapolation, because kinetics are a major source of species differences, such that a reduced uncertainty factor would be required.

7.3 Derivation of health-based guidance values

Traditionally, a health-based guidance value for threshold effects has been derived from a NOAEL or LOAEL divided by an appropriate composite uncertainty factor, either default values or CSAFs (IPCS, 2005), on the assumption that the NOAEL represents an intake close to the threshold for the adverse effect. In practice, the limit of detection for the incidence of adverse effects in animal experiments depends on the sample size, and more than 100 animals may be needed to achieve confidence intervals in the range of ±5%.

Many studies have shown that the BMDL for a 5% response is similar to the experimental NOAEL (Allen et al., 1994). Fowles et al. (1999) came to a somewhat different conclusion. They examined acute inhalation lethality data and compared NOAELs with BMDs corresponding to 1%, 5%, and 10% response incidences. Similarly to the "quantal" parts of the results of the Allen et al. (1994) studies, BMDLs based on 10% incidence corresponded approximately to NOAELs. However, because the dose–response for lethality is so steep, BMDLs for 5% and 1% incidences were very close to those for 10% incidence. As a result, the BMDLs for a 1% incidence were on average only about 1.6 or 3.6 times smaller than a NOAEL, depending on whether a log-probit or Weibull model was used. This possibly can be explained by the smaller sample sizes in these experiments, not by the difference in end-points.

Given the uncertainty in the relationship of the NOAEL and the threshold of the adverse effect, finding a BMR such that the resulting BMD and BMDL correspond numerically (on average) to a NOAEL may not be relevant and is certainly not necessary for the application of BMD approaches. Also, the use of the BMDL to set a health-based guidance value would need to take into account the same uncertainties as when a NOAEL is used as the basis for establishing an ADI/TDI.

7.4 Estimation of the margin of exposure

The normal default uncertainty factor of 100 has a long history of use for threshold effects and can be regarded as the margin between two points—the NOAEL or BMDL from the experimental data and a level of human intake/exposure that would be without appreciable health risk. Because this is based on a NOAEL or BMDL, the ratio is equivalent to a margin of safety, and there would be negligible risk providing that the intake was at or less than the ADI/TDI.

In the case of adverse effects that are considered not to show a biological threshold in their dose–response, the BMDL could not be considered to represent an intake close to a threshold, but is simply the confidence interval on the BMD. Consequently, the margin between the BMDL and the estimated human intake/exposure would not be a margin or safety and is therefore termed an MOE. The MOE is calculated as the ratio between two experimental estimates, the BMDL and the predicted or estimated human intake/exposure. Calculation of an MOE does not require extrapolation of the data beyond the range of observations (IPCS, 1999; Edler et al., 2002).

Uncertainties related to interspecies differences and human variability, which are the basis for the usual 100-fold uncertainty factor used in the derivation of an ADI/TDI, would be equally applicable to an MOE based on animal data, but there would be additional uncertainties related to the nature of the dose–response relationship below the experimental/observable range, the impact of genetic polymorphisms in the processes critical to the production of a mutated cell, and the subsequent clonal expansion and progression into a cancer. Consequently, an MOE of 100 would be inadequate to reflect the fact that the starting point (the BMDL) cannot be regarded as a threshold or the additional uncertainties related to the mode of action.

Application of linear low-dose extrapolation using the BMDL for a 5% response (see below) to estimate a one in a million lifetime risk is equivalent to an MOE of 50 000.

7.5 Quantitative estimations of the magnitude of the risk at levels of human exposure

The results of a dose–response model can be used to estimate the possible risks at intakes/exposures above a health-based guidance value such as the ADI and at very low levels of human exposure or to estimate intakes/exposures associated with predefined levels of risk, such as a one in a million lifetime risk of cancer.

Estimation of risks of intakes/exposures above a health-based guidance value, derived by the application of uncertainty factors to a BMDL from a study in either animals or humans, would need to use the slope characteristics in the dose–response model. For example, if an intake is of concern because it is above the health-based guidance value, then the extent of any risk could be estimated by reference back to the modelled animal dose–response relationship. Traditionally, an estimate of the possible risk has not been made, and intakes above the ADI/TDI have been considered to have eroded the uncertainty factor. However, if one assumes that the dose–response relationship in humans has a similar shape to that in the animal study, the ADI is set with default uncertainty factors that will obscure any quantitative estimates of the risk above the ADI (the risk at the ADI is assumed to be negligible). More accurate estimates of differences in sensitivity between humans and animals would be required for such calculations.

Estimation of risks at very low levels of human exposure or of exposures associated with responses below the BMR requires extrapolation outside the data used to generate the dose–response model. Extrapolation outside the observed range—for example, from an incidence of about 5% to one in a million—will require extrapolation over many orders of magnitude. Low-dose extrapolation may be undertaken using the dose–response relationship defined by the model that was fitted to the experimental data or by application of a standardized mathematical approach, such as linear extrapolation, to the starting point. An advantage of using a model is that the risk estimates can be compared across different compounds. The major uncertainty associated with such estimates is the biological relevance of the model in the region of extrapolation.

7.6 Presentation of results

In a scientific or logical sense, the risk assessment is finished when the conclusions have been drawn. However, when the conclusions are simulation results, some distillation or condensation is often necessary in order to make the results comprehensible. Since there is always some danger that crucial information may be lost, care must be exercised to ensure that the summary process does not omit information that is important for the decision.

7.6.1 Tables

Precise communication of quantitative information requires numbers. More numbers will portray more information than fewer numbers, but will take longer to assimilate. Tables 6–10 give examples of the range of options, from high to low complexity, that may be considered, all taken from the same simulation results for exposure. It is recommended that, in case of effects of concern or a single effect found in several studies, all quantitative results be summarized in a table. The risk assessor should sort out the most relevant results and present the data to the risk manager in a clear and understandable way.

7.6.2 Graphs

Although they may allow quick comparison, tables inherently compare one value at a time. Graphing or visualization is in some ways a better means of digesting the entire distribution. A one-dimensional simulation will produce a frequency distribution (when simulating variability) or a likelihood distribution (when representing uncertainty). There are two ways of plotting frequency or likelihood curves (see Figure 10). The first is to plot density against value, which emphasizes the values that are the most common or likely. The second is to plot cumulative percentiles against value, which allows the percentile corresponding to a particular value to be read from the plot. A graphical presentation of the dose modelling in relation to the experimental data may also be helpful in deciding which dose descriptor should be used for lifetime risk.

Table 6. Population percentiles from a two-dimensional simulation

		Uncertainty												
		Average	SD	Minimum	P1	P5	P10	P25	Median	P75	P90	P95	P99	Maximum
Variability	Average	0.457	0.063	0.234	0.236	0.366	0.403	0.456	0.462	0.497	0.502	0.503	0.510	0.874
	Minimum	0.047	0.061	0.000	0.000	0.000	0.000	0.016	0.055	0.076	0.076	0.076	0.076	0.874
	P1	0.094	0.065	0.000	0.000	0.000	0.007	0.072	0.101	0.129	0.129	0.130	0.131	0.874
	P5	0.146	0.068	0.000	0.000	0.000	0.069	0.144	0.148	0.178	0.179	0.180	0.180	0.874
	P10	0.188	0.074	0.000	0.000	0.000	0.116	0.187	0.205	0.216	0.216	0.217	0.218	0.874
	P25	0.274	0.083	0.000	0.000	0.119	0.207	0.287	0.291	0.317	0.320	0.320	0.327	0.874
	Median	0.401	0.105	0.000	0.000	0.267	0.352	0.399	0.404	0.471	0.476	0.476	0.484	0.874
	P75	0.586	0.064	0.388	0.394	0.519	0.531	0.561	0.568	0.651	0.657	0.657	0.667	0.874
	P90	0.808	0.030	0.760	0.762	0.774	0.776	0.784	0.790	0.843	0.847	0.848	0.858	0.874
	P95	0.949	0.024	0.874	0.923	0.930	0.931	0.941	0.944	0.953	0.963	1.014	1.056	1.058
	P99	1.247	0.086	0.874	1.138	1.142	1.147	1.149	1.287	1.296	1.321	1.403	1.462	1.473
	Maximum	2.192	0.483	0.875	1.573	1.579	1.584	1.599	2.559	2.592	2.608	2.619	2.663	2.670

Pxx = xxth percentile; SD = standard deviation.

Table 7. Population percentiles with confidence intervals

Percentile	Average (confidence interval)
Average	0.457 (0.366, 0.503)
Minimum	0.047 (0.000, 0.076)
1st percentile	0.094 (0.000, 0.130)
5th percentile	0.146 (0.000, 0.180)
10th percentile	0.188 (0.000, 0.217)
25th percentile	0.274 (0.119, 0.320)
Median	0.401 (0.267, 0.476)
75th percentile	0.586 (0.519, 0.657)
90th percentile	0.808 (0.774, 0.848)
95th percentile	0.949 (0.930, 1.014)
99th percentile	1.247 (1.142, 1.403)
Maximum	2.192 (1.579, 2.619)

Table 8. Population percentiles with standard deviations

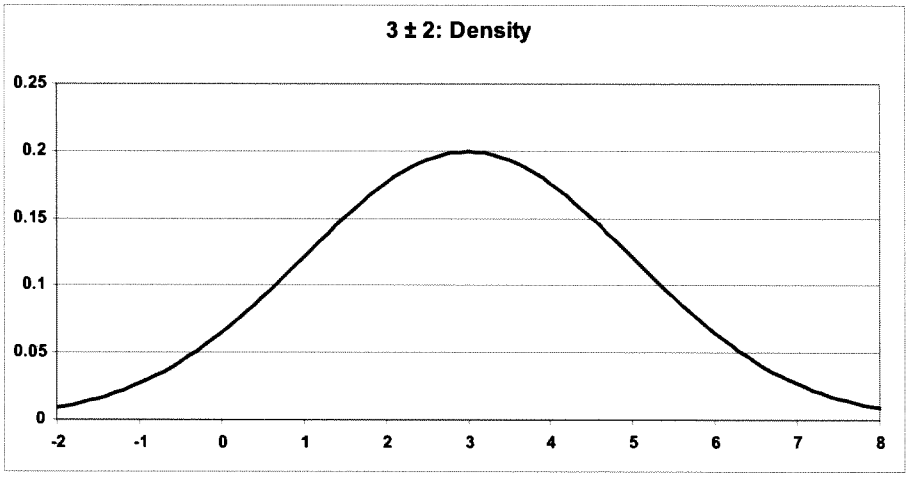
Percentile	Average ± standard deviation
Average	0.457 ± 0.063
Minimum	0.047 ± 0.061
1st percentile	0.094 ± 0.065
5th percentile	0.146 ± 0.068
10th percentile	0.188 ± 0.074
25th percentile	0.274 ± 0.083
Median	0.401 ± 0.105
75th percentile	0.586 ± 0.064
90th percentile	0.808 ± 0.030
95th percentile	0.949 ± 0.024
99th percentile	1.247 ± 0.086
Maximum	2.192 ± 0.483

Table 9. Selected population percentiles with confidence intervals

Percentile	Average (confidence interval)
Average	0.457 (0.366, 0.503)
Median	0.401 (0.267, 0.476)
90th percentile	0.808 (0.774, 0.848)
95th percentile	0.949 (0.930, 1.014)
99th percentile	1.247 (1.142, 1.403)

Table 10. Population mean with uncertainty estimate

	Average ± standard deviation
Average	0.457 ± 0.063



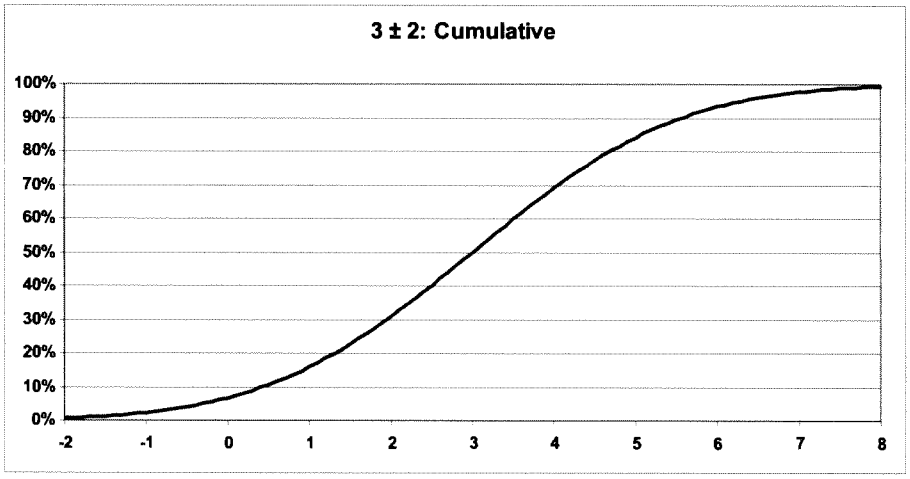
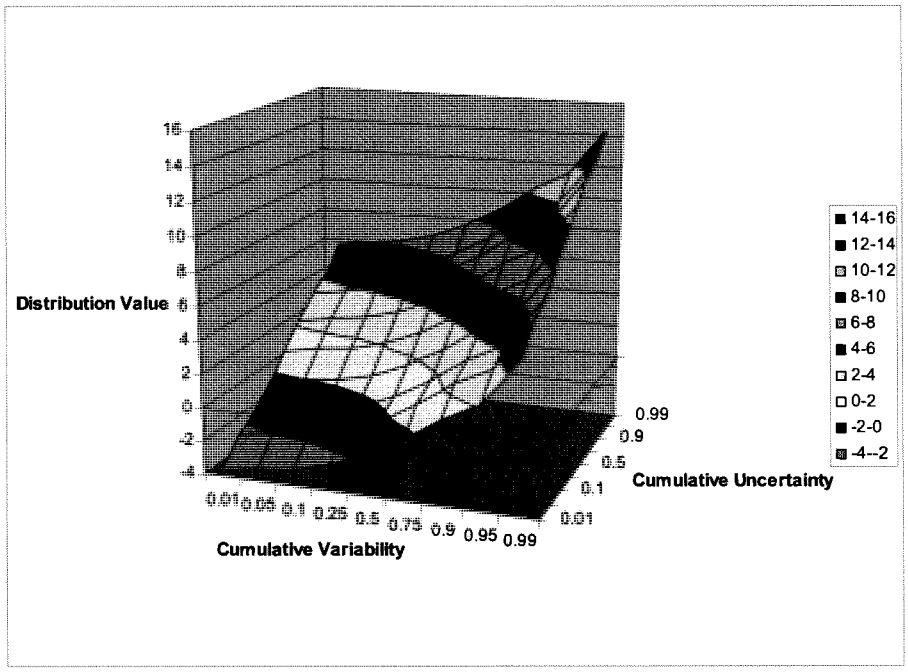


Fig. 10. Plotting frequency distributions.

Two-dimensional results are more difficult to display. Two strategies for adding an extra dimension are illustrated in Figure 11. The first uses three-dimensional perspective to portray the third dimension. The second uses shading, where darker hues are used to represent either higher density or more central values. This is particularly of use for displaying uncertainty, as the less well defined (more uncertain) parts of a curve appear fuzzy.



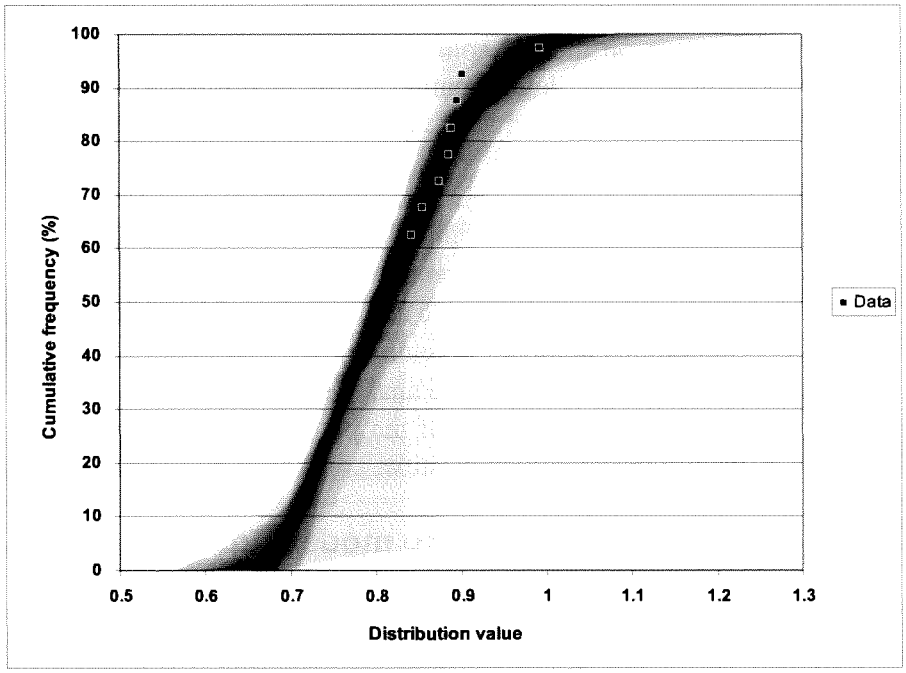


Fig. 11. Plotting results of three-dimensional simulations.

7.7 Risk assessment context and questions

The output of the DRM should be directed towards addressing specific questions about the likelihood of adverse health effects in response to exposure to chemicals. This would build on conventional risk assessment procedures that have been accepted internationally as the indicator for determining acceptable levels of exposure. These rely on the identification of a NOAEL/no-observed-effect level (NOEL) for a critical end-point in the effect data and incorporation of uncertainty factors to allow for interspecies and interindividual variation.

DRM offers the potential to provide additional information for the risk manager, specifically a more scientifically robust method for determining the health-based guidance values (e.g. ADI) using the BMD and better information on the likelihood of effects at low doses that are below the levels observed in biological systems. The mathematical models will also provide estimates of the statistical uncertainty surrounding estimates of likely effect.

Whether traditional safety-based assessments or DRM assessments are carried out, the risk manager will still require information on the toxicology of the adverse health effect and the robustness of the determination of the health-based guidance value to help inform the management options. This may include the following:

- a discussion of the strength and weight of evidence;
- uncertainties and gaps in the data;
- information on the nature and severity of the (critical) effect;
- limitations in the interpretation;
- assumptions made in the analysis; and
- qualitative assessment of the potential effects of exceeding the health-based guidance value.

7.8 Synopsis of approach to modelling

DRM involves six basic steps: data selection, model selection, statistical linkage, parameter estimation, implementation, and evaluation (see chapter 4, Table 1). In undertaking a DRM exercise,

two factors that will impact the types of outputs and that may be of importance to the risk manager are briefly described below.

7.8.1 Data sets

Traditional safety assessments focus primarily on a single critical end-point, whereas DRM gives the potential for separating out multiple end-points. Modelling outcomes may be based on data from single or multiple experiments. In the latter situation, meta-analysis may integrate the results of several independent studies that are considered to be "combinable".

The risk manager could see four types of data from the modelling evaluations: namely, quantal, count, continuous, and ordinal categorical data. The risk manager will need to understand what data sets were modelled and, if quantitative information from more than one data set is presented, will need guidance on the rationale for forwarding the additional data set information and for synthesizing this additional information. This guidance may include information about the consistency (or inconsistency) of the quantitative response across the end-points. Such information could be used by the risk manager to strengthen (or weaken) his or her confidence in the quantitative evaluation of the potential for health impacts.

If DRM information is available from human epidemiological evaluation, then an understanding of both the strengths as well as the possible limitations (often in the quantitative exposure information) of the data set may also temper or strengthen the qualitative or quantitative assessment from the animal studies.

7.8.2 Uncertainty

DRM should capture the relative uncertainties in the estimates of risk. This information will allow the generation of confidence limits on health-based guidance values. However, such confidence limits will still capture only one part of the uncertainty inherent in these estimates. The risk manager will need to know what uncertainty is accounted for in the information provided, and the risk assessment information will need to clearly indicate what uncertainty is not accounted for in a quantitative assessment.

One approach that has been used to capture variability in population response is calculation of population percentiles. Availability of dose–response functions when linked with population-based exposure assessments has allowed risk managers to calculate percentiles of populations above target exposure or intake levels. Likewise, dose–response functions have also been utilized to calculate percentiles of the population above target risk levels.

One of the advantages of DRM is that the confidence limit around the BMD can be calculated. From the conservative point of view, the lower limit of the dose is most important. However, this is not the same as to say that the confidence limit of the health-based guidance values can be calculated, as the uncertainty factors will obscure such estimates.

7.9 Explaining/interpreting the output of the dose–response analysis

Advice to the risk manager should describe the uncertainties inherent in such an approach to the use of dose–response data, such as uncertainties in the slope estimate in animals, the relevance of this slope to humans (such an approach is more appropriate if the response is a continuous variable, rather than quantal), and the appropriateness of the uncertainty factor applied to allow for species differences and human variability.

7.9.1 Outputs in the observable biological range

The output of the analysis takes the form of a numerical quantity—at present, commonly a TDI or ADI derived from a NOAEL, which is a single point in the dose–response relationship. The dose–response analysis uses more of the available information by fitting a mathematical model to all the data in the observable biological range and then determining the dose associated with a specified response level. A statistical lower bound (e.g. the 95% lower bound on the dose) is often used to account for statistical uncertainties (a BMDL) and for the level of health protection required by the risk manager. As with the NOAEL, the BMDL can be used as the starting point for deriving a health-based guidance value and/or MOE. However, unlike the NOAEL, the BMD

approach uses the whole range of experimental dose–response data, and therefore it is not limited by the doses selected by the investigators.

7.9.1.1 Health-based guidance values

On the basis of current practice, it appears that the BMD approach leads to doses that are usually quite similar to NOAELs for the studies in question (see section 7.3). In the same way as for the derivation of the ADI/TDI, uncertainty factors, for example 100, are applied to the BMDL to obtain the health-based guidance value. However, the confidence intervals that are possible in the case of the BMD-derived health-based guidance value provide the risk manager with an increased understanding of the uncertainty associated with the risk assessment. This allows a more informed decision to be made when choosing among risk management options.

7.9.1.2 Margin of exposure

An MOE is determined by comparing the point of departure (the BMDL) with the actual or estimated human exposure. The MOE is used when limited toxicological or human data exist but the hazard identification and characterization data are insufficient to set a health-based guidance value. Alternatively, the MOE approach is used when it is inappropriate to derive a health-based guidance value owing to the nature of the effect, such as for substances that are genotoxic and carcinogenic.

The acceptability of an MOE depends on its magnitude and is ultimately a risk management decision. To aid that decision, the risk assessor should provide information on the nature of the toxicity involved and nature and magnitude of the uncertainties, from both the toxicological and exposure perspectives. Although the risk assessor should not provide an assessment of the acceptability of the MOE, guidance on its adequacy, taking into account the severity/nature of the toxicity, uncertainties, and variability, should be given—for example, in terms of high, medium, or low concern. The use of all the data by the dose–response analysis enables the uncertainties to be better defined. The MOE can also be used by the risk manager for priority setting.

There is no internationally accepted value for an MOE for a genotoxic and carcinogenic compound such that the exposure would not be a significant health risk. However, several institutions and countries have used the MOE approach, and their conclusions provide examples of MOE values that have been considered acceptable:

- The National Health and Medical Research Council in Australia concluded that a guideline dose for carcinogens present in soil could be calculated by application of uncertainty factors up to 50 000 to the BMD (not BMDL). The factor applied in any particular case would depend on the nature of the effects (National Health and Medical Research Council, 1999).

- The reciprocal of the MOE, the exposure potency index (EPI), has been used by Health Canada for genotoxic and carcinogenic compounds in their *Human Health Risk Assessment for Priority Substances* under the Canadian Environmental Protection Act (Health Canada, 1994). MOE values of <5000, 5000–500 000, and >500 000 indicate high, medium, and low priority, respectively.

- The Committee on Carcinogenicity in the United Kingdom considered derivation of the minimal risk level for a genotoxic and carcinogen compound. One proposal was that an adequate MOE for carcinogenicity might be 10 000 (Gaylor et al., 1999; Gold et al., 2003). A particular carcinogenic impurity posed a negligible carcinogenic risk if an uncertainty factor of 10 000 was applied to the estimated 5% BMD (BMD_5) (Committee on Carcinogenicity, 2003). The MOE for average intakes for acrylamide in men in Norway has been estimated using the T_{25} value[1] and the LED_{10} (the lower bound on the effective dose for a 10% increase in risk) (approximately equivalent to $BMDL_{10}$) methods. These approaches result in MOE values of 1306 and 1225 for T_{25} and LED_{10}, respectively.

[1] The tumorigenic descriptor T_{25} is the chronic daily dose that will give 25% of the animals tumours above background at a specific tissue site. The T_{25} is determined by linear interpolation from the lowest dose giving a statistically significant increase in tumours (Dybing et al., 1997).

- The 64th (WHO, 2006) and 67th (WHO, 2007a) meetings of JECFA used MOE approaches for the evaluation of several substances that were genotoxic and carcinogenic. The 64th JECFA developed general considerations for the formulation of advice on compounds that are both genotoxic and carcinogenic. This meeting established MOEs for acrylamide, ethyl carbamate, polybrominated diphenyl ethers, and polycyclic aromatic hydrocarbons. The 67th JECFA established an MOE for 1,3-dichloro-2-propanol.

- A joint European Food Safety Authority/WHO conference on the risk assessment of substances that are both genotoxic and carcinogenic (Barlow et al., 2006) compared the approaches that are currently used. "This conference concluded that the MOE approach was a useful and pragmatic option...."

- O'Brien et al. (2006) presented a critical appraisal of the approaches to the risk assessment of genotoxic carcinogens in food and concluded that "Overall, MOE is the most appropriate default approach because it combines information on potency and exposure, without the generation of numerical risk estimates of unknown reliability." They presented case-studies on the calculation of MOEs for acrylamide, aflatoxin B_1, benzo(*a*)pyrene, dimethylnitrosamine, ethyl carbamate, and 2-amino-1-methyl-6-phenylimidazo(4,5*b*)pyridine.

7.9.2 Outputs outside the observable biological range

DRM evaluations can produce information in several formats, including dose–response functions that allow, along with estimates of exposure, the prediction of risks at specified exposure levels and functions that allow the estimation of exposure levels resulting in specified risks. In addition, DRM exercises can provide uncertainty analyses. The availability of such outputs from DRM exercises can provide both opportunities for additional assessment as well as challenges in interpretation for the risk manager.

Three different methods have been used or proposed for quantitative risk assessment by regulatory authorities in the United States and Europe for non-threshold (genotoxic) carcinogens. In the area of food safety, the United States Food and Drug Administration has used a simple, direct method for low-dose

cancer risk assessment. A point on the dose–response curve is chosen below which the data no longer appear to be reliable (e.g. 1–10% tumour incidence), and a straight line is drawn from the upper confidence limit on risk at that point to the origin (Gaylor et al., 1997). The linearized multistage model was previously extensively used by the United States Environmental Protection Agency (USEPA, 1986). The LED_{10} method was later proposed by the USEPA (1996), and the T_{25} (Dybing et al., 1997; Sanner et al., 2001) method has been used in Europe (European Commission, 1999; SCCNFP, 2003). Lifetime cancer hazards may be estimated by linear extrapolation using LED_{10} and T_{25} as starting points. The results obtained with these extrapolation methods are in most cases nearly indistinguishable (Sanner et al., 2001). A measure for an assessment of concern may be arrived at by comparing the calculated risks for some specific scenario of human exposure to such substances, with some default policy-determined risk level.

7.9.2.1 Prediction of risks at specified exposure levels

One type of output from DRM is the prediction of risks at specified exposure levels. This output can take the generic form of predicting "X number of health-impacted individuals at exposure Y". Examples of such estimates have been used to predict the number of excess lung cancer deaths due to smoking two packs of cigarettes per day, the number of excess skin cancers from arsenic-contaminated water, and the number of excess mortality cases due to air pollution. In the optimal case, such estimates are supported by parallel assessments that describe the uncertainty in such estimates, by providing additional information on the range of estimates, rather than a single value. The risk manager can then make such statements as "Up to X individuals may be impacted by exposure Y". This same information can allow the risk manager to see how low the estimates of the health impact may be; when confidence limits are included in such estimates, many uncertain health impacts can be shown to include the potential for no health impacts. Assumptions inherent in such estimates that can impact interpretation by the risk manager include choice of models, choice of end-points, and limitations in initial data sets that were extrapolated.

One use of such information has been to evaluate the effect of different maximum limits for a chemical on risks. This type of consideration was included when JECFA evaluated aflatoxin B_1 and the impact of different maximum limits on risk (WHO, 1999, 2007b). Similar assessments have also been performed for lead and fumonisins B_1 and B_2 (Carrington et al., 1996; Humphreys et al., 2001). For example, the health impacts of current particulate standards (WHO, 2000, 2003) have been estimated. Availability of such estimates can provide additional information for risk managers to conduct cost–benefit analyses, risk–benefit assessments, and evaluations of public health interventions.

7.9.2.2 Prediction of exposure levels producing specified risk levels

Another type of output from DRM is risk level estimates. In these estimates, a specific level of risk is evaluated and the amount of exposure that would be estimated to result in that risk is determined. For example, a common level of risk related to carcinogen exposures that has been evaluated in the United States has been 10^{-6} over a lifetime. Estimates of exposure that would result in that level of risk have been determined, and such estimates have been made for approximately 100 environmental pollutants (http://www.epa.gov/iris). For the risk manager, availability of such estimates can allow for development of risk-based consistency in proposed regulatory actions.

7.9.2.3 Uncertainty analyses

A third type of output from DRM is that linked with uncertainty analysis. One example of such approaches is when the DRM output is linked with distributions of population effects with confidence intervals. The result from such analysis is a distribution of potential population risks. For example, in Figure 12, the outputs for three models and two data sets were used to generate a set of 3000 different model parameters (two data sets, three models, 500 bootstraps).

One approach that has been used to extrapolate dose–response models beyond bioassay data has focused on the use of biomarker data to extend the dose–response curve 1–2 orders of magnitude closer to environmentally relevant exposures. Such approaches can be facilitated when DRM data are available.

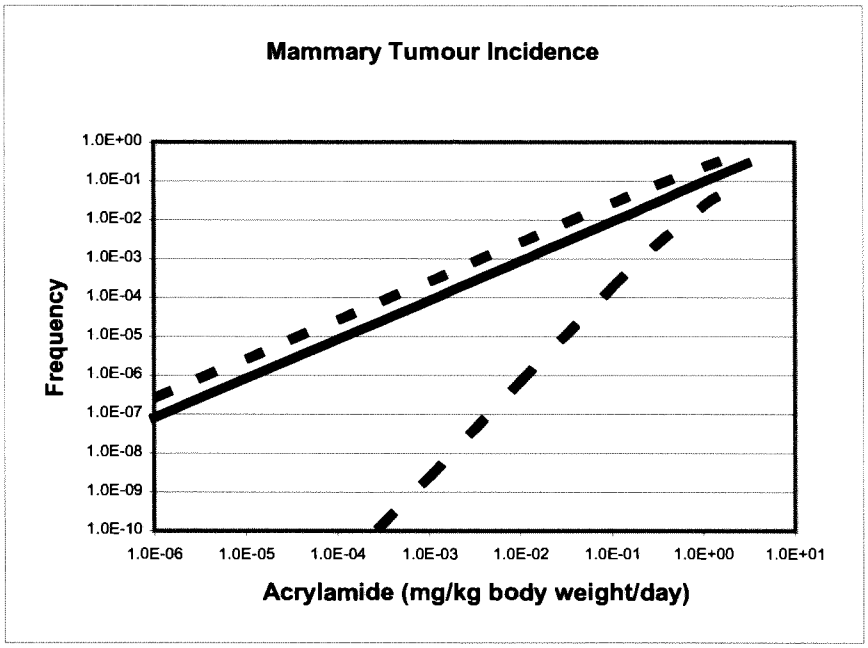


Fig. 12. Integrated uncertainty analysis for mammary tumours. The dark line is the central (median) estimate, and the dotted lines are the 5% and 95% confidence limits.

All these modelling approaches exhibit similar limitations and difficulties. A benefit is that DRM allows for the transfer of more quantitative toxicological data into risk manager assessment methods such as cost–benefit and risk–benefit analyses. The limitation is the question of whether the model outputs are accurate and representative of public health impacts.

7.10 Issues for risk managers

7.10.1 Risk assessment issues

7.10.1.1 Population versus individual effects

The potential health effect at the population level can be informed by DRM. However, as the behaviour, environment, or biological characteristics may vary among individuals, a dose–response model may need to describe or model these characteristics to produce a prediction of adverse health effects in the population.

The output of the dose–response model should identify the degree of any subpopulation effects.

7.10.1.2 Risk characterization

The actual risk to the population of an adverse health effect requires consideration of both the likelihood and severity of the effect, as determined from the dose–response model when combined with the exposure to the chemical in the population under consideration. The exposure may be determined from consumption surveys, measurement of environmental media, direct contact information, or biomarkers (e.g. IPCS, 2000; Kroes et al., 2002).

Consideration of the DRM data together with exposure data will help identify populations at risk. This information, together with knowledge about the severity of the adverse health effects, will inform the risk management options.

7.10.2 Risk management issues

7.10.2.1 Risk management options

A risk assessment can be used to establish that a risk is of a sufficient magnitude that regulation or other type of intervention may be warranted. DRM can then be used to evaluate the consequences of possible interventions that aim to reduce the risk. That is, a model may be used to estimate change in the likelihood of the adverse health effect occurring following implementation of a particular intervention. To date, alternative risk management options have been evaluated using DRM in a limited number of cases. For example, at the request of the Codex Committee on Food Additives and Contaminants, the 49th JECFA analysed the application of two hypothetical standards for aflatoxin contamination in food in model populations (WHO, 1999).

A range of risk management interventions are available, with the types of interventions varying from a ban on a particular product (e.g. carcinogenic antibiotics, DDT), establishing regulatory limits (e.g. aflatoxins), advice on consumption or use patterns (e.g. consumption of predatory fish that accumulate high levels of methylmercury), and control at source of production (e.g. emissions of dioxins).

7.10.2.2 Cost–benefit and risk–benefit analyses

While health risk management decisions should be based on risk assessments, a number of other factors will influence the final decisions. In particular, it may also be necessary to undertake a cost–benefit analysis (e.g. health costs to the community from exposure to aflatoxins versus the cost of implementation of a management strategy) and/or risk–benefit analysis (e.g. risk associated with methylmercury in fish versus nutritional benefits of fish consumption) and to assess the feasibility of the intervention, availability of alternatives, and loss of products of economic value. These factors are beyond the scope of the assessment of the risks and will need to reflect wider societal factors.

7.10.2.3 Acceptable level of risk

Different institutions and countries may make different risk management decisions based on different perceptions of the risk that is deemed to be acceptable to society. The ADI, which usually incorporates a composite uncertainty factor of 100 when based on animal studies, has been accepted by international institutions and countries as a health-based guidance value. Although DRM can give a prediction of the risk at various exposures, there is no international agreement on how to interpret this new information, the appropriate criteria for making or reversing risk decisions, or the acceptable level of risk determined using this technique.

A predicted risk level, such as 10^{-6}, determined from dose–response analysis has been used by some countries and institutions as being not appreciable or negligible (virtually safe dose). Variations around the calculated risk by a factor of about 10 trigger further consideration of the qualitative aspects of the risk assessment, such as variability and uncertainty (Sanner et al., 2001; SCCNFP, 2003). In the case of compounds in drinking-water considered to be genotoxic carcinogens, WHO has assigned guideline values associated with an estimated upper-bound excess lifetime cancer risk of 10^{-5} determined by a mathematical model (WHO, 2004b). The United States Occupational Safety and Health Administration has considered a lifetime cancer risk for workers higher than 10^{-3} to represent an unacceptably high risk, and its goal is to reduce this risk to less than 10^{-5} (OSHA, 1983, 1984).

Proposals for the application of lifetime risk estimates in establishing tolerable risk levels have also been published in Europe (Bos et al., 2004).

8. CONCLUSIONS AND RECOMMENDATIONS

8.1 Conclusions

- Full DRM can be considered a more sophisticated or robust alternative to the NOAEL approach in all cases where suitable dose–response data are available (e.g. several dose groups with different response levels).

- For quantal dose–response data, the interest is often in low response (incidence) levels. This may call for low-dose extrapolation by several orders of magnitude (e.g. for tumour incidences). However, equally plausible dose–response risk models may result in highly divergent low estimates. A currently applied approach is to estimate a BMD_{10} and linearly extrapolate from that point downwards, as a conservative approach. Another option, currently under development, is to apply a Bayesian approach that considers the various models all together.

- For continuous dose–response data, two approaches of DRM exist. One is to transform the continuous data into quantal data. The other is to consider continuous dose–response data as information on the severity of the effect and therefore as a function of dose. In the latter approach, measurable changes of effect are often close to response levels considered as adverse (e.g. 10% inhibition of cholinesterase), and the low-dose extrapolation problem is minor or non-existent.

- For the purpose of deriving an ADI, TDI, or RfD, DRM may be used for deriving a BMD, to be used as a point of departure in the same way as the NOAEL is used (i.e. the same uncertainty factors would be applied to the BMD as to the NOAEL).

- DRM may also be used for estimating risks at a given (human) exposure level. For risks in terms of incidences (quantal data), this may involve low-dose extrapolation.

- DRM exercises can provide information on uncertainties associated with the data and identify factors contributing to uncertainties in risk estimates.

- Application of DRM for all end-points can be cost prohibitive, so it is efficient to pre-select the apparently more sensitive end-points. In some

cases, however, it is not easy to identify the most sensitive end-points by visual inspection, so all of the end-points may need to be modelled.

- The BMD and the BMDL should always be reported, so that the quality of the data and the model fit are clear and potencies can be compared on the basis of the BMD.

- The output of the different models used in DRM should be presented.

8.2 Recommendations

- Toxicity testing protocols (e.g. Organisation for Economic Co-operation and Development guidelines) should be reviewed for optimization for BMD and other DRM approaches, including optimal designs for the number of animals and number of doses for different dose–response curves. Additional research is needed for the development of optimal study designs. Guidance should be developed for combining existing studies with a view to DRM.

- Better guidance needs to be developed for combined analysis of different data sets for more precisely estimating BMDs.

- Better understanding of when and how to use the BMR needs to be developed.

- Better understanding of the shape of the dose–response curve at low doses needs to be developed. Additional research is needed to determine the biological basis for extrapolation (e.g. by using biomarkers, tumour precursors, genetically modified animals, and toxicokinetics for target dose estimation).

- Improved guidance needs to be developed for risk communication based on the results of DRM and probabilistic assessment techniques. This should include communication of the types of uncertainty and the relation to statistical variability, imprecision, and the use of confidence intervals.

- The use of DRM should be reviewed and additional general principles for its use developed when more experience becomes available.

REFERENCES

Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov BN & Csaki F ed. Proceedings of the second international symposium on information theory. Budapest, Akademiai Kiado, pp 267–281.

Allen BC, Kavlock RJ, Kimmel CA, & Faustman EM (1994) Dose–response assessment for developmental toxicity: II. Comparison of generic benchmark dose estimates with NOAELs. Fundam Appl Toxicol, **23**: 487–495.

Allen BC, Strong PL, Price CJ, Hubbard SA, & Daston GP (1996) Benchmark dose analysis of developmental toxicity in rats exposed to boric acid. Fundam Appl Toxicol, **32**: 194–204.

Barlow S, Renwick AG, Kleiner J, Bridges JW, Busk L, Dybing E, Edler L, Eisenbrand G, Fink-Gremmels J, Knaap A, Kroes R, Liem D, Müller DJG, Page S, Schlatter J, Tritscher A, Rolland V, Tueting W, & Würtzen, G (2006) Risk assessment of substances that are both genotoxic and carcinogenic. Food Chem Toxicol, **44**(10): 1636–1650.

Barnes DG & Dourson ML (1988) Reference dose (RfD): Description and use in health risk assessments. Regul Toxicol Pharmacol, **8**: 471–486.

Barnes DG, Daston GP, Evans JS, Jarabek AM, Kavlock RJ, Kimmel CA, Park C, & Spitzer HL (1995) Benchmark dose workshop: Criteria for use of a benchmark dose to estimate a reference dose. Regul Toxicol Pharmacol, **21**: 296–306.

Bos PMJ, Baars BJ, & van Raaij MTM (2004) Risk assessment of peak exposure to genotoxic carcinogens: A pragmatic approach. Toxicol Lett, **151**: 43–50.

Bozdogan H (1987) Model-selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, **52**: 345–370.

Burnham KP & Anderson DR (2002) Model selection and multimodel inference, 2nd ed. New York, Springer.

Carrington CD & Bolger PM (2000) A pooled analysis of the Iraqi and Seychelles methylmercury studies. Hum Ecol Risk Assess, **6**: 323–340.

Carrington CD, Bolger PM, & Scheuplein RJ (1996) Risk analysis of dietary lead exposure. Food Addit Contam, **13**: 61–76.

Cogliano VJ, Baan RA, Straif K, Grosse Y, Secretan MB, El Ghissassi F, & Kleihues P (2004) The science and practice of carcinogen identification and evaluation. Environ Health Perspect, **112**(13): 1269–1274 (http://ehp.niehs.nih.gov/docs/2004/6950/abstract.html).

Committee on Carcinogenicity (2003) Minutes of United Kingdom Committee on Carcinogenicity of Chemicals in Food, Consumer Products and the Environment meeting, 26 June 2003 (http://www.advisorybodies.doh.gov.uk/coc/meetings/coc032.htm).

Crump KS (1984) A new method for determining allowable daily intakes. Fundam Appl Toxicol, **4**: 854–871.

Crump KS (1995) Calculation of benchmark doses from continuous data. Risk Anal, **15**(1): 79–89.

Crump K (2002) Critical issues in benchmark calculations from continuous data. Crit Rev Toxicol, **32**(3): 133–153.

Davidian M & Giltinan DM (1995) Nonlinear models for repeated measurement data. London, Chapman and Hall.

Dourson ML & Stara JF (1983) Regulatory history and experimental support of uncertainty (safety) factors. Regul Toxicol Pharmacol, **2**: 224–238.

Dourson ML, Hertzberg RC, Hartung R, & Blackburn K (1985) Novel approaches for the estimation of acceptable daily intake. Toxicol Ind Health, **1**: 23–41.

Dybing E, Sanner T, Roelfzema H, Kroese D, & Tennant RW (1997) A simplified carcinogenic potency index. Description of the system and study of correlations between carcinogenic potency and species/site specificity and mutagenicity. Pharmacol Toxicol, **80**: 272–279.

Edler LK, Dourson M, Kleiner J, Mileson B, Nordmann H, Renwick A, Slob W, Walton K, & Würtzen G (2002) Mathematical modelling and quantitative methods. Food Chem Toxicol, **40**: 283–326.

Efron B (1987) Better bootstrap confidence intervals. J Am Stat Soc, **82**: 171–185.

Efron B & Tibshirani R (1993) An introduction to the bootstrap. London, Chapman and Hall.

European Commission (1999) European Commission guidelines for setting specific concentration limits for carcinogens in Annex 1 of Directive 67/548/EEC. Inclusion of potency considerations. Brussels, European Commission, Working Group on the Classification and Labelling of Dangerous Substances.

European Commission (2000) White paper on food safety. Commission of the European Communities, 12 January 2000, Comm (1999) 719 final (http://ec.europa.eu/dgs/health_consumer/library/pub/pub06_en.pdf).

Evans JS, Graham JD, Gray GM, & Sielken RL (1994) A distributional approach to characterizing low-dose cancer risk. Risk Anal, **14**: 25–34.

Evans M, Hastings N, & Peacock B (1993) Statistical distributions, 2nd ed. New York, John Wiley & Sons.

Fitzgerald DJ, Robinson NI, & Pester BA (2004) Application of benzo(*a*)pyrene and coal tar tumor dose–response data to a modified benchmark dose method of guideline development. Environ Health Perspect, **112**: 1341–1346.

Fowles JR, Alexeeff GV, & Dodge D (1999) The use of benchmark dose methodology with acute inhalation lethality data. Regul Toxicol Pharmacol, **29**: 262–278.

Garvin T (2001) Analytical paradigms: The epistemological distances between scientists, policy makers, and the public. Risk Anal, **21**(3): 443–455.

Gaylor DW (1989) Quantitative risk analysis for quantal reproductive and developmental effects. Environ Health Perspect, **79**: 243–246.

Gaylor D & Aylward L (2004) An evaluation of benchmark dose methodology for non-cancer continuous-data health effects in animals due to exposures to dioxin (TCDD). Regul Toxicol Pharmacol, **40**: 9–17.

Gaylor D & Slikker W (1994) Modelling for risk assessment of neurotoxic effects. Risk Anal, **14**(3): 333–338.

Gaylor D & Slikker W (2004) Role of the standard deviation in the estimation of benchmark doses with continuous data. Risk Anal, **6**: 1683–1687.

Gaylor DW, Axelrad JA, Brown RP, Cavagnaro JA, Cyr WH, Hulebak KI, Lorentzen RJ, Miller MA, Mulligan LT, & Schwetz BA (1997) Health risk assessment practices in the U.S. Food and Drug Administration. Regul Toxicol Pharmacol, **26**: 307–321.

Gaylor DW, Kodell RL, Chen JJ, & Krewski D (1999) A unified approach to risk assessment for cancer and noncancer endpoints based on benchmark doses and uncertainty/safety factors. Regul Toxicol Pharmacol, **29**: 151–157.

Gelman A, Carlin JB, Stern HS, & Rubin DB (2004) Bayesian data analysis, 2nd ed. Boca Raton, Florida, Chapman and Hall/CRC.

Ghani AC, Ferguson NM, Donnelly CA, & Anderson RM (2000) Predicted vCJD mortality in Great Britain. Nature, **406**: 583–584.

Gold LS, Gaylor DW, & Slone TH (2003) Comparison of cancer risk estimates based on a variety of risk assessment methodologies. Regul Toxicol Pharmacol, **237**: 45–53.

Hacking I (1976) The great decision. In: The emergence of probability. A philosophical study of early ideas about probability, induction and statistical inference. Cambridge, Cambridge University Press, pp 63–72.

Hasselblad V & Jarabek AM (1995) Dose–response analysis of toxic chemicals. In: Berry DA & Stangl DK ed. Bayesian biostatistics. New York, Marcel Dekker.

Health Canada (1994) Human health risk assessment for priority substances. Ottawa, Ontario, Minister of Supply and Services Canada (http://www.hc-sc.gc.ca/ewh-semt/pubs/contaminants/existsub/approach/index_e.html).

Hertzberg RC (1991) Quantitative extrapolation of toxicological finds. In: Krewski D & Franklin C ed. Statistics in toxicology. New York, Gordon & Breach Science.

Hertzberg RC & Miller M (1985) A statistical model for species extrapolation using categorical response data. Toxicol Ind Health, **1**(4): 43–57.

Hertzberg RC & Wymer L (1991) Modelling the severity of toxic effects. In: Proceedings of the 84th annual meeting and exhibition of the Air and Waste Management Association, 16–21 June 1991, Vancouver, British Columbia. Pittsburgh, Pennsylvania, Air and Waste Management Association.

Humphreys SH, Carrington CD, & Bolger PM (2001) A quantitative risk assessment for fumonisins B_1 and B_2 in corn. Food Addit Contam, **18**: 211–220.

IPCS (1987) Principles for the safety assessment of food additives and contaminants in food. Geneva, World Health Organization, International Programme on Chemical Safety (Environmental Health Criteria 70; http://www.inchem.org/documents/ehc/ehc/ehc70.htm).

IPCS (1994) Assessing human health risks of chemicals: Derivation of guidance values for health-based exposure limits. Geneva, World Health Organization, International Programme on Chemical Safety (Environmental Health Criteria 170; http://www.inchem.org/documents/ehc/ehc/ehc170.htm).

IPCS (1999) Principles for the assessment of risks to human health from exposure to chemicals. Geneva, World Health Organization, International Programme on Chemical Safety (Environmental Health Criteria 210; http://www.inchem.org/documents/ehc/ehc/ehc210.htm).

IPCS (2000) Human exposure assessment. Geneva, World Health Organization, International Programme on Chemical Safety (Environmental Health Criteria 214; http://www.inchem.org/documents/ehc/ehc/ehc214.htm).

IPCS (2004) IPCS risk assessment terminology. Geneva, World Health Organization, International Programme on Chemical Safety (Harmonization Project Document No. 1; http://www.who.int/ipcs/methods/harmonization/areas/ipcsterminologyparts1and2.pdf).

IPCS (2005) Chemical-specific adjustment factors for interspecies differences and human variability: Guidance document for use of data in dose/concentration–response assessment. Geneva, World Health Organization, International Programme on Chemical Safety (Harmonization Project Document No. 2; http://whqlibdoc.who.int/publications/2005/9241546786_eng.pdf).

Kim AH, Kohn MC, Portier CJ, & Walker NJ (2002) Impact of physiologically based pharmacokinetic modeling on benchmark dose calculations for TCDD-induced biochemical responses. Regul Toxicol Pharmacol, **36**: 287–296.

Kimmel CA & Gaylor DW (1988) Issues in qualitative and quantitative risk analysis for development toxicology. Risk Anal, **8**: 15–20.

Kodell RL, Chen JJ, & Gaylor DW (1995) Neurotoxicity modeling for risk assessment. Regul Toxicol Pharmacol, **22**(1): 24–29.

Kroes R, Müller D, Lambe J, Löwik MRH, van Klavern J, Kleiner J, Massey R, Mayer S, Urieta I, Verger P, & Visconti A (2002) Assessment of intake from the diet. Food Chern Toxicol, **40**: 327–385.

Lehman AJ & Fitzhugh OG (1954) Ten-fold safety factor studies. Assoc Food Drug Off US Q Bull, **XVIII**(1): 33–35.

Moerbeek M, Piersma AH, & Slob W (2004) A comparison of three methods for calculating confidence intervals for the benchmark dose. Risk Anal, **24**(1): 31–40.

Murrell JA, Portier CJ, & Morris RW (1998) Characterizing dose–response. I. Critical assessment of the benchmark dose concept. Risk Anal, **18**: 13–26.

National Health and Medical Research Council (1999) Toxicity assessment for carcinogenic soil contaminants. Canberra, Commonwealth of Australia.

NRC (1983) Risk assessment in the federal government: Managing the process. Washington, DC, National Research Council, National Academy of Science Press.

NRC (1994) Science and judgement in risk assessment. Washington, DC, National Research Council, National Academy of Science Press.

NRC (1996) Understanding risk: Informing decisions in a democratic society. Washington, DC, National Research Council, National Academy of Science Press.

NTP (2002) Report on carcinogens, 10th ed. Research Triangle Park, North Carolina, United States Department of Health and Human Services, National Institutes of Health, National Toxicology Program (http://ntp-server.niehs.nih.gov/index.cfm?objectid=72016262-BDB7-CEBA-FA60E922B18C2540).

O'Brien J, Renwick AG, Constable A, Dybing E, Müller DJG, Schlatter J, Slob W, Tueting W, van Benthem J, Williams GM, & Wolfreys A (2006) Approaches to the risk assessment of genotoxic carcinogens in food: A critical appraisal. Food Chem Toxicol, **44**(10): 1613–1635.

OSHA (1983) Occupational exposure to inorganic arsenic: Supplemental statement of reasons for the final rule, 48 F.R. 1,864. Washington, DC, United States Department of Labor, Occupational Safety and Health Administration.

OSHA (1984) Occupational exposure to ethylene oxide: Final standard, 49 F.R. 46,936. Washington, DC, United States Department of Labor, Occupational Safety and Health Administration.

Piersma AH, Verhoef A, te Biesebeek JD, Pieters MN, & Slob W (2000) Developmental toxicity of butyl benzyl phthalate in the rat using a multiple dose study design. Reprod Toxicol, **14**: 417–425.

Pieters MN, Bakker M, & Slob W (2004) Reduced intake of deoxynivalenol in the Netherlands: A risk assessment update. Toxicol Lett, **153**: 145–153.

Portier CJ (1994) Biostatistical issues in the design and analysis of animal carcinogenicity experiments. Environ Health Perspect, **102**(Suppl 1): 5–8.

Portier C & Kohn M (1996) A biologically-based model for the carcinogenic effects of 2378-TCDD in female Sprague-Dawley rats. Organohalogen Compds, **29**: 222–227.

Presidential Commission (1997) Risk assessment and risk management in regulatory decision-making, Vol 2. Washington, DC, The Presidential/Congressional Commission on Risk Assessment and Risk Management.

Raftery A, Madigan D, & Hoeting J (1997) Bayesian model averaging for linear regression models. J Am Stat Assoc, **92**: 179–191.

Renwick AG & Lazarus NR (1998) Human variability and noncancer risk assessment—an analysis of the default uncertainty factor. Regul Toxicol Pharmacol, **27**: 3–20.

Renwick AG, Barlow SM, Hertz-Picciotto I, Boobis AR, Dybing E, Edler L, Eisenbrand G, Grieg JB, Kleiner J, Lambe J, Müller DJG, Smith MR, Tritscher A, Tuijtelaars S, van den Brandt PA, Walker R, & Kroes R (2003) Risk characterization of chemicals in food and diet. Food Chem Toxicol, **41**: 1211–1271.

Rescher N (1969) Many-valued logic. New York, McGraw-Hill.

Sand S, Falk Filipsson A, & Victorin K (2002) Evaluation of the benchmark dose method for dichotomous data: Model dependence and model selection. Regul Toxicol Pharmacol, **36**: 184–197.

Sand S, von Rosen D, & Falk Filipsson A (2003) Benchmark calculations in risk assessment using continuous dose response information: The influence of variance and the determination of a cut-off value. Risk Anal, **23**: 1059–1068.

Sanner T, Dybing E, Willems MI, & Kroese ED (2001) A simple method for quantitative risk assessment of non-threshold carcinogens based on the dose descriptor T25. Pharmacol Toxicol, **88**: 331–341.

SCCNFP (2003) The SCCNFP's notes of guidance for the testing of cosmetic ingredients and their safety evaluation, 5th rev. Adopted by the SCCNFP during the 25th plenary meeting of 20 October 2003. Brussels, The Scientific Committee on Cosmetic Products and Non-Food Products Intended for Consumers.

Sherman CD & Portier CJ (1998) Multistage carcinogenesis models. In: Encyclopedia of biostatistics. Sussex, Wiley & Sons, pp 2808–2814.

Slob W & Pieters MN (1998) A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: General framework. Risk Anal, **18**: 787–798.

Slob W, Moerbeck M, Rauniomaa E, & Piersma AH (2005) A statistical evaluation of toxicity designs for the estimation of the benchmark dose in continuous endpoints. Toxicol Sci, **84**: 167–185.

Thompson KM (2002) Variability and uncertainty meet risk management and risk communication. Risk Anal, **22**(3): 647–654.

Timchalk C, Nolan RJ, Mendrala AL, Dittenber DA, Brzak KA, & Mattsson JL (2002) A physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) model for the organophosphate pesticide chlorpyrifos in rats and humans. Toxicol Sci, **66**: 34–53.

USEPA (1986) Guidelines for carcinogen risk assessment. United States Environmental Protection Agency. Fed Regist, **51**: 33992–34003.

USEPA (1996) Proposed guidelines for carcinogen risk assessment. United States Environmental Protection Agency. Fed Regist, **61**: 17960–180011.

USEPA (2005) Guidelines for carcinogen risk assessment 2005. Washington, DC, United States Environmental Protection Agency, Risk Assessment Forum (EPA/630/P-03/001F; http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=116283).

WHO (1997) Food consumption and exposure assessment of chemicals. Report of a FAO/WHO consultation, 10–14 February 1997 (WHO/FSF/97.5; available from Department of Food Safety, World Health Organization, 20 Avenue Appia, CH-1211 Geneva 27, Switzerland).

WHO (1999) Evaluation of certain food additives and contaminants. Forty-ninth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization.

WHO (2000) Air quality guidelines for Europe, 2nd ed. Copenhagen, World Health Organization Regional Office for Europe.

WHO (2002) Evaluation of certain food additives and contaminants. Fifty-seventh report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (Technical Report Series 909).

WHO (2003) Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide. Report of a Working Group. Copenhagen, World Health Organization Regional Office for Europe (http://www.euro.who.int/document/e79097.pdf).

WHO (2004a) Risk assessment of *Listeria monocytogenes* in ready-to-eat foods. Geneva, World Health Organization (Microbiological Risk Assessment Series, No. 5, Technical Report; http://www.who.int/foodsafety/publications/micro/mra_listeria/en/index.html).

WHO (2004b) Guidelines for drinking-water quality, 3rd ed. Vol. 1. Recommendations. Geneva, World Health Organization (http://www.who.int/water_sanitation_health/dwq/gdwq3rev/en/).

WHO (2006) Evaluation of certain food contaminants. Sixty-fourth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series No. 930; http://whqlibdoc.who.int/trs/WHO_TRS_930_eng.pdf).

WHO (2007a) Evaluation of certain food additives and contaminants. Sixty-seventh report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series No. 940; http://whqlibdoc.who.int/trs/WHO_TRS_940_eng.pdf).

WHO (2007b) Evaluation of certain food additives and contaminants. Sixty-eighth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series No. 947; http://whqlibdoc.who.int/trs/WHO_TRS_947_eng.pdf).

Williams GM & Iatropoulos MJ (2002) Alteration of liver cell function and proliferation: differentiation between adaptation and toxicity. Toxicol Pathol, **30**(1): 41–53.

Woutersen RA, Jonker D, Stevenson H, te Biesebeek JD, & Slob W (2001) The benchmark 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.

ANNEX 1: TERMINOLOGY

Acceptable daily intake (ADI)/tolerable daily intake (TDI)/reference dose (RfD): Estimated maximum amount of an agent, expressed on a body mass basis, to which an individual in a (sub)population may be exposed daily over the individual's lifetime without appreciable health risk.

Acceptable risk: A risk management term. The acceptability of risk depends on scientific data, on social, economic, and political factors, and on the perceived benefits arising from exposure to an agent.

Additional risk (extra risk): The additional proportion of total animals that respond in the presence of the dose, or the probability of response at dose d, P(d), minus the probability of response in the absence of exposure, P(0).

Adverse effect: Change in the morphology, physiology, growth, development, reproduction, or lifespan 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.

Akaike information criterion: A statistical procedure that provides a measure of the goodness of fit of a dose–response model to a set of data.

Assessment factor: Numerical adjustment to extrapolate from experimentally determined (dose–response) relationships to estimate the exposure to an agent below which an adverse effect is not likely to occur (see *Safety factor* and *Uncertainty factor*).

Benchmark concentration (BMC): The concentration of a substance that is associated with a specified low incidence of risk of a health effect, or the concentration associated with a specified measure or change of a biological effect.

Benchmark dose (BMD): A dose of a substance associated with a specified low incidence of risk, generally in the range of 1–10%, of a health effect; or the dose associated with a specified measure or change of a biological effect.

Benchmark dose lower confidence limit (BMDL): A lower one-sided confidence limit on the BMD.

Benchmark response (BMR): The response, generally expressed as in excess of background, at which a benchmark dose or concentration is desired.

Bernoulli distribution: A theoretical distribution of the number of successes in a finite set of independent trials with a constant probability of success. It is a discrete distribution having two possible outcomes labelled by $n = 0$ and $n = 1$, in which $n = 1$ ("success") occurs with probability p and $n = 0$ ("failure") occurs with probability $q \equiv 1 - p$, where $0 < p < 1$.

Binomial distribution: The statistical distribution of the probabilities of observing 0, 1, 2, … , n events in a sample of n independent trials each with the same individual probability that the event occurs.

Bootstrap: A statistical technique based on multiple resampling with replacement of the sample values or resampling of estimated distributions of the sample values that is used to calculate confidence limits or perform statistical tests for complex situations or where the distribution of an estimate or test statistic cannot be assumed.

Cancer potency (cancer slope factor): A number that estimates the cancer risk (incidence) for a lifetime exposure to a substance per unit of dose, which is generally expressed as mg/kg body weight per day.

Categorical data: Results obtained where observations or measurements on individuals or samples are stratified according to degree or severity of an effect (e.g. none, mild, moderate, or severe).

Categorical default factor: A factor based on common characteristics of a group of compounds (e.g. physical/chemical properties or pathways of metabolism).

Chemical-specific adjustment factor (CSAF): A factor based on quantitative chemical-specific toxicokinetic or toxicodynamic data, which replaces some or all of the default uncertainty factor.

Chi-square test: A statistical test used to examine the deviation of an observed number of events from an expected number of events.

Clustered data: Measurements collected on some grouping of individuals (e.g. litters in reproductive and developmental studies).

Confidence interval (one-sided): An interval below the estimated upper confidence limit, or an interval above the estimated lower confidence limit, that is expected to include the true value of an estimated parameter with a specified confidence (percentage of the time).

Confidence interval (two-sided): An estimated interval from the lower to upper confidence limit of an estimate of a parameter. This interval is expected to include the true value of the parameter with a specified confidence percentage (e.g. 95% of such intervals are expected to include the true values of the estimated parameters).

Confidence limit: An estimated value below (or above) which the true value of an estimated parameter is expected to lie for a specified percentage of such estimated limits.

Constrained dose–response model: Estimates of one or more parameters of the model restricted to a specified range (e.g. equal to or greater than zero).

Continuous data: Effects measured on a continuum, e.g. organ weight or enzyme concentration, as opposed to quantal or categorical data, where effects are classified by assignment to a class.

Convergence: A parameter approach that estimates a single value with increasing sample size or increasing number of computer iterations.

Covariate: An independent variable other than dose that may influence the outcome of an effect (e.g. age, body weight, or polymorphism).

Critical effect: The adverse effect, or its known precursor, that is relevant to human risk assessment and that occurs in the dose/concentration scale in the most sensitive animal species.

Degrees of freedom: For dose–response model fitting, the number of data points minus the number of model parameters estimated from the data.

Default value: Pragmatic, fixed, or standard value used in the absence of relevant data.

Dichotomous data: Quantal data where an effect for an individual may be classified by one of two possibilities (e.g. dead or alive), with or without a specific type of tumour.

Dispersion: Variation (differences) from a central (mean or median) value.

Dose: Total amount of an agent administered to or taken up 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.

Dose–response assessment: Analysis of the relationship between the total amount of an agent administered to, taken up by, or absorbed by an organism, system, or (sub)population and the changes developed in that organism, system, or (sub)population in reaction to that agent, and inferences derived from such an analysis with respect to the entire population.

Dose–response model: A mathematical relationship (function) that relates (predicts) a measure of an effect to a dose.

Dose–response trend: Relationship between incidence or severity of a biological effect and a function of dose. Simply the slope for a linear dose–response.

ED_x: Effective dose associated with a biological effect in x% of the individuals. Dose may be the external exposure often expressed in milligrams of the substance per day per kilogram body weight raised to a power (generally 1, 3/4, or 2/3) or area under the curve (AUC) in blood or target tissue where the substance remains in the body over a period of time.

Estimate: An empirical value derived from data for a parameter.

Exposure: Concentration or amount of a particular agent that reaches a target organism, system, or (sub)population in a specific frequency for a defined duration.

Gamma distribution: A unimodal statistical distribution (relative proportion of responders as a function of some measure) restricted to effects greater than or equal to zero that can describe a wide variety of shapes (e.g. flat, peaked, asymmetrical).

Gaussian (normal) distribution: A unimodal symmetrical (bell-shaped) distribution where the most prevalent value is the mean (average) and the spread is measured by the standard deviation. Mathematically, the distribution varies from minus infinity with zero probability to plus infinity with zero probability.

Goodness of fit: A statistic that measures the dispersion of data about a dose–response curve in order to provide a test for rejection of a model due to lack of an adequate fit (e.g. a p-value < 0.1).

Hazard identification: The identification of the type and nature of adverse effects that an agent has an inherent capacity to cause in an organism, system, or (sub)population.

Hill equation: A dose–response curve, frequently used for enzyme kinetics, that monotonically approaches an asymptote (maximum value) as a function of dose raised to a power.

Hybrid model: For continuous data, establishes abnormal values based on the extremes in controls (unexposed individuals or animals) and estimates the risk of abnormal levels as a function of dose.

Incidence: Proportion or probability of individuals or animals exhibiting an effect that varies from zero to one, sometimes expressed as a percentage from 0% to 100%.

Independence: The result in one animal or individual does not influence the result in another animal or individual.

Intercept term: The estimated value at zero dose or the dose corresponding to a zero effect.

Least squares: A statistical procedure that estimates the values of dose–response parameters such that the sum of squares of deviations of data points from their estimated values is minimized (i.e. minimizes the estimated variance).

Likelihood function: Relative probabilities that various values of population parameters would arise from the sample observations.

Likelihood ratio: Ratio of the probability that the observed data arise from a set of model parameters relative to the maximum probability that arises from the set of maximum likelihood estimates.

Linear dose–response model: The amount of change in a response is proportional to the amount of change in some function of dose.

Linearized multistage model: Dose–response model based on the multistage model of carcinogenesis that is restricted to a form that is approximately linear at low doses.

Local maximum: Mathematical solution that maximizes a function in a region that may not be the overall global maximum.

Logistic model: A sigmoidal (S-shaped) function that relates the proportion of individuals with a specified characteristic to an independent variable.

Lognormal distribution: A mathematical description where the natural logarithm of a random variable has a normal distribution.

Log transformation: Logarithm of raw data.

Lowest-observed-adverse-effect level (LOAEL): The lowest concentration or dose of a substance, found by experiment or observation, that causes an adverse alteration of morphology, functional capacity, growth, development, or lifespan of the target organisms distinguishable from normal (control) organisms of the same species and strain under the same defined conditions of exposure.

Lowest-observed-effect level (LOEL): The lowest concentration or dose of a substance, found by experiment or observation, that causes any alteration of morphology, functional capacity, growth, development, or lifespan of the target organisms distinguishable from normal (control) organisms of the same species and strain under the same defined conditions of exposure.

Margin of exposure (MOE): Ratio of the no-observed-adverse-effect level (NOAEL) or benchmark dose lower confidence limit (BMDL) for the critical effect to the theoretical, predicted, or estimated exposure dose or concentration.

Maximum likelihood estimate: Estimate of a population parameter most likely to have produced the sample observations.

Mechanism of action: A detailed description of the precise chain of events from the molecular level to gross macroscopic or histopathological toxicity.

Michaelis-Menten equation: A dose–response curve, frequently used for enzyme kinetics, with maximum slope at zero dose that approaches a maximum asymptote at increasing dose.

Mode of action: A series of events that may lead to induction of the relevant end-point of toxicity for which the weight of evidence supports plausibility.

Monotonic dose–response: A dose–response that never decreases as dose increases. A monotonic function may be flat (constant) up to a threshold dose or may be flat at high doses if a biological limit (e.g. saturation) is attained.

Multinomial: Animals or individuals may be classified by more than two (binomial) categories (e.g. in a reproductive study, fetuses may be dead, alive normal, or alive abnormal).

Negligible risk: A risk management term. In cases where a quantitative risk estimate has been made, it is any risk less than an upper-bound incremental lifetime risk calculated using conservative risk assessment techniques such as the benchmark dose.

Non-linear dose–response model: Mathematical relationship that cannot be expressed simply as the change in response being proportional to the amount of change of some function of dose.

No-observed-adverse-effect level (NOAEL): The highest concentration or dose of a substance, found by experiment or observation, that causes no detectable adverse alteration of morphology, functional capacity, growth, development, or lifespan of the target organisms under defined conditions of exposure.

No-observed-effect level (NOEL): The highest concentration or dose of a substance, found by experiment or observation, that causes no detectable alteration of morphology, functional capacity, growth, development, or lifespan of the target organisms under defined conditions of exposure.

Normal distribution: A mathematical description where a continuous random variable x with a mean μ and a variance σ^2 has a probability density function:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$$

Objective function: Choice of function that is optimized for maximum likelihood estimation.

Ordinal data: Integers designating the rank, order, or counts.

Parameter: A value used to numerically describe a population of values (e.g. the mean and standard deviation); or a value used to describe a dose–response curve (e.g. the intercept and the slope of a linear dose–response).

Point of departure: The point on a dose–response curve established from experimental data (e.g. the benchmark dose), generally corresponding to an estimated low effect level (e.g. 1–10% incidence of an effect). Depending on the mode of action and available data, some form of extrapolation below the point of departure may be employed for low-dose risk assessment, or the point of departure may be divided by a series of uncertainty factors to arrive at a reference dose. Points of departure include the BMD, BMDL, LOAEL, and carcinogenic potency estimates, such as the T_{25}.

Polynomial: A mathematical function of the sum of a constant, linear term, quadratic term, cubic term, etc.

Probability: The proportion (on a scale of 0 to 1) of cases for which a particular event occurs. Zero indicates the event never occurs, and one indicates the event always occurs.

Probability distribution: A mathematical description of the relative probabilities of all possible outcomes of a measurement.

Probit function: Assumes that the relative probabilities of effects as a function of dose are described by a normal distribution. The cumulative probability as a function of dose has a sigmoidal shape.

Profile likelihood: A plot of the likelihood function versus the estimated value of a parameter.

P-value: In testing a hypothesis, the probability of a type I error (false positive). The probability that the sample (experimental) results are compatible with a specific hypothesis.

Quadratic term: A quantity in a mathematical formula that is raised to the second power (squared).

Quantal data: Dichotomous (binomial) classification where an individual or animal is placed in one of two categories (e.g. dead or alive, with or without a particular type of tumour, normal or abnormal level of a hormone).

Quantile: Percentile (cumulative probability) of a distribution that ranges from zero to the 100th percentile.

Regression analysis: A statistical process that produces a mathematical function (regression equation) that relates a dependent variable (biological effect) to an independent variable (e.g. dose rate, duration of exposure, age).

Repeated measures: A biological end-point is measured for the same individual or animal at different times (ages).

Response: Change developed in the state or dynamics of an organism, system, or (sub)population in reaction to exposure to an agent.

Residual variance: The variance in experimental measurements remaining after accounting for the variance due to the independent variables (e.g. dose rate, duration of exposure, age). Typically referred to as the inherent unaccountable experimental variation.

Risk: The probability of an adverse effect in an organism, system, or (sub)population caused under specified circumstances by exposure to an agent.

Risk assessment: A process intended to calculate or estimate the risk to a given target organism, system, or (sub)population, including the identification of attendant uncertainties, following exposure to a particular agent, taking into account the inherent characteristics of the agent of concern as well as the characteristics of the specific target system.

Risk characterization: The qualitative and, wherever possible, quantitative determination, including attendant uncertainties, of the

probability of occurrence of known and potential adverse effects of an agent in a given organism, system, or (sub)population, under defined exposure conditions.

Safety factor: Composite (reductive) factor by which an observed or estimated no-observed-adverse-effect level (NOAEL) is divided to arrive at a criterion or standard that is considered safe or without appreciable risk (see *Assessment factor* and *Uncertainty factor*).

Severity: The degree to which an effect changes and impairs the functional capacity of an organ system.

Shape parameter: The exponent on dose in a dose–response function that dictates the curvature of the function.

Threshold: Dose or exposure concentration of an agent below which a stated effect is not observed or expected to occur.

Threshold of toxicological concern: An exposure threshold value below which there is a very low probability of an appreciable risk to human health.

Toxicodynamics: The process of interaction of chemical substances with target sites and the subsequent reactions leading to adverse effects.

Toxicokinetics: The process of the uptake of potentially toxic substances by the body, the biotransformation they undergo, the distribution of the substances and their metabolites in the tissues, and the elimination of the substances and their metabolites from the body. Both the amounts and the concentrations of the substances and their metabolites are studied. The term has essentially the same meaning as *pharmacokinetics*, but the latter term should be restricted to the study of pharmaceutical substances.

Uncertainty: Imperfect knowledge concerning the present or future state of an organism, system, or (sub)population under consideration.

Uncertainty factor: Reductive factor by which an observed or estimated no-observed-adverse-effect level (NOAEL) is divided to

arrive at a criterion or standard that is considered safe or without appreciable risk (see *Assessment factor* and *Safety factor*).

Unconstrained dose–response model: No restrictions imposed on the estimates of parameters.

Upper-tail probability: Probability that a variable exceeds a specified value.

Validation: Process by which the reliability and relevance of a particular approach, method, process, or assessment is established for a defined purpose.

Variability: Observable diversity in biological sensitivity or response and in exposure parameters.

Variance: Measure of variability, standard deviation squared.

Weibull: Form of a dose–response curve characterized by a relatively shallow slope at low doses that increases sharply as dose increases before levelling off at high doses.

Weighted least squares estimate: Parameter estimate obtained by minimizing the sum of squares of observed and estimated values weighted by a function, frequently the reciprocal of the variance of an observation.

RESUME, CONCLUSIONS ET RECOMMANDATIONS

1. Résumé

La modélisation de la relation dose-réponse (DRM), destinée à évaluer quantitativement les risques et enfin à étayer les décisions de santé publique concernant les expositions à des produits chimiques, peut être décrite comme un processus en six étapes. Les quatre premières étapes – sélection des données, choix du modèle, mise en relation statistique et estimation des paramètres – constituent une analyse de la relation dose-réponse. Ces étapes composent le processus permettant d'obtenir une description mathématique des données, en vue de prédire les réponses à des doses connues ou d'établir des estimations de doses à partir d'une réponse donnée. La cinquième étape réalise une synthèse des résultats de l'analyse dose-réponse et des estimations de l'exposition afin de guider les décisions de santé publique. L'étape finale, qui peut éventuellement intervenir plus tôt dans la DRM, évalue la qualité de la relation dose-réponse et la sensibilité des prédictions aux hypothèses ayant servi à l'analyse.

La caractérisation de la relation dose-réponse dans les études chez l'homme et chez l'animal constitue une composante majeure de la caractérisation des dangers et sert à extrapoler les incidences des effets nocifs sur la gamme de niveaux d'exposition humaine. Au cours des années, diverses méthodes ont été mises au point pour traiter de telles relations, améliorer l'extrapolation pour les faibles doses et dériver des valeurs guides reposant sur des considérations sanitaires telles que les doses journalières admissibles (DJA), les doses journalières tolérables (DJT) et les doses de référence (D_{ref}). La DRM peut s'avérer utile dans les évaluations des risques en permettant un meilleur usage des données disponibles et en fournissant des outils pour évaluer la qualité des données et les incertitudes résultantes sur les estimations de la relation dose-réponse.

D'une manière générale, les estimations obtenues par DRM sont établies à partir de données provenant de l'ensemble de la courbe dose-réponse pour l'effet critique. L'approche standard reposant sur la dose sans effet nocif observé (DSENO) peut être considérée comme un cas spécial et simplifié d'analyse de la

relation dose-réponse, dans la mesure où elle identifie une dose unique supposée ne pas avoir d'effet nocif appréciable. La modèle DRM reflète les caractéristiques de la courbe dose-réponse, en permettant notamment d'estimer sa pente. Dans le cas d'un cadre de régression, il indique l'écart-type et l'intervalle de confiance pour les paramètres de modélisation. L'un des inconvénients de l'approche DSENO réside dans l'impossibilité de quantifier les degrés de variabilité et d'incertitude, alors que d'autres modèles de la relation dose-réponse peuvent faciliter l'analyse de ces grandeurs. L'utilisation d'un modèle dose-réponse peut permettre d'optimiser la conception de l'étude et de préciser les besoins en matière d'études supplémentaires. L'approche DSENO intègre des informations biologiques à travers l'application d'un jugement d'expert, néanmoins subjectif. Une DRM complète est en mesure de fournir une analyse plus « riche en éléments scientifiques » grâce à l'inclusion quantitative plus formelle dans les modèles de facteurs et de covariables, par exemple. Les estimations obtenues à partir de cette DRM facilitent la comparaison dans un cadre commun entre des expériences, des effets et des composés qui diffèrent sur le plan quantitatif. La modélisation DRM peut aussi conduire à de meilleures évaluations des risques et de l'innocuité et offre la possibilité d'étudier les probabilités d'effets se manifestant en dehors de la plage observable.

Le choix des modèles à utiliser dépend du type de donnée. Il faut sélectionner un modèle dose-réponse et un modèle décrivant la variabilité des données. Une fois qu'on dispose de modèles adaptés à un jeu de données, on peut évaluer leur degré de représentativité pour ces données par des mesures de la qualité de l'ajustement. On peut en outre comparer leur capacité à décrire les données. Les incertitudes portant sur les inférences tirées de ces modèles se répartissent dans quatre catégories : incertitudes statistiques sur les inférences dues à la variabilité des réponses entre les sujets des expériences, erreurs expérimentales (randomisation imparfaite, erreurs de dosage, localisation défavorable de la dose, par exemple), variabilité d'une expérience à l'autre due aux différences inévitables dans l'exécution de l'expérience et incertitude due au fait que l'on ne connaît pas le « vrai modèle » décrivant les données. L'analyse de la relation dose-réponse doit, dans la mesure du possible, prendre en compte l'ensemble de ces quatre sources de variabilité et d'incertitude.

Le calcul des doses de référence (BMD) est une application particulièrement importante de la DRM. Les BMD sont les doses pour lesquelles on détermine par déduction qu'il se produira un niveau donnée de réponse. Lorsqu'on dispose de données appropriées, les BMD offrent une alternative à l'approche DSENO pour le calcul de valeurs guides reposant sur des considérations sanitaires. Lorsqu'une extrapolation s'avère nécessaire, il convient de représenter l'incertitude associée à la prédiction. Il est dans ce cas particulièrement important d'indiquer l'incertitude liée au modèle.

Une DRM complète peut apporter des informations supplémentaires au gestionnaire de risques. La sortie du modèle doit être conçue pour répondre à certaines questions concernant la probabilité d'effets sanitaires nocifs. Elle peut être présentée essentiellement de trois façons. Premièrement, elle peut servir à établir des valeurs guides reposant sur des considérations sanitaires telles que les DJA, les DJT ou les D_{ref}, d'une manière analogue aux procédures actuellement appliquées à partir de la DSENO ou de la dose minimale avec effets nocifs observés (DMENO). La DRM peut être une méthode plus sûre sur le plan scientifique pour déterminer ces valeurs guides. Deuxièmement, la sortie de la DRM peut être utilisée en gestion des risques pour estimer une marge d'exposition (ME), par détermination du rapport de la dose correspondant à une limite donnée de la réponse à un niveau d'exposition humaine. Troisièmement, sur la base de la relation dose-réponse modélisée, cette sortie peut être une estimation quantitative de l'ampleur du risque ou de l'effet sanitaire pour un niveau d'exposition humaine, moyennant l'hypothèse généralement acceptée que les facteurs d'incertitude utilisés couvrent les incertitudes associées aux différences de sensibilité entre individus et espèces. La DRM peut fournir de meilleures informations sur la probabilité des effets pour les doses faibles et inférieures aux niveaux observés dans les systèmes biologiques, ainsi que de meilleures estimations des incertitudes statistiques entachant les estimations des effets probables.

La multiplicité des jeux de données et des incertitudes peut influer sur le type de sortie des exercices de DRM et avoir de l'importance pour les gestionnaires de risques. On peut utiliser la DRM avec des données d'exposition pour identifier les sous-

populations à risque. Elle peut aussi aider les gestionnaires dans la détermination des priorités et dans l'évaluation des conséquences d'interventions proposées pour réduire les risques. Pour la communication à propos des risques, l'application des techniques de DRM offre des opportunités, mais comporte aussi des difficultés. Les évaluations par DRM peuvent produire des informations sous plusieurs formats, et notamment sous forme de fonctions dose-réponse permettant, avec les estimations de l'exposition, de prédire les risques pour des niveaux d'exposition donnés, ainsi que de fonctions permettant inversement d'estimer les niveaux d'exposition à l'origine de risques donnés. On obtient ainsi notamment des estimations du risque potentiel d'absorption plus importante qu'une valeur guide reposant sur des considérations sanitaires, DJA par exemple. Ces évaluations offrent aussi des approches pour comparer les risques ou les bénéfices concurrents et s'intéresser aux incertitudes susceptibles d'influer sur les risques prédits. Toutefois, à moins que la situation en termes de risque ne soit envisagée à l'échelle de la population, il existe un problème de communication à propos du risque car lorsqu'on présente le niveau de risque dans des situations où aucun niveau d'exposition n'est dépourvu de risque, la modélisation prévoit qu'un certain pourcentage de la population subira des effets jugés nocifs. Il faut reconnaître que l'utilisation de la DRM impose aux données des exigences en termes de qualité et de quantité et nécessite des compétences spécifiques.

L'utilisation courante des estimations tirées der la DRM pourrait, du point de vue de la gestion des risques, permettre une meilleure caractérisation avant la prise de décisions en :

- apportant des informations sur les valeurs guides (ampleur et types des impacts sanitaires) ;
- montrant les bénéfices de différentes actions réglementaires ;
- fournissant au décideur une appréciation des données « plus que ponctuelle » ;
- favorisant la cohérence dans les décisions, moyennant des ajustements appropriés pour tenir compte des différences entre les effets, les niveaux d'effet, les espèces et les types d'étude ; et en
- permettant en continu et en permanence des interactions itératives entre l'évaluateur et le gestionnaire de risques.

L'utilisation de la DRM et des techniques d'évaluation probabiliste pour décrire quantitativement la variabilité et l'incertitude génère de nouvelles difficultés dans la communication à propos des risques. Ces difficultés résident notamment dans :

- l'explication de la prévision, pour un certain pourcentage de la population, d'un dépassement du niveau de sécurité et/ou de l'apparition d'effets nocifs ;
- l'explication du niveau de risque dans les cas où on suppose qu'aucun niveau d'exposition n'est dépourvu de risque ;
- la comparaison entre risques ou bénéfices concurrents ;
- la mise en lumière d'incertitudes influant sur le risque prédit ; et
- l'explication du fait qu'en matière de risque, une estimation indique ce qui peut se passer au niveau d'une population, plutôt qu'à celui d'un individu, ce qui, notons le, vaut aussi pour l'approche DJA/DJT.

2. Conclusions

- La DRM complète peut être considérée comme une alternative plus élaborée et plus robuste à l'approche DSENO dans tous les cas où l'on dispose de données appropriées sur la relation dose-réponse (pour plusieurs groupes de dose et différents niveaux d'exposition, par exemple).

- Pour les données dose-réponse ponctuelles, on s'intéresse souvent aux faibles niveaux de réponse (d'incidence). Il est parfois nécessaire, dans cette perspective, d'extrapoler sur plusieurs ordres de grandeur (pour l'incidence des tumeurs, par exemple). Cependant, des modèles également plausibles de la relation dose-réponse peuvent fournir des estimations fortement divergentes pour les faibles valeurs. Une approche actuellement appliquée en tant que méthode prudente consiste à estimer la BMD_{10} (dose pour un risque de 10 %) et à extrapoler linéairement à partir de ce point vers les valeurs descendantes. Une autre solution, actuellement en cours de développement, applique une approche bayésienne, considérant globalement les divers modèles.

- Pour les données dose-réponse continues, il existe deux approches de type DRM. L'une comprend la transformation des

données continues en données ponctuelles. L'autre considère les données dose-réponse continues comme des informations sur la gravité de l'effet et donc comme une fonction de la dose. Dans cette dernière approche, des variations mesurables de l'effet sont souvent proches des niveaux de réponses considérés comme nocifs (par exemple, inhibition de 10 % de la cholinestérase) et le problème de l'extrapolation à faible dose est mineur ou ne se pose pas.

- Pour dériver une DJA, une DJT ou une Dref, on peut faire appel à la DRM pour déterminer une BMD, qui sera utilisée comme point de départ de la même façon qu'une DSENO (c'est-à-dire qu'on appliquera les mêmes facteurs d'incertitude à la BMD qu'à la DSENO).

- La DRM peut aussi être employée pour estimer les risques correspondant à un niveau d'exposition (humaine) donné. Pour évaluer les risques en termes d'incidence (données ponctuelles), cette opération peut devoir inclure une extrapolation aux faibles doses.

- Les exercices de DRM peuvent apporter des informations sur les incertitudes associées aux données et identifier des facteurs contribuant aux incertitudes sur les estimations des risques.

- L'application de la DRM à tous les points finaux peut être extrêmement onéreuse, il est donc plus efficace de présélectionner les points finaux apparemment les plus sensibles. Dans certains cas cependant, il n'est pas facile d'identifier visuellement ces points de sorte qu'il peut être nécessaire de modéliser tous les points finaux.

- La BMD et la borne inférieure de l'intervalle de confiance de la BMD (BMDL) doivent toujours être indiquées de manière à ce que la qualité des données et de l'ajustement du modèle apparaisse clairement et que l'on puisse procéder à une comparaison de puissances à partir de la BMD.

- Il convient de présenter la sortie des différents modèles de DRM.

3. Recommandations

- Les protocoles d'évaluation de la toxicité (par exemple les Lignes directrices de l'Organisation de coopération et de développement économiques) doivent être examinés pour optimiser l'approche utilisant la BMD et d'autres approches de type DRM, notamment pour choisir au mieux les nombres d'animaux et de doses pour les différentes courbes dose-réponse. Des recherches supplémentaires sont nécessaires pour développer des types d'étude optimaux. Il faut également élaborer des conseils pour combiner les études existantes en vue d'une DRM.

- Il faut mettre au point des recommandations pour l'analyse combinée de différents jeux de données en vue d'une estimation plus précise des BMD.

- Il faut aussi parvenir à mieux comprendre quand et comment utiliser la réponse de référence (BMR).

- La forme de la courbe dose-réponse aux faibles doses doit être mieux interprétée. Des recherches supplémentaires sont nécessaires pour déterminer la base biologique de l'extrapolation (en faisant appel, par exemple, à des marqueurs biologiques, à des précurseurs de tumeur, à des animaux génétiquement modifiés ou à la toxico-cinétique, pour estimer la dose cible).

- Il faut élaborer de meilleures recommandations pour la communication à propos des risques sur la base des résultats de la DRM et des techniques d'évaluation probabiliste. Cette communication devra couvrir les types d'incertitude, leur relation avec la variabilité statistique, l'imprécision et l'utilisation des intervalles de confiance.

- L'utilisation de la DRM doit faire l'objet d'un bilan et des principes généraux supplémentaires devront être développés à mesure que l'on disposera de plus d'expérience.

RESUMEN, CONCLUSIONES Y RECOMENDACIONES

1. Resumen

La creación de modelos de la relación dosis-respuesta, para su utilización en la evaluación cuantitativa del riesgo y en último término para documentar las decisiones en materia de salud pública, se puede describir como un proceso en seis etapas. Las cuatro primeras etapas—selección de datos, selección del modelo, vinculación estadística y estimación de los parámetros—constituyen el análisis de la relación dosis-respuesta. Estas etapas están relacionadas con el proceso mediante el cual se obtiene una descripción matemática de los datos, a fin de evaluar respuestas previstas para dosis conocidas u obtener estimaciones de la dosis cuando lo que interesa es una respuesta determinada. La quinta etapa consiste en la integración de los resultados del análisis de la relación dosis-respuesta en las estimaciones de la exposición, con el objetivo de orientar las decisiones relativas a la salud pública. La última etapa, que se puede elegir aplicar antes, consiste en una evaluación de la calidad del análisis de la relación dosis-respuesta y de la sensibilidad de las predicciones de los modelos con respecto a las hipótesis utilizadas en el análisis.

La caracterización de las relaciones dosis-respuesta en estudios realizados en animales y personas ha sido un componente importante de la caracterización del peligro y se ha utilizado en la extrapolación de incidencias de efectos adversos en la gama de los niveles de exposición humana. Durante años se han elaborado diversos métodos para ajustar dichas relaciones, mejorar la extrapolación a dosis bajas y obtener valores guía basados en la salud, como la ingesta diaria admisible (IDA), la ingesta diaria tolerable (IDT) y las dosis de referencia. La creación de modelos puede ser útil en las evaluaciones del riesgo para utilizar mejor los datos disponibles y para suministrar instrumentos de evaluación de la calidad de los datos y las consiguientes incertidumbres en las estimaciones de la relación dosis-respuesta.

En general, las estimaciones de los modelos de la relación dosis-respuesta se basan en los datos obtenidos de la totalidad de la

curva correspondiente a dicha relación para el efecto crítico. El método normalizado de la concentración sin efectos adversos observados (NOAEL) se puede considerar como un caso especial simplificado de análisis de la relación dosis-respuesta, puesto que identifica una dosis única que se supone que no tiene un efecto adverso apreciable. El modelo de la relación dosis-respuesta refleja las características de la curva de dicha relación, en particular porque proporciona estimaciones de la pendiente. En el caso de un marco de regresión, proporciona el error estándar y los intervalos de confianza para los parámetros del modelo. La utilización del método de la NOAEL tiene el inconveniente de que no es posible cuantificar el grado de variabilidad e incertidumbre que puede haber, mientras que otros modelos de la relación dosis-respuesta pueden facilitar el análisis de la sensibilidad y la incertidumbre. El examen de un modelo de dosis-respuesta puede mejorar al máximo la formulación del estudio y aclarar la necesidad de estudios adicionales. El método de la NOAEL incorpora información biológica mediante la aplicación de un parecer "experto", pero subjetivo. La creación de modelos de la relación dosis-respuesta completos permitiría un análisis más "científico", por ejemplo mediante la inclusión cuantitativa más oficial de factores y covariantes en los modelos. Las estimaciones derivadas de los modelos de dosis-respuesta mejoran la capacidad para comparar experimentos, efectos y compuestos con diferencias cuantitativas en el ámbito de un marco común. Los modelos pueden mejorar las evaluaciones del riesgo y de la inocuidad, ofreciendo al mismo tiempo oportunidades para examinar la probabilidad de los efectos fuera de la gama observable.

La elección de los modelos que se van a utilizar depende del tipo de datos. Dichos modelos deben incluir un patrón para la relación dosis-respuesta y otro para la variabilidad de los datos. Una vez ajustados los modelos a una serie de datos, se puede evaluar el grado en que los describen individualmente utilizando medidas de la precisión del ajuste. Además, se puede comparar entre ellos la capacidad para describir los datos. Las incertidumbres sobre las consecuencias que puedan derivarse de dichos modelos entran en cuatro categorías principales: incertidumbre estadística de las consecuencias debida a la variabilidad entre las respuestas de los sujetos objeto de experimentación, errores experimentales (por ejemplo, distribución al azar imperfecta, errores de dosificación,

localización desfavorable de las dosis), variabilidad entre experimentos debida a diferencias inevitables en su realización e incertidumbre debida al hecho de que no se conoce el "verdadero modelo" para los datos. Siempre que sea posible, en el análisis de la relación dosis-respuesta hay que abordar las cuatro fuentes de variabilidad e incertidumbre.

Una aplicación particularmente importante de los modelos de dosis-respuesta es el cálculo de las dosis de referencia. Son las dosis con las cuales se deduce que se producirá un determinado nivel de respuesta. Cuando se dispone de datos apropiados, las dosis de referencia son una alternativa al método de la NOAEL para calcular los valores guía basados en la salud. Cuando es necesaria una extrapolación, se debe representar la incertidumbre asociada con una predicción. En este caso es particularmente importante incluir la incertidumbre del modelo.

La creación de modelos de la relación dosis-respuesta completos ofrece la posibilidad de proporcionar información adicional a los gestores del riesgo. Los resultados de los modelos se deben orientar hacia el examen de cuestiones específicas relativas a la probabilidad de efectos adversos en la salud. Se pueden presentar de tres maneras principales. En primer lugar, se pueden utilizar para el establecimiento de un valor guía basado en la salud, por ejemplo una IDA, una IDT o unas dosis de referencia, de manera análoga a los procedimientos actuales basados en la NOAEL o la concentración más baja con efectos adversos observados (LOAEL). Los modelos de dosis-respuesta pueden ser un método más sólido desde el punto de vista científico para determinar valores guía basados en la salud. En segundo lugar, los resultados de dichos modelos se pueden utilizar en la gestión del riesgo para estimar un margen de exposición, mediante el cálculo de la relación entre la dosis correspondiente a un límite determinado de respuesta y un nivel de exposición humana. En tercer lugar, sobre la base de la relación dosis-respuesta obtenida mediante el modelo, el resultado puede ser una estimación cuantitativa de la magnitud del riesgo/efecto en la salud para el nivel de exposición humana, con la hipótesis generalmente aceptada de que los factores de incertidumbre utilizados incluyen las incertidumbres relativas a las diferencias de sensibilidad intraespecíficas e interespecíficas. Los modelos de dosis-respuesta pueden proporcionar mejor información sobre la probabilidad de efectos con dosis bajas, inferiores a los

niveles observados en los sistemas biológicos, y pueden proporcionar asimismo mejores estimaciones de las incertidumbres estadísticas de los efectos probables.

Dos factores que pueden influir en el tipo de resultados obtenidos de la aplicación de los modelos de dosis-respuesta y que pueden ser importantes para el gestor del riesgo son las series de datos múltiples y las incertidumbres. Los modelos se pueden utilizar con datos de exposición para identificar las subpoblaciones en situación de riesgo. También se pueden emplear para ayudar a los gestores del riesgo a establecer prioridades y evaluar las consecuencias de las intervenciones propuestas encaminadas a reducir el riesgo. Para la comunicación del riesgo, la utilización de técnicas con modelos de dosis-respuesta ofrece oportunidades y retos. Las evaluaciones con estos modelos pueden generar información de varios tipos, como funciones de la relación dosis-respuesta que permiten, junto con las estimaciones de la exposición, la predicción de los riesgos con niveles específicos de exposición y funciones que permiten la estimación de los niveles de exposición que dan lugar a riesgos determinados. Esto incluye las estimaciones del posible riesgo de ingestas por encima de un valor guía basado en la salud, por ejemplo la IDA. Las evaluaciones con los modelos de dosis-respuesta también ofrecen métodos para comparar riesgos o beneficios competitivos y permiten concentrar la atención en las incertidumbres que pueden influir en el riesgo pronosticado. Sin embargo, salvo que la situación del riesgo se examine en la población, su comunicación presenta el problema de que, aun explicando el nivel de riesgo en esas circunstancias en las que no hay un nivel inocuo de exposición, cabe predecir que cierto porcentaje de la población va a registrar algunos efectos considerados adversos. Hay que reconocer que la utilización de los modelos de la relación dosis-respuesta requiere cierta cantidad y calidad de datos, así como conocimientos técnicos específicos.

El uso potencial "continuo" de las estimaciones derivadas de los modelos de dosis-respuesta puede, desde una perspectiva de gestión del riesgo, mejorar la caracterización para la adopción de decisiones, porque:

- facilita información sobre lo que ocurre por encima del valor guía basado en la salud (magnitud y tipos de efectos en la salud);
- demuestra los beneficios de distintas medidas normativas;
- ofrece a los encargados de la adopción de decisiones una apreciación de los datos desde más de un punto de vista;
- promueve la coherencia en las decisiones, si se hacen ajustes apropiados para las diferencias en los efectos, el nivel de los efectos, las especies y la formulación del estudio; y
- facilita una interacción iterativa entre el asesor del riesgo y el gestor del riesgo de manera continua e ininterrumpida.

La utilización de modelos de la relación dosis respuesta y de técnicas de evaluación probabilística para describir de manera cuantitativa la variabilidad y la incertidumbre incorpora nuevos retos a la comunicación del riesgo. Algunos de ellos son los siguientes:

- explicar que se prevé que un cierto porcentaje de la población superará el nivel de inocuidad y/o sufrirá un efecto adverso;
- explicar el nivel de riesgo en esas circunstancias en las que se supone que no hay un nivel inocuo de exposición;
- comparar los riesgos o los beneficios en pugna;
- prestar una atención especial a las incertidumbres que influyen en el riesgo pronosticado; y
- explicar que una estimación del riesgo se refiere a lo que puede ocurrir a la población, más que a nivel individual, y señalar que esto es lo que ocurre también con el enfoque de la IDA/IDT.

2. Conclusiones

- La creación de modelos de la relación dosis-respuesta completos se puede considerar un método alternativo más complejo o válido que el de la NOAEL en todos los casos en que se disponga de datos apropiados de la relación dosis-respuesta (por ejemplo, para varios grupos de dosis con distintos niveles de respuesta).

- Para los datos cuantales de la relación dosis-respuesta, el interés radica con frecuencia en los niveles bajos de respuesta (incidencia). Esto puede exigir una extrapolación a dosis más

bajas en varios órdenes de magnitud (por ejemplo, para las incidencias de tumores). Sin embargo, los modelos del riesgo de la relación dosis-respuesta que sean igualmente admisibles pueden dar lugar a estimaciones bajas muy divergentes. Un método aplicado actualmente, considerado prudente, consiste en estimar una dosis de referencia$_{10}$ (dosis con un riesgo del 10%) y hacer una extrapolación de manera lineal descendente desde ese punto. Otra opción, todavía en preparación, consiste en aplicar un método bayesiano, que examina los distintos modelos en conjunto.

- Para la obtención de datos continuos de la relación dosis-respuesta hay dos sistemas de utilización de los modelos. Uno consiste en transformar los datos continuos en datos cuantales. El otro en considerar los datos continuos de la relación dosis-respuesta como información de la gravedad del efecto y, por consiguiente, como una función de la dosis. En el segundo sistema, los cambios mensurables de los efectos suelen estar cerca de los niveles de respuesta considerados adversos (por ejemplo, la inhibición del 10% de la colinesterasa) y el problema de la extrapolación a dosis bajas es insignificante o inexistente.

- Con el fin de obtener un valor de la IDA, la IDT o las dosis de referencia, se pueden utilizar los modelos de dosis-respuesta para derivar una dosis de referencia, que se utilizará como punto de partida de la misma manera que se utiliza la NOAEL (es decir, se aplicarían a la dosis de referencia los mismos factores de incertidumbre que a la NOAEL).

- También se pueden utilizar modelos de la relación dosis-respuesta para estimar los riesgos en un determinado nivel de exposición (humana). Para los riesgos expresados como incidencias (datos cuantales) puede ser necesaria la extrapolación a dosis bajas.

- El uso de modelos de dosis-respuesta puede proporcionar información sobre las incertidumbres asociadas con los datos e identificar los factores que contribuyen a ellas en las estimaciones del riesgo.

- La aplicación de modelos de la relación dosis-respuesta a todos los efectos finales puede tener un costo prohibitivo, de manera que sería útil realizar una selección previa de los efectos finales aparentemente más sensibles. Sin embargo, en algunos casos no es fácil identificar los más sensibles mediante una inspección visual, de manera que hay que aplicar el modelo a todos ellos.

- Se debería notificar siempre la dosis de referencia y su límite inferior de confianza, de manera que la calidad de los datos y el ajuste del modelo sean claros y se puedan comparar sus potencias basándose en la dosis de referencia.

- Se deben presentar los resultados de los distintos métodos utilizados en los modelos de dosis-respuesta.

3. Recomendaciones

- Se deben examinar los protocolos de las pruebas de toxicidad (por ejemplo, las directrices de la Organización de Cooperación y Desarrollo Económicos) para conseguir unos resultados óptimos de las dosis de referencia y demás métodos basados en modelos de la relación dosis-respuesta, por ejemplo las formulaciones óptimas correspondientes al número de animales y el número de dosis para diferentes curvas de la relación dosis-respuesta.

- Hay que elaborar mejores orientaciones para el análisis combinado de distintas series de datos, a fin de estimar las dosis de referencia con mayor precisión.

- Es necesario fomentar un mayor conocimiento de cuándo y cómo se ha de utilizar la respuesta de referencia.

- Hay que tratar de conocer mejor la forma de la curva de la relación dosis-respuesta a dosis bajas. Se requieren nuevas investigaciones para determinar la base biológica de la extrapolación (por ejemplo, utilizando biomarcadores, precursores de tumores, animales modificados genéticamente y la tóxicocinética para la estimación de dosis específicas).

- Es necesario elaborar orientaciones mejores para la comunicación del riesgo basada en los resultados de los modelos dosis-respuesta y de las técnicas de evaluación probabilística. Deben incluir la comunicación de los tipos de incertidumbre y la relación con la variabilidad estadística, la imprecisión y la utilización de intervalos de confianza.

- Se debe examinar la utilización de modelos de la relación dosis-respuesta y se han de elaborar principios generales adicionales para su uso cuando se disponga de más experiencia.

THE ENVIRONMENTAL HEALTH CRITERIA SERIES (continued)

Acetaldehyde (No. 167, 1995)
Acetone (No. 207, 1998)
Acetonitrile (No. 154, 1993)
Acrolein (No. 127, 1991)
Acrylamide (No. 49, 1985)
Acrylic acid (No. 191, 1997)
Acrylonitrile (No. 28, 1983)
Aged population, principles for evaluating the effects of chemicals (No. 144, 1992)
Aldicarb (No. 121, 1991)
Aldrin and dieldrin (No. 91, 1989)
Allergic hypersensitization associated with exposure to chemicals, principles and methods for assessing (No. 212, 1999)
Allethrins (No. 87, 1989)
Aluminium (No. 194, 1997)
Amitrole (No. 158, 1994)
Ammonia (No. 54, 1986)
Anticoagulant rodenticides (No. 175, 1995)
Arsenic (No. 18, 1981, 1st edition)
Arsenic and arsenic compounds (No. 224, 2001, 2nd edition)
Asbestos and other natural mineral fibres (No. 53, 1986)
Assessment of risks to human health from exposure to chemicals, principles for the (No. 210, 1999)
Autoimmunity associated with exposure to chemicals, principles and methods for assessing (No. 236, 2006)
Bacillus thuringiensis (No. 217, 1999)
Barium (No. 107, 1990)
Benomyl (No. 148, 1993)
Benzene (No. 150, 1993)
Beryllium (No. 106, 1990)
Biomarkers and risk assessment: concepts and principles (No. 155, 1993)
Biomarkers in risk assessment: validity and validation (No. 222, 2001)
Biotoxins, aquatic (marine and freshwater) (No. 37, 1984)
Boron (No. 204, 1998)
Brominated diphenylethers (No. 162, 1994)
Butanols – four isomers (No. 65, 1987)
Cadmium (No. 134, 1992)
Cadmium – environmental aspects (No. 135, 1992)
Camphechlor (No. 45, 1984)
Carbamate pesticides: a general introduction (No. 64, 1986)
Carbaryl (No. 153, 1994)
Carbendazim (No. 149, 1993)
Carbon disulfide (No. 10, 1979)
Carbon monoxide
(No. 13, 1979, 1st edition)
(No. 213, 1999, 2nd edition)
Carbon tetrachloride (No. 208, 1999)
Carcinogens, summary report on the evaluation of short-term *in vitro* tests (No. 47, 1985)
Carcinogens, summary report on the evaluation of short-term *in vivo* tests (No. 109, 1990)
Chlordane (No. 34, 1984)
Chlordimeform (No. 199, 1997)
Chlordecone (No. 43, 1984)
Chlorendic acid and anhydride (No. 185, 1996)
Chlorinated paraffins (No. 181, 1996)
Chlorine and hydrogen chloride (No. 21, 1982)
Chloroalkyl ethers, selected (No. 201, 1998)
Chlorobenzenes other than hexachlorobenzene (No. 128, 1991)
Chlorofluorocarbons, fully halogenated (No. 113, 1990)
Chlorofluorocarbons, partially halogenated (ethane derivatives) (No. 139, 1992)
(methane derivatives) (No. 126, 1991)
Chloroform (No. 163, 1994)
Chlorophenols (No. 93, 1989)
Chlorothalonil (No. 183, 1996)
Chromium (No. 61, 1988)
Chrysotile asbestos (No. 203, 1998)
Copper (No. 200, 1998)
Cresols (No. 168, 1995)
Cyhalothrin (No. 99, 1990)
Cypermethrin (No. 82, 1989)
Cypermethrin, alpha- (No. 142, 1992)
DDT and its derivatives (No. 9, 1979)
DDT and its derivatives – environmental aspects (No. 83, 1989)
Deltamethrin (No. 97, 1990)
Demeton-S-methyl (No. 197, 1997)
Dermal absorption (No. 235, 2006)
Diaminotoluenes (No. 74, 1987)
Diazinon (No. 198, 1997)
1,2-Dibromoethane (No. 177, 1996)
Di-*n*-butyl phthalate (No. 189, 1997)
1,2-Dichloroethane
(No. 62, 1987, 1st edition)
(No. 176, 1995, 2nd edition)
2,4-Dichlorophenoxyacetic acid (2,4-D) (No. 29, 1984)
2,4-Dichlorophenoxyacetic acid – environmental aspects (No. 84, 1989)
1,3-Dichloropropene, 1,2-dichloropropane and mixtures (No. 146, 1993)
Dichlorvos (No. 79, 1988)
Dinitro-ortho-cresol (No. 220, 2000)
Diesel fuel and exhaust emissions (No. 171, 1996)
Diethylhexyl phthalate (No. 131, 1992)
Diflubenzuron (No. 184, 1996)
Dimethoate (No. 90, 1989)
Dimethylformamide (No. 114, 1991)
Dimethyl sulfate (No. 48, 1985)
Diseases of suspected chemical etiology and their prevention, principles of studies on (No. 72, 1987)
Disinfectants and disinfectant by-products (No. 216, 1999)
Dithiocarbamate pesticides, ethylenethiourea, and propylenethiourea: a general introduction (No. 78, 1988)
Electromagnetic fields (No. 137, 1992)
Elemental speciation in human health risk assessment (No. 234, 2006)
Endosulfan (No. 40, 1984)
Endrin (No. 130, 1992)
Environmental epidemiology, guidelines on studies in (No. 27, 1983)
Epichlorohydrin (No. 33, 1984)
Essential trace elements: Principles and

THE ENVIRONMENTAL HEALTH CRITERIA SERIES (continued)

methods for the assessment of risk (No. 228, 2001)
Ethylbenzene (No. 186, 1996)
Ethylene oxide (No. 55, 1985)
Extremely low frequency (ELF) fields (No. 36, 1984)
(No. 238, 2007)
Fenitrothion (No. 133, 1992)
Fenvalerate (No. 95, 1990)
Flame retardants: a general introduction (No. 192, 1997)
Flame retardants: tris(chloropropyl) phosphate and tris(2-chloroethyl) phosphate (No. 209, 1998)
Flame retardants: tris(2-butoxyethyl) phosphate, tris(2-ethylhexyl) phosphate and tetrakis(hydroxymethyl) phosphonium salts (No. 218, 2000)
Fluorides (No. 227, 2001)
Fluorine and fluorides (No. 36, 1984)
Food additives and contaminants in food, principles for the safety assessment of (No. 70, 1987)
Formaldehyde (No. 89, 1989)
Fumonisin B_1 (No. 219, 2000)
Genetic effects in human populations, guidelines for the study of (No. 46, 1985)
Glyphosate (No. 159, 1994)
Guidance values for human exposure limits (No. 170, 1994)
Health risks in children associated with exposure to chemicals, principles for evaluating (No. 237, 2006)
Heptachlor (No. 38, 1984)
Hexachlorobenzene (No. 195, 1997)
Hexachlorobutadiene (No. 156, 1994)
Alpha- and beta-hexachlorocyclohexanes (No. 123, 1992)
Hexachlorocyclopentadiene (No. 120, 1991)
n-Hexane (No. 122, 1991)
Human exposure assessment (No. 214, 2000)
Hydrazine (No. 68, 1987)
Hydrogen sulfide (No. 19, 1981)
Hydroquinone (No. 157, 1994)
Immunotoxicity associated with exposure to chemicals, principles and methods for assessment (No. 180, 1996)
Infancy and early childhood, principles for evaluating health risks from chemicals during (No. 59, 1986)
Isobenzan (No. 129, 1991)
Isophorone (No. 174, 1995)
Kelevan (No. 66, 1986)
Lasers and optical radiation (No. 23, 1982)
Lead (No. 3, 1977)[a]
Lead, inorganic (No. 165, 1995)
Lead – environmental aspects (No. 85, 1989)
Lindane (No. 124, 1991)
Linear alkylbenzene sulfonates and related compounds (No. 169, 1996)
Magnetic fields (No. 69, 1987)
Man-made mineral fibres (No. 77, 1988)
Manganese (No. 17, 1981)
Mercury (No. 1, 1976)[a]
Mercury – environmental aspects (No. 86, 1989)
Mercury, inorganic (No. 118, 1991)
Methanol (No. 196, 1997)
Methomyl (No. 178, 1996)
2-Methoxyethanol, 2-ethoxyethanol, and their acetates (No. 115, 1990)
Methyl bromide (No. 166, 1995)
Methylene chloride (No. 32, 1984, 1st edition)
(No. 164, 1996, 2nd edition)
Methyl ethyl ketone (No. 143, 1992)
Methyl isobutyl ketone (No. 117, 1990)
Methylmercury (No. 101, 1990)
Methyl parathion (No. 145, 1992)
Methyl *tertiary*-butyl ether (No. 206, 1998)
Mirex (No. 44, 1984)
Modelling dose–response for the risk assessment of chemicals, Principles for (No. 239, 2009)
Morpholine (No. 179, 1996)
Mutagenic and carcinogenic chemicals, guide to short-term tests for detecting (No. 51, 1985)
Mycotoxins (No. 11, 1979)
Mycotoxins, selected: ochratoxins, trichothecenes, ergot (No. 105, 1990)
Nephrotoxicity associated with exposure to chemicals, principles and methods for the assessment of (No. 119, 1991)
Neurotoxicity associated with exposure to chemicals, principles and methods for the assessment of (No. 60, 1986)
Neurotoxicity risk assessment for human health, principles and approaches (No. 223, 2001)
Nickel (No. 108, 1991)
Nitrates, nitrites, and *N*-nitroso compounds (No. 5, 1978)[a]
Nitrobenzene (No. 230, 2003)
Nitrogen oxides (No. 4, 1977, 1st edition)[a]
(No. 188, 1997, 2nd edition)
2-Nitropropane (No. 138, 1992)
Nitro-and nitro-oxypolycyclic aromatic hydrocarbons, selected (No. 229, 2003)
Noise (No. 12, 1980)[a]
Organophosphorus insecticides: a general introduction (No. 63, 1986)
Palladium (No. 226, 2001)
Paraquat and diquat (No. 39, 1984)
Pentachlorophenol (No. 71, 1987)
Permethrin (No. 94, 1990)
Pesticide residues in food, principles for the toxicological assessment of (No. 104, 1990)
Petroleum products, selected (No. 20, 1982)
Phenol (No. 161, 1994)
d-Phenothrin (No. 96, 1990)
Phosgene (No. 193, 1997)
Phosphine and selected metal phosphides (No. 73, 1988)
Photochemical oxidants (No. 7, 1978)
Platinum (No. 125, 1991)
Polybrominated biphenyls (No. 152, 1994)
Polybrominated dibenzo-*p*-dioxins and dibenzofurans (No. 205, 1998)
Polychlorinated biphenyls and terphenyls (No. 2, 1976, 1st edition)[a]
(No. 140, 1992, 2nd edition)
Polychlorinated dibenzo-*p*-dioxins and dibenzofurans (No. 88, 1989)
Polycyclic aromatic hydrocarbons, selected non-heterocyclic (No. 202, 1998)
Progeny, principles for evaluating health risks associated with exposure to chemicals during pregnancy (No. 30, 1984)
1-Propanol (No. 102, 1990)
2-Propanol (No. 103, 1990)
Propachlor (No. 147, 1993)
Propylene oxide (No. 56, 1985)
Pyrrolizidine alkaloids (No. 80, 1988)
Quintozene (No. 41, 1984)
Quality management for chemical safety testing (No. 141, 1992)
Radiofrequency and microwaves (No. 16, 1981)
Radionuclides, selected (No. 25, 1983)
Reproduction, principles for evaluating health risks associated with exposure to chemicals (No. 225, 2001)
Resmethrins (No. 92, 1989)

---

[a] Out of print

THE ENVIRONMENTAL HEALTH CRITERIA SERIES (continued)

Selenium (No. 58, 1986)
Static fields (No. 232, 2006)
Synthetic organic fibres, selected (No. 151, 1993)
Styrene (No. 26, 1983)
Sulfur oxides and suspended particulate matter (No. 8, 1979)
Tecnazene (No. 42, 1984)
Tetrabromobisphenol A and derivatives (No. 172, 1995)
Tetrachloroethylene (No. 31, 1984)
Tetradifon (No. 67, 1986)
Tetramethrin (No. 98, 1990)
Thallium (No. 182, 1996)
Thiocarbamate pesticides: a general introduction (No. 76, 1988)
Tin and organotin compounds (No. 15, 1980)
Titanium (No. 24, 1982)
Tobacco use and exposure to other agents (No. 211, 1999)
Toluene (No. 52, 1986)
Toluene diisocyanates (No. 75, 1987)
Toxicity of chemicals (Part 1), principles and methods for evaluating the (No. 6, 1978)
Toxicokinetic studies, principles of (No. 57, 1986)
Transgenic animal mutagenicity assays (No. 233, 2006)
Tributyl phosphate (No. 112, 1991)
Tributyltin compounds (No. 116, 1990)
Trichlorfon (No. 132, 1992)
1,1,1-Trichloroethane (No. 136, 1992)
Trichloroethylene (No. 50, 1985)
Tricresyl phosphate (No. 110, 1990)
Triphenyl phosphate (No. 111, 1991)
Tris- and bis(2,3-dibromopropyl) phosphate (No. 173, 1995)
Ultrasound (No. 22, 1982)
Ultraviolet radiation
(No. 14, 1979, 1st edition)
(No. 160, 1994, 2nd edition)
Vanadium (No. 81, 1988)
Vinyl chloride (No. 215, 1999)
Vinylidene chloride (No. 100, 1990)
White spirit (No. 187, 1996)
Xylenes (No. 190, 1997)
Zinc (No. 221, 2001)

CONCISE INTERNATIONAL CHEMICAL ASSESSMENT DOCUMENT SERIES

CICADs are IPCS risk assessment documents that provide concise but critical summaries of the relevant scientific information concerning the potential effects of chemicals upon human health and/or the environment

Acrolein (No. 43, 2002)
Acrylonitrile (No. 39, 2002)
Arsine: human health aspects (No. 47, 2002)
Asphalt (bitumen) (No. 59, 2004)
Azodicarbonamide (No. 16, 1999)
Barium and barium compounds (No. 33, 2001)
Benzoic acid and sodium benzoate (No. 26, 2000)
Benzyl butyl phthalate (No. 17, 1999)
Beryllium and beryllium compounds (No. 32, 2001)
Biphenyl (No. 6, 1999)
Bromoethane (No. 42, 2002)
1,3-Butadiene: human health aspects (No. 30, 2001)
2-Butenal (No. 74, 2008)
2-Butoxyethanol (No. 10, 1998)
2-Butoxyethanol (update) (No. 67, 2005)
Butyl acetates (No. 64, 2005)
Carbon disulfide (No. 46, 2002)
Chloral hydrate (No. 25, 2000)
Chlorinated naphthalenes (No. 34, 2001)
Chlorine dioxide (No. 37, 2002)
4-Chloroaniline (No. 48, 2003)
Chlorobenzenes other than hexachlorobenzene: environmental aspects (No. 60, 2004)
Chloroform (No. 58, 2004)
Coal tar creosote (No. 62, 2004)
Cobalt and inorganic cobalt compounds (No. 69, 2006)
Crystalline silica, quartz (No. 24, 2000)
Cumene (No. 18, 1999)
1,2-Diaminoethane (No. 15, 1999)
3,3'-Dichlorobenzidine (No. 2, 1998)
1,2-Dichloroethane (No. 1, 1998)
1,1-Dichloroethene (Vinylidene chloride) (No. 51, 2003)
2,2-Dichloro-1,1,1-trifluoroethane (HCFC-123) (No. 23, 2000)
Diethyl phthalate (No. 52, 2003)
Diethylene glycol dimethyl ether (No. 41, 2002)
Dimethylformamide, N,N- (No. 31, 2001)
Diphenylmethane diisocyanate (MDI) (No. 27, 2001)
Elemental mercury and inorganic mercury compounds; human health aspects (No. 50, 2004)
Ethylenediamine (No. 15, 1999)
Ethylene glycol: environmental aspects (No. 22, 2000)
Ethylene glycol: human health aspects (No. 45, 2002))
Ethylene oxide (No. 54, 2003)
Formaldehyde (No. 40, 2002)
2-Furaldehyde (No. 21, 2000)
Glyoxal (No. 57, 2004)
Heptachlor (No. 70, 2006)

CONCISE INTERNATIONAL CHEMICAL ASSESSMENT DOCUMENT SERIES (continued)

HCFC-123 (No. 23, 2000)
Hydrogen cyanide and cyanides: human health aspects (No. 61, 2004)
Hydrogen sulfide: human health aspects (No. 53, 2003)
Limonene (No. 5, 1998)
Manganese and its compounds (No. 12, 1999)
Manganese and its compounds: environmental aspects (No. 63, 2004)
Mercury, elemental, and inorganic mercury compounds: human health aspects (No. 50, 2003)
Methyl and ethyl cyanoacrylates (No. 36, 2001)
Methyl chloride (No. 28, 2001)
Methyl methacrylate (No. 4, 1998)
N-Methyl-2-pyrrolidone (No. 35, 2001)
Methyltin, butyltin, and octyltin compounds, Mono- and disubstituted (No. 73, 2006)
Mononitrophenols (No. 20, 2000)
N-Nitrosodimethylamine (No. 38, 2002)
Phenylhydrazine (No. 19, 2000)
N-Phenyl-1-naphthylamine (No. 9, 1998)
Polychlorinated biphenyls: human health aspects (No. 55, 2003)
Resorcinol (No. 71, 2006)
Silver and silver compounds: environmental aspects (No. 44, 2002)
1,1,2,2-Tetrachloroethane (No. 3, 1998)
Tetrachloroethene (No. 68, 2006)
1,1,2,2-Tetrafluoroethane (No. 11, 1998)
Thiourea (No. 49, 2003)
Tin and inorganic tin compounds (No. 65, 2005)
o-Toluidine (No. 7, 1998)
2,4,6-Tribromophenol and other simple brominated phenols (No. 66, 2005)
Tributylin oxide (No. 14, 1999)
1,2,3-Trichloropropane (No. 56, 2003)
Triglycidyl isocyanurate (No. 8, 1998)
Triphenyltin compounds (No. 13, 1999)
Vanadium pentoxide and other inorganic vanadium compounds (No. 29, 2001)

To order further copies of monographs in these series, please contact WHO Press, World Health Organization, 1211 Geneva 27, Switzerland (Fax: +41-22-791 4857; E-mail: bookorders@who.int)

ISBN 978 92 4 157239 2

CH018.qxd 4/10/09 12:16 PM Page 1

18

Mouse models of human B lymphoid neoplasms

HERBERT C MORSE III, JERROLD M WARD AND MICHAEL A TEITELL

Introduction 00
Precursor B cell lymphoblastic lymphoma/leukemia 00
Definition 00
Synonyms 00
Epidemiology of spontaneous disease 00
Epidemiology of disease in genetically engineered mice 00
Histologic features 00
Immunologic features 00
Molecular features 00
Postulated cell of origin 00
Comments 00
Small B cell lymphoma/leukemia 00
Definition 00
Synonyms 00
Epidemiology of spontaneous disease 00
Epidemiology of disease in genetically engineered mice 00
Histologic features 00
Immunologic features 00
Molecular features 00
Postulated cell of origin 00
Comments 00
Splenic marginal zone B cell lymphoma 00
Definition 00
Synonyms 00
Epidemiology of spontaneous disease 00
Epidemiology of disease in genetically engineered mice 00
Histologic features 00
Immunologic features 00
Molecular features 00
Postulated cell of origin 00
Comments 00
Follicular B cell lymphoma 00
Definition 00
Synonyms 00
Epidemiology of spontaneous disease 00
Epidemiology of disease in genetically engineered mice 00
Histologic features 00
Immunologic features 00
Molecular features 00
Postulated cell of origin 00
Diffuse large B cell lymphoma 00
Definition 00
DLBCL – centroblastic 00
DLBCL – immunoblastic 00
DLBCL – histiocyte associated 00
DLBCL in genetically engineered mice 00
Diffuse high-grade blastic B cell lymphoma/leukemia 00
Definition 00
Synonyms 00
Epidemiology of spontaneous disease 00
Epidemiology of disease in genetically engineered mice 00
Histologic features 00
Immunologic features 00
Molecular features 00
Presumed cell of origin 00
Comments 00
Key points 00
References 00

INTRODUCTION

The identification of T cells and B cells as subsets of lymphocytes readily distinguished by phenotype and function heralded a new era for the field of immunology. This milestone also marked the end of an epoch in which hematologic malignancies were categorized primarily on the basis of their cytology, histology, and clinical presentation. Classification schemes founded on recognition of microscopic features were gradually replaced by those emphasizing cell lineage and state of differentiation as defined by immunophenotyping. Identification of distinctive chromosomal abnormalities in metaphase spreads foresaw the wealth of molecular genetic approaches that now belong to the armamentarium of a modern department of human hematopathology. The sum of approaches to diagnosis provided by clinical features, morphology, immunophenotype, and genetic characteristics was brought to bear in the 2001 World Health

EXHIBIT
15-29
9/5/17

CH018.qxd 4/10/09 12:16 PM Page 2

Table 18.1 Relationships between classification of mouse and human B cell-lineage lymphomas

Human B cell neoplasm	Mouse B cell neoplasm
Precursor B cell neoplasm	Precursor B cell neoplasm
Precursor B lymphoblastic leukemia/lymphoma	Precursor B lymphoblastic leukemia/lymphoma
Mature B cell neoplasms	Mature B cell neoplasms
Chronic lymphocytic leukemia/small lymphocytic lymphoma	Small B cell lymphoma/leukemia
B cell prolymphocytic leukemia	
Lymphoplasmacytic lymphoma	
Splenic marginal zone lymphoma	Splenic marginal zone lymphoma
Hairy cell leukemia	
Plasma cell myeloma	
Solitary plasmacytoma of bone	
Extraosseus plasmacytoma	Extraosseus plasmacytoma
Extranodal marginal zone B cell lymphoma of mucosal-associated lymphoid tissue	
Nodal marginal zone lymphoma	
Follicular lymphoma	Follicular B cell lymphoma
Mantle cell lymphoma	
Diffuse large B cell lymphoma	Diffuse large B cell lymphoma
Morphologic variants:	Morphologic variants:
Centroblastic	Centroblastic
Immunoblastic	Immunoblastic
T cell/histiocyte-rich	Histiocyte associated
Subtypes:	Subtypes:
Mediastinal (thymic) large B cell lymphoma	Mediastinal (thymic) large B cell lymphoma
Intravascular large B cell lymphoma	
Primary effusion lymphoma	
Burkitt lymphoma/leukemia	
	Diffuse high-grade B cell blastic lymphoma/leukemia

Organization (WHO) classification of human hematopoietic neoplasms,[1] which has recently been updated.

An appreciation of the diversity and pathogenesis of mouse hematopoietic neoplasms has historically lagged far behind that of seemingly similar human malignancies. The spur to changing the *status quo* came from the remarkable progress made in manipulating the mouse genome. The use of transgenic (TG), knockout, knockin, and mutagenized mice to model human diseases has yielded a wealth of new strains, many of which unexpectedly develop hematologic malignancies. In addition, the rate at which the scientific community is generating mouse models of human cancer that require validation has grown almost exponentially. To optimize the utility of mouse models of human lymphoma and leukemia, community standards were developed for classifying both lymphoid and nonlymphoid neoplasms using systems that parallel the WHO schemes for human diseases.[2] These 'Bethesda proposals' were recognized to have inherent limitations due to lack of information on specific disorders and were designed to be open to revision with the development of new data.

This chapter will focus exclusively on B cell–lineage lymphomas, except for plasma cell–related neoplasms, which are covered in Chapter 14. The classification scheme to be used here differs from that published in the Bethesda proposals, with changes suggested by information garnered from the literature and the experiences of our laboratories. The classification defines individual disease entities based on morphology, immunophenotype, genetic features, and state of differentiation in relation to a presumed cell of origin.

Table 18.1 lists the mouse lymphoma types distinguished by these characteristics and their closest human counterpart. This comparison shows that many disease entities recognized in humans do not have parallels in the mouse for either spontaneous diseases or those appearing in genetically engineered mice. Several factors are likely to have contributed to this discrepancy. First, much of our detailed knowledge about mouse B cell lymphomas comes from studies of a limited number of strains. The best-studied strains are NFS.V$^+$ congenic mice[3] and the AKXD recombinant inbred (RI) strains.[4–6] The NFS congenics and most of the RI strains express ecotropic murine leukemia viruses (MuLV) from endogenous germ-line loci at high levels from early in life. When MuLV infect somatic cells, their insertion into genomic DNA is inherently mutagenic.

Some of these mutagenic insertions result in the unprogrammed activation of cellular genes, some of which are protooncogenes that can participate in the process of transformation.[7] These contributions to disease pathogenesis in the mouse have no direct parallels in human B lymphomas that instead feature activation of protooncogenes induced by balanced chromosomal translocations. Translocations resulting in protooncogene activation do occur in the mouse, being found in almost all pristane-induced plasmacytomas of BALB/c mice[8] and rarely in diffuse large B cell lymphomas,[9] but they are certainly not the rule.

Second, there are significant differences between mouse and human immune systems in their development, structure, phenotype, and function.[10,11] For example, expression of CD23 and CD5 is mutually exclusive on mouse but not on human B cell subsets, and CD38 is downregulated on mouse germinal center (GC) B cells but upregulated in humans.[12] More important, perhaps, is that the spleen is the major secondary lymphoid organ of the mouse, whereas lymph nodes fill that niche in humans. In addition, the mouse spleen is responsible for extramedullary hematopoiesis throughout the animal's lifespan, while hematopoiesis in humans is confined to the bone marrow.

Finally, the genetic and epigenetic alterations required for neoplastic transformation sometimes differ for mouse and human.[13] As one example, programmed telomere shortening is viewed as a tumor-suppressor pathway in humans, with activation of telomerase being a requirement for averting replicative senescence. In contrast, telomerase is not repressed in murine somatic cells, and the cellular response to telomere damage differs for mice and humans.[14] A second example comes from studies of mice with targeted mutations of tumor-suppressor genes in which the phenotypes of mutant mice are sometimes highly discordant with the effects of the mutation in humans.[15] Species-specific differences in the immune system and the molecular circuitry required for transformation could thus make it difficult to model some human diseases in mice.

This difficulty is exemplified by efforts to develop a mouse model of Burkitt lymphoma by generating 'simple' *MYC* transgenics,[16–19] yeast artificial chromosome (YAC) transgenics with *MYC* inserted into the IgH locus[20,21] and, most recently, a knockin mouse with *MYC* inserted into the IgH locus in an orientation characteristic of most translocations in endemic Burkitt lymphoma and mouse plasmacytomas.[22] All these mice developed lymphomas with varying latency. Analyses of immunoglobulin (Ig) gene rearrangements and phenotypic analyses showed that the Eμ-Myc lymphomas included pre-B as well as IgM^+ IgD^- lymphomas, a phenotype consistent with immature or transitional B cells. The λ-Myc lymphomas were also found to have features of immature or transitional B cells and had non-mutated Ig genes. Approximately 25 percent of the Myc knockin lymphomas were lymphoblastic with a starry-sky appearance and expression of Bcl-6, a phenotype very much like human Burkitt lymphoma;[22] however, the Ig genes of these lymphomas did not contain somatic hypervariable region mutations,[23] whereas all human Burkitt lymphoma Ig genes contain such mutations to varying extents. Could some mouse B cells passage the GC without being mutated? Can Ig gene mutation occur outside the GC? The answer to the latter question has recently been shown to be 'yes.'[24] Answers to questions such as this will be critical for fully assessing the validity of a mouse model for a human disease.

These concerns are important to those involved in basic research efforts to understand the pathogenesis of mouse lymphomas or to develop mouse models of human cancers. Other scientific communities need only determine whether mice presented to them have a neoplasm and, if so, how to classify that tumor. In the former circumstance, many approaches can be applied to make diagnoses that contribute to understanding B cell pathogenesis. In the latter, histologic analysis of fixed sections stained with hematoxylin and eosin will sometimes be the only material available for diagnosis. The classification system described below for each lymphoma class is intended to be useful for both the basic research and applied science communities.

PRECURSOR B CELL LYMPHOBLASTIC LYMPHOMA/LEUKEMIA

Definition

Precursor B cell lymphoblastic lymphoma/leukemia (pre-B LBL) is a high-grade neoplasm of small to medium-sized, round, uniform precursor B cells arising in the bone marrow. A leukemic phase and spread to peripheral lymphoid and nonlymphoid tissues is frequently present.

Synonyms

Lymphoblastic lymphoma; lymphoblastic leukemia

Epidemiology of spontaneous disease

SL/Kh is a recombinant congenic mouse strain, derived from proto-SL and AKR strains, that develops pre-B LBL.[25,26] Tumors arise from a combination of host genetic factors and reintegration of endogenous MuLV provirus into several loci that include *Stat5a, Stat5b, Evi3, c-Myc,* and *n-Myc.*

Crosses between AKR and DBA/2 mice created 24 distinct AKXD RI strains with distinct MuLV proviral integration sites, most of which develop lymphoma.[4–6] Tumors classified as pre-B LBL on the basis of lymphoblastic cytology and the presence of clonal IgH rearrangements without detectable IgL clonality morphology arise in at least ten different AKXD RI strains, although analysis of additional phenotypic features of pre-B LBL have not yet been performed.[6]

CH018.qxd 4/10/09 12:16 PM Page 4

Epidemiology of disease in genetically engineered mice

BCR-ABL^{p210} *retrovirus-transduced bone marrow cells*:[27–29] Infection of bone marrow (BM) cells from mice treated with 5-fluorouracil with a BCR-ABL^{p210}-expressing retrovirus followed by injection of lethally irradiated hosts yielded pre-B LBL in DBA/2 recipients, T cell acute lymphocytic leukemia (T-ALL) in C57BL/6 recipients, and other hematopoietic tumors in BALB/c recipients.

Metallothionein-1 promoter-BCR-ABL^{p190} *TG*:[30,31] Transgenic mice developed either chronic myeloid leukemia (CML) in blast crisis or pre-B LBL.

IgV$_H$-promoter-Eμ or retroviral LTR-bcr/v-abl TG:[32] Transgenic mice expressing a synthetic BCR-ABL gene created by fusing *bcr* with *v-abl* sequences developed either pre-B LBL or probably pre-T LBL.

BCR-ABL^{p190} knockin:[33] BCR-ABL inserted in-frame into exon 1 of the endogenous *bcr* locus by homologous recombination mimics half of the reciprocal translocation seen in Philadelphia (Ph)-positive cases. These *BCR-ABL^{p190}* 'knockin' mice developed pre-B LBL.

Eμ-myc TG:[16] These mice developed pre-B LBL together with more mature lymphomas. Most tumors lacked IgL gene rearrangements and were surface immunoglobulin (sIg)$^-$. Crossing E*μ-myc* TG with E*μ-pim1* TG mice, which develop low-frequency pre-T LBL, resulted in congenital pre-B LBLs that were B220$^+$ and sIg$^-$.[34]

Eμ-Blk TG:[35] Blk is a Src family tyrosine kinase that associates with and signals from the pre-BCR and BCR.[36] Mice with a constitutively active Blk (Y495F) flanked by a H-2K promoter and E*μ* developed pro-B and pre-B LBL along with intermediate single positive–stage thymic pre-T LBL. B cell–lineage tumors occurred in 45 percent of mice by 6 months. Activating lesions for the *Blk* homolog in human B and T cell tumors have not yet been described.

LTR-TEL/AML1 retrovirus-transduced BM cells:[37–39] A t(12;21)(p13;q22) occurs in 25 percent of pediatric and 3 percent of adult B-ALL cases, resulting in a *TEL/AML1* oncogene. TEL/AML1 retains the amino portion of the TEL protein fused to the AML1 DNA binding domain. TEL/AML1 may cause B-ALL by direct repression of AML1 target genes or by TEL inhibition of other ETS family proteins via binding through its pointed domain.[40–42] Transduction of wild-type or *p16^{Ink4a}*/p19^{Arf} null BM retrovirus was followed by adoptive transfer into lethally irradiated C57BL/6 mice. A low frequency of pre-B-ALL and T-ALL from the transduced wild-type donor cells and a higher frequency of an undetermined leukemia type was seen with the *p16^{Ink4a}*/p19^{Arf}-deficient cells.

Aiolos knockout:[43] Aiolos is a member of a kruppel-like zinc finger transcription factor family that binds DNA and regulates lymphocyte development and function. A knockout of Aiolos causes hyperproliferation and constitutive B cell activation followed by lymphadenopathy and evolution into probable B-LBL in 20 percent of mice. Alternatively, spliced Aiolos isoforms lacking a full complement of DNA binding domains in normal and leukemic human B cells have been reported, but their role in B cell transformation is not resolved.[44,45]

p16^{INK4a}/p19(p14)^{ARF} *knockouts*:[46] The *p16^{INK4a}/p19(p14)^{ARF}* genes overlap in the genome and encode tumor suppressor proteins that inhibit cancer formation.[47] Mice deficient in *p16^{Ink4a}/p19^{Arf}* developed mainly soft tissue sarcomas or B220$^+$ B cell lymphomas with prominent nucleoli that effaced lymph nodes. Although not further characterized, these features suggest a pre-B LBL phenotype. Deletion of *p16^{INK4a}*/p14^{ARF} occurs in 20 percent of B-ALL and 60 percent of T-ALL arising during childhood.[48] Also, DNA methylation of the *p16^{INK4a}* (but not of the *p14^{ARF}*) locus occurs in multiple types of non-Hodgkin lymphoma (NHL) and acute myeloid leukemia (AML), accompanied by loss of *p16^{INK4a}* expression.[49]

IL-7 TG:[50–53] Interleukin-7 (IL-7) and the IL-7/γcR regulate lymphopoiesis. Four distinct transgenic strains were generated to aberrantly express IL-7. A major histocompatibility complex (MHC)-II E$_\alpha$ promoter–IL-7 transgene causes λ5 and EBF-expressing pre/pro-B/pro-B LBL with germ-line Ig genes.[53] Clonality is difficult to establish, and femurs are packed with lymphoid blast cells. There is accompanying lymphadenopathy. The tumor incidence varies by strain, with almost 100 percent of BALB/c and 25 percent of C57BL/6 mice developing morbid disease. Human lymphoid malignancies from IL-7 dysregulation have not been reported.

Histologic features

Pre-B LBL forms sheets of small to medium-sized blast cells in the bone marrow or infiltrated organs. The nuclei appear round to oval with condensed chromatin and a central, prominent nucleolus. There is a high nuclear:cytoplasmic ratio, with scant eosinophilic cytoplasm. A high mitotic index and abundant apoptosis can provide a 'starry-sky' appearance from macrophages stuffed with apoptotic debris (tingible body macrophages) that is histologically indistinguishable from that seen with pre-T LBL or more mature sIg$^+$ diffuse high-grade blastic B cell lymphoblastic lymphomas (DBLL), previously referred to as Burkitt-like.[2] Distinction is made from pre-T LBL, which are CD3$^+$ and B220$^-$, and DBLL, which are sIg$^+$ and contain rearranged IgL. Gross presentation may include lymphadenopathy without splenomegaly, a relatively unique occurrence for B cell lineage mouse tumors.

Immunologic features

Precursor B cells that by flow cytometry (fluorescence-activated cell sorting; FACS) are sIg$^-$, B220$^+$, CD19$^+$, CD43$^{+/-}$, Ly6d (ThB)$^+$, and by immunohistochemistry (IHC) are TdT$^+$.

CH018.qxd 4/10/09 12:16 PM Page 5

Molecular features

Clonal population of precursor B cells with IgH gene rearrangement, typically of one allele and germ-line configuration of κ and λ IgL.

Postulated cell of origin

Bone marrow precursor B cell lymphoblast.

Comments

Mice containing BCR-ABL–infected BM or mice expressing either form of BCR-ABL (or bcr-v-abl) using transgenic, tetracycline-inducible or knockin technologies develop pre-B LBL with variable penetrance and strain dependence.[54] The p210 isoform of BCR-ABL generally does not result in B-ALL in humans. The human counterpart malignancies are pro-B and pre-B acute lymphoblastic leukemia (B-ALL) or B lymphoblastic lymphoma (B-LBL). Pre-B LBL shows similar histologic, cytologic, phenotypic, and molecular features in mice and humans.

SMALL B CELL LYMPHOMA/LEUKEMIA

Definition

A neoplasm of small, round, monomorphic B cells in the spleen and lymph nodes, sometimes admixed with prolymphocytes and paraimmunoblasts in pseudofollicles. Cases without leukemia (i.e., lymphoma) are more common than cases with leukemia.

Synonyms

Lymphocytic lymphoma; small lymphocytic lymphoma; chronic lymphocytic leukemia; well-differentiated lymphocytic lymphoma.

Epidemiology of spontaneous disease

Frequency was ~1 percent in AKXD RI strains and ~12 percent of all B cell lineage lymphomas in NFS.V$^+$.[55] The average age at diagnosis was ~400 days in NFS.V$^+$ mice. Uncommon in most commonly used mouse strains and stocks.

Epidemiology of disease in genetically engineered mice

Eμ-TCL1 TG:[56] Mice exhibited involvement of bone marrow at 2 months, spleen at 5 months, and leukemia at 5 months with CD5$^+$CD11b$^+$Bcl-6$^-$ clonal B cells. Spleens often exhibited marked expansion of the marginal zone.

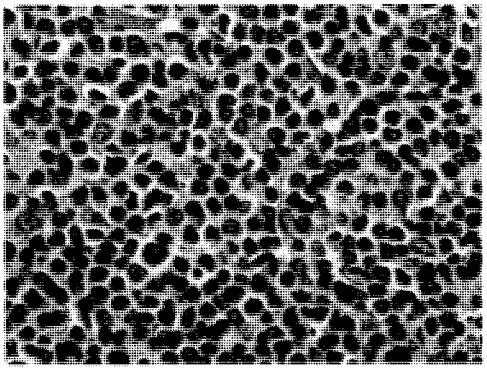


Figure 18.1 Small B cell lymphoma showing a uniform population of small lymphocytes similar to those of the normal mantle zone. There are few mitotic figures, apoptotic bodies, or tingible body macrophages.

TRAF2DN/BCL-2 TG:[57] Transgenic mice exhibited splenic involvement at 10–16 months often associated with leukemia, ascites, and pleural effusion with CD5$^{+/-}$ CD11b$^+$Bcl-6$^-$ B cells. Splenic marginal zone expansion was often present.

IL-5 TG:[58] Forty percent of mice presented with clonal CD5$^+$ B cell leukemia by 22 months of age.

Histologic features

Enlarged splenic white pulp filled with small lymphocytes and some larger prolymphocytes, with few mitotic or apoptotic figures (Fig. 18.1). Isolated accumulations of [AQ1] cells with features of immunoblasts can form proliferation centers similar to those seen in human chronic lymphocytic leukemia (CLL). A readily evident blood phase with predominantly small lymphocytes is seen in about 25 percent of cases. Extensive, diffuse infiltration of lymph nodes, lung, liver, and kidney can occur in the presence or absence of leukemia. Progression to immunoblastic lymphoma is rare, with immunoblasts and small lymphocytes cohabiting isolated areas, a possible parallel to Richter's transformation in human CLL.

Immunologic features

NFS.V$^+$ lymphomas were usually IgM$^+$B220^d CD5^d CD11b$^{+/-}$ by FACS and PAX-5$^+$Bcl-6$^-$ by IHC.

Molecular features

Clonal IgH rearrangements. No major structural rearrangements of protooncogenes or tumor suppressor genes were identified in mice with spontaneous disease.

CH018.qxd 4/10/09 12:16 PM Page 6

Postulated cell of origin

Not known. Immunoglobulin V-region sequencing required to determine whether pre- or post-GC.

Comments

Strains NZB and NZW, and genetically engineered strains including those listed above under induced diseases, regularly exhibit age-related expansions of populations of $CD5^+$ B cells in the peritoneum, spleen, and blood, sometimes sequentially. Expression of CD5 on mouse B cells is usually confined to a substantial proportion of B cells in the peritoneum and ~2 percent of B cells in the spleen.

SPLENIC MARGINAL ZONE B CELL LYMPHOMA

Definition

A neoplasm of cells from the splenic marginal zone presenting as low-, intermediate-, or high-grade disease that arises in the spleen and is almost always restricted to the spleen.

Synonyms

Splenic marginal zone lymphoma.

Epidemiology of spontaneous disease

The frequency of marginal zone lymphoma (MZL) was ~7 percent in NZB mice.[59] Marginal zone lymphoma accounted for ~35 percent of all neoplasms in $NFS.V^+$ and ~18 percent in AKXD strains.[55] The mean latency in NZB mice was ~450 days. In $NFS.V^+$ mice, the average age at diagnosis for low-, intermediate-, or high-grade disease was ~400 days.[60] In AKXD RI strains, the mean latency of predominantly low-grade disease was ~500 days.[6] A significant number of one-year-old TAN mice, a strain derived from a cross between NZB and NZW, also had MZL. Frequency is less than 1 percent in most commonly used mouse strains and stocks.

Epidemiology of disease in genetically engineered mice

In *p53* knockouts,[61] MZL was diagnosed in ~35 percent of ex-breeder mice.

Histologic features

Three levels of progression have been described: MZL, [AQ1] MZL+, and MZL++ (Fig. 18.2, Fig. 18.3, and Fig. 18.4). Marginal zone lymphoma is characterized by a substantially widened marginal zone comprising cells with normal cytology yielding a halo-like appearance around the follicles. Early lesions include hyperplastic marginal zone B cells with abundant eosinophilic cytoplasm and uniformly shaped nuclei. At this stage, mitoses are rare. With progression to MZL+, there is a merging of the marginal zones and the cells have taken on centroblastic features. Infiltration of the red pulp is often extensive, but follicles are usually intact. By the MZL++ stage, the spleen can be almost completely taken over by cells with centroblastic or, more rarely, immunoblastic morphology. The white pulp is usually rudimentary, consisting only of a periarteriolar lymphoid sheath (PALS). Mitoses are common. The lymphoma occasionally extends to the splenic node and rarely

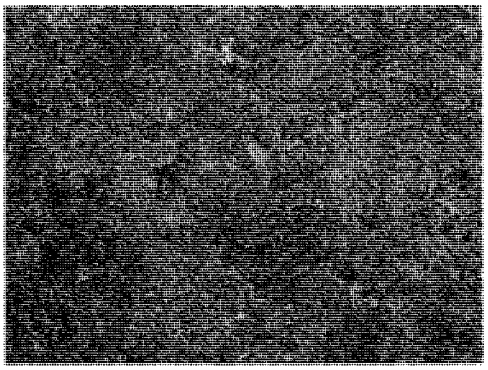


Figure 18.2 Splenic marginal zone lymphoma. Note atrophic white pulp and a halo of normal-appearing marginal zone lymphoid cells fingering out into the red pulp. Mitoses and apoptotic bodies are uncommon at this stage.

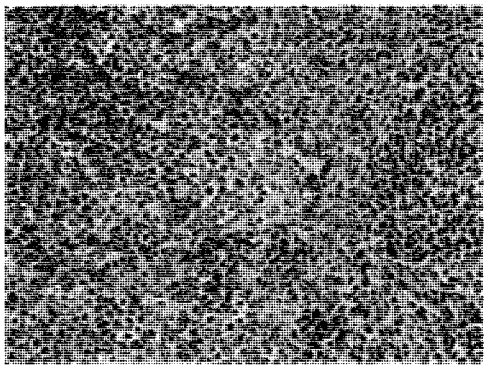


Figure 18.3 Splenic marginal zone lymphoma. The marginal zone is replaced by neoplastic lymphocytes adjacent to atrophic white pulp. Cells retain the features of normal marginal zone cells with an ovoid nucleus featuring fine chromatin, small, often inconspicuous nucleoli, and plentiful cytoplasm.

CH018.qxd 4/10/09 12:16 PM Page 7

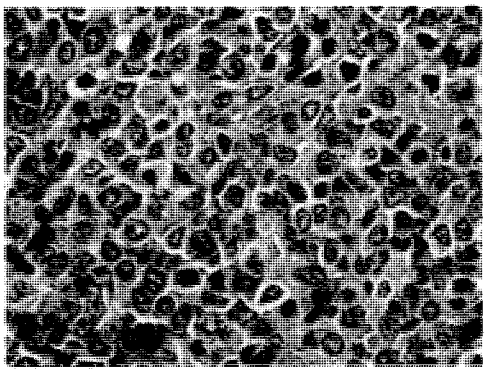


Figure 18.4 Marginal zone lymphoma with pleiomorphic population of lymphocytes with moderate amount of eosinophilic cytoplasm.

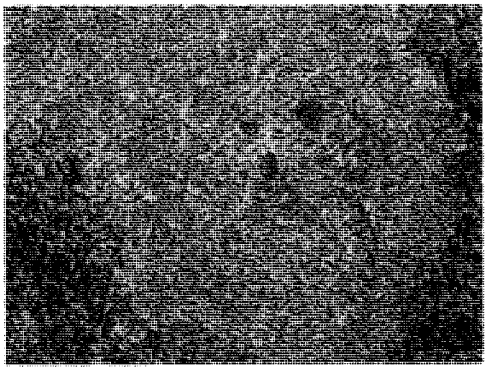


Figure 18.5 Follicular lymphoma. Depletion of normal darker T lymphocytes in the periarteriolar lymphoid sheath and replacement of white pulp by lymphoma.

disseminates to the liver or other tissues. A leukemic phase is extremely rare.

Immunologic features

By FACS analysis, MZL – regardless of grade – expressed IgM at varying levels and little to no IgD.[60] Almost all cells expressed CD5 and CD45R(B200) at low levels and CD38 at high levels but were CD23$^-$. Nearly half of cases expressed CD11B at low levels. By IHC, cases were consistently IgM$^+$CD45R(B220)^{dull} IgD$^-$CD5$^-$.

Molecular features

Clonal for Ig gene rearrangements, with oligoclonality being more common among MZL than MZL+ or MZL++.[60,61] Rearrangements of cellular loci identified in MZL of AKXD RI strains included *Nmyc1*, *Pim1*, and *Evi1* in ~16 percent of cases.[6] Marginal zone lymphoma of NFS.V$^+$ mice studied for common proviral integration sites identified three or more integrations at *Stk10*, *Abcc5*, *Sox4*, *Mela*, and *Edh2*.[62] *Gfi1* was also identified as a common integration site (CIS) and was expressed at elevated levels in all grades of MZL.[62]

Postulated cell of origin

Splenic marginal zone B cell.

Comments

Non-splenic MZL in humans is associated with autoimmune diseases including, most prominently, Sjögren syndrome and Hashimoto thyroiditis.[63] Marginal zone lymphoma also occurs in association with infections caused by *Helicobacter pylori*, *Chlamydia psittaci*, and other agents. The observation that MZL occurs in autoimmune NZB mice and in the TAN strain, derived from a cross between NZB and NZW, suggests a possible etiologic link with human MZL. Shared low-affinity reactivity with autoantigens or repetitive bacterial antigens could be that tie. Autoimmune disease is generally not associated with these lymphomas in mice.

FOLLICULAR B CELL LYMPHOMA

Definition

Follicular B cell lymphoma (FBL) is a neoplasm comprising a mixture of B-lineage cells with cytologic features of centrocytes and centroblasts in varying proportions (Fig. 18.5 and Fig. 18.6). [AQ1]

Synonyms

Follicular lymphoma; centroblastic/centrocytic lymphoma; follicular center cell lymphoma, mixed; reticulum cell sarcoma, type B; lymphoma-pleiomorphic.

Epidemiology of spontaneous disease

This is the most common neoplasm in aging mice of many inbred strains. It accounts for 6 percent of cases in NFS.V$^+$ mice, 35 percent of cases in CFW strains, and 8 percent of cases in AKXD RI strains.[55] The average age at diagnosis is ~435 days for NFS.V$^+$ and ~400 days for the AKXD RI strains but ranges from 18 to 24 months in most other strains and stocks.

CH018.qxd 4/10/09 12:16 PM Page 8

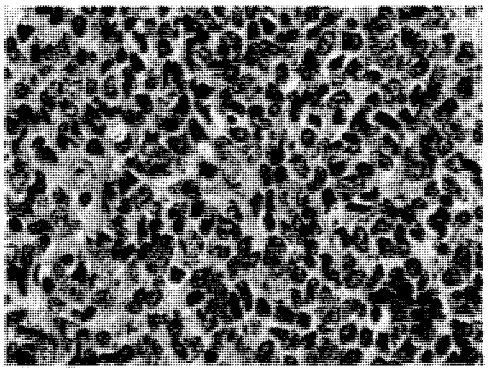


Figure 18.6 Follicular lymphoma showing a pleiomorphic population of centroblasts and centrocytes. Centrocytes feature irregularly shaped, often elongated, relatively dense nuclei with little cytoplasm. Centroblasts are marked by larger, vesicular nuclei with more prominent nucleoli, often with two appended to the nuclear membrane.

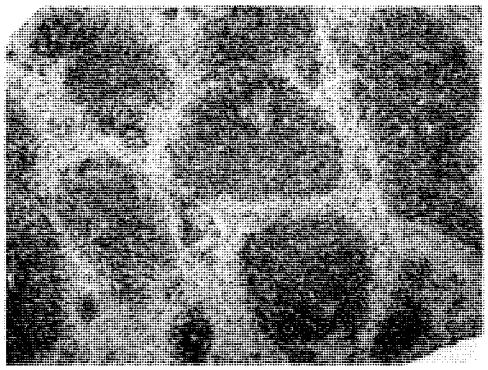


Figure 18.7 Immunohistochemistry of follicular lymphoma B cells expressing CD45R(B220) on the cell surface.

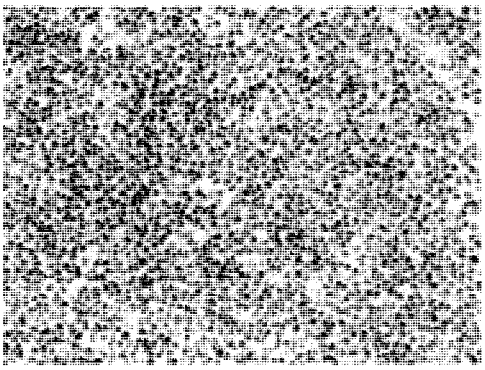


Figure 18.8 Immunohistochemistry of follicular lymphoma expressing PAX-5 (B cell lineage specific activator protein; BSAP) at high levels in the nucleus of lymphoma cells.

Epidemiology of disease in genetically engineered mice

Eμ-BCL-2 TG:[64,65] Both TG lines of mice were reported to develop follicular hyperplasia that progressed to overt lymphoma. No increase in lymphoma incidence was found in later studies of the same strains.

Mcl1 *TG*:[66] About 65 percent of mice presented with lymphoma by two years of age, most presenting with splenomegaly and lymphadenopathy. Around 20 percent of cases were FBL.

VavP-Bcl-2 TG:[67] Approximately 60 percent of mice not dying previously with autoimmune disease succumbed to lymphoma between 10 and 18 months of age. Lymphoma development was preceded by increases in the number and size of GC and class-switched B cells with mutated Ig genes. The expanded GC phenotype was totally dependent on the presence of $CD4^+$ T cells.

Histologic features

Animals present most commonly with gross lesions of splenomegaly and varying degrees of enlargement of the mesenteric lymph node and Peyer's patches. In the spleen, the white pulp is expanded, appearing as white nodules that coalesce as the disease progresses. Occasionally, there is only a solitary large white nodule in the spleen. The smaller nodules are enlarged, sometimes with fused follicles with centrocytes and centroblasts as the chief cellular constituents. Normal small follicular B cells are pushed to the periphery, and the T cell zone is reduced or eliminated. In some mouse strains, tumors may arise within the PALS or inactive white pulp with no apparent relation to the GC. The proportions of centrocytes and centroblasts vary from case to case. The proportion of blasts should be less than 50 percent to differentiate the condition from diffuse large B cell lymphoma (DLBCL).

Immunologic features

By FACS, both the centrocytes and centroblasts in spontaneous cases were most often $IgM^+IgD^-CD5^{dull}CD45R(B220)^{dull\ to\ normal}$. By IHC, both populations were IgM^+B220^+(Fig. 18.7 and Fig. 18.8).[68] [AQ1]

Molecular features

All cases examined in $NFS.V^+$ and AKXD mice were mono- or oligoclonal for Ig gene rearrangements.[3,6] No rearrangements in Bcl-2 were seen in large panels of FBL

from NFS.V$^+$ mice or other inbred strains.[2] Genomic rearrangements of oncogenes or tumor suppressor genes were seen only for *Evi1* in AKXD mice.[6]

Postulated cell of origin

Germinal center centrocytes and centroblasts.

DIFFUSE LARGE B CELL LYMPHOMA

Definition

Diffuse large B cell lymphoma is a B cell neoplasm characterized by a diffuse proliferation of cells with large nuclei and distinctive cytologic features characteristic of each of the variants seen in spontaneous disease: centroblastic (CBL), immunoblastic (IBL), and histiocyte associated (HA) DLBCL.

DLBCL – centroblastic

SYNONYMS

Large cleaved follicular center lymphoma; centroblastic lymphoma.

EPIDEMIOLOGY OF SPONTANEOUS DISEASE

Centroblastic DLBCL accounted for ≈12 percent of cases in NFS.V$^+$ mice and 17 percent in CFW mice but were not observed among AKXD RI strain lymphomas.[55] The average age at diagnosis was ≈400 days in NFS.V$^+$ mice. The disease is uncommon in most commonly used strains and stocks of inbred mice.

HISTOLOGIC FEATURES

The white pulp is greatly expanded, with a population of cells with round nuclei, often with two nucleoli attached to the nuclear membrane, and a moderate amount of basophilic cytoplasm (Fig. 18.9). These are admixed to varying extents with smaller cells with characteristics of centrocytes. The diagnosis is readily made when over 70 percent of the cells are blasts. When the proportions of centrocytes and blasts range from 40 percent to 70 percent and the ratio varies in different fields, a distinction between FBL and DLBCL is difficult. In the classification system of Fredrickson and Harris,[68] CBL is subdivided into follicular, marginal zone, and diffuse to identify distinct regional origins for the first two and CBL of uncertain origin for the last. The lymphoma frequently infiltrates the lung, liver, and kidney and, less frequently, the bone marrow. [AQ1]

IMMUNOLOGIC FEATURES

By FACS, the tumors are IgM$^+$ or IgG$^+$, CD19$^+$, and B220$^+$ and often express CD5 at low levels. The tumors are almost uniformly Bcl-6$^+$ PAX-5$^+$IRF-8$^+$PU.1$^+$ by IHC, but CD138$^-$XBP-1$^-$IRF-4$^-$BLIMP$^-$.

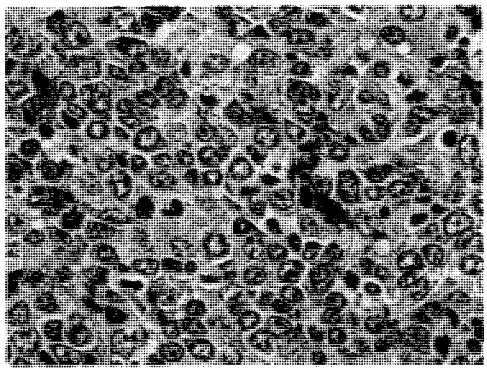


Figure 18.9 Diffuse large B cell lymphoma (DLBCL), centroblastic. Centroblasts have chromatin along the nuclear membrane and prominent nucleoli.

MOLECULAR FEATURES

The tumors are clonal for Ig gene rearrangements, and rare cases have structural changes in Bcl-6.[9] Oligonucleotide microarray analyses have shown no clear distinctions between CBL subclassified as follicular or diffuse, suggesting similar origins from follicular B cells for both. By this methodology, both are readily distinguished from the centroblastic form of MZL, MZL++.

POSTULATED CELL OF ORIGIN

Germinal center dark zone centroblasts.

DLBCL – immunoblastic

SYNONYMS

Immunoblastic lymphoma.

EPIDEMIOLOGY OF SPONTANEOUS DISEASE

Immunoblastic lymphoma comprised ~8 percent of cases in NFS.V$^+$ mice and 4 percent in CFW mice, but was not observed among AKXD strain lymphomas.[55]The average age at diagnosis in NFS.V$^+$ mice was ~400 days.

HISTOLOGIC FEATURES

These are highly aggressive tumors populated by cells with large nuclei having dispersed chromatin and large, frequently bar-shaped nucleoli (Fig. 18.10). The cells are frequently admixed with centroblasts and centrocytes [AQ1]

CH018.qxd 4/10/09 12:16 PM Page 10

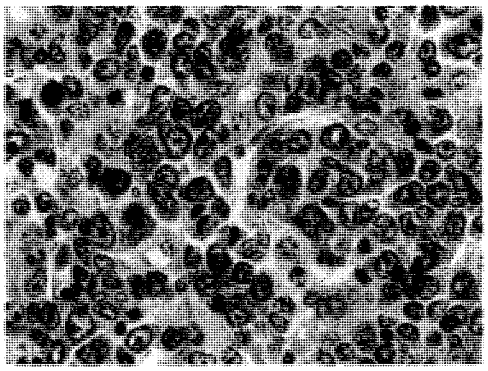


Figure 18.10 Diffuse large B cell lymphoma (DLBCL), centroblastic-immunoblastic. Central nucleoli can be seen in some of the large blast cells. Immunoblasts often have large, magenta, bar-shaped nucleoli appended to the nuclear membrane on one side.

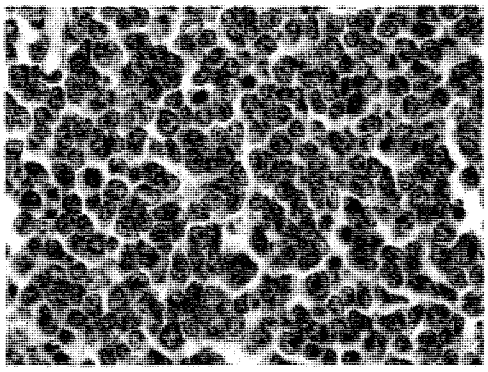


Figure 18.11 Diffuse high-grade lymphoblastic lymphoma of mature B cells. Cells are of intermediate size with scant cytoplasm and often having a prominent central nucleolus. Mitoses and apoptotic bodies are common and tingible body macrophages in many cases can give a starry-sky appearance.

reflecting their probable origin in FBL or from a post-GC immunoblast.

IMMUNOLOGIC FEATURES

By FACS, surface Ig levels are very low and they are B220^{dull}. By IHC, almost all cases are Bcl-6$^+$ and PAX-5$^+$ but IRF-8$^-$XBP-1$^-$IRF-4$^-$PU.1$^-$BLIMP$^-$.

MOLECULAR FEATURES

Clonal for Ig gene organization.

PRESUMED CELL OF ORIGIN

Post-GC immunoblast.

DLBCL – histiocyte associated

SYNONYMS

Diffuse large cell lymphoma, histiocyte associated (DLCL[HS]).

EPIDEMIOLOGY OF SPONTANEOUS DISEASE

In AKXD RI strains, ~20 percent of lymphomas were HA, while only ~1 percent of NFS.V$^+$ lymphomas were diagnosed as such.[55] The average age at diagnosis for the AKXD RI strains was ~500 days. The disease is uncommon in commonly used strains and stocks of mice.

HISTOLOGIC FEATURES

All mice present with splenomegaly associated with a marked expansion of macrophages (histiocytes) with frothy eosinophilic cytoplasm and, usually, quite limited numbers of lymphocytes. The histiocytes often occupy the entire white pulp, obliterating the PALS and normal follicular structure, leaving patches of B cells pushed to the periphery. Most often, the lymphocytes have features characteristic of FBL or CBL although rare cases with features of MZL, small B cell lymphoma (SBL), or IBL have been observed. Lymphadenopathy was seen in about half of the cases and was associated with diffuse histiocytic infiltration.

IMMUNOLOGIC FEATURES

The histiocyte population is frequently positive for EMR1 (F4/80) and LGALS3 (Mac-2) by IHC, with the B cell populations clearly delineated by exclusive expression of PAX-5.

MOLECULAR FEATURES

Frequently, the diagnosis can be made only after evaluating Ig gene organization and IHC staining. The presence of clonal Ig gene rearrangements and PAX-5$^+$ populations of cells associated with a histologic picture dominated by histiocytes is diagnostic. The differential diagnosis is true histiocytic sarcoma. In some cases, clonal rearrangements of T cell receptor (TCR)β have also been observed.

PRESUMED CELL OF ORIGIN

Germinal center B cells for the B cell component, in most cases; tissue macrophages for the histiocytic component.

DLBCL in genetically engineered mice

T and B cell–lineage lymphomas and autoimmunity are among the most common unexpected features of transgenic

CH018.qxd 4/10/09 12:16 PM Page 11

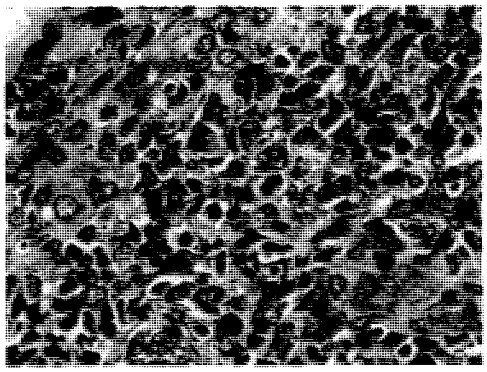


Figure 18.12 Diffuse large B cell lymphoma (DLBCL), histiocyte-associated, showing large eosinophilic histiocytes in association with large blastic B lymphoma cells and normal erythroid cells with small dark nucleoli. The B cell component is most often centroblastic or follicular, but cases with malignant immunoblasts, plasma cells, marginal zone B cells or, rarely, small B cells, are seen.

and knockout mice.[69] The B cell neoplasms appearing in individual lines of such mice may span several histologic types, ranging from FBL to CBL and IBL and histiocyte-associated DLBCL (Fig. 18.12). They may resemble the spontaneous lymphomas of older mice of commonly used strains while developing in younger animals than in wild-type mice and having a more aggressive phenotype. It is important to distinguish induced from spontaneous tumors in genetically engineered mice. The genetic manipulation may simply accelerate the development of a disease seen in wild-type mice or may induce novel diseases not seen in conventional strains and stocks.

[AQ1]

FEATURES OF GENETICALLY ENGINEERED MICE WITH DLBCL

Rizl knockout:[70] About 37 percent of −/− and 19 percent of +/− mice killed between 18 and 22 months of age had clonal B cell lymphomas with histologic and cytologic features of centroblastic DLBCL.

H2-L^d-Il6 *TG:*[71] Between 6 and 19 months of age, many TG mice developed FBL or CBL that sometimes coexisted with plasmacytomas. The lymphomas were IgM$^+$CD19$^+$ CD45R(B220)$^+$. Several of the lymphomas had t(12;15) IgH/Myc translocations detectable by polymerase chain reaction (PCR) and/or Southern blotting.

Eμ-TCL1 *TG:*[72] Transgenic mice averaging ~12 months of age presented with splenomegaly and often with lymphadenopathy due to lymphomas. The tumors were classified as FBL, DLBCL(HA), CBL, and DBLL. The neoplasms were clonal, had mutated Ig genes, and were Bcl-6$^+$.

Bad *knockouts:*[73] About 20 percent of mice homozygous for a null allele developed clonal B cell lymphomas between 18 and 24 months of age. The lymphomas were surface IgM$^+$ or IgG$^+$ and were Bcl-6$^+$.

*Iμ*BCL-6 *knockins:*[74] Mice with *Bcl-6* knocked into the IgH locus developed DLBCL and FBL beginning at 13 months of age. The lymphomas were clonal with mutated IgG genes and variably expressed some late GC markers such as IRF-4. Trisomy 15 was common.

*Eμ*Myc *knockins:*[22,23] Mice with *Myc* knocked into the IgH locus 5′ of Eμ developed DLBCL between 6 and 21 months of age. Other B cell-lineage neoplasms included FBL, plasmacytomas, and DBLL. The DLBCL were clonal and Bcl-6$^+$IgM$^+$CD45R(B220)$^+$CD19$^+$.

PRESUMED CELL OF ORIGIN

Germinal center B cells.

DIFFUSE HIGH-GRADE BLASTIC B CELL LYMPHOMA/LEUKEMIA

Definition

Diffuse high-grade lymphoblastic lymphoma/leukemia is a highly aggressive neoplasm characterized by a uniform population of medium-size B cells usually associated with a high mitotic rate and extensive apoptosis, sometimes with a leukemic phase.

Synonyms

Lymphoblastic lymphoma; Burkitt and Burkitt-like lymphoma; DLBCL of lymphoblastic lymphoma subtype (DLBCL[LL]).

Epidemiology of spontaneous disease

Diffuse high-grade lymphoblastic lymphoma/leukemia accounted for ~20 percent of cases in NFS.V$^+$ mice and ~30 percent in AKXD RI strains and CFW mice.[55] The average age at diagnosis for NFS.V$^+$ mice was ~280 days and for AXKD RI strains was ~350 days. This disease is not frequent among commonly used inbred strains and stocks.

Epidemiology of disease in genetically engineered mice

Eμ-Myc *TG:* All TG mice with a particular construct[16] died with clonal lymphoma by 84 weeks of age, with a median time to diagnosis of ~90 days, while mice with a different construct[18] died with clonal lymphoma, with a mean time to diagnosis of ~115 days.

IgH/c-myc YAC and EΔIgH/c-myc YAC TG:[20,21] All mice died with clonal lymphoma before 20 weeks of age.

CH018.qxd 4/10/09 12:16 PM Page 12

λ-MYC TG:[19] All mice died with clonal lymphoma with an average age at diagnosis of ~125 days.

Histologic features

Cases present with lymphadenopathy and variable involvement of the spleen. Affected tissues are infiltrated with lymphoblasts: uniform cells of medium size with little cytoplasm, moderately dispersed chromatin, and indistinct nucleoli. These are associated with many mitotic figures and often with large numbers of tingible body macrophages ingesting apoptotic cells and leading to a starry-sky appear-
[AQ1] ance (Fig. 18.11). In lymph nodes, early infiltration of the deep cortex progresses to total replacement of normal cells, with growth outside the capsule and into the fat. The spleen, when involved, exhibits diffuse infiltration of the red and white pulp. Perivascular and peribronchial infiltrates of the lungs and periportal live infiltrates are common.

Immunologic features

Diffuse high-grade lymphoblastic lymphoma/leukemia has a spectrum of surface phenotypes ranging from patterns similar to those of normal immature or transitional B cells – $IgM^+IgD^-CIQR1(AA4.1)^+$ – to that of cells that are Ig class switched and mutated for Ig V-region sequences. The less mature phenotype is characteristic of most lymphomas of E*μ-Myc*, IgH/*c-myc* YAC, EΔIgH/c-myc YAC, and λ-MYC transgenic mice. More mature phenotypes are characteristic of many spontaneous DBLL of NFS.V^+ mice and some lymphomas of E*μ-TCL1* TG and E*μMyc* knockin mice along with many other genetically engineered mice.

Molecular features

The lymphomas are clonal for IgH and IgL rearrangements. Structural rearrangements of cellular genes, most due to proviral insertions, were found in pooled studies of NFS.V^{+3} and AKXD RI^6 lymphomas for *Zfp521* (*Evi3*) (11.9 percent), *Pim1* (5.6 percent), *Evi1* (4.8 percent), and *Myc* (0.8 percent). Lymphomas of λ-MYC were characterized by chromosomal instability and frequent biallelic deletions of *Cdkn2a* (p16).[19] Lymphomas of E*μ-Myc* mice had frequent changes in the $p19^{ARF}$-MDM2-p53 tumor suppressor axis.[22]

Presumed cell of origin

Immature or transitional B cells for E*μ-Myc*, IgH/*c-myc* YAC, EΔIgH/c-myc YAC, and λ-MYC transgenic mice. Probable GC or early post-GC cells for those with features similar to the DBLL of E*μ-TCL1* transgenic mice.

Comments

B cell-lineage lymphomas with lymphoblastic cytology but distinct from precursor B lymphoblastic neoplasms are seen at low to high frequency in many strains of genetically engineered mice and a number of conventional inbred strains. The lymphomas of λ-MYC transgenic mice were previously designated Burkitt lymphoma,[2] but the findings that Ig genes are not mutated and that they have a surface phenotype of transitional or immature B cells indicate that they differ from most human Burkitt lymphoma cases, suggesting that another designation is warranted. Mouse cases with similar histology and cytology occurring in mice other than the λ-MYC transgenic mice were previously designated Burkitt-like.[2] The findings that they rarely have structural alterations in *Myc* and do not overexpress *Myc* distinguish them from human Burkitt-like lymphomas.[1]

KEY POINTS

- Mouse B cell-lineage lymphomas comprise a spectrum of tumor types that can be distinguished through studies of their histologic, immunophenotypic, and molecular features and their relationship to a presumed cell of origin.
- Most detailed information on these lymphomas has come from studies of a limited number of mouse strains that express murine leukemia viruses at high levels, with the viruses acting as insertional mutagens.
- Some mouse lymphoma classes identified in high-virus mice exhibit many similarities to distinct lymphoma types in humans, whereas others appear to be species specific. There are many human B cell-lineage neoplasms that have no known counterpart in mice.
- Efforts to model human lymphomas in the mouse through genetic engineering have generally failed to produce accurate replicas but have generated invaluable information on the functions of oncogenes and tumor suppressor genes in normal B cell biology and their contributions to the transformed state in mice.
- The true utility of any mouse model will only be known when it has been rigorously dissected using the combined powers of histologic, immunophenotypic, genetic, and epigenetic analyses.

ACKNOWLEDGMENTS

This work was supported in part by the Intramural Research Program of the NIH, National Institute of Allergy and Infectious Diseases.

CH018.qxd 4/10/09 12:16 PM Page 13

REFERENCES

● = Key primary paper
◆ = Major review article

◆1. Jaffe ES, Harris NL, Stein H, Vardiman JW. World Health Organization classification of tumors. Pathology and genetics. Tumors of haematopoietic and lymphoid tissues. *Lyon: IARC Press*, 2001.

◆2. Morse HC, Anver MR, Fredrickson TN *et al.* Bethesda proposals for classification of lymphoid neoplasms in mice. *Blood* 2002; **100**: 246–58.

●3. Hartley JW, Chattopadhyay SK, Lander MR *et al.* Accelerated appearance of multiple B cell lymphoma types in NFS/N mice congenic for ecotropic murine leukemia viruses. *Lab Invest* 2000; **80**: 159–69.

4. Mucenski ML, Taylor BA, Jenkins NA. AKXD recombinant inbred strains: models for studying the molecular genetic basis of murine lymphomas. *Mol Cell Biol* 1986; **6**: 4236–43.

5. Gilbert DJ, Neuman PE, Taylor BA, *et al.* Susceptibility of AKXD recombinant inbred mouse strains to lymphomas. *J Virol* 1993; **67**: 2083–90.

●6. Morse HC, Qi CF, Chattopadhyay SK *et al.* Combined histiologic and molecular features reveal previously unappreciated subsets of lymphoma in AKXD recombinant inbred mice. *Leuk Res* 2001; **25**: 719–33.

●7. Suzuki T, Shen HF, Akagi K *et al.* New genes involved in cancer identified by retroviral tagging. *Nat Genet* 2002; **32**: 166–74.

8. Potter M, Wiener F. Plasmacytomagenesis in mice – model of neoplastic development dependent upon chromosomal translocations. *Carcinogenesis* 1992; **13**: 1681–97.

9. Qi CF, Hori M, Coleman AE *et al.* Genomic organisation and expression of BCL6 in murine B-cell lymphomas. *Leuk Res* 2000; **24**: 719–32.

10. Gordon CJ, Grafton G, Wood PM *et al.* Modelling the human immune response: can mice be trusted? *Curr Opin Pharmacol* 2001; **1**: 431–8.

◆11. Mestas J, Hughes CCW. Of mice and not mend. Differences between mouse and human immunology. *J Immunol* 2004; **172**: 2731–8.

12. Oliver AM, Martin F, Kearney JF. Mouse Cd38 is down-regulated on germinal center B cells and mature plasma cells. *J Immunol* 1997; **158**: 1108–15.

◆13. Hahn WC, Weinberg RA. Modelling the molecular circuitry of cancer. *Nat Rev Cancer* 2002; **2**: 331–40.

14. Smogorzewska A, deLange T. Different telomere damage signaling pathways in human and mouse cells. *EMBO J* 2002; **21**: 4338–48.

15. Hooper ML. Tumor suppressor gene mutations in humans and mice: parallels and contrasts. *EMBO J* 1998; **17**: 6783–9.

●16. Adams JM, Harris AW, Pinkert CA *et al.* The c-*myc* oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature* 1985; **318**: 533–8.

17. Yukawa K, Kikutani H, Inomoto T *et al.* Strain dependency of B and T lymphoma development in immunoglobulin heavy chain enhancer (Eμ)-*myc* transgenic mice. *J Exp Med* 1989; **170**: 711–26.

18. Schmidt EV, Pattengale PK, Weir Leder P. Transgenic mice bearing the human c-myc gene activated by an immunoglobulin enhancer: a pre-B-cell lymphoma model. *Proc Natl Acad Sci U S A* 1988; **85**: 6047–51.

19. Kovalchuk AL, Qi CF, Torrey TA *et al.* Burkitt lymphoma in the mouse. *J Exp Med* 2000; **192**: 1183–90.

20. Butzler C, Zou X, Popov AV, Bruggemann M. Rapid induction of B-cell lymphomas in mice carrying a human IgH/c-*myc* YAC. *Oncogene* 1997; **14**: 1383–8.

21. Palomo C, Zou X, Nicholson IC *et al.* B-cell tumorigenesis in mice carrying a yeast artificial chromosome-based immunoglobulin heavy/c-*myc* translocus is independent of the heavy chain intron enhancer (Eμ). *Cancer Res* 1999; **59**: 5625–8.

22. Park SS, Kim JS, Han SS *et al.* Insertion of Myc into IgH induces B-cell lymphomas and plasmacytomas in mice. *Cancer Res* 2005; **65**: 1306–15.

23. Zhu D, Qi CF, Morse HC *et al.* Deregulated expression of the Myc cellular oncogene drives development of mouse 'Burkitt-like' lymphomas from naive B cells. *Blood* 2005; **105**: 2135–7.

24. William J, Euler C, Christensen S, Shlomchik MJ. Evolution of autoantibody responses via somatic hypermutation outside of germinal centers. *Science* 2002; **297**: 2066–70.

25. Yamada Y, Matsushiro H, Ogawa MS *et al.* Genetic predisposition to pre-B lymphomas in SI Kh strain mice. *Cancer Res* 1994; **54**: 403–7.

26. Hiai H, Tsuruyama T, Yamada Y. Pre-B lymphomas in SI/Kh mice: a multifactorial disease model. *Cancer Sci* 2003; **94**: 847–50.

27. Elefanty AG, Hariharan IK, Cory S. BCR-ABL, the hallmark of chronic myeloid-leukemia in man, induces multiple hematopoietic neoplasms in mice. *EMBO J* 1990; **9**: 1069–78.

28. Daley GQ, Vanetten RA, Baltimore D. Induction of chronic myelogenous leukemia in mice by the p210bcr/ABL gene of the Philadelphia chromosome. *Science*. 1990; **247**: 824–30.

29. Kelliher MA, McLaughlin J, Witte ON, Rosenbergn N. Induction of a chronic myelogenous leukemia-like syndrome in mice with V-ABL and BCR/ABL. *Proc Natl Acad Sci U S A* 1990; **87**: 6649–53.

●30. Heisterkamp N, Jenster G, Tenhoeve J *et al.* Acute leukemia in BCR/ABL transgenic mice. *Nature* 1990; **344**: 251–3.

31. Voncken JW, Kaartinen V, Pattengale PK *et al.* BCR/ABL p210 and p190 cause distinct leukemia in transgenic mice. *Blood* 1995; **86**: 4603–11.

32. Hariharan IK, Harris AW, Crawford M *et al.* A BCR-V-ABL oncogene induces lymphomas in transgenic mice. *Mol Cell Biol* 1989; **9**: 2798–805.

33. Castellanos A, Pintado B, Weruaga E *et al.* A BCR-ABL^{p190} fusion gene made by homologous recombination causes B-cell acute lymphoblastic leukemias in chimeric mice

CH018.qxd 4/10/09 12:16 PM Page 14

with independence of the endogenous BCR product. *Blood* 1997; **90**: 2168–74.

34. Verbeek S, Vanlohuizen M, Vandervalk M *et al.* Mice bearing the E-μ-Myc and E-μ-Pim-1 transgenes develop pre-B-cell leukemia prenatally. *Mol Cell Biol* 1991; **11**: 1176–9.
35. Malek SN, Dordai DI, Reim J *et al.* Malignant transformation of early lymphoid progenitors in mice expressing an activated Blk tyrosine kinase. *Proc Natl Acad Sci U S A* 1998; **95**: 7351–6.
36. Lowell CA, Soriano P. Knockouts of Src-family kinases: stiff bones, wimpy T cells, and bad memories. *Genes Dev* 1996; **10**: 1845–57.
37. Golub TR, Barker GF, Bohlander SK *et al.* Fusion of the Tel gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A* 1995; **92**: 4917–21.
38. Romana SP, Poirel H, Leconiat M *et al.* High-frequency of t(12-21) in childhood B-lineage acute lymphoblastic leukemia. *Blood* 1995; **86**: 4263–9.
39. Romana SP, Mauchauffe M, Leconiat M *et al.* The t(12-21) of acute lymphoblastic leukemia results in a TEL-AML1 gene fusion. *Blood* 1995; **85**: 3662–70.
40. Hiebert SW, Sun WH, Davis JN *et al.* The t(12;21) translocation converts AML-1B from an activator to a repressor of transcription. *Mol Cell Biol* 1996; **16**: 1349–55.
41. Fears S, Gavin M, Zhang DE *et al.* Functional characterization of ETV6 and ETV6/CBFA2 in the regulation of the MCSFR proximal promoter. *Proc Natl Acad Sci U S A* 1997; **94**: 1949–54.
42. Bernardin F, Yang YD, Cleaves R *et al.* TEL-AML1, expressed from t(12;21) in human acute lymphocytic leukemia, induces acute leukemia in mice. *Cancer Res* 2002; **62**: 3904–8.
43. Wang JH, Avitahl N, Cariappa A *et al.* *Aiolos* regulates B cell activation and maturation to effector state. *Immunity* 1998; **9**: 543–53.
44. Liippo J, Nera KP, Veistinen E *et al.* Both normal and leukemic B lymphocytes express multiple isoforms of the human *Aiolos* gene. *Eur J Immunol* 2001; **31**: 3469–74.
45. Schmitt C, Tonnelle C, Dalloul A *et al.* *Aiolos* and Ikaros: regulators of lymphocyte development, homeostasis and lymphoproliferation. *Apoptosis* 2002; **7**: 277–84.
46. Serrano M, Lee HW, Chin L *et al.* Role of the INK4a locus in tumor suppression and cell mortality. *Cell* 1996; **85**: 27–37.
47. Lowe SW, Sherr CJ. Tumor suppression by INK4a-ARF: progress and puzzles. *Curr Opin Genet Devel* 2003; **13**: 77–83.
48. Takeuchi S, Bartram CR, Seriu T *et al.* Analysis of a family of cyclin-dependent kinase inhibitors – p15/MTS2/INK4b, p16/MTS1/INK4a, and p18 genes in acute lymphoblastic leukemia of childhood. *Blood* 1995; **86**: 755–60.
49. ◆ French SW, Dawson DW, Miner MD *et al.* DNA methylation profiling: a new tool for evaluating hematologic malignancies. *Clin Immunol* 2002; **103**: 217–30.
50. Samaridis J, Casorati G, Traunecker A *et al.* Development of lymphocytes in interleukin 7-transgenic mice. *Eur J Immunol* 1991; **21**: 453–60.
51. Uehira M, Matsuda H, Hikita I *et al.* The development of dermatitis infiltrated by $\gamma\Delta$ T-cells in IL-7 transgenic mice. *Int Immunol* 1993; **5**: 1619–27.
52. Rich BE, Campos Torres J, Tepper RI *et al.* Cutaneous lymphoproliferation and lymphomas in interleukin-7 transgenic mice. *J Exp Med* 1993; **177**: 305–16.
53. Fisher AG, Burdet C, Bunce C *et al.* Lymphoproliferative disorders in IL-7 transgenic mice – expansion of immature B cells which retain macrophage potential. *Int. Immunol* 1995; **7**: 415–23.
54. Wong S, Witte ON. Modeling Philadelphia chromosome positive leukemias. *Oncogene* 2001; **20**: 5644–59.
55. ◆ Morse HC III, McCarty T, Qi CF *et al.* B lymphoid neoplasms of mice: characteristics of naturally occurring and engineered diseases and relationships to human disorders. *Adv Immunol* 2004; **81**: 97–121.
56. Bichi R, Shinton SA, Martin ES *et al.* Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression. *Proc Natl Acad Sci U S A* 2002; **99**: 6955–60.
57. Zapata JM, Krajewska M, Morse HC *et al.* TNF receptor-associated factor (TRAF) domain and BCL-2 cooperate to induce small B cell lymphoma/chronic lymphocytic leukemia in transgenic mice. *Proc Natl Acad Sci U S A* 2004; **101**: 16600–5.
58. Wen XS, Zhang DQ, Kikuchi Y *et al.* Transgene-mediated hyper-expression of IL-5 inhibits autoimmune disease but increases the risk of B cell chronic lymphocytic leukemia in a model of murine lupus. *Eur J Immunol* 2004; **34**: 2740–9.
59. Yumoto T, Yoshida Y, Yoshida H *et al.* Prelymphomatous and lymphomatous changes in splenomegaly of New Zealand black mice. *Acta Pathol Jpn* 1980; **30**: 171–86.
60. Fredrickson TN, Lennert K, Chattopadhyay SK *et al.* Splenic marginal zone lymphomas of mice. *Am J Pathol* 1999; **154**: 805–12.
61. Ward JM, Tadesse-Heath L, Perkins SN *et al.* Splenic marginal zone B-cell and thymic T-cell lymphomas in p53-deficient mice. *Lab Invest* 1999; **79**: 3–14.
62. Shin MS, Fredrickson TN, Hartley JW *et al.* High-throughput retroviral tagging for identification of genes involved in initiation and progression of mouse splenic marginal zone lymphomas. *Cancer Res* 2004; **64**: 4419–27.
63. Morse HC III, Kearney JF, Isaacson PG *et al.* Cells of the marginal zone – origins, function and neoplasia. *Leuk Res* 2001; **25**: 169–78.
64. McDonnell TJ, Korsmeyer SJ. Progression from lymphoid hyperplasia to high-grade malignant lymphoma in mice transgenic for the t(14,18). *Nature* 1991; **349**: 254–6.
65. Strasser A, Harris AW, Cory S. Eμ-BCL-2 transgene facilitates spontaneous transformation of early pre-B and immunoglobulin-secreting cells but not T cells. *Oncogene* 1993; **8**: 1–9.
66. Zhou P, Levy NB, Xie HY *et al.* MCL1 transgenic mice exhibit a high incidence of B-cell lymphoma manifested as a spectrum of histologic subtypes. *Blood* 2001; **97**: 3902–9.

CH018.qxd 4/10/09 12:16 PM Page 15

67. Egle A, Harris AW, Bath ML *et al.* VavP-Bcl2 transgenic mice develop follicular lymphoma preceded by germinal center hyperplasia. *Blood* 2004; **103**: 2276–83.
68. Fredrickson TN, Harris AW. Atlas of mouse hematopathology. Amsterdam: Harwood Academic, 2000.
69. Taddesse-Heath L, Morse HC III. Lymphoma in genetically engineered mice. In: *Pathology of genetically engineered mice.* Ames: Iowa State University Press 2000: 365–81.
70. Steele-Perkins G, Fang W, Yang XH *et al.* Tumor formation and inactivation of *Riz1*, an Rb-binding member of a nuclear protein-methyltransferase superfamily. *Genes Dev* 2001: **15**: 2250–62.
71. Kovalchuk AL, Kim JS, Park SS *et al.* IL-6 transgenic mouse model for extraosseous plasmacytoma. *Proc Natl Acad Sci U S A* 2002; **99**: 1509–14.
●72. Hoyer KK, French SW, Turner DE *et al.* Dysregulated TCL1 promotes multiple classes of mature B cell lymphoma. *Proc Natl Acad Sci U S A* 2002; **99**: 14392–7.
73. Ranger AM, Zha JP, Harada H *et al.* *Bad*-deficient mice develop diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A* 2003; **100**: 9324–9.
●74. Cattoretti G, Pasqualucci L, Ballon G *et al.* Deregulated Bcl6 expression recapitulates the pathogenesis of human diffuse large B cell lymphomas in mice. *Cancer Cell* 2005; **7**: 445–55.

Detail of Query

Figure call-outs have been added through the text (marked with [AQ1]); please ensure that these are all located in the required position

NOTE Based on text order, should Figures 18.11 and 18.12 be renumbered so that the image currently labeled 18.12 appears with text for DLBCL, histiocyte-associated; then the image currently labeled 18.11 appears last, close to text on diffuse high-grade blastic B cell lymphoma/leukemia?

Expert Report
Christopher J. Portier, Ph.D.

Charge

Glyphosate acid is a colorless, odorless, crystalline solid. Glyphosate is the term used to describe the salt that is formulated by combining the deprotonated glyphosate acid and a cation (isopropylamine, ammonium, or sodium). This expert report is intended to review the available scientific evidence relating to the potential of glyphosate and glyphosate-based formulations (GBFs), including Roundup®, to cause Non-Hodgkin's Lymphoma (NHL) in humans.

Qualifications

I received an undergraduate degree in mathematics in 1977 from Nicholls State University and a Master's degree and Ph.D. in biostatistics from the University of North Carolina School of Public Health in 1979 and 1981 respectively. My Ph.D. thesis addressed the optimal way to design a two-year rodent carcinogenicity study to assess the ability of a chemical to cause cancer[1, 2]; the optimal dosing pattern from my thesis is still used by most researchers. My first employment following my doctoral degree was a joint appointment at the National Institute of Environmental Health Sciences (NIEHS) and the National Toxicology Program (NTP) to conduct research on the design and analysis of experiments generally employed in toxicology. After 5 years with NIEHS/NTP, I developed my own research group which eventually became the Laboratory of Quantitative and Computational Biology and then the Laboratory of Computational Biology and Risk Assessment (LCBRA). One highlight during this period was the development of the Poly-3 Test for survival adjustment of data from two-year carcinogenicity studies in rodents[3, 4]; this test is used as the main method of analysis of these studies by the NTP and many others. We also did a complete analysis of the historical controls animals from the NTP studies[5, 6]. The LCBRA focused on the application of computational tools to identify chemicals that are toxic to humans, to develop tools for understanding the mechanisms underlying those toxicities and to quantify the risks to humans associated with these toxicities. The main toxicological focus of the LCBRA was cancer and my laboratory developed many methods for applying multistage models to animal cancer data and implemented the use of these models in several experimental settings[7-19]. In my last few years at the NIEHS/NTP, my research focus expanded to the development of tools for evaluating the response of complex experimental and human systems to chemicals[20-24] and the name of the laboratory shifted to Environmental Systems Biology.

Over my 32 years with the NIEHS/NTP, I was involved in numerous national priority issues that went beyond my individual research activities. After Congress asked NIEHS to work with the Vietnamese government to address the hazards associated with Agent Orange use during the Vietnamese War, I was given the responsibility of working with

EXHIBIT
15-30
9/5/17

my counterparts in Vietnam to build a research program in this area[25]. Congress also tasked NIEHS with developing a research program (EMF-RAPID) to address concerns about the risks to humans from exposure to power lines and to report back to Congress on what we found. I was in charge of evaluating all research developed under this program and was responsible for the final recommendations to Congress on this issue[26-28].

While at the NIEHS/NTP, I also had administrative positions that relate to my qualifications. From 2000 to 2006 I was the Director of the Environmental Toxicology Program (ETP) at NIEHS. The ETP included all of the toxicology research laboratories within the NIEHS Intramural Research Program. It was my responsibility to ensure the research being done was pertinent to the mission of the NIEHS, addressing high priority concerns about toxic substances and human health and that the NIEHS had adequate resources to complete this research.

During this time I was also Associate Director of the NTP, a position in which I was the scientific and administrative director of the NTP (The Director of the NTP was also the NIEHS Director and gave me complete autonomy in the management and science of the NTP). These two positions were historically always combined at the NIEHS and the NTP so that one person was in charge of all toxicological research at the NIEHS/NTP. The NTP is the world's largest toxicology program, routinely having 15 to 25 active two-year carcinogenicity studies, numerous genetic toxicology studies and many other toxicological studies being conducted at any given time. The NTP two-year carcinogenicity studies and their technical reports are also considered the "gold standard" of cancer studies due to their extreme high quality, their tremendous utility in evaluating human health hazards and the rigor and transparency they bring to the evaluation of the data. All data from NTP two-year cancer studies are publicly available including data on individual animals and images from the pathology review of each animal. The NTP is also home to the Report on Carcinogens, the US Department of Health and Human Services official list of what is known or reasonably anticipated to be carcinogenic to humans. It was my responsibility to decide what items eventually went onto this list while I was Associate Director of the NTP. In 2006, I became an Associate Director of the NIEHS, a senior advisor to the director and the director of the Office of Risk Assessment Research (ORAR). ORAR focused on stimulating new research areas on the evaluation of health risks from the environment and addressed major risk assessment issues on behalf of the NIEHS/NTP. For example, in this capacity, I lead a multiagency effort to understand the health risks to humans from climate change and to develop a research program in this area[29].

I left the NIEHS/NTP in 2010 to become the Director of the National Center for Environmental Health (NCEH) at the Centers for Disease Control and Prevention and simultaneously Director of the Agency for Toxic Substances and Disease Registry (ATSDR). NCEH does research and supports activities aimed at reducing the impact of environmental hazards on public health. One well-respected research effort of the NCEH is the National Biomonitoring Program. This program tests for the presence of hundreds of chemicals in human blood and urine in a national sample of people in the

United States. ATSDR advices the Environmental Protection Agency (EPA) and communities on the potential health impacts from toxic waste dump sites (superfund sites). ATSDR is required by law to produce ToxProfiles. These are comprehensive reviews of the scientific literature for specific chemicals generally found at superfund sites. They also provide an assessment of the safety of these chemicals. As part of my activities at ATSDR, I began a modernization of the ToxProfiles to use systematic review methods in their assessments; this effort was linked to a similar effort that I had helped to implement at the NIEHS/NTP.

Aside from my official duties in my various federal jobs, I also served on numerous national and international science advisory panels. Most notable, for my qualifications for this statement, are my serving as Chair from 2005 to 2010 of the Subcommittee on Toxics and Risk of the President's National Science and Technology Council, member and chair of EPA'S Science Advisory Panel from 1998 to 2003 (focused specifically on advising their pesticides program) and chair of the International Agency for Research on Cancer (IARC) advisory group that updated and improved its rules for reviewing scientific data to ensure that conclusions on the carcinogenicity of human exposures are the best possible (Preamble)[30]. As part of my work on science advisory panels, I have served on EPA's Science Advisory Board, as an advisor to the Australian Health Council on risk assessment methods, as an advisor to the Korean Food and Drug Administration on toxicological methods, and served on several World Health Organization (WHO) International Program on Chemical Safety scientific panels dealing with risk assessment. Besides the guidelines for evaluating cancer hazards used by the IARC, I have either chaired or served as a member of scientific panels developing guidance documents for other organizations including the EPA.

I have received numerous awards, most notably the Outstanding Practitioner Award from the International Society for Risk Analysis and the Paper of the Year Award (twice) from the Society of Toxicology Risk Assessment Specialty Section. I am a fellow of the American Statistical Association, the International Statistical Institute, the World Innovation Foundation and the Ramazinni Institute. I have published over 250 peer-reviewed scientific papers, book chapters and technical documents on topics in toxicology and risk assessment.

Finally, I have served on numerous national and international committees tasked with evaluating the risk and/or hazard of specific environmental chemicals, including glyphosate. For example, I have contributed to risk assessments for EPA, the Food and Drug Administration, the Centers for Disease Control and Prevention, the National Institutes of Health, the WHO and IARC.

Reliance List

During the course of my preparation for this report, I have reviewed the following materials:

a. All epidemiological data relating to the ability of glyphosate formulations to cause NHL in humans.

b. Scientific papers on the cellular origins of NHL
c. Peer-reviewed scientific data relating to the carcinogenicity, genotoxicity and oxidative stress caused by glyphosate
d. Technical reports relating to the carcinogenicity of glyphosate provided by the defendant to the lawyers for the plaintiff
e. The USEPA, the European Food Safety Authority (EFSA), the German Federal Institute for Risk Assessment, the European Chemical Agency, the IARC and the WHO/Food and Agriculture Organization Joint Meeting on Pesticide Residues reviews of the scientific literature relating to the potential for glyphosate to cause cancer.
f. Technical documents available from EFSA regarding animal carcinogenicity data on glyphosate prepared by organizations other than the defendant
g. Various other documents produced in the litigation

A complete list of my reliance materials is at the end of this report.

Methodology for Causality Evaluation

The evaluation of whether glyphosate and/or GBFs can cause NHL in humans requires the review and synthesis of scientific evidence from studies of human populations (epidemiology), animal cancer studies, and studies investigating the mechanisms through which chemicals cause cancer. Many different approaches[31, 32] are used to synthesize these three areas of science to answer the question "Does this chemical cause cancer in humans?" In any of these three science areas, the quality of the individual studies has to be assessed and summarized to make certain the studies included in the overall assessment are done appropriately. Once the quality of the individual studies has been assessed, a judgment needs to be made concerning the degree to which the studies support a finding of cancer in humans. To do this, the EPA, IARC, the European Chemical Agency (EChA), the US Report on Carcinogens, and many others use guidelines[30, 33-35] that rely upon aspects of the criteria for causality developed by **Hill (1965)**[36].

Hill listed nine (9) aspects of epidemiological studies and the related science that one should consider in assessing causality. The presence or absence of any of these aspects is neither sufficient nor necessary for drawing inferences of causality. Instead, the nine aspects serve as means to answer the question of whether other explanations are more credible than a causal inference. As noted by Hill:

> *"None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a sine qua non. What they can do, with greater or less strength, is to help us to make up our minds on the fundamental question — is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?"*

The nine aspects cited by Hill include consistency of the observed association, strength

of the observed association, biological plausibility, biological gradient, temporal relationship of the observed association, specificity of the observed association, coherence, evidence from human experimentation and analogy. These are briefly described below.

An inference of causality is strengthened when several of the studies show a **consistent positive association** between cancer and the exposure. This addresses the key issue of replication of studies which is critical in most scientific debates. If studies are discordant, differences in study quality, potential confounding, potential bias and statistical power are considered to better understand that discordance.

An inference of causality is strengthened when the **strength of the observed association** in several studies are large and precise. These large, precise associations lessen the possibility that the observed associations are due to chance or bias. A small increase in risk of getting cancer does not preclude a causal inference since issues such as potency and exposure level may reduce the ability of a study to identify larger risks. Meta-analyses provide an objective evaluation of the strength of the observed association across several studies with modest risks to help clarify strength of the observed associations.

An inference of causality is strengthened when there is data supporting **biological plausibility** demonstrated through experimental evidence. Animal carcinogenicity studies, in which tumor incidence is evaluated in experimental animals exposed to pure glyphosate, play a major role in establishing biological plausibility. There are numerous types of mechanisms that can lead to cancer[37], most of which can be demonstrated through experimental studies in animals, human cells, animal cells, and/or other experimental systems. Occasionally, occupational, accidental or unintended exposures to humans allow researchers to evaluate mechanisms using direct human evidence.

An inference of causality is strengthened when there is a **biological gradient** showing a reasonable pattern of changing risk with changes in exposure (e.g. risk increases with increasing exposure or with longer exposure). In many epidemiological studies, this aspect cannot be examined due to limitations in the study design or due to a lack of clarity in the presentation of the results. When a study does address an exposure-response relationship, failure to find a relationship can be due to a small range of exposures, insufficient sample size or a changing exposure magnitude over time that has not been accounted for.

An inference of causality is strengthened when there is a **temporal relationship** in which the exposure comes before the cancer. This aspect is necessary to show causality; if it is not present, a causal inference is not plausible. Because the latency period for cancers can be long (years), evaluation of studies should consider whether the exposure occurred sufficiently long ago to be associated with cancer development.

An inference of causality is strengthened when the exposure is **specific** for a given cancer. This would mean that the disease endpoint being studied is only due to the cause being assessed. This issue is seldom applicable and, since NHL has other causes, specificity is not applicable to the determination of causality for glyphosate.

An inference of causality is strengthened when other lines of experimental evidence are **coherent** with a causal interpretation of the association seen in the epidemiological evidence. To evaluate coherence, information from animal carcinogenicity studies, mechanistic investigations and information on the metabolism of the chemical being studied would be considered.

An inference of causality is strengthened when there is **experimental evidence in humans** supporting a causal interpretation. Seldom is this type of information available when addressing the toxicity of chemicals. However, experiments in which an individual reduces or limits exposures and the risk of cancer is reduced would carry considerable weight in the evaluation (e.g. studies evaluating the cancer risks of people who stop cigarette smoking compared with continuing smoking have demonstrated reduced lung cancer risks). No such data are available for glyphosate.

Finally, an inference of causality is strengthened when there are other chemical agents with **analogous** structures showing similar effects in humans and/or animals and/or showing similar biological impacts in mechanistic studies. No such data are available for glyphosate.

The most logical approach to developing an inference of causality is to step through each of the aspects of causality developed **by Hill (1965)**[36] and apply them to the available data for glyphosate and for glyphosate formulations. This is done in the sections that follow.

Consistency of the Associations seen in Human Epidemiological Studies

Relevant Epidemiology Studies

In their meta-analysis, **Chang and Delzell (2016)**[38] performed a systematic literature search of all scientific literature up to June, 2015, to identify all epidemiological studies that were pertinent to evaluating an association between glyphosate and NHL. They identified 12 relevant epidemiology studies[39-50]. Their search agrees with all current reviews of glyphosate and I will use their findings from the literature up until 2015. To cover from June 2015 to the present (April 1, 2017), I used their searching algorithm and identified 117 additional published studies, none of which were new epidemiology studies. These same 12 studies will be considered for use in this evaluation. Other experts will be discussing the studies as well as their strengths and their weaknesses; I will focus on using the results of these studies in evaluating causality so I will only briefly describe each study.

Cantor et al. (1992)[39] did an in-person interview study comparing 622 white men, newly diagnosed with NHL, to 1245 population-based controls in Iowa and Minnesota. They originally identified 780 cases, of which 694 (89%) were interviewed. After pathology review, only 622 were found to have NHL, the remaining cases having leukemia or other diseases. Three different sources of controls were used, random digit dialing (76.7% response rate), Health Care Financing Administration rolls (79% response

rate) and deceased controls with eligible proxies (77% response rate). Both cases and controls were questioned regarding their use of agricultural products including Roundup® and any other glyphosate-based formulations. For deceased or incompetent controls (184) and cases (number not given), proxy interviews were done with a close relative. When cases in farmers were compared to cases in non-farmer controls, 26 cases (out of 266) and 49 controls (out of 547) had handled herbicides containing glyphosate yielding an odds ratio[1] (OR) of 1.1 (95% confidence interval 0.7-1.9). This analysis controlled for vital status, age, state, cigarette smoking status, family history of lymphopoietic cancer, high-risk occupations and high-risk exposures in a logistic analysis. The authors noted there was "minimal evidence for confounding of results for any single pesticide by exposure to pesticides belonging to other chemical families." Because the exposure is determined based on interviews in cases and controls, this study has the potential for recall bias[2]. However, the authors note that the bias could both increase or decrease the OR because of non-differential exposure misclassification[3] because of difficulties in accurate recall of past pesticide exposures for both controls and treated individuals. This study will not be included separately into the evaluation since it overlaps with **De Roos et al. (2003)**[43]

Two additional studies conducted by **Zahm et al. (1990)**[51] in Nebraska and **Hoar et al. (1986)**[52] in Kansas collected information on pesticide and herbicide use, but did not report specifically on the effects of glyphosate. **De Roos et al. (2003)**[43] pooled the data from these two studies with the data from **Cantor et al. (1992)**[39] to examine pesticide exposure to glyphosate in farming as risk factors for NHL. The three case-control studies[39, 51, 52] had slightly different designs. The design for the Minnesota study[39] is

---

[1] The odds ratio (OR) is calculated as the proportion of exposed cases with disease to exposed controls divided by the proportion of non-exposed cases to non-exposed controls. For rare diseases, this value approximates the population risk ratio (PRR) which is the probability of having the disease in exposed individuals divided by the probability of having the disease in non-exposed individuals. If the PRR is 1, then there is no difference in the probability of having the disease regardless of your exposure. Values of PRR greater than 1 imply the risk is higher in the exposed population. Because the OR is an estimate of the PRR for rare diseases, it is usually accompanied by a 95% confidence interval that describes the probable range of the estimate. If the OR is greater than 1, then the exposure is associated with the disease. If the lower 95% confidence bound for the OR is greater than 1, this is typically used to say the association is statistically significant.

[2] Recall bias occurs when cases are more likely to say they are exposed to glyphosate than controls or when controls are more likely to say they are exposed to glyphosate than cases. The recall must be different for the cases than the controls for this to cause a bias; errors in recalling past exposures that happen for both cases and controls would not be recall bias.

[3] Non-differential exposure misclassification occurs when the probability of an error in determining whether an individual is exposed or not is the same for both cases and controls.

provided directly above. In Nebraska[51], the cases were identified through the Nebraska Lymphoma Study Group and area hospitals for 66 counties and included all white men and women diagnosed with NHL between July 1, 1983 and June 30, 1986. Controls were obtained by random-digit dialing, Medicare records or state mortality files depending upon age and vital status. All study participants were over age 21 and even though this study included a few women, they were excluded from the **De Roos et al. (2003)** analysis. The response rates for cases and controls were 91% and 87% respectively. In Kansas[52], cases were randomly sampled from a registry at the University of Kansas of white men, over age 21, diagnosed between 1979 and 1981. The response rates for cases and controls were 96% and 94% respectively. Controls were population-based matched on age and vital status. As for the Nebraska study, controls for live cases were obtained from Medicare records for cases 65+ and by random-digit dialing for cases <65 years; controls for deceased patients came from state mortality records. The resulting pooled case-control study had 870 cases and 2569 controls (for analyzing the relationship between glyphosate and NHL, there were only 650 cases and 1933 controls following exclusion of subjects with missing data). For any glyphosate exposure, there were 36 exposed cases and 61 exposed controls with an OR (95% confidence interval) of 2.1 (1.1-4.0) in a logistic regression analysis controlling for all other pesticides reported, age and study site. The authors also analyzed the data using a Bayesian hierarchical regression analysis yielding an OR (95% confidence interval) of 1.6 (0.9-2.8) controlling for the same parameters as the logistic regression. They also conducted an analysis of "potentially carcinogenic" pesticides which included glyphosate. When just one of these pesticides was used by subjects, the logistic regression OR was 1.6 (0.8-3.1), two to four pesticides yielded an OR of 2.7 (0.7 to 10.8) and when more than five were used, the OR was 25.9 (1.5-450.2) in the logistic regression analysis and 1.1 (0.8-1.7), 1.3 (0.7-2.3) and 2.0 (0.8-5.2) respectively for the Bayesian analysis. Removing glyphosate from the list of "potentially carcinogenic" pesticides yielded equivalent ORs of 1.2 for one pesticide, 1.2 for two to four pesticides and 1.1 for five or more pesticides. The authors note that the positive results seen in their study are not likely due to recall bias since there were few associations seen over the 47 pesticides they studied. Also, although some of the positive results could be due to chance, the use of the hierarchical regression analysis theoretically decreases the chance of false positive findings. In the Kansas study[52], suppliers for 110 subjects with farming experience were identified and provided information on the subjects' crops and pesticide purchases. In general, the suppliers reported less pesticide use than the subjects of the study with no consistent differences in agreement rates between cases and controls. The agreement between suppliers and subjects improved when pesticide use during the last 10 years was considered. This supports a reduced role of recall bias in these studies and a possible role of non-differential exposure misclassification. The reduced ORs when using the Bayesian analysis as compared to the logistic regression is not surprising because the authors used a non-informative prior rather than a less conservative prior. In addition, adjustment for 47 pesticides is also likely to reduce the significance of the observed ORs for pesticides that are associated with NHL as demonstrated by the analysis of "potentially carcinogenic" pesticides (this model is possibly over-parameterized since it

includes over 47 dependent variables for only 36 exposed cases; this can significantly reduce the ORs and increase the confidence bounds). This pooled case-control study is the strongest study with sufficient power (3.8% of subjects exposed) and will be included in the evaluation of causation.

Lee et al. (2004)[44] pooled data from **Zahm et al. (1990)**[51] and **Cantor et al. (1992)**[39] (previously described) to evaluate whether asthma acts as an effect modifier of the association between glyphosate exposure and NHL. Women were included in this analysis whereas **De Roos et al. (2003)**[43] excluded women. The final study published by Lee included 872 cases and 2336 controls of which 45 cases and 132 controls had been told by their doctors they had asthma. The OR of association between glyphosate and NHL in non-asthmatics was 1.4 (0.98-2.1) and 1.2 (0.4-3.3) in asthmatics when controlling for age, vital status and state (geographical location). This study completely overlaps with the study by **De Roos et al. (2003)**[43] with the exception of the inclusion of the few women in the study by **Zahm et al. (1990)**[51]. Since this study only looks at effect modification due to asthma, it does not contribute to the overall evaluation of causality and it will be excluded from further evaluations.

Nordstrom et al. (1998)[40] conducted a population-based case-control study of hairy cell leukemia (HCL); a subtype of B-cell NHL) in Sweden that included an evaluation of exposures to glyphosate. The study included 111 men with NHL reported to the Swedish Cancer Registry between 1987 and 1992 (with one patient from 1993 accidentally included). Controls (400 in total) were drawn from the National Population Registry matched for age and county with the cases. The response rates were 91% for cases (10 refused to participate out of the original 121) and 83% (84 controls refused to participate out of 484 selected). Almost all questionnaires were answered by the subject of the study (4 cases and 5 controls were answered by proxies). The study reported an OR for glyphosate exposure and HCL of 3.1 (0.8-12) controlling only for age. This study had very limited power for detecting an association because there were only four cases and five controls with glyphosate exposure (1.8% of the total study population). In addition, because they failed to adjust for other exposures, the potential for confounding in this study is greater than those presented previously. The authors noted that they attempted to minimize recall bias by only using living cases in the analysis. Also, even though matching was performed to identify the controls, this matching was not used in the final analysis. This study was later used in a pooled analysis of HCL and NHL[42] and will not be considered independently in the evaluation for causation but will be used in the context of the pooled analysis.

Hardell and Eriksson (1999)[41] conducted a population-based case-control study of all male patients older than 25 years diagnosed with NHL between 1987 and 1990 in the four most northern counties of Sweden. After excluding misdiagnosed cases, they included 442 cases of which 404 answered their questionnaire (most by proxy) for a response rate of 91%; 192 of these cases were deceased. For each living case, two male matched controls were chosen from the National Population Registry and matched on age and county. For each deceased case, two male controls were chosen from the National Registry for Causes of Death, matched for age and year of death. The response

rate for the controls was 84% (741 out of 884 identified). Study subjects were sent a detailed questionnaire and, in most cases, this was supplemented with a phone interview. A complete working history was obtained with questions regarding exposure to numerous chemicals to avoid a focus on pesticides and organic solvents, the focus of the study. Exposure was defined as at least one full day of exposure more than one year before diagnosis. For glyphosate exposure, the authors identified four cases and three controls with exposures and a univariate OR of 2.3 (0.4-13). A multivariate analysis of both glyphosate and phenoxy herbicides produced an OR of 5.8 (0.6-54). The study has limited power for detecting an effect because the exposure frequency is very low (0.6% exposed). This study was later used in a pooled analysis of HCL and NHL[42] and will not be considered independently in the evaluation for causation but will be used in the context of the pooled analysis.

Hardell et al. (2002)[42] conducted a pooled analysis of NHL and HCL by combining the studies of **Nordstrom et al. (1998)**[40] and **Hardell and Eriksson (1999)**[41]. This study fully overlaps with the previous two studies. The analysis controlling for age, study, county and vital status yielded an OR of 3.04 (1.08-8.52) based on eight exposed cases and eight exposed controls. A more extensive analysis additionally controlled for other pesticides and yielded a smaller OR of 1.85 (0.55-6.20). As for the study by **De Roos et al. (2003)**, the analysis may be over-parameterized (more than eight dependent variables with only eight exposed cases) which could lead to a reduction in the ORs and larger confidence bounds. Even with the pooled data, **Hardell et al. (2002)** had limited power to detect an effect because the exposure frequency for cases and controls was very low (1% exposed). This study is a valid case-control study and will be used in the evaluation of causality.

In a later study, **Eriksson et al. (2008)**[46] conducted a population-based case-control study where cases were identified as NHL patients aged 18-74 years diagnosed in four major hospitals in Sweden from December 1, 1999 until April 30, 2002. In total, 995 cases were identified as matching the study parameters with 910 (91%) answering the questionnaire shortly after diagnosis. All cases were classified into subgroups with 810 B-cell, 53 T-cell, and 38 unspecified lymphomas. Controls (1,108) were randomly selected from the population registry and matched on health service, region, sex and age and interviewed in several periods during the conduct of the study; 1,016 controls responded to the questionnaire (92% response rate). Study subjects were sent a detailed questionnaire and, in many cases, a phone interview followed. Exposure was defined as at least one full day of exposure more than one year before diagnosis. The univariate analysis, adjusting for age, sex and year of diagnosis (cases) or enrollment (control) yielded an OR of 2.02 (1.10-3.71) based on 29 exposed cases and 18 exposed controls. When cases and controls were divided into those with ≤10 days per year exposure and those with >10 days per year exposure, the ORs were 1.69 (0.70-4.07) and 2.36 (1.04-5.37) respectively. When diagnoses were grouped into various subtypes of NHL, the results did not change dramatically except for small lymphocytic lymphoma and chronic lymphocytic lymphoma which showed an increased OR of 3.35 (1.42-7.89). A multivariate analysis of glyphosate controlling for other agents with statistically

increased odds ratios and/or odds ratios greater than 1.5 yielded an OR of 1.51 (0.77-2.94). In a similar analysis to the multivariate analysis, latency periods of one to ten years showed an OR of 1.11 (0.24-5.08) and >10 years had an OR of 2.26 (1.16-4.40). This study was much larger than the previous Swedish studies (2.3% exposed) and, although there may have been confounding from other pesticides, this was addressed in the multivariate analysis and the latency analysis. This study is a valid case-control study and will be used in the evaluation of causality.

McDuffie et al. (2001)[50] recruited incidence cases of NHL in men 19 years or older from six Canadian provinces with a first diagnosis between September 1, 1991 and December 31, 1994. Each provincial Cancer Registry or, in the case of Quebec, hospital, had a target number of cases and ended recruitment when the case number was reached. Controls were men 19 years or older selected at random from provincial health insurance records, computerized telephone listings or voter registration lists, depending upon the province. Cases and controls were sent questionnaires with surrogates ineligible to answer the questionnaires for deceased cases or controls. Each subject who reported 10 hours per year or more of pesticide exposure and a random sample of 15% who reported less exposure were interviewed by telephone to obtain details on pesticide use. A pilot study was conducted to obtain an improved version of the telephone interview questionnaire used by **Hoar et al. (1986)**[52] and **Zahm et al. (1990)**[51] that would provide accurate pesticide exposure assessment in the form of a screening questionnaire and a telephone interview questionnaire. This was followed by a validation study (27 farmers) where the final questionnaires used to screen and include potential cases and controls were administered and the answers regarding pesticide usage showed excellent concordance with purchases through their local agrochemical supplier. The screening questionnaire was returned by 517 cases of NHL (67.1% response rate) and 1506 controls (48% response rate). Following analysis of the screening questionnaire, the telephone interview was administered to 179 cases and 456 controls to obtain more detailed exposure information. The OR for glyphosate exposure and NHL was 1.26 (0.87-1.80) stratified by age group and province of residence and the OR was 1.20 (0.83-1.74) when the analysis also controlled for significant medical variables (51 exposed cases and 133 exposed controls). An exposure-response evaluation was performed where the OR for exposure between zero to two days per year was 1.0 (0.63-1.57) and for greater than two days per year was 2.12 (1.20-3.73) with the latter group having 23 exposed cases and 36 exposed controls. This study had excellent sample size and power (8.1% of subjects exposed), but a low response rate to the screening questionnaire. Also, by adjusting for significant medical variables, this study ruled out many confounders but did not adjust for other pesticide exposures. The effort to validate the recall of pesticide usage for farmers supports a lack of recall bias in the study. This study is a valid case-control study and will be used in the evaluation of causality.

Hohenadel et al. (2011)[48] re-analyzed the data of **McDuffie et al. (2001)**[50] to specifically investigate the impact of exposure to multiple pesticides on NHL. Four cases of NHL were excluded from this evaluation following a pathology review. They reported associations with the use of glyphosate with and without malathion but not with

glyphosate overall. The OR for glyphosate (ever used) without malathion (ever used) was 0.92 (0.54-1.55) and the OR for glyphosate (ever used) with malathion (ever used) was 2.1 (1.31-3.37). **Chang and Delzell (2016)**[38] combined the ORs from the glyphosate only analysis with the glyphosate and malathion analyses using random-effects meta-analysis to get a combined OR for glyphosate of 1.4 (0.62-3.15). This study was specifically targeted to interactions of various pesticides and does not substantively contribute to an evaluation of glyphosate. Since it is a refined analysis of **McDuffie et al. (2001)**[50], it will be included in the evaluation of causation only in the context of the combined analysis provided by **Chang and Delzell (2016).**

Orsi et al. (2009)[47] conducted a hospital-based case-control study of men and women diagnosed with lymphoid neoplasms in five hospitals in France between 2000 and 2004 who were aged 20-75 years (the abstract gives the age range as 18-75 years). All diagnoses were cytologically or histologically confirmed. The evaluation only included men and questionnaires/interviews were completed by 491 cases (95.7% response rate) which included 244 cases with NHL. Controls were patients in the same hospital (mostly orthopedic or rheumatological patients) with no prior history of lymphoid neoplasms and excluding patients admitted to the hospital for cancer or a disease directly related to occupation, smoking or alcohol abuse. The controls were matched to cases by hospital and age. Of the 501 candidate controls, 456 participated (91% response). Exposure was evaluated differently for subjects who had non-occupational exposures from those who had occupational exposures. For both, the subjects had to fill out a questionnaire/interview on occupations and home gardening pesticide exposures. For those who had worked professionally as farmers or gardeners for at least 6 months, a specific agricultural occupational questionnaire/interview was administered and exposure was determined on the basis of this extra data. The OR for occupational use of glyphosate and NHL was 1.0 (0.5-2.2) with 12 exposed cases and 24 exposed controls stratified by age and center category. A further analysis was done by individual subtypes of NHL with an OR of 1.0 (0.3-2.7) for diffuse large cell lymphoma, 1.4 (0.4-5.2) for follicular lymphoma, 0.4 (0.1-1.8) for chronic lymphocytic leukemia (CLL) and 1.8 (0.3-9.3) for HCL. No separate analysis of non-occupational use of glyphosate was provided, nor does it seem specific data on glyphosate usage was ascertained for subjects who were not professional farmers or gardeners. This could lead to non-differential misclassification of exposure which could reduce the ORs of the study. Barring this, the sample size was sufficient to detect an effect (5.3% with occupational exposure) and this study will be included in the evaluation of causality.

Cocco et al. (2013)[49] evaluated data from a multi-center case-control study of lymphoid neoplasms in six European countries from 1998 to 2004. Cases included only adult patients diagnosed with lymphoma during the study period drawn from participating centers. Controls were either selected by sampling from the general population on sex, age group, and residence area (Germany, Italy), or from hospital controls matched to the patient excluding patients with cancer, infectious diseases, and immunodeficiency diseases (Czech Republic, France, Ireland, Spain). The study included 2348 lymphoma cases (88% participation) and 2462 controls (81% response rate in hospital-based controls and 52% in population-based controls). Exposures were derived using an

occupational exposure matrix developed by industrial hygienists and occupational experts from the research centers. Only 35 individuals (cases and controls not broken out) in the study were exposed to carbamates (glyphosate was grouped with the carbamates). No results were provided for NHL and the only OR provided for glyphosate was for B-cell lymphoma where the OR was 3.1 (0.6-17.1) based on four exposed cases and two exposed controls. No information was provided on the total number of cases for each type of lymphoma evaluated. This study has very limited power to evaluate an association between NHL and glyphosate and provides only information on B-cell lymphomas with very few exposed cases and controls. As has been done by most researchers evaluating these data, this study will receive very little weight in the evaluation of causality.

De Roos et al. (2005)[45] reported results on the association of glyphosate and cancer incidence from the Agricultural Health Study (AHS), a prospective cohort study in Iowa and North Carolina, which included 57,311 private and commercial applicators who were licensed to apply restricted-use pesticides at the time of enrollment. Recruitment occurred between 1993 and 1997 and cohort members were matched to cancer registry files to identify cases and the National Death Index (1999) to ascertain vital status. Incident cancers were identified from the date on enrollment until 31 December, 2001, with the average follow-up time being 6.7 years. Comprehensive use data was obtained by self-administered questionnaire for 22 pesticides, ever/never use for 28 additional pesticides, and general information on work practices. Applicators were given a second self-administered questionnaire on occupational exposures and lifestyle factors. They used three exposure metrics in their analyses: a) ever personally mixed or applied pesticides containing glyphosate; b) cumulative exposure days of use of glyphosate (years of use times days per year); and c) intensity weighted cumulative exposure days (years of use times days per year times intensity of use). Persons whose first primary tumor occurred before the time of enrollment (1074) were excluded from the analysis as were those who were lost to follow-up (298), did not provide age information (7) or information on glyphosate use (1678) leaving 54,315 subjects for inclusion. There were 92 cohort members with a diagnosis of NHL during the study period of which 77.2% had ever used glyphosate resulting in a rate ratio[4] (RR) of 1.2 (0.7-1.9) when controlling for age and an RR of 1.1 (0.7-1.9) when controlling for age, lifestyle factors, demographics and five other pesticides for which cumulative-exposure-day variables were most highly associated with glyphosate cumulative-exposure-days (2,4-D, alachlor, atrazine, metalochlor, and trifluralin) or, for chemicals with only ever/never exposure information that were most highly associated with glyphosate ever/never use (benomyl, maneb, paraquat, carbaryl and diazinon). When cumulative exposure days in exposed individuals are divided into tertiles and RRs examined using the lowest exposed tertile as

[4] The rate ratio (RR) is estimated as the incidence in the exposed population divided by the incidence in the unexposed population. Incidence is calculated as the number of events in a fixed period of time divided by the person years at risk. Unlike the OR, the RR does not require the assumption of a rare disease to serve as a good estimate of the population risk ratio (PRR).

the reference group, the RRs drop with values of 0.7 (0.4-1.4) and 0.9 (0.5-1.6) for tertiles 2 and 3 respectively controlling for demographic and lifestyle factors and other pesticides (30,699 subjects). When intensity-weighted exposure days are examined again using exposed tertile 1 as the reference group, the RRs drop with values of 0.6 (0.3-1.1) and 0.8 (0.5-1.4) for tertiles 2 and 3 intensity-weighted exposure days respectively controlling for demographic and lifestyle factors and other pesticides (30,699 subjects). Analyses are not shown for the evaluation of the exposed tertiles against never exposed because the authors felt that never exposed and exposed subjects differed in terms of socio-economic factors and other exposures like smoking[45].

This is a typical cohort study, but has some limitations in terms of its interpretation. The majority (75.5%) of subjects in the cohort reported having ever personally mixed or applied products containing glyphosate and was composed primarily of male, middle-aged, private applicators. For glyphosate, reliability of the answers by subjects on the use of glyphosate between the first and second questionnaire were evaluated in the AHS[53]: 82% agreement for whether they had ever mixed or applied glyphosate, 53% agreement on years mixed or applied, and 62% agreement on days per year mixed or applied and 62% agreement on decade first applied. They saw no differences in over versus under reporting between the two questionnaires suggesting this could lead to non-differential exposure bias and reduce the RRs in this study. Another weakness, noted by the authors, is that the small number of incident cases during follow-up period hindered precise effect estimates. Also, the high frequency of exposure to many pesticides (e.g. 73.8% were exposed to 2,4-D) means subjects unexposed to glyphosate were likely to be exposed to other agents that may also induce NHL, reducing the RRs. Also, as noted by the EPA's FIFRA Science Advisory Panel (SAP)[54] in their review of the EPA's issue paper on the carcinogenicity of glyphosate and as noted in a critique[55] of the European Food Safety Agency's risk assessment for glyphosate, the follow-up time in this cohort study may not be long enough to produce a sufficient sample size for evaluation of the association between NHL and glyphosate. Like other studies, this study has few exposed cases and controls, but the authors adjust their analysis for many other pesticides which could reduce ORs and increase confidence bounds limiting the ability of the study to show positive results. This study could also suffer from a survival bias because pesticide applicators were recruited as case participants after their exposure had begun and those with a cancer prior to enrollment were excluded.

This study will be included in the evaluation of causality.

Consistency of Associations

Hill (1965)[36] defines consistency as the answer "yes" to the question "Has it repeatedly been observed by different persons, in different places, circumstances and times?" For these studies, the answer is indeed yes.

If the population relative risk (PRR) for an association of glyphosate with NHL were equal to 1 (no effect), then one would expect very few statistically significant results in multiple studies and that about half of the studies would have ORs or RRs below one

and half above one. As noted by both the **IARC Monograph 112 (2015)**[56] and by **Chang and Delzell (2016)**[38], when comparing studies, the most reasonable comparison is to use the most-fully-adjusted risk estimates. I will mostly limit my comments to these most-fully-adjusted risk estimates.

Consistency of the associations across several epidemiology studies is not simply a matter of seeing how many were statistically significant and how many were not but must also address the consistency of the direction of the responses. Figure 1 shows a forest plot of all ORs and RRs from the epidemiology studies discussed previously. Each horizontal line in the forest plot shows the mean estimate of the OR/RR as a black square and the 95% confidence interval around this estimate as whiskers extending left and right from the black square.

The first obvious conclusion to be drawn from Figure 1 is that all of the mean OR/RR estimates (black squares) are consistently ≥1. This implies that all of the studies are pointing in the same direction toward a positive effect. In their meta-analyses, **Schinasi and Leon (2014)**[57], **IARC (2015)**[56] and **Chang and Delzell (2016)**[38] all identified 6 papers (highlighted in red in Figure 1) as being the most reliable for evaluation of the ability for glyphosate to induce NHL in people: **McDuffie et al. (2001)**[50], **Hardell et al. (2002)**[42], **De Roos et al. (2003)**[43] and **(2005)**[45], **Eriksson et al. (2008)**[46] and **Orsi et al. (2009)**[47]. I will refer to these papers as the six core epidemiology studies. As noted above, if the true underlying risk ratio was 1 (no effect), you would expect about half of the findings to be below 1 and half to be equal to 1 or greater. Using only the results from the 6 core studies, you can see that all are ≥1; the probability of this happening is $(0.5)^6$ or 0.016, strongly suggesting the studies do not agree with an underlying PRR=1 and that they consistently support a positive effect.

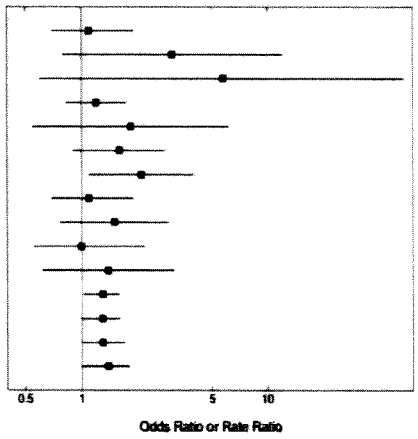
A second way in which consistency can be evaluated is to combine the individual studies using meta-analysis to obtain a combined analysis using both the ORs and the RR (CRR) and test for heterogeneity in the studies. The meta-analysis done by **Chang and Delzell (2016)** includes the same analysis as that done by the **IARC (2015)** and is an improvement over **Schinasi and Leon (2014),** so I will focus my comments on using the **Chang and Delzell (2016)** meta-analysis. **Chang and Delzell (2016)** did four separate meta-analyses on the glyphosate epidemiology studies using two different methods (random-effects and fixed-effects models). In their first analysis (model 1)[5], they combined the most-fully-adjusted risk estimates from the six core studies to yield a CRR of 1.27 (1.01-1.59) for both random-effects and fixed-effects models supporting an association between NHL and glyphosate exposure in these studies. In a second analysis (model 2), they replace the results of the Bayesian analysis in **De Roos et al. (2003)** with the results of the logistic regression analysis and get the same CRR of 1.30 (1.03-1.64) for both random-effects and fixed-effects models. In a third analysis (model 3), they replace from model 1 the **McDuffie et al. (2001)** results in with a combined meta-

[5] **Chang and Delzell (2016)** provided only one significant digit to the right of the decimal point in their confidence bounds; the EPA SAP (2017) re-calculated models 1-4 of **Chang and Delzell (2016)** to provide two significant digits – these are presented here.

analytic result they derived from analyses by **Hohenadel et al. (2011)** (this study reanalyzed the same data as **McDuffie et al. (2001)**, splitting results between asthmatics and non-asthmatics) resulting in a CRR of 1.32 (1.00-1.73) for both random-effects and fixed-effects models. Finally, in a fourth analysis (model 4), they use model 3 but replaced the Bayesian analysis in **De Roos et al. (2003)** with the logistic regression analysis yielding a CRR of 1.37 (1.04-1.82) for both random-effects and fixed-effects models. In essence, none of the different meta-analyses rejected the notion of a combined, statistically significant positive effect.

Figure 1: Odds Ratios and Rate Ratios from the most-fully-adjusted risk estimates from selected epidemiology studies and from the meta-analyses of **Chang and Delzell (2016)**[38]. "RR" refers to the OR or RR from the study, "Lower" refers to the 95% lower bound, "Upper" to the 95% upper bound and "Weight" refers to the weight applied to that specific study in Model 1 of the meta-analysis (Table 3 in Chang and Delzell). For **De Roos et al. (2003)**, the first row is for the Bayesian model analysis and the second row, labelled "logistic regression" is from the logistic model analysis.

Study	RR	Lower	Upper	Weight (model 1)
Cantor et al. (1992)	1.10	0.70	1.90	0.0
Nordstrom et al. (1998)	3.10	0.80	12.00	0.0
Hardell and Eriksson (1999)	5.60	0.60	54.00	0.0
McDuffie et al. (2001)	1.20	0.83	1.74	38.1
Hardell et al. (2002)	1.85	0.55	6.20	3.6
De Roos et al. (2003)	1.60	0.90	2.80	16.2
logistic regression	2.10	1.10	4.00	0.0
De Roos et al. (2005)	1.10	0.70	1.90	21.0
Eriksson et al., (2008)	1.51	0.77	2.94	11.6
Orsi et al. (2009)	1.00	0.55	2.20	3.6
Hohenadel et al. (2011)	1.40	0.62	3.15	0.0
Meta-Analysis: Model 1	1.30	1.03	1.60	
Meta-Analysis: Model 2	1.30	1.00	1.60	
Meta-Analysis: Model 3	1.30	1.00	1.70	
Meta-Analysis: Model 4	1.40	1.00	1.80	



As stated above, another way to evaluate consistency in the epidemiological data would be to evaluate the heterogeneity in the studies. Heterogeneity may be due to differences in participants, outcomes, exposure metrics, methods for questioning study subjects, sex of the subjects, etc. **Chang and Delzell (2016)** formally tested for heterogeneity of the responses from the six core studies using Cochran's Q statistic and the I^2 statistic[58]. For models 1 to 4, the p-values from Cochran's Q test are 0.84, 0.59, 0.85, and 0.63 respectively (typically you reject the concept of homogenous studies in favor of heterogeneous studies if $p<0.10$). The I^2 statistic for all four models are 0.0% (values for I^2 can range from 0-100% with concern for heterogeneity above 50%). The fact that the fixed-effects models and random-effects models gave the same results also supports a lack of heterogeneity in the data. There is no indication of heterogeneity in these six core studies. Lack of heterogeneity supports the interpretation of the meta-analyses as showing a positive association and strong consistency of the findings across the six core studies.

Chang and Delzell (2016) also evaluated the association between subtypes of NHL and glyphosate exposure where possible. For B-cell lymphomas, they combined the results of **Eriksson et al. (2008)**[46] with those of **Cocco et al. (2013)**[49] and saw a CRR (random-effects and fixed-effects) of 2.0 (1.1-3.6) with an I^2 of 0 and a Cochran's Q test p-value of 0.58. For diffuse large B-cell lymphomas, they combined the results of **Eriksson et al. (2008)**[46] with those of **Orsi et al. (2009)**[47] and saw a CRR (random-effects and fixed-effects) of 1.1 (0.5-2.3) with an I^2 of 0 and a Cochran's Q test p-value of 0.79. For combined chronic lymphocytic leukemia and small lymphocytic lymphoma, they combined the results of **Eriksson et al. (2008)**[46] with those of **Orsi et al. (2009)**[47] and saw a CRR using the random-effects model of 1.3 (0.2-10) and for the fixed effects model 1.9 (0.9-4.0) with an I^2 of 83.7% and a Cochran's Q test p-value of 0.01. For follicular lymphomas, they combined the results of **Eriksson et al. (2008)**[46] with those of **Orsi et al. (2009)**[47] and saw a CRR (random-effects and fixed-effects) of 1.7 (0.7-3.9) with an I^2 of 0 and a Cochran's Q test p-value of 0.73. And finally, for HCL, they combined the results of **Nordstrom et al. (1998)**[40] with those of **Orsi et al. (2009)**[47] and saw a CRR (random-effects and fixed-effects) of 2.5 (0.9-7.3) with an I^2 of 0 and a Cochran's Q test p-value of 0.63. These subtype analyses are based upon small numbers of cases and only two studies making them unreliable, when considered individually, to address the question of consistency in the data. However, when they are combined with the results for the meta-analyses of the core studies of NHL, these studies add support to the conclusion that these data are consistent.

Chang and Delzell (2016) also performed a sensitivity analysis by only doing meta-analyses on studies with similar characteristics. Using only the five case-control studies, the CRR was 1.3 (1.0-1.7). Breaking them into the type of control used, there were four studies using population controls with a CRR of 1.4 (1.0-1.8). There were four studies with males only with a CRR of 1.3 (1.0-1.7) and two studies with males and females with a CRR of 1.2 (0.8-1.8). Three studies were done in North America with a CRR of 1.2 (1.0-1.6), three in Europe with a CRR of 1.3 (0.8-2.1); two of the three studies were in Sweden with a CRR of 1.6 (0.9-2.8). All of the resulting meta CRRs were the same for the fixed-effects model and the random-effects model. This sensitivity analysis shows that the results do not differ significantly from the main CRR for the six core studies combined adding support to the findings being consistent across the different studies.

In case-control studies, selection bias arises when the reasons cases and controls choose to participate in the study could lead to systematic biases that might result in a positive or negative finding independent of the exposure being studied. For example, if cases with exposure are more likely to participate than controls with exposure, the result would be higher OR values; however, this difference has to be differential and not simply a difference in participation rates. It is possible that in a few of these studies, the method by which controls were selected could contribute to selection bias that might lead to increased ORs. However, given the diverse types of cases and controls used in the five core case-control studies, this is unlikely to explain the consistent findings seen from these studies. It is also possible that the lack of complete data on cases versus controls could result in selection bias if the reasons for not completing the questionnaire/interview are different between cases and controls and relates to

exposure. There is no indication of this type of selection bias in these reports, and this is unlikely to explain the consistency seen in these data.

Exposure misclassification can lead to increases or decreases in the OR or RR values seen in both case-control and cohort studies. For example, in case-control studies, if cases are more likely to say they were exposed to glyphosate than controls, this would inflate the OR values; this is one type of recall bias. This type of bias is less likely in cohort studies. In all six of the core studies, this issue was discussed by the authors. In every case, they concluded there was bound to be some exposure misclassification, but that it was most likely non-differential, meaning that the misclassification was random; this would likely reduce the OR/RRs seen in the studies rather than increase them.

Confounding occurs when there is an exposure or some other factor that is tightly associated with both glyphosate exposure and NHL diagnosis that, if controlled for, could explain the results. The most likely source of confounding in these studies would be exposures to other pesticides. Four[42, 43, 45, 46] of the six core studies controlled for exposure to other pesticides and saw basically the same findings as the other two studies. Another concern for confounding would be if the cases had immune deficiencies that could be linked to NHL; in all of the case-control studies, such cases were excluded. Finally, other agricultural exposures (e.g. animals, other chemicals, infectious agents) could be correlated with glyphosate exposure and may be linked to NHL; none of the studies controlled for these factors. However, not all exposed cases were farmers; if confounding via other agricultural exposures is occurring, it is not possible to determine the magnitude or direction of such an effect from these data.

In conclusion, we have six core epidemiology studies done on two different continents by four different research groups using different designs, questionnaires and study populations that are highly consistent with no obvious bias or confounding that would explain the results. **There is a consistency of associations across the six core studies.**

Strength of the Association seen in Human Epidemiological Studies

To explain strength of association, **Hill (1965)** gives the classic example of John Snow and the cholera epidemic of 1855 where the risk ratio of dying if you drank water from the Southwark and Vauxhall Company (polluted by sewage) compared to drinking from the Lambeth Company water (sewage free) was 14. Yet, for the six core studies, the OR/RR ranges from 1.0 to 1.85 for the most-fully-adjusted risk estimates and to 2.1 if you include the fully adjusted risk estimate from De Roos et al. (2003)[45] using logistic regression. These are moderate OR/RR estimates making it conceivable they are individually due to either chance or bias. Thus, with the exception of the logistic regression analysis in **De Roos et al. (2003)**[45], none of the core studies demonstrate large, precise risks as envisioned by **Hill (2016)**[36]. However, **Hill (1965)** was not expressing himself in statistical terms where the significance of an association is dependent upon the precision of the observations. If the statistical variation around an OR/RR estimate is large relative to the estimate itself, the estimate is not very precise

and generally would not be statistically significant. The result from the study by **Hardell and Eriksson (1999)** shown in Figure 1 is an example of an estimate with very large statistical variation. On the other hand, a very small (in value), precise OR or RR estimate could be statistically significant and prove important in deciding causation. The meta-analyses shown in Figure 1 all demonstrate estimates of OR/RR that are significantly different from 1 rejecting the concept that the overall association is due to chance. The statistically significant estimate of the OR/RR for B-cell lymphomas in the meta-analysis support this finding as well.

In summary, we have six core epidemiology studies that all show approximately the same, modest increase in OR/RR that, when combined, demonstrate a significant strength of association. **There is a strong association across the six core studies**

Biological Plausibility

The range of data one can use to determine biological plausibility is quite diverse and can be exceptionally complicated. For simplicity, it can be divided into the types of assays that can be used in this evaluation: animal cancer bioassays, toxicokinetic studies, studies from accidental exposures in humans, and studies of specific biological mechanisms in animals or cells derived from humans or animals. Animal cancer bioassays are intended to test whether glyphosate can cause cancers in mammals, thus supporting the concept that the chemical could cause cancer in humans. Toxicokinetic studies provide insight into the degree to which glyphosate is absorbed by humans, distributed to various organs in the body, what happens to the chemical once it is in the body (metabolism), and, finally, how it is eliminated from the body. Studies from accidental exposures in humans can provide some information on the effects of glyphosate through changes in the chemistry and cellular structure of human blood. Studies of biological mechanisms are generally addressing what effects the chemical may have on human and animal cells under controlled, laboratory conditions. Some of the studies in this section were done with technical grade (virtually pure) glyphosate and some with the glyphosate formulations that humans encounter in occupational and environmental settings. I will summarize the literature in each of these areas and offer an opinion to their support of biological plausibility of NHL in humans.

Animal Cancer Bioassays

Typical animal cancer bioassays will expose animals (generally rats or mice) to a chemical for a substantial proportion of the animal's life (generally 2 years) then kill the animal and examine its organs and tissues for tumors. There are guidelines on how to conduct and analyze these studies. Typically, chemical registrants conduct cancer bioassays for pesticide approval pursuant to guidelines developed under the guidance of the Organization for Economic Cooperation and Development (OECD[59]). Other groups[30, 33, 34] provide guidance on how to analyze these studies based upon methodology papers from the published literature. These studies are conducted in a way that controls for everything in the animal's environment (e.g., food type, water quality, how often the animals are handled) leaving only the exposure to explain

differences in tumor formation between control and exposed animals. Even then, non-cancer endpoints can also be modified by the chemical and these may have an impact on tumor rates in the animals (e.g., survival, death from some other toxic effect of the chemical); these must be accounted for when reaching conclusions from the study.

Studies generally use four groups of animals, one group receiving no exposure (control) and the remaining three groups are test animals, with each group receiving different dose exposures to the chemical[60]. Doses generally above human experience are used in animal carcinogenicity studies because only relatively small numbers of animals are being used to evaluate risk for a large human population and because even the best known human carcinogens do not cause cancer in large fractions (say 20%) of the human population. The basic underlying premise of this design consideration is that, as the dose increases, so does the risk of getting a tumor. By exposing animals to the highest dose possible, you increase the ability of the study to identify a risk if one is present. However, one must be careful not to use a dose that is so high it will cause cancers by processes that would never work at lower doses. To avoid this, studies are designed around a maximum tolerated dose (MTD) or limit dose. This dose is generally determined based upon a subchronic study (90 days) in the same animals and is usually the maximum dose that can be tolerated by the animals without any signs of significant toxicity in the exposed animals (e.g., weight loss, tissue damage). The OECD and EPA provide guidelines[33, 59] on how to choose this top dose. These guidelines are in general agreement with the scientific literature[60].

The guidelines also address the methods by which the data should be analyzed. For example, the EPA guidelines[61] state that:

"A trend test such as the Cochran-Armitage test (Snedecor and Cochran, 1967) asks whether the results in all dose groups together increase as 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 that of 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."

In fact, most guidelines and peer-reviewed publications come to the same conclusion[30, 59, 60, 62] on what tests to use, as did EPA's FIFRA Scientific Advisory Panel (SAP) in their review of the EPA's issue paper of the carcinogenicity of glyphosate[54]. The US National Toxicology Program (NTP) uses both a trend test[3, 4, 63] and Fisher's exact test for analyzing carcinogenicity data. Unless otherwise noted in this document, all p-values presented in this section on animal cancer studies were recalculated on my computer and are the exact one-sided p-values for the Fisher test (p_{Fisher}) and/or the Cochran-Armitage linear trend test (p_{Trend}) where appropriate. In cases where the data is pooled and the numbers of tumors are large, the approximate p-value based upon the normal distribution is used for the trend test to avoid excessive computation time; these are noted as p_{TrendA}. The approximation (p_{TrendA}) is generally equivalent to the exact p-value (p_{Trend}) when there are more than 10 animals with tumors[64].

To avoid doing large numbers of tests and over-analyzing the data, my comments will generally rely upon the use of the trend test with the results from Fisher's exact test serving as a descriptive discussion of the findings. This is in agreement with SAP comments[54] and is generally accepted in the evaluation of animal cancer studies.

Even with the high doses used in these studies, it is sometimes necessary to use "historical controls" to evaluate a given response. Historical controls are generally the historical collection of tumor responses from untreated control groups from studies in the same laboratory within two to three years of the study being evaluated[30, 34, 59, 65, 66]. Evaluation of the data using the historical controls should be done rigorously to correctly evaluate the responses seen in a given study. Where a valid historical control dataset was available, I used the mean tumor response in the controls to calculate the probability of observing the trend seen in the study or a more significant trend if the true probability of response is the historical control average; this is labeled p_{Hist}. In all cases, the guidelines and literature support the use of the control in the current study as the most appropriate control group to use unless there is a specific need to address historical responses. Many guidelines[30, 33, 34, 67] suggest historical controls be used for evaluating rare tumors and findings in assays that appear to be unusual. It is explicitly noted that significant increases in tumors over what is seen in the concurrent control should not be rejected simply because the tumors are in the range of the historical controls[30]. Nor is it recommended to reject significant increases in tumor responses because the control response is on the low end of the historical range. Animals are randomly assigned to control and exposure groups and any low response in controls is likely to also reflect similar response patterns in treated animals. This is in agreement with SAP comments[54] on the EPA issue paper on glyphosate[61] and with all guidelines for analyzing animal carcinogenicity data.

There are 13 animal carcinogenicity studies in rats[68-80] and eight in mice[81-88]. Only two studies[71, 77] appear in the peer-reviewed literature; the remaining studies are partially available through several sources. For three of the rat studies[70, 74, 78] and two mouse studies[83, 86], technical reports from the performing laboratory are available from documents provided by the registrant. For the remaining unpublished studies, data was obtained from the EPA review of glyphosate[61], the European Food Safety Authority review of glyphosate[89, 90] and supplemental material from a review of the carcinogenicity of glyphosate by a panel of scientists on behalf of Monsanto[91].

Many additional endpoints, other than cancer incidence and related toxicities, were evaluated in these studies; I will only provide comments on the tumor incidence data and related data where relevant to the cancer findings.

It is unusual to have multiple carcinogenicity studies in the same experimental animal model arising from different laboratories. Methods for the combined analysis of multiple animal cancer bioassays are not available in the scientific literature. However, pooled analyses, as conducted in epidemiology[92, 93] are applicable for combining animal carcinogenicity studies. The basic concept is to pool all data from the same sex/species/strain into one study and analyze it appropriately. The basic steps are: 1) select the studies to be pooled; 2) merge the data for analysis; 3) estimate study specific

effects; 4) estimate pooled effects; 5) explain the differences between the pooled effects and the individual study effects; 6) do a sensitivity analysis if possible. These steps will be used to analyze pooled data from animal carcinogenicity studies where pooling is done by sex, species, strain and duration of exposure to limit heterogeneity across pooled studies. In their recommendations to the EPA regarding EPA's issue paper on the carcinogenicity of glyphosate[54], the FIFRA Science Advisory panel strongly supported the use of a pooled analysis to address the question of consistency citing my comments to the EPA[94].

Rat Studies

Reyna and Gordon (1974)[76] exposed Albino rats (probably Sprague-Dawley) to ammonium salt of glyphosate (13.85% purity) in a two-year chronic feeding study. Only EPA[61] reported on this study and provided no details other than to report there were approximately 70 animals per group and there was insufficient reporting on the histopathology findings. Insufficient detail is available on this study.

This study is inadequate for use in deciding on causality.

Burnett et al. (1979)[70] exposed male and female albino rats to an aqueous monosodium salt solution of glyphosate by oral intubation (purity not given). There were 90 animals per group and doses were 0, 3, 10 and 30 mg/kg/day for 24 months. EPA[61] reported that no histopathological alterations were observed; no additional information was available on this study. This study had severely reduced sensitivity to observe any cancer findings because the highest dose used in this study is very low compared to the MTDs in the other rat studies. This study does not contribute to the evaluation of cancer causation in laboratory animals and will be excluded from any further discussion.

Lankas et al. (1981)[74] exposed groups of 50 male and 50 female Sprague-Dawley rats to glyphosate (98.7% purity) in feed (see Table 1 for doses) for 26 months. This study is not in concordance with OECD guidelines (they were not available at the time of this study), but as noted by EFSA[89], it was in general accordance with the 1981 OECD guidelines. Information on this study was available from EPA[61], EFSA[89], Greim et al.[91], the original study report from Bio/dynamics Inc.[95] and memos from Monsanto to EPA provided by Monsanto.

There were no survival differences in this study and there was no indication that the highest dose used exceeded the maximum-tolerated dose.

Table 1 shows the statistically significant trend in testicular interstitial cell tumors that was observed (p_{Trend}=0.009). Historical controls were provided in the study report for five studies with response rates of 4/116, 5/75, 4/113, 6/113 and 5/118 for a mean response of 4.5% (24/535). Comparing this historical control mean to the observed response yields p_{Hist}=0.006, showing that this result is significant, even when comparing it to the historical control dataset. **Lankas et al. (1981)** argued that the tumor rates at sacrifice were not statistically significant from control suggesting this finding is not related to glyphosate. However, by reducing the numbers of animals to only those at

terminal sacrifice, the power to find an effect was significantly reduced. Also, if the tumor increases the animal's chances of dying, then some animals with tumors will die early, which could bias results only seen at terminal sacrifice. This type of analysis is simply never done; it appears to have been developed for this case to dismiss the effects seen in the study. **Lankas et al. (1981)** also suggested the control response was low compared to the historical rates, but the concurrent control is always the best control group to use unless it is clearly flawed[33, 34, 59]; in this case, there was no apparent problem with the controls because the probability of seeing 0/50 if the true background response is 4.5% is about 10% and this control group is not significantly different than the historical controls. **EFSA**[89] noted rates for interstitial cell hyperplasia (a potential precursor for the interstitial cell tumors) and saw no dose-response trend (Table 1). However, these very low rates would suggest that the tumors arising in the 10 animals that did get interstitial cell tumors are independent of a mechanism involving interstitial cell hyperplasia. The tumor response for interstitial cell tumors was not monotonic (tumor rates increasing as dose increases), but was still within statistical variation. The EPA SAP agrees, concluding that "requiring visual confirmation of a monotonic trend in scatter plots of data ... is known to be a poor way of assessing trend"[54].

An increase in Thyroid C-cell carcinomas (Table 1) was observed in female rats (p_{Trend}=0.003) but combining adenomas and carcinomas was only marginally significant (p_{Trend}=0.072). Independent pathologists brought in by Monsanto argued these tumors were not treatment related. The authors provided historical control data for both carcinomas and carcinomas combined with adenomas from nine control groups with mean responses of 4/453=0.9% for carcinomas and 46/453=10.2% for the combined tumors. The significance of both results was unchanged using the historical control data.

The authors also mentioned that the incidence of lymphocytic hyperplasia in the thymus and lymph nodes were slightly elevated above controls (p_{Trend}=0.143). The middle dose group was significantly different from controls (p_{Fisher}=0.018).

This study also had a statistically significant increase in pancreatic islet cell tumors in the lowest dose (p_{Fisher}=0.028) in males (Table 1), but not any of the other doses; the trend test was not significant (p_{Trend}=0.312).

The highest dose used in this study in Sprague-Dawley rats is far below the MTD. Even though EFSA[89] noted that this study was in general accordance with the 1981 OECD guidelines, they dismissed it for not meeting current guidelines due to the low-doses used. EPA[61] also excluded this study from consideration. However, the study saw an increase in testicular tumors in males and Thyroid C-cell carcinomas in females that should be carefully evaluated in determining causality. Also, this is the study with the longest exposure (26 months) and provides unique information to the overall evaluation.

Additional tumors seen to have significant increases in other studies using Sprague-Dawley Rats are also included in Table 1.

Table 1: Tumors of interest in male and female Sprague-Dawley rats the 26-month feeding study of **Lankas (1981)**[74]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	3.05	10.30	31.49	
	Female	0	3.37	11.22	34.02	
Testicular interstitial cell tumors	Male	0/50	3/50	1/50	6/50**	P_{Trend}=0.009 P_{Hist}=0.006
Interstitial cell hyperplasia	Male	1/50	1/50	1/50	0/50	P_{Trend}=0.830
Thyroid C-cell Carcinomas	Female	1/47	0/49	2/50	6/47	P_{Trend}=0.003 P_{Hist}=<0.001
Thyroid C-cell Adenomas and Carcinomas	Female	6/47	3/49	8/50	9/47	P_{Trend}=0.072 P_{Hist}=0.072
Pancreas Islet Cell Tumors	Male	0/50	5/50*	2/50	3/50	P_{Trend}=0.312
lymphocytic hyperplasia, thymus and lymph nodes	Female	27/50	35/50	38/50*	35/50	P_{Trend}=0.143
Thyroid C-cell Adenomas and Carcinomas	Male	1/47	2/49	4/49	4/49	P_{Trend}=0.122
Thyroid Follicular-cell Adenoma	Male	5/47	1/49	2/49	2/49	P_{Trend}=0.748
Liver Neoplastic Nodule	Male	3/50	5/50	1/50	3/10	P_{Trend}=0.630
Kidney Adenoma	Male	1/50	5/50	0/50	0/50	P_{Trend}=0.979

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01

In conclusion, this study shows positive result for testes interstitial cell tumors and hepatocellular adenomas in male Sprague-Dawley rats and a positive response for thyroid c-cell carcinomas in female Sprague-Dawley rats and will be included in the overall evaluation of causation.

Stout and Ruecker (1990)[78] exposed groups of 50 male and 50 female Sprague-Dawley rats to glyphosate (98.7% purity) in feed (see Table 2 for doses) for 24 months. This study was done under OECD guidelines.

There were no survival differences in this study and there was no indication that the highest dose used exceeded the maximum-tolerated dose.

Pancreatic islet cell tumors were increased in all dose groups relative to the controls in male rats and statistically significant for the lowest (p_{Fisher}=0.015) and highest (p_{Fisher}=0.032) dose groups (Table 2). However, these rates include the 10 animals that were sacrificed at one year. Due to the short duration of exposure, the rats terminated at one year were likely not at risk of developing this tumor; it is very unusual to include these animals in the final tumor counts (EPA[61] also excluded these animals). In the pathology tables for this study, there were no tumors in any of the 10 animals at the interim sacrifice. Removing these 10 animals does not alter the p-values for trend or

Fisher's exact test. Historical control data for this tumor in this laboratory was reported as 23/432 or 5.3%[96] and a trend comparison against this control rate was not significant (p_{hist}=0.15). The lack of a trend is driven by the up and down nature of the response. Assuming the historical rate of 5.3% is correct, the chances of seeing eight or more tumors in 47 animals is 0.003. Similarly, for the mid- and high-doses, this probability is 0.124 and 0.014, respectively. Females did not show an increase in this tumor. The authors provided a table with the combined results for pancreatic islet-cell adenomas and carcinomas from this study with the tumor counts from the **Lankas et al. (1981)**[74] study arguing the results do not show a dose-related increase. Animals studied for 26 months versus 24 months can have very different responses to the same chemical and very different control incidence.

In male rats, there was a statistically significant trend (p_{Trend}=0.015) after removal of interim-sacrificed animals for hepatocellular adenomas but a significant increase for adenomas and carcinomas combined (p_{Trend}= 0.05, Table 2) and not in females (not shown). Liver carcinomas are generally also provided in a separate analysis, but these data were not provided by the authors (the data would suggest the hepatocellular carcinomas would have a negative trend).

There was also a significant increase in thyroid C-cell adenomas in the female rats (p_{Trend}=0.049) and a marginal increase[6] in adenomas and carcinomas combined (p_{Trend}=0.052) regardless of whether interim sacrificed animals are included (Table 2). In males, the trend for adenomas was p_{Trend}=0.084 and for adenomas and carcinomas was p_{Trend} =0.091. Adenomas were seen in male rats at the interim sacrifice demonstrating that male rats at the interim sacrifice were at risk for this tumor. If these animals are added back into the analysis, the trend test in males has p_{Trend}=0.063 for adenomas and p_{Trend}=0.068 for adenomas and carcinomas combined.

Several other tumors demonstrating significant findings in other studies of Sprague-Dawley rats are included in Table 2 and do not show significant effects.

In conclusion, the finding of an increased incidence of pancreatic islet-cell tumors in this study cannot easily be ruled out as a chance finding. Findings of significant increases in liver adenomas in male rats with no increases in carcinomas could be due to chance. The findings of significant increases in thyroid c-cell tumors in males and females should be compared with other studies. This study will be included in the overall evaluation of causation.

---

[6] In statistics, it is common to refer to p-values in the range of 0.10>p-value>0.05 as marginal when the target p-value is ≤0.05; this is done to avoid missing trends in data reflected by almost significant findings

Table 2: Tumors of interest in male and female Sprague-Dawley rats from the 24-month feeding study of Stout and Ruecker (1990)[78]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	89	362	940	
	Female	0	113	457	1183	
Pancreas Islet Cell Tumors (with interim sacrifice)	Male	1/58	8/57*	5/60	7/59*	P_{Trend}=0.147 P_{Hist}=0.140
Pancreas Islet Cell Tumors (without interim sacrifice)	Male	1/48	8/47*	5/50	7/49*	P_{Trend}=0.147 P_{Hist}=0.150
Hepatocellular adenomas (without interim sacrifice)	Male	3/50	2/50	3/50	8/50	P_{Trend}=0.015
Hepatocellular Adenomas and Carcinomas (without interim sacrifice)	Male	6/50	4/50	4/50	10/50	P_{Trend}=0.050
Thyroid C-Cell Adenomas (with interim sacrifice)	Female	2/60	2/60	6/60	6/60	P_{Trend}=0.050
Thyroid C-Cell Adenomas (without interim sacrifice)	Female	2/50	2/50	6/50	6/50	P_{Trend}=0.049
Thyroid C-Cell Adenomas and Carcinomas (with interim sacrifice)	Female	2/60	2/60	7/60	6/60	P_{Trend}=0.053
Thyroid C-Cell Adenomas and Carcinomas (without interim sacrifice)	Female	2/50	2/50	7/50	6/50	P_{Trend}=0.052
Thyroid C-Cell Adenomas (with interim sacrifice)	Male	2/60	4/60	8/60	7/60	P_{Trend}=0.063
Thyroid C-Cell Adenomas (without interim sacrifice)	Male	0/50	4/50	8/50**	5/50*	P_{Trend}=0.084
Thyroid C-Cell Adenomas and Carcinomas (with interim sacrifice)	Male	2/60	6/60	8/60*	8/60*	P_{Trend}=0.068
Thyroid C-Cell Adenomas and Carcinomas (without interim sacrifice)	Male	0/50	6/50*	8/50**	6/50*	P_{Trend}=0.091
Testis Interstitial Cell Tumors	Male	2/50	0/50	3/50	2/50	P_{Trend}=0.296
Kidney Adenomas	Males	0/50	2/50	0/50	0/50	P_{Trend}=0.813
Thyroid Follicular Adenoma/Carcinoma	Males	2/50	1/48	3/48	3/50	P_{Trend}=0.225

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01

Atkinson et al. (1993)[68] conducted a combined chronic toxicity/carcinogenicity study of glyphosate (98.9% pure). They used 50 Sprague-Dawley rats in each group for both sexes with dietary exposures given in Table 3. An additional 35 rats/sex/dose were included for interim sacrifices.

There were no survival differences in this study and there was no indication that the highest dose used exceeded the maximum-tolerated dose.

Table 3: Tumors of interest in male and female Sprague-Dawley rats from the 24-month feeding study of Atkinson et al. (1993)[68]

Tumor	Sex	Doses (mg/kg/day)					p-values
	Male	0	11	112	320	1147	
	Female	0	12	109	347	1134	
Thyroid Follicular Adenomas and Carcinomas	Male	0/50	0/21	0/17	2/21	2/49	P_{Trend}=0.099
Thyroid Follicular Adenomas and Carcinomas (adding terminal sacrifice animals to denominator)	Male	0/50	0/50	0/50	2/50	2/49	P_{Trend}=0.034
Thyroid C-cell Adenomas and Carcinomas	Female	8/50	1/27	1/29	1/29	7/49	P_{Trend}=0.197
Thyroid C-cell Adenomas and Carcinomas	Male	9/50	1/21	1/17	2/21	9/49	P_{Trend}=0.183
Testes Interstitial Cell Tumors	Male	3/50	1/25	0/19	0/21	2/50	P_{Trend}=0.580
Kidney Adenomas	Males	1/50	0/50	0/50	0/50	0/50	p_{Trend}=1
Hepatocellular Adenomas	Males	2/50	1/50	1/50	2/50	3/50	P_{Trend}=0.155
Pancreas Islet-Cell Adenoma	Male	0/50	0/50	0/50	0/50	1/50	P_{Trend}=0.200

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01

The authors reported no significant effects, as do **EPA**[61] and **EFSA**[89]. The study did not do detailed histopathological examination on all animals in all groups for every tumor type, but did examine all control and high dose animals, all animals that died before study termination and animals showing macroscopic tumors at study termination; liver, kidney and lungs were examined for all animals. This severely weakens the study for addressing dose-response trends. However, in reviewing the pathology tables provided in **Greim et al. (2015)**[91], thyroid follicular adenomas and carcinomas were found to be marginally significant (p_{Trend}=0.099) by the trend test. If the three middle exposure groups had seen no other tumors and the denominators were the entire 50 animals on study, the trend analysis becomes significant (p_{Trend}=0.034).

Without examination of the animals free of gross tumors at terminal sacrifice, the findings from this study will be given less weight in the overall evaluation of causation.

Brammer (2001)[69] conducted a two-year carcinogenicity study in Wistar rats in which groups of 52 animals were exposed to glyphosate (97.6% pure) at doses provided in

Table 4 . An additional 12 animals were sacrificed at one-year.

A significant positive trend in survival was noted by the EPA (p=0.03), however this trend was not accomplished using a Kaplan-Meir test[97] (the appropriate test), but simply a test relating to the percent surviving to terminal sacrifice. There was no indication that the highest dose used exceeded the maximum-tolerated dose.

EPA[61], but not **EFSA**[89], noted there was a statistically significant trend of hepatocellular adenomas in male rats with the highest dose also being statistically significant from the control. Trend analysis gives p_{Trend}=0.008 and the Fisher's exact test comparison of high dose to control is p_{Fisher}=0.027. EPA dismissed this finding as potentially due to a slight difference in the number of animals at the terminal sacrifice in this study versus controls. However, no formal statistical evaluation of survival is provided and it cannot be assumed from these numbers that survival was significantly impacted in these animals. **Greim et al. (2015)**[91] used slightly different numbers for this tumor because three animals (one in the control group, one in the low-dose group and one in the mid-dose group) in the interim sacrifice group died before their sacrifice time and, from the pathology tables provided in their paper, these could not be separated from others. These numbers have been included in Table 4, but it does not change the significance of the findings. **Greim et al. (2015)**[91] dismissed these findings, partly because of the same survival argument used by the EPA and partly because they had a historical control dataset where the range of historical response was from 0-11.5%; they did not provide the mean response or the individual tumor responses for these historical controls. As mentioned earlier, dismissing results because they are in the range of the historical controls is an unacceptable method for using historical controls to evaluate a study, and in this case, there is no reason to question the concurrent controls.

Table 4: Tumors of interest in male and female Wistar rats from the 24-month feeding study of Brammer (2001)[69]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	121	361	1214	
	Female	0	145	437	1498	
Hepatocellular Adenoma	Male	0/52	2/52	0/52	5/52*	P_{Trend}=0.008
Hepatocellular Adenoma (from Greim et al., 2015[91])	Male	0/53	2/53	0/53	5/52*	P_{Trend}=0.008 P_{Hist}=0.006
Mammary Gland Adenomas and Adenocarcinomas	Female	3/51	2/51	0/51	2/51	P_{Trend}=0.575
Skin Keratocanthoma	Male	1/51	0/51	1/51	1/51	P_{Trend}=0.392

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01

I obtained historical control data from 16 control groups in Wistar rats from Charles River Laboratories for the years 2003 to 2011[98]. Although these are outside of the optimal time range for the animals used in the **Brammer (2001)** study, they can serve as an illustration of why using a range can be misleading. There were 52 liver adenomas

seen in 1217 control animals for a mean response of 4.27% with a range of 0% to 17.5% (individual study findings of 6/100, 0/60, 1/60,1/50,1/80, 14/112, 1/65, 0/60, 21/120, 0/50, 1/50, 2/60, 0/50, 1/100, 1/150, 2/50; 13 studies with ≤2% response). Assuming the underlying probability of having a tumor in controls is 4.27%, p_{Hist}=0.006 (Table 4). Thus, even though the responses seen in **Brammer (2001)** are in the range of the historical controls, the trend is highly significant when historical controls are used appropriately. **Greim et al. (2015)** also mentioned findings of increased toxicity at the high dose for which they provided numbers for only hepatocyte fat vacuolation and hepatitis; none of these findings were statistically significant by any test.

In conclusion, this study shows a positive result for hepatocellular adenomas in male Wistar rats and will be included in the overall evaluation of causation.

Pavkov and Wyand (1987)[75] exposed Sprague-Dawley rats to glyphosate trimesium salt (sulfosate, 56.2% pure) in feed for two years. Eighty animals/sex were tested in the control, low-dose and mid-dose groups, and 90/sex were tested in the high dose group. Doses of 0, 4.2, 21.2 and 41.8 mg/kg/day were used in males and 0, 5.4, 27, and 55.7 mg/kg/day in females. This study showed no significant findings according to EPA[61]. No details were given beyond that simple statement and no others reported on this study. The doses in this study are far below the MTD so this study would have reduced sensitivity to detect an effect if one existed. This study also used a different chemical than the other Sprague-Dawley rat studies and is not comparable on that basis.

This study is not acceptable for use in the evaluation of causality due to the lack of details about the study.

Suresh, (1996)[79] exposed Wistar rats to glyphosate (96.8% pure) in feed for two years. Fifty animals/sex were tested in four exposure groups shown in Table 5.

There were no survival differences in this study and there was no indication that the highest dose used exceeded the maximum-tolerated dose.

EPA[61] concluded there were no tumors increased due to glyphosate exposure in this study and **EFSA**[89] concluded that, "[n]one of the significant microscopic changes, increased and decreased incidences (in liver, spleen, lymph nodes, adrenals, thymus, gonads, uterus, mammary gland) observed have shown dose relationship, hence appeared to be incidental and not related to the treatment with the test compound." (page 491). **Greim et al. (2015)**[91] provided data on hepatocellular adenomas and carcinomas in both sexes but none of these showed significant trends or pairwise tests (Table 5). However, there was another study with a strong significant trend in hepatocellular adenomas in Wistar rats[69] so these are also included in Table 5 for comparison. No other tumors were mentioned by any other group and an examination of the grouped pathology tables provided by **Greim et al. (2015)** show an increase in mammary gland adenomas at the mid-dose (p_{Fisher}=0.017) but no significant trend. However, there was another study with a strong significant trend in mammary gland adenomas and adenocarcinomas combined in Wistar rats[80] so these are also included in Table 5 for comparison. Like the **Atkinson et al. (1993)**[68] study, **Suresh (1996)** did not do full pathology on all of the animals in the interim exposure groups making

interpretation of this study problematic.

This study will be included in the overall evaluation of causation.

Table 5: Tumors of interest in male and female Wistar rats from the 24-month feeding study of Suresh(1996)[79]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	6.3	59.4	595.2	
	Female	0	8.6	88.5	886	
Mammary Gland Adenoma and Carcinoma	Female	5/40	3/28	8/33	2/48	P_{Trend}=0.970
Hepatocellular Adenoma	Male	24/50	22/50	10/50	21/50	P_{Trend}=0.374
Skin Keratocanthoma	Male	0/50	0/50	0/50	0/50	P_{Trend}=1

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01

Enemoto (1997)[72] exposed Sprague-Dawley rats to glyphosate (95.7% pure) in feed for two years. Fifty animals/sex were tested in four exposure groups (see Table 6). In addition, 10 animals per exposure group were exposed for 1 year and another 10 for 18 months at which point they were sacrificed and examined. These interim sacrifice animals (1 year and 18 months) are included in the analysis if tumors were seen in these groups.

There were no survival differences in this study and there was no indication that the highest dose exceeded the maximum-tolerated dose.

EPA and EFSA both found no significant changes in tumors in any group. **Greim et al. (2015)** again provide tables for a number of tumors, none of which show significant effects except for the incidence of kidney adenomas in male rats (p_{Trend}=0.004, Table 6). Examining the pathology tables provided in **Greim et al. (2015)** reveals no additional tumors showing an increase in tumor incidence with dose. A different study[74] in Sprague-Dawley rats demonstrated a strong significant trend in mammary gland adenomas, thyroid C-cell carcinomas, skin Keratocanthomas and testicular interstitial cell tumors so these are also included in Table 6 for comparison.

This study showed a significant increase in kidney adenomas and will be included in the overall evaluation of causation.

Table 6: Tumors of interest in male and female Sprague-Dawley rats from the 24-month feeding study of Enemoto (1997)[72]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	104	354	1127	
	Female	0	115	393	1247	
Mammary Gland Adenoma	Female	23/50	27/50	24/50	30/50	P_{Trend}=0.106
Kidney Adenoma	Male	0/50	0/50	0/50	4/50	P_{Trend}=0.004
Thyroid C-cell Adenomas/Carcinomas	Female	4/60	7/60	8/60	4/60	P_{Trend}=0.692
Thyroid C-cell Adenomas/Carcinomas	Male	8/70	10/70	6/70	7/70	P_{Trend}=0.697
Thyroid Follicular-cell Adenomas/Carcinomas	Male	4/70	2/70	1/70	0/70	P_{Trend}=0.990
Testes Interstitial Cell Tumors	Male	3/49	2/50	0/50	2/50	P_{Trend}=0.594
Hepatocellular Adenomas	Male	1/60	0/60	2/60	1/60	P_{Trend}=0.371
Skin Keratocanthoma	Male	3/50	3/50	0/50	6/50	P_{Trend}=0.065
Pancreas Islet-Cell Adenoma	Male	4/50	1/50	2/50	1/50	P_{Trend}=0.844

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01

Wood et al. (2009)[80] exposed Wistar rats to glyphosate (94.7% to 97.6% pure) in feed for two years. Fifty-one animals/sex were tested in four exposure groups at doses shown in Table 7.

No survival differences were seen in this study.

EFSA[89] found no dose-related tumor increases while **EPA**[61] noted an increase in mammary gland adenomas and adenocarcinomas combined with p_{Trend}=0.062 for adenomas, p_{Trend}=0.042 for adenocarcinomas and p_{Trend}=0.007 for the combined tumors (Table 7). EPA concluded there was no progression from adenoma to adenocarcinoma and argued the increase was not glyphosate related. This conclusion is contradicted by the fact that 6 animals in control and the lower dose groups got carcinomas with no adenomas in any of the animals in these groups. It seems likely that, in this case, mammary gland adenocarcinomas can arise without the presence of any adenomas. **Greim et al (2015)**[91] also noted an increase in skin keratoacanthoma in males (p_{Trend}=0.030). Review of the pathology tables identified no other tumors with increased tumor rates as a function of dose. There was another study with a strong significant trend in hepatocellular adenomas in Wistar rats[69] so this tumor is also included in Table 7 for comparison.

This study showed an increase in mammary tumors in females and skin keratoacanthomas in males and will be used in the evaluation of causality.

Table 7: Tumors of interest in male and female Wistar rats from the 24-month feeding study of Wood et al. (2009)[80]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	85.5	285.2	1077.4	
	Female	0	104.5	348.6	1381.9	
Mammary Gland Adenomas	Female	0/51	0/51	0/51	2/51	P_{Trend}=0.062
Mammary Gland Adenocarcinomas	Female	2/51	3/51	1/51	6/51	P_{Trend}=0.042
Mammary Gland Adenomas and Adenocarcinomas	Female	2/51	3/51	1/51	8/51*	P_{Trend}=0.007
Skin Keratocanthoma	Male	2/51	3/51	0/51	6/51	P_{Trend}=0.030
Hepatocellular Adenoma	Male	0/51	2/51	1/51	1/51	P_{Trend}=0.418

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01

Excel (1997)[73] exposed Sprague-Dawley rats to glyphosate (purity not given) in feed for two years. Fifty-one animals/sex were tested in four exposure groups at doses of 0, 150, 780 and 1290 mg/kg/day in males and 0, 210, 1060 and 1740 mg/kg/day in females. **EPA**[61], **EFSA**[89] and **Greim et al. (2015)**[91] had concerns with the quality of this study, the characterization of the chemical being used and with tumor rates in this strain of animals being too low. The Supplemental Material from **Greim et al. (2015)** on this study shows no significant increase in any tumor and virtually all animals having no tumors in controls and treated animals.

This study is inadequate for use in deciding on causality for the same reasons given by the **EPA, EFSA** and **Greim et al. (2015)**.

Chruscielska, K. (2000)[71] exposed Wistar rats to glyphosate as a 13.8% solution (purity not given) in drinking water for two years. According to **Greim et al. (2015)**[91], this appears to be the glyphosate formulation Perzocyd. Eighty-five animals/sex were tested in four exposure groups. The authors listed the doses as control, 300 mg/L, 900 mg/L and 2700 mg/L in drinking water. **Greim et al. (2015)**[91] estimated the intake of glyphosate to be 0, 1.9, 5.7 and 17 mg/kg/day for females and 0, 2.2, 6.5, and 19 mg/kg/day in males. There was a slight increase in malignant adenomas of the pituitary gland and an opposite decrease in pituitary adenomas suggesting no effect or potentially a promotional effect in which adenomas are promoted to carcinomas by glyphosate. No other increased tumor responses were reported in the manuscript. Because of the low exposures, this study is an inadequate challenge to the animals (the highest dose is far below the MTD). The reporting of this study is very limited and it the overall quality of the work cannot be evaluated.

This study is inadequate for use in deciding on causality.

Seralini, G. E., et al. (2014)[77] exposed Sprague-Dawley rats to the glyphosate formulation Roundup in drinking water for two years as part of a broader experiment on

Roundup-Ready Corn. Ten animals/sex were tested in four exposure groups at doses of 0, 0.00005, 400 and 22500 mg/L in females. The authors reported an increase in the incidence of mammary gland tumors (mainly fibroadenomas and adenocarcinomas) in female rats with incidences of 5/10 for control and 9/10, 10/10, 9/10 (p_{Fisher}=0.016) in the low-, mid- and high-doses groups respectively. It is difficult to assess the quality of this study due to limited reporting on the histopathological descriptions of the tumors and the very small sample size.

This study will not be used in the evaluation of causality.

Joint Analysis - Rats

Table 8 summarizes the significance for all tumors of interest in rats.

Brammer (2001)[69] saw a significant increase in hepatocellular adenomas in male Wistar rats with increasing dose (p_{Trend}=0.008, Table 4). The other two acceptable studies in Wistar rats (**Wood et al. (2009)**[80] and **Suresh (1996)**[79] did not see significant increases (Tables 5 and 7). On the basis of statistical significance, these studies are inconsistent. To reject these findings based upon only 1/3 being positive is the same as rejecting a coin as being fair if, in three flips of the coin, the result is one head and two tails; it simply is not possible and there is a better way to address these findings. Given different doses and different sample sizes, we need to formally test for consistency in these studies. **Suresh (1996)** saw 48% response for hepatocellular adenomas in controls whereas the other two studies saw no tumors in the control animals. Thus, although all three studies are in Wistar rats, **Suresh (1996)** has a significantly different control response from the other two. **Suresh (1996)** did not give a substrain for the Wistar rats used, but **Brammer (2001)** and **Wood et al. (2009)** used different substrains. All three studies used different diets and were conducted in different facilities. Thus, there is no obvious explanation for the dramatically different rates in **Suresh (1996).** It is known that the same strain of rats from different laboratories can have markedly different control tumor responses. Because they have similar control response, **Brammer (2001)** and **Wood et al. (2009)** can be pooled into a single study to ask the question "Does the significant trend for **Brammer (2001)** disappear when it is pooled with the negative study of **Wood et al. (2009)**?" The analysis of the pooled studies yields p_{Trend}=0.013 supporting the conclusion that glyphosate causes hepatocellular adenomas in Wistar rats with similar background responses.

Wood et al. (2009)[80] saw a significant increase in mammary gland adenomas and adenocarcinomas (p_{Trend}=0.007, Table 7) in females that was not seen in the other two studies (Tables 4 and 6). The background rates in these studies differ only slightly and a pooled analysis of all three studies yields p_{TrendA}=0.459, suggesting that combining the data eliminates the dose-response trend seen in **Wood et al. (2009)**. However, if the Wistar rats used in **Suresh (1996)** differed in their response for hepatocellular adenomas, they may differ for this tumor as well. Combining only **Wood et al. (2009)** with **Brammer (2001)** results in p_{Trend}=0.037. Given the mixed results from the pooling for this tumor I conclude there is limited support for the notion that glyphosate can cause mammary gland adenomas and adenocarcinomas in Wistar rats.

Wood et al. (2009)[80] saw a significant increase in skin keratocanthomas (p_{Trend}=0.030, Table 7) in males that was not seen in the other two studies (Tables 4 and 6). The background rates in these studies differ only slightly and a pooled analysis of all three studies yields p_{TrendA}=0.010, suggesting that combining the data does not eliminate the dose-response trend seen in **Wood et al. (2009)**. Combining only **Wood et al. (2009)** with **Brammer (2001)** results in p_{Trend}=0.053. Given the results from the pooling for this tumor I conclude there is support for the notion that glyphosate can cause skin keratocanthomas in Wistar rats.

In Sprague-Dawley rats, there were four studies that were acceptable for inclusion in the evaluation of causality with one[74] yielding strong positive responses for thyroid C-cell carcinomas in females and testicular interstitial tumors and hepatocellular adenomas in males and another[72] yielding a strong result for kidney adenomas in males. **Lankas (1981)**[74] saw a significant increase in thyroid C-cell carcinomas in female rats exposed to glyphosate (p_{Trend}=0.003, Table 1) and a marginal increase in C-cell adenomas and carcinomas combined (p_{Trend}=0.072, p_{hist}=0.072, Table 1; two of the other three studies also saw marginal results for thyroid C-cell adenomas and carcinomas in females (Tables 2 and 3). A pooled analysis using all four studies yields p_{TrendA}=0.390. This pooled analysis does not support the results seen in **Lankas (1981)**. However, the **Lankas (1981)** study was for 26 months and the other three were for 24 months; the C-cell carcinomas could be a result of the longer exposure period even though the dose is substantially lower in this study compared to the other two. From these data, I conclude that the evidence is weak that glyphosate causes thyroid C-cell tumors in female Sprague-Dawley rats.

Thyroid C-cell adenomas and carcinomas combined, in males, show marginally significant dose-response trends in **Stout and Ruecker (1990**, Table 2) but not in the remaining three studies. Pooling all four studies yields a significant trend of p_{TrendA}=0.041. From these data, I conclude that there is evidence is that glyphosate causes thyroid C-cell tumors in male Sprague-Dawley rats.

Thyroid follicular-cell adenomas and carcinomas combined, in males, show a significant dose-response trend in **Atkinson et al. (1993**, Table 3) but not in the remaining three studies;. Pooling all four studies yields no significant trend with p_{TrendA}=0.618. From these data, I conclude that there is no evidence that glyphosate causes thyroid follicular-cell tumors in male Sprague-Dawley rats.

Hepatocellular adenomas, in males, show a significant dose-response trend in **Stout and Ruecker (1990**, Table 2) but not in the remaining three studies. Pooling all four studies yields a marginally significant trend with p_{Trend}=0.073. From these data, I conclude that there is limited evidence that glyphosate causes thyroid follicular-cell tumors in male Sprague-Dawley rats.

Table 8: Summary of significance tests for 5 tumors from 7 studies in Rats

Brammer (2001)[69]	Wistar	+++[1]	-						
Wood (2009)[80]		-	+++	++					
Suresh (1996)[79]		-	-						
Pooled Wistar Rats		++[2]	++[2]	+++					
Lankas (1981)[74]	Sprague Dawley	-[3]			+	-	-	+++	-
Enemoto (1997)[72]		-			-	-	-	-	+++
Atkinson et al. (1993)[68]		-			-	-	++	-	-
Stout and Ruecker (1990)		++			-	+	-	-	-
Pooled Sprague-Dawley Rats		+			-	++	-	-	++[4]

[1]entries are p_{Trend}/p_{Hist} with values: – p>0.1, + 0.1≥p>0.05, ++ 0.05≥p>0.01, +++ p≤0.01; [2]pooling results from **Brammer (2001)** and **Wood (2009)** only; [3]liver neoplastic nodules; [4]excluding **Lankas (1981)**

Another significant trend seen in Sprague-Dawley rats is the finding of testes interstitial cell tumors from **Lankas (1981)**[74] (P_{Trend}=0.009, Table 1); the other three studies were negative for this tumor (Tables 2, 3 and 6). Combining the other three studies with that of **Lankas (1981)** for testes interstitial tumors results in a p-value for trend that is clearly non-significant (p_{TrendA}=0.608). However, as noted above, the **Lankas (1981)** study was for 26 months and the other two were for 24 months; the tumors could be a result of the longer exposure period even though the dose is substantially lower in this study compared to **Stout and Ruecker (1990)**, **Atkinson et al.(1993)** and **Enemoto (1997)**.

The final tumor in Sprague-Dawley rats showing a strong significant trend is kidney

adenomas in males from the study by **Enemoto (1997)**[72] (P_{Trend}=0.004, Table 6). The kidney tumor data is not significant for the studies by **Lankas (1981)**[74] (Table 1), **Atkinson et al. (1993)**[99] (Table 3) and **Stout and Ruecker (1990)**[78] (Table 2). Pooling the **Enemoto (1997)** study with that of **Lankas (1981)**[74], **Stout and Ruecker (1990)** and **Atkinson et al. (1993)** yields p_{TrendA}=0.201. Removing the 26-month study by **Lankas (1981)**[74] yields a p-value for the three combined 24-month studies of p_{Trend}=0.031; thus, the association between glyphosate and kidney adenomas in male Sprague-Dawley rats is supported by these data, even with the difficulty associated with interpreting the results in the low- and mid-doses in the **Atkinson et al. (1993)** study. There is evidence to support an increase in kidney tumors in male Sprague-Dawley rats exposed to glyphosate.

In summary, there is evidence that glyphosate causes hepatocellular adenomas and skin keratocanthomas in male Wistar rats, mammary gland adenomas and adenocarcinomas in female Wistar rats and kidney adenomas and thyroid C-cell adenomas and carcinomas in male Sprague-Dawley rats. There is limited evidence glyphosate causes hepatocellular adenomas in male Sprague-Dawley rats.

Mouse Studies

Reyna and Gordon (1974)[86] exposed Swiss White mice to glyphosate (>97% purity) in feed for 16 months in males and 18 months in females. Fifty animals/group/sex were tested in three exposure groups; control, 17 mg/kg and 50 mg/kg. Only 10 animals per group were examined for histopathological changes.

There was no impact on survival of administration of glyphosate and no indication that the high dose exceeded the MTD.

No significant increases were seen in any tumor from this study. However, given the small sample size for histopathological evaluation and the low doses used for this study, this study is inadequate.

This study will not be used in the evaluation of causality.

Knezevich and Hogan, (1983)[83] exposed CD-1 mice to glyphosate (99.8% pure) in feed for two years. Fifty animals/group/sex were tested in four exposure groups (see Table 9).

There were no survival differences in this study and there was no indication that the highest dose used exceeded the MTD.

EPA[100] found a significant increase in kidney tubular cell adenomas in male mice based upon the original pathology done from the study and this analysis is shown in Table 9 (p_{Trend}=0.019). Kidney tubular cell adenomas are very rare tumors in CD-1 mice so it is important to compare these results with the historical controls. No historical controls were available from the laboratory that conducted **Knezevich and Hogan, (1983)** so IARC, EPA and EFSA all used historical control databases from published studies in the

literature[101-103]. These studies have virtually identical rates for the important tumors seen in CD-1 mice; I will use the study by **Giknis and Clifford (2000)**[102] since it best covers the range of studies we have for CD-1 mice. For studies of approximately two years, the mean historical tumor response in controls is 0.27%. Applying this control response rate to the kidney adenomas yields p_{Hist}=0.005, strengthening the significance of the evaluation against the concurrent control. EPA originally used a similar analysis and reached the same conclusions. However, in 1985, the registrant had a group of pathologists review the kidney slides. Using additional kidney sections from this study, the pathologists identified an additional adenoma in the control animals and changed the classification for three adenomas to carcinomas (Table 9). With these changes, the adenomas no longer have a significant trend (P_{Trend}=0.442, P_{Hist}=0.121) but carcinomas have a marginally significant trend against concurrent controls and a clearly significant trend using historical controls (p_{Trend}=0.063, p_{Hist}=0.002, historical control rate of 0.15%). These historical control rates may not apply to this analysis because the reevaluation of the kidney tumors considered additional sections and no information is available on how additional sections affect historical control rates in this strain of mice; differences have been seen in other settings[104]. The incidence of combined carcinomas and adenomas has the same marginal significance against the concurrent control and significance against the historical controls (p_{Trend}=0.065, p_{Hist}=0.011, historical control rate of 0.44%). However, there was considerable disagreement on whether the one adenoma in the control group was correctly diagnosed[105]. Removing this one adenoma from the control group results in p_{Trend}=0.019 and p_{Hist}=0.005.

Other CD-1 mouse studies have seen increases in malignant lymphomas, hemangiosarcomas and lung adenocarcinomas (males) and hemangiomas (females). Evaluations of those tumors for this study yields results that are not significant; for malignant lymphoma, p_{Trend}=0.754, p_{Hist}=0.767, with the historical control rate equal 6.2%, for hemangiosarcomas p_{Trend}=0.503, p_{Hist}=0.591, with the historical control rate equal to 2.5%, for lung adenocarcinomas p_{Trend}=0.918, p_{Hist}=0.899, with the historical control rate equal to 9.2% and for hemangiomas p_{Trend}=0.631. No other tumors were found in this study.

The **EPA**[61] has produced many different arguments to dismiss the findings of renal tumors from this study. One argument is that the pathology working group requested by the EPA in 1986 concluded these lesions were not glyphosate related because *"1) renal tubular cell tumors are spontaneous lesions for which there is a paucity of historical control data for this mouse stock; 2) there was no statistical significance in a pairwise comparison of treated groups with the concurrent controls and there was no evidence of a statistically significant linear trend; 3) multiple renal tumors were not found in any animal; and 4) compound-related nephrotoxic lesions, including pre-neoplastic changes, were not present in male mice in this study."* Reason number one no longer exists as there are two very good historical control databases for CD-1 mice[101, 102]. The second reason, while technically correct, is not supportable since the Agency's own guidelines for evaluating carcinogenicity studies state that *"Significance in*

either kind of test [trend or pair-wise] *is sufficient to reject the hypothesis that chance accounts for the result."* The third reason is also weak since one would not expect (nor require) multiple tumors to appear when dealing with a rare tumor. For the fourth point, EPA provides data on the rate of bilateral chronic interstitial nephritis in the study which it considers to show no statistically significant results although the trend test is highly significant (p_{Trend}=0.006, Table 9). EPA then states, without reference, that *"chronic interstitial nephritis is not considered to be a precursor lesion for tubular neoplasms"*. I could find no published research to either support or refute this statement. However, chronic interstitial nephritis is an inflammation of the interstitial tissue surrounding the glomeruli and tubules in the kidney. Inflammation is well known

Table 9: Tumors of interest in male and female CD-1 mice from the 24-month feeding study of **Knezevich and Hogan (1983)**[83]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	157	814	4841	
	Female	0	190	955	5874	
Kidney Adenoma[1] (original pathology)	Male	0/49	0/49	1/50	3/50	P_{Trend}=0.019 P_{Hist}=0.005
Kidney Adenoma (EPA pathology)	Male	1/49	0/49	0/50	1/50	P_{Trend}=0.442 P_{Hist}=0.121
Kidney Carcinoma[2] (EPA pathology)u	Male	0/49	0/49	1/50	2/50	P_{Trend}=0.063 P_{Hist}=0.002
Kidney Adenoma and Carcinoma Combined[3] (EPA pathology)	Male	1/49	0/49	1/50	3/50	P_{Trend}=0.065 P_{Hist}=0.011
Malignant Lymphoma[4]	Male	2/49	5/49	4/50	2/50	P_{Trend}=0.754 P_{Hist}=0.767
Hemangiosarcoma[5]	Male	0/50	0/49	1/50	0/50	P_{Trend}=0.503 P_{Hist}=0.591
Bilateral Chronic Interstitial Nephritis	Male	5/49	1/49	7/50	11/50	P_{Trend}=0.006
Hemangiooma[6]	Female	0/49	1/49	1/50	0/50	P_{Trend}=0.631
Lung Adenocarcinoma[7]	Male	4/48	3/50	2/50	1/50	P_{Trend}=0.918 P_{Hist}=0.899

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01, [1]historical rate=0.27%, [2]historical rate=0.15%, [3]historical rate=0.44%, [4]historical rate=6.2%, [5]historical rate=2.5%, [6]No Historical Controls, [7]Historical rate=9.2%

to play an important role in kidney cancer[106] and many other cancers so this argument also fails to support rejection of these findings.

In summary, this study shows a positive result for kidney tumors in male CD-1 mice and will be included in the overall evaluation of causation.

Atkinson, et al., (1993)[81] exposed CD-1 mice to glyphosate (>97% purity) in feed for two years. Fifty animals/group/sex were tested in four exposure groups (see Table 10).

There was no impact on survival of administration of glyphosate and no indication that the high dose exceeded the MTD.

Table 10: Tumors of interest in male and female CD-1 mice from the 24-month feeding study of **Atkinson et al. (1993)**[81]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	98	297	988	
	Female	0	102	298	1000	
Kidney Adenoma and Carcinoma Combined[1]	Male	2/50	2/50	0/50	0/50	P_{Trend}=0.981 P_{Hist}=1
Malignant Lymphoma[2]	Male	4/50	2/50	1/50	6/50	P_{Trend}=0.087 P_{Hist}=0.085
Hemangiosarcoma[3]	Male	0/50	0/50	0/50	4/50	P_{Trend}=0.004 P_{Hist}=0.001
Hemangioma[4]	Female	0/50	0/50	0/50	0/50	P_{Trend}=1
Lung Adenocarcinoma[5]	Male	10/50	7/50	8/50	9/50	P_{Trend}=0.456 P_{Hist}=0.449

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01, [1]historical rate=0.44%, [2]historical rate=6.2%, [3]historical rate=2.5%, [4]No historical control rate, [5]Historical rate=9.2%

Hemangiosarcomas were the only tumors showing a significant trend in this study (P_{Trend}=0.004, P_{Hist}=0.001, Table 10). Also shown in Table 10 are the results for malignant lymphomas, kidney tumors and lung adenocarcinomas (males) and hemangioma (females); there is a marginal trend for malignant lymphomas (P_{Trend}=0.087, P_{Hist}=0.085) and no trend for kidney tumors.

The **EPA**[61] concluded the findings in this study were not treatment related based upon the tumors appearing only in the high dose group, a lack of statistical significance between the response in this group and control response and that these tumors are commonly observed in mice as both spontaneous and treatment related effects. There is no scientific support for excluding positive findings in the highest dose group, a view also held by the **SAP**[54]. I have already commented on how EPA's guidelines treat trend tests and Fisher's Exact test results, although in this case, the value of the comparison of the highest exposure group to controls, p_{Fisher}=0.059, is marginally significant. The argument regarding the frequency of this tumor in controls is addressed directly by the evaluation against the historical control rates; if these rates were high enough to exclude this finding, P_{Hist} would have be above 0.05 instead of 0.001. The mean

historical control incidence of hemangiosarcomas in controls from two-year cancer bioassays in CD-1 mice is 2.5% and the response seen in the high-dose group is 8.9%. The **SAP**[54] stated very clearly that the practice, being used by the EPA, of negating a positive finding because of historical control data was not acceptable[54]. (page 63). The EPA Cancer Guidelines[33] state this very clearly *"...statistically significant increases in tumors 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 somewhat lower than average."*

In summary, this study shows a positive result for hemangiosarcomas in male CD-1 mice and will be included in the overall evaluation of causation.

Wood et al., (2009)[88] exposed CD-1 mice to glyphosate (95.7% pure) in feed for 80 weeks. Fifty-one animals/groups/sex were tested in four exposure groups (see Table 11).

There was no effect on survival and no information suggesting the study exceeded the MTD.

No increase in kidney tumors or hemangiosarcomas (males) or hemangiomas (females) were seen in this study. There was a monotonic increase in lung adenocarcinomas (p_{Trend}=0.028, p_{Hist}=0.031) in males and a monotonic increase in malignant lymphomas (p_{Trend}=0.007, p_{Hist}=0.007) in males. The historical control incidence for this study is different from the earlier studies because this study is only for 80 weeks instead of 104 weeks (two years); the historical control rate for malignant lymphomas in CD-1 mice after 80 weeks is 2.6% instead of 6.2%, the historical control rate at two years[102].

For lung adenocarcinomas, the **EPA**[61] again argued a lack of significance for pairwise comparisons (in violation of its guidelines) and that there was no evidence of progression from adenomas to carcinomas. Even though there was no increase in lung adenomas as a function of exposure, it is possible to have an increase in lung adenocarcinomas without an associated increase in adenomas[107]. For malignant lymphomas, EPA notes that there was a statistically significant response and that the high dose was significantly different from control (p_{Fisher}=0.028), but then uses an argument based upon the number of analyses done in this study to adjust the Fisher Exact test p-value to 0.082 (an adjustment for multiple comparisons is indeed warranted in evaluating the outcomes of these animal cancer studies, this will be addressed later in my report in the evaluation of all of the studies combined).

The **EPA**[61] uses historical control data[103, 108] to exclude the malignant lymphomas and cite a mean response of 4.5% and a range of 1.5% to 21.7%. **Son and Gopinath (2004)**[108] saw 21 animals out of 1453 examined prior to 80 weeks with lung adenocarcinomas (1.4%). **Giknis and Clifford (2005)**[103] saw a mean rate of 4.5% with a range of 0% to 21.7% in 52 studies which included mostly 78 week controls (26 studies) and 104 week controls (21 studies). Including only studies of 80 weeks or less, the rate in **Giknis and Clifford (2005)** is 37/1372=2.7% with a range of 0% to 14%. **Giknis and Clifford (2000)**[102] (the reference I have been citing) did a similar evaluation, using mostly the same data as their 2005 paper and saw an average tumor incidence before

80 weeks of 2.6% with a range of 0% to 14%. Based upon its flawed interpretation of the **Giknis and Clifford (2005)** historical controls, EPA argues that the incidence of concurrent controls in the study was low (it was 0%) and rejected the positive finding. In fact, of the 26 studies in the 18-month control groups evaluated by **Giknis and Clifford (2005),** eight (31%) had response of 0% and eight (31%) had only one tumor. The evaluation used by the EPA is incorrect. In addition, as noted earlier, the use of historical control data to negate a positive finding is not supported by **EPA's guidelines**[33, 54] or its **SAP**[54].

There was an increase in the number of animals with multiple malignant tumors (P_{Trend}=0.046)

In summary, this study shows a positive result for malignant lymphomas and lung adenocarcinomas in male CD-1 mice and will be included in the overall evaluation of causation.

Table 11: Tumors of interest in male and female CD-1 mice from the 18-month feeding study of **Wood et al. (2009)**[88]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	71.4	234.2	810	
	Female	0	97.9	299.5	1081.2	
Kidney Adenoma[1]	Male	0/51	0/51	0/51	0/51	P_{Trend}=1
Malignant Lymphoma[2]	Male	0/51	1/51	2/51	5/51*	P_{Trend}=0.007 P_{Hist}=0.007
Hemangiosarcoma	Male	0/51	0/51	0/51	0/51	P_{Trend}=1
Lung Adenocarcinoma[3]	Male	5/51	5/51	7/51	11/51	p_{Trend}=0.028 P_{Hist}=0.031
Hemangioma[4]	Female	0/51	2/51	0/51	1/51	p_{Trend}=0.438
Animals with Malignant Neoplasms	Male	14/51	20/51	17/51	20/51	P_{Trend}=0.203
Animals with Malignant Neoplasms	Female	23/51	15/51	17/51	18/51	P_{Trend}=0.628
Animals with multiple malignant tumors	Male	1/51	2/51	3/51	5/51	P_{Trend}=0.046

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01, [1]historical rate=0.44%, [2]historical rate=2.6%, [3]Historical rate=2.5%, [4]No Historical Control Rate

Sugimoto (1997)[87] exposed CD-1 mice to glyphosate (94.61-95.67% pure) in feed for two years. Fifty animals/group/sex were tested in four exposure groups (see Table 12).

There were no effects of treatment on survival and no indication the highest dose had exceeded the MTD.

Kidney adenomas (p_{Trend}=0.062, p_{Hist}=0.005), malignant lymphomas (p_{Trend}=0.016,

p_{Hist}=0.017) and hemangiosarcomas (p_{Trend}=0.062, p_{Hist}=0.004) in male mice and hemangiomas (p_{Trend}=0.002) in female mice all showed increased tumor incidence with increasing dose. The evaluation of lung adenocarcinomas in males showed no significant dose-related trend (p_{Trend}=0.148, p_{Hist}=0.140). This study also had an increase in animals with any malignancy in males (p_{Trend}=0.001) but not in females (p_{Trend}=0.362). Note that no hemangiosarcomas were seen in the 26 control groups evaluated by **Giknis and Clifford (2000)** so the development of an estimate of the historical control response is difficult (if the historical control rate is 0, then any observed response other than 0 has a p-value of 0). The fact that this tumor was never seen in the historical controls should strongly support any positive finding as being significant. However, to still allow for a test using historical control data, I used the historical control estimate of the mean response that would result in a 5% chance of seeing no tumors in 1149 animals. This estimated historical control response value was 0.0026. This value was used in the analysis for hemangiosarcomas in male CD-1 mice exposed for 18 months (p_{Hist} <0.001).

Table 12: Tumors of interest in male and female CD-1 mice from the 18-month feeding study of **Sugimoto (1997)**[87]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	165	838.1	4348	
	Female	0	153.2	786.8	4116	
Kidney Adenoma[1]	Male	0/50	0/50	0/50	2/50	P_{Trend}=0.062 P_{Hist}=0.005
Malignant Lymphoma[2]	Male	2/50	2/50	0/50	6/50	P_{Trend}=0.016 P_{Hist}=0.017
Hemangiosarcoma[3]	Male	0/50	0/50	0/50	2/50	P_{Trend}=0.062 P_{Hist}=0.004
Hemangioma[4]	Female	0/50	0/50	2/50	5/50*	P_{Trend}=0.002
Lung Adenocarcinoma[5]	Male	1/50	1/50	6/50	4/50	P_{Trend}=0.148 P_{Hist}=0.140
Number of animals with Malignant Neoplasms	Male	5/50	5/50	11/50	16/50**	P_{Trend}=0.001
Number of animals with Malignant Neoplasms	Female	9/50	13/50	16/50	13/50	P_{Trend}=0.362

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01, [1]historical rate=0.44%, [2]historical rate=2.6%, [3]historical rate=0/1424 (0.26% - 95% confidence limit), [4]No Historical Control Rate, [5]Historical rate=2.5%

EPA[61] only addressed the hemangiomas in the female mice and did not note any other significant effects. For the females, EPA argued that the high dose was approximately four times higher than the current recommended high dose from the **OECD guidelines**[109]. This study was correctly designed under the previous guidelines (the limit was <5% in feed) and there is no indication that this dose exceeded the MTD. The EPA also argued that when the p-value for Fisher's Exact test was adjusted for multiple comparisons, the new p-value for the high-dose group for hemangiomas was 0.055.

For the hemangiosarcomas in males, none of the 26 historical control groups examined by **Giknis and Clifford (2000)** had hemangiosarcomas, making this a very rare tumor in males prior to 80 weeks on study. The malignant lymphomas in males are statistically significant against both the concurrent controls and the historical controls. Finally, there is clearly an overall increase of malignancies in the males.

In summary, this study shows a positive result for kidney adenomas, malignant lymphomas and hemangiosarcomas in male CD-1 mice, hemangiomas in female CD-1 mice and an overall increase in malignancies as a function of exposure in male CD-1 mice. This study will be included in the overall evaluation of causation.

Kumar (2001)[84] exposed Swiss Albino mice to glyphosate (>95% purity) in feed for two years. Fifty animals/group/sex were tested in four exposure groups (see Table 13).

The survival was decreased in the highest exposure group but this was not statistically significant and there was no other data indicating the MTD was exceeded for this study.

Kidney adenomas (p_{Trend}=0.062) and malignant lymphomas (p_{Trend}=0.064, p_{Hist} =0.070) in male mice demonstrated marginal statistical significance and hemangiosarcomas (p_{Trend}=0.500) in male mice demonstrated no statistical significance. In this study, not all animals in the low- and mid- dose groups were evaluated for kidney tumors, so a second analysis was done based on only the animals examined in these two groups (p_{Trend}=0.088). No historical control data was available for hemangiosarcomas and kidney adenomas in Swiss Albino mice. For the malignant lymphomas, EFSA provided a historical control data set showing a mean response of 46/250=0.184 (18.4%) with a range of 6% to 30%. Using this historical control data, the trend is only marginally significant (p_{Hist}=0.070). I have some concern that the responses at two of the doses are outside of the historical control range and the third dose is at the upper limit of the historical control range. However, this is a small historical control dataset for a tumor with a relatively high background tumor rate, thus placing too much emphasis on this historical control population is not warranted.

In a recent memo, Martens (2017) [110] asserts that the incidence counts for malignant lymphomas and kidney adenomas appearing in Greim et al. (2015)[91] and EFSA (2013)[89] are incorrect and provides different rates (shown in Table 13). The p-values for both of these tumors are reduced using the incidence counts from the Martens memo. However, it should be noted that if the counts for malignant lymphomas in the Martens (2017) memo are correct, then all three exposure groups have responses outside of the range of the historical controls. It is unclear from Greim et al. (2015), EFSA or Martens (2017) which tumor incidence counts are correct.

<u>There was a significant increase in hemangiomas (any tissue) in female mice</u> (p_{Trend}=0.004).

In summary, this study shows support for an increase for malignant lymphomas and kidney adenomas as a function of exposure in male Swiss Albino mice and an increase in hemangiomas in female Swiss Albino mice. This study will be included in the overall

Table 13: Tumors of interest in male and female Swiss Albino mice from the 18-month feeding study of **Kumar (2001)**[84]

Tumor	Sex	Doses (mg/kg/day)				p-values
	Male	0	14.5	149.7	1453	
	Female	0	15	151.2	1466.8	
Kidney Adenoma (only tissues examined microscopically)	Male	0/50	0/26	1/22	2/50	P_{Trend}=0.088
Kidney Adenoma (as reported by Greim et al.)	Male	0/50	0/50	1/50	2/50	P_{Trend}=0.062
Kidney Adenoma (as reported by Martens)	Male	0/50	0/50	0/50	1/50	P_{Trend}=0.250
Malignant Lymphoma[1] (as reported by Greim et al.)	Male	10/50	15/50	16/50	19/50	P_{Trend}=0.064 P_{Hist}=0.070
Malignant Lymphoma[1] (as reported by Martens)	Male	10/50	16/50	18/50	19/50*	P_{Trend}=0.141 P_{Hist}=0.150
Hemangiosarcoma	Male	0/50	0/50	2/50	0/50	P_{Trend}=0.500
Hemangioma (any tissue)	Female	1/50	0/50	0/50	5/50	P_{Trend}=0.004

*- p_{Fisher}<0.05, **- p_{Fisher}<0.01, [1]Historical control rate=0.184 (46/250 mice)

evaluation of causation.

Pavkov and Turner (1987)[85] exposed CD-1 mice to glyphosate trimesium salt (56.2%) and 1% propylene glycol (wet weight vehicle) in feed for two years. Eighty animals/sex/group were tested in control, low- and mid-dose groups and 90 animals/sex were tested at the high dose. Exposure levels were 0, 11.7, 118 and 991 mg/kg/day in males and 0, 16, 159 and 1341 mg/kg/day in females. **EPA**[61] lists this study as completely negative for any cancer findings. No details on this study are provided by the EPA nor is it listed in the **Greim et al. (2015)**[91] manuscript. There was limited information on this study in a Data Evaluation Report from EPA (accession number 4021 40-06) that discussed findings from this study. EPA noted that body weight and food consumption were reduced in the highest exposure group, but the actual amounts of these reductions were not available. They also noted that the authors failed to make it clear that the tumors reported in the study had been histopathologically validated. Data was presented for tumors in the livers and lungs of male mice and the lungs of female mice. No other data is provided.

This study is not acceptable for inclusion in the evaluation of causation due to the lack of information on the tumor incidence in tissues other than liver and lung.

George et al. (2010)[82] exposed groups of 20 male Swiss Albino mice to a glyphosate

formulation (Roundup Original, 36g/L glyphosate) at a dose of 25 mg/kg (glyphosate equivalent dose) topically three times per week, topically once followed one week later by 12-o-tetradecanoylphorbol-13-acetate (TPA) three times per week, topically three times per week for three weeks followed one week later by TPA three times per week, or a single topical application of 7,12-dimethyl-benz[a]anthracene (DMBA) followed one week later by topical application of glyphosate three times per week for a total period of 32 weeks. Appropriate untreated, DMBA-treated, and TPA-treated controls were included. The group exposed to DMBA followed by glyphosate demonstrated a significant increase ($p<0.05$) in the number of animals with tumors (40% of the treated animals versus no tumors in the controls) indicating glyphosate has a promotional effect on carcinogenesis in the two-stage model in skin. This study addresses the question of whether glyphosate is more likely to cause skin tumors through initiation (starting the cancer process) or promotion (moving the process along after it starts). This study supports the overall concept that glyphosate can have an impact on tumor incidence.

EPA[61] discounted this study because it included only 20 animals per group, tested only males and did not conduct a histopathological analysis. It is hard to understand how EPA could reject a positive finding using 20 mice; typically one would ignore a negative study that had too few animals as not having sufficient statistical power to see an effect but never reject positive findings for this reason. Also, 20 animals per group is common for skin-painting initiation-promotion studies like the one presented here. Doing a study in only males is not a reason to ignore the positive findings in a study. Finally, in initiation-promotion studies of mouse skin, histopathological evaluation would be done if one were interested in separating papillomas from carcinomas. It is highly unlikely that the lesions seen in 40% of the DMBA/glyphosate treated mice were not papillomas or carcinomas.

Some members of the EPA SAP noted[54] that the rodent data were consistent with glyphosate acting as a tumor promoter but, because *"[t]here has been no direct test of this hypothesis (such as in a standard initiation-promotion bioassay)...,"* this *"conclusion was speculative."* (page #). Because the EPA dismissed this study without any discussion, the SAP did not recognize there was an initiation-promotion supporting a promotional effect of glyphosate.

This study is included in the evaluation of causality as support for a promotional effect of glyphosate on some tumors.

Joint Analysis - Mouse

In their evaluation of the mouse studies, **EPA**[61] and **EFSA**[89] chose to challenge the results in each study separately, dismiss the studies as showing no effect, and never compared results across the various studies. In response to the evaluation done by the **IARC**[30], **EFSA**[90] extracted the original data and did trend tests on kidney tumors, malignant lymphomas and hemangiosarcomas in male mice in five of the mouse studies, the same five studies I consider acceptable for a causation analysis. Rather than formally evaluate these cancer responses for consistency by pooling the data where appropriate, EPA and EFSA simply produced a table with the responses for each dose

group in each study and concluded (subjectively) they were inconsistent. In addition, EPA and EFSA argued that doses above 1000 mg/kg/day (there are only two of these) were outside the range of what would be tested today under OECD guidelines and should be excluded. I will now address both points.

In CD-1 mice, there are four useful animal carcinogenicity studies and one study in Swiss Albino mice. As with the rats, consistency across studies can be addressed in two ways. The first is by simply looking at the overall findings to evaluate where they agree or disagree in terms of statistical significance. Table 14 summarizes the positive and negative findings for all five cancers in which at least one study in CD-1 mice showed a significant trend. It is clear that not every tumor shows a positive trend with glyphosate exposure in every study. For hemangiosarcomas in males, there are clear positive findings in the studies by **Sugimoto (1997)** and **Atkinson et al. (1993)** and non-significant responses in **Wood et al. (2009)** and **Knezevich and Hogan (1983)**. In females, hemangiosarcomas are only present in the study by **Sugimoto (1997)**. Malignant lymphomas in males are clearly positive in two studies[87, 88] and marginally positive in a third[81] but negative in the fourth[83]. Both of the strong positive studies exposed animals for 18 months. Kidney tumors in males are positive in two studies[83, 87] and negative in the remaining two[81, 88]. Lung adenocarcinomas in males are only positive in the study by **Wood et al. (2009)**. **Sugimoto (1997)** had four clearly positive associations between tumors and glyphosate while the others had two or less.

Table 14: Summary of significance tests for 5 tumors from 4 studies in CD-1 Mice

Sugimoto 1997[87]	18	+/+++[1]	+++	++/++	+/+++	-/-
Wood 2009[88]	18	-/-	-	+++/+++	-/-	++/++
Sugimoto & Wood Pooled		++/+++	+++	+++/+++	++/+++	-/-
Atkinson 1993[81]	24	+++/+++	-	+/+	-/-	-/-
Knezevich 1983[83]	24	-/-	-	-/-	+/++	-/-
Atkinson & Knezevich Pooled		-/-	-	-/-	+/+	-/-
All CD-1 Studies Pooled		++/++	++/++	+/+	+++/+++	-/-

[1]entries are p_{Trend}/p_{Hist} with values: – $p>0.1$, + $0.1 \geq p > 0.05$, ++ $0.05 \geq p > 0.01$, +++ $p \leq 0.01$

As seen for the rat studies, this simple evaluation of the positive versus negative findings fails to resolve the issue of which findings are driving the overall responses in these data. To do this, I will again pool the studies. Table 14 summarizes the pooled analyses.

For kidney tumors in males, pooling the two 18-month studies yields significant increases in incidence (p_{Trend}=0.015, p_{Hist}=0.003) and pooling of the two year studies shows marginal significance (p_{Trend}=0.081, p_{Hist}=0.054). Pooling all four studies results in (p_{Trend}=0.005, p_{Hist}=0.007), thus the positive trend remains. **Knezevich and Hogan (1983)** saw a 4% response for kidney carcinomas in their highest exposure group. The largest response seen for kidney carcinomas in controls in 48 studies by **Giknis and Clifford (2000)** and in 52 studies by **Giknis and Clifford (2005)** was 2% and in the control groups from 11 two-year cancer studies, **Chandra and Frith (1992)**[101] saw only one animal out of 725 with a kidney carcinoma. In 46 control datasets, **Giknis and Clifford (2000)** saw 39 control groups with no adenomas, five with one adenoma and two with two adenomas; both 24-month studies saw two adenomas in the highest exposure group, a very rare finding. To better illustrate, there are 16 groups of animals in the four studies. For any one group, there is a 2/44 or 4.3% chance of getting a response 4% or larger. The chances of randomly getting 3 or more such responses in 16 groups is 2.9% and the chances of two of these being in any two of the four highest exposure groups is 0.01. In summary, the strong finding in two of the four studies, the positive finding when all four studies are pooled and the very low probability that this is due to chance when compared to historical controls support the conclusion that glyphosate causes kidney tumors in male mice.

For malignant lymphomas in males, pooling the two 18-month studies, **Sugimoto (1997)** and **Wood et al. (2009)**, results in a significant trend (p_{Trend}=0.005, p_{Hist}=0.006). Pooling the two 24-month studies, **Knezevich and Hogan (1983)** and **Atkinson et al. (1993)**, yields (p_{Trend}=0.653, p_{Hist}=0.649). The main differences between these two findings is in the control response; the pooled control response at 24 months is 6/99 (6%) versus 2/101 at 18 months (2%). This is expected since, in the absence of any exposure, tumor rates increase as a function of age[5]. **Giknis and Clifford (2000)** show a control response at 18 months of 4% and a control response at 24 months of 6% (matching the value for the pooled studies). Pooling all four studies results in (p_{TrendA}=0.073, p_{Hist}=0.080). However, the responses seen for malignant lymphomas in controls by **Giknis and Clifford (2000)** show only one historical control group in twenty-six 18-month groups with 10% or higher response. The responses at the high doses (10% and 12%) in the two 18-month studies are very unlikely to have arisen by chance. There are eight groups of animals in the two studies. For any one group, there is a 1/26 or 3.8% chance of getting a response of at least 10% based on the 26 control groups from **Giknis and Clifford (2000)**. The chances of getting two or more such responses in eight groups is 0.035 and the chances of these being in three of the four highest exposure groups is 0.004. For the 24-month studies, the higher background rate makes it difficult to identify a small change in incidence, thus the findings in the 24-month studies and the 18-month studies are not inconsistent. In summary, the very strong findings in the 18-month studies, the very strong positive findings when the two 18-month studies are pooled, the low probability that the responses seen in the 18-month studies are due to chance, and the

marginal increase in malignant lymphomas in the 18-month study in Swiss Albino mice[84] support the conclusion that glyphosate causes malignant lymphoma in male mice.

For hemangiosarcomas in males, pooling the two 18-month studies results in a significant trend (p_{Trend}=0.015, p_{Hist}=0.002). Pooling the two 24-month studies yields (p_{Trend}=0.490, p_{Hist}=0.429). The main difference between these two findings is the 0/50 response in animals exposed at 4841 mg/kg/day in the study by **Knezevich and Hogan (1983)**. Removing this one exposure group in the pooled 24-month analysis yields (p_{Trend}<0.001, p_{Hist}<0.001). Pooling all four studies results in (p_{Trend}=0.045, p_{Hist}=0.043). No hemangiomas were seen in controls groups from twenty-six 18-month studies by **Giknis and Clifford (2000)** so the two hemangiosarcomas seen in the high dose group in the study by **Sugimoto (1997)** are biologically very significant. For the 24-month historical controls, only two out of 20 control groups had a response greater than 8%. In summary, the very strong findings in the 18-month studies, the positive finding when all four studies are pooled and the low probability that the responses seen in the 18-month studies are due to chance support the conclusion that glyphosate causes hemangiosarcomas in male CD-1 mice.

For hemangiomas in females, pooling the two 18-month studies results in a significant trend (p_{Trend}=0.001). Pooling the two-year studies results in p_{Trend}=0.424. Pooling all four studies results in p_{Trend}=0.018. In summary, the very strong findings in one 18-month study, the positive finding when all four studies are pooled and the low probability that the responses seen in the **Sugimoto (1997)** study are due to chance, support the conclusion that glyphosate causes hemangiomas in female CD-1 mice.

For lung adenocarcinomas in male CD-1 mice, pooling the two 18-month studies results shows no significant trend (p_{Trend}=0.417, p_{Hist}0.126). Pooling the two 24 month studies yields (p_{TrendA}=0.985, p_{Hist}=0.993). Pooling all four studies results in (p_{TrendA}=0.937, p_{Hist}=0.744). In summary, the moderate findings in one 24 month study, and the negative finding when any studies are pooled suggest that the linkage between glyphosate and lung adenocarcinomas in male CD-1 mice is due to chance.

The one study in Swiss Albino mice[84] was effectively negative for all endpoints except malignant lymphomas and kidney adenomas where marginally significant tumor responses were seen. Considering the findings for kidney adenomas in CD-1 mice, glyphosate may also cause kidney adenomas in male Swiss Albino mice from the study of **Kumar (2001)**.

To summarize the findings in mice, glyphosate causes hemangiosarcomas, kidney tumors and malignant lymphomas in male CD-1 mice and hemangiomas in female CD-1 mice after 18 months of exposure, kidney tumors in male CD-1 mice after 24 months exposure and possibly kidney adenomas in male Swiss albino mice. When 18-month and 24-month studies are pooled, there is a significant increase in hemangiosarcomas in male mice, hemangiomas in female mice and kidney tumors in male mice.

Discussion and Summary Animal Carcinogenicity Studies

As noted earlier, there has been a suggestion that using doses substantially larger than 1000 mg/kg/day exceeds the current limit dose set by the OECD. The only place in the **OECD guidance**[67] that addresses a dose of 1000 mg/kg/day is in paragraph 23 which reads:

> *"For the chronic toxicity phase of the study, a full study using three dose levels may not be considered necessary, if it can be anticipated that a test at one dose level, equivalent to at least 1000 mg/kg body weight/day, is unlikely to produce adverse effects. This should be based on information from preliminary studies and a consideration that toxicity would not be expected, based upon data from structurally related substances. A limit of 1000 mg/kg body weight/day may apply except when human exposure indicates the need for a higher dose level to be used."*

This language does not preclude the use of a dose exceeding 1000 mg/kg/day nor does it advocate ignoring such doses when evaluating the results of an animal carcinogenicity study. In fact, the reasons for excluding a dose in an animal carcinogenicity study are clearly outlined in paragraph 90 within **OECD guidance**[59] and reads:

> *"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."*

While one study has a slight decrease in body-weight gain, there are no indications in any other studies of an exceedance in dose that would support ignoring the findings from any exposure group.

EPA[33] uses a slightly different criteria to determine which dose to include or exclude based on an earlier OECD document. These are spelled out in EPA's guideline document for carcinogenicity risk assessment[33]

> *"Other signs of treatment-related toxicity associated with an excessive high dose may include (a) significant reduction of body weight gain (e.g., greater than 10%), (b) significant increases in abnormal behavioral and clinical signs, (c) significant changes in hematology or clinical chemistry, (d) saturation of absorption and detoxification mechanisms, or (e) marked changes in organ weight, morphology, and histopathology. It should be noted that practical upper limits have been established to avoid the use of excessively high doses in long-term carcinogenicity studies of environmental chemicals (e.g., 5% of the test substance in the feed for dietary studies or 1 g/kg body weight for oral gavage studies [OECD, 1981])."* As before, this applies to only one study presented in this review.

Both of these guidelines make good scientific sense. In the 12 acceptable rodent carcinogenicity studies included in this evaluation, no study had sufficient toxicity at the highest dose to justify removing the highest dose from the analysis. Hence, the analyses presented here did not drop the doses >1000 mg/kg/day. This is also supported by one member of the **EPA's SAP**[54].

Twenty chronic rodent carcinogenicity studies have been done using glyphosate as the test compound. Eight of these studies are unacceptable for use in an evaluation of causality leaving seven studies in rats and five studies in mice. Because of the large number of evaluations done in an individual animal carcinogenicity study, there is concern that the false-positive rates could be exaggerated. For example, if 20 evaluations are done and a finding is deemed significant if $p_{Trend}<0.05$, then you would expect that 20*0.05=1 evaluation would be positive simply due to chance.

Table 15: Observed versus expected tumor sites with significant trends in the 12 acceptable rodent carcinogenicity studies using glyphosate.

Species	Strain	Sex	Total Sites[1]	Exp. <0.05	Obs. <0.05	Tumors[2] p<0.05	Exp. <0.01	Obs. <0.01	Tumors p<0.01
Rat (7 studies)	Sprague-Dawley (4 studies)	M	86	4.3	4	TICT, TFAC, KA, HA	0.9	2	TICT, KA
		F	102	5.1	1	TCCC	1.0	1	TCCC
	Wistar (3 studies)	M	64.5	3.2	2	HA, SK	0.6	1	HA
		F	76.5	3.8	2	MC, MAC	0.8	1	MAC
Mouse (5 studies)	CD-1 (4 studies)	M	42	2.1	8	KA, KC, KAC, HS(2)[3], ML(2), LAC	0.4	5	KA,KC, HS(2), ML
		F	60	3	1	H	0.6	1	H
	Albino (1 study)	M	10.5	0.5	0		0.1	0	
		F	15	0.8	1	H	0.2	1	H
Rats (7 studies)	All (7 studies)	M	150.5	7.5	6	TICT, KA, HA(2), TFAC, SK	1.5	3	TICT, KA, HA
		F	178.5	8.9	3	TCCC, MC, MAC	1.8	2	TCCC, MAC
		Both	329	16.5	9	TICT, KA, HA(2), TFAC, SK, TCCC, MC, MAC	3.3	5	TICT, KA, HA, TCCC, MAC
Mice (5 studies)	All (5 studies)	M	52.5	2.6	8	KA, KC, KAC, HS(2), ML(2), LAC	0.5	5	KA,KC, HS(2), ML
		F	75	3.8	2	H(2)	0.7	2	H(2)
		Both	127.5	6.4	10	KA, KC, KAC, HS(2)[3], H(2), ML(2), LAC	1.3	7	KA,KC, HS(2), H(2), ML
All (12 studies)	All (12 studies)	M	203	10.1	14	TICT, KA(2), HA(2), TFAC, SK, KC, KAC, HS(2), ML(2), LAC	2.0	8	TICT, HA, KA(2),KC, HS(2), ML
		F	253.5	12.7	5	TCCC, MC, MAC, H(2)	2.5	4	TCCC, MAC, H(2)
		Both	456.5	22.8	19	TICT, KA(2), HA(2), TFAC, SK, KC, KAC, HS(2), H, ML(2), LAC, TCCC, MC, MAC	4.6	12	TICT, HA, KA(2), KC, HS(2), H(2), ML, TCCC, MAC

[1]Number of sites examined is based upon suggestions by Dr. J. Haseman in his written testimony to the EPA; male mice – 10.5 sites; female mice – 15 sites; male rats – 21.5 sites; female rats – 25.5 sites

[2]Tumor abbreviations are: KA – kidney adenoma; KC – kidney carcinoma; KAC – kidney adenoma or carcinoma; HS – hemangiosarcoma; H – hemangioma; HA – hepatocellular adenoma; LAC – lung adenoma or adenocarcinoma; ML – malignant lymphoma; MC – mammary gland carcinoma; MAC – mammary gland adenoma or carcinoma; TCCC – thyroid C-cell carcinoma; TFAC

– thyroid follicular cell adenoma or carcinoma; TICT – testes interstitial cell tumor; SK – skin keratocanthoma
[3](x): x studies with this result

The EPA asked the SAP to comment on its evaluation of glyphosate[61] at a meeting in Washington, DC in December 2016[54]. Many comments were received from outside experts at this meeting; one such set of comments came from Dr. J. K. **Haseman (2016)**[111]. **Haseman (2016)** directly addressed the false-positive error rate and concluded that the results seen in these studies were due to chance. He did this by deciding how many evaluations were likely for each study (broken into sex-by-species groups) and then aggregating the findings. He concluded that the effective number of analyses were 10.5 in male mice, 15 for female mice, 21.5 for male rats, and 25.5 for female rats. **Haseman (2016)** made two assumptions in his analysis that are not valid. The first was that all of the possible trend tests had been done on all of the sites he considered reasonable for such an evaluation. He identified eight positive findings. However, EPA had not evaluated all of the sites nor had they considered doing a formal analysis using historical control data. EPA identified eight sex/species groups that had at most one positive tumor finding using the trend test with $p_{Trend} \leq 0.05$. In Tables 1-14 above, I have identified 19 tumors with $p_{Trend} \leq 0.05$ or $p_{Hist} \leq 0.05$ and 12 with $p_{Trend} \leq 0.01$ or $p_{Hist} \leq 0.01$ (Table 15). Secondly, Dr. Haseman assumed one could aggregate all the studies into one large analysis of Type-1 error. However, inference in these studies is always made by sex/species/strain (e.g. glyphosate causes hemangiosarcomas in male CD-1 mice; not glyphosate causes cancer in rodents), and the analysis should have been done by grouping each separately. Table 15 shows these analyses as well as the aggregated analysis for all of the acceptable studies.

With the exception of male Sprague-Dawley rats, the observed number of tumors are at or near the expected number for the different sex/strain groups in rats (Table 15). For male Sprague-Dawley rats, 0.8 cases with $p_{Trend} \leq 0.01$ or $p_{Hist} \leq 0.01$ are expected and two were observed ($p=0.21$). In female CD-1 mice and Swiss Albino mice, the expected and observed numbers are approximately equal. However, in male CD-1 mice, there were 2.1 tumors expected for $p_{Trend} \leq 0.05$ or $p_{Hist} \leq 0.05$ and eight were observed ($p<0.001$) and there were 0.4 expected for $p_{Trend} \leq 0.01$ or $p_{Hist} \leq 0.01$ and five were observed ($p<0.001$). This clearly could not have occurred by chance alone. Even if one incorrectly groups all sexes and species together, there are 4.6 expected responses for $p_{Trend} \leq 0.01$ or $p_{Hist} \leq 0.01$ and 12 observed ($p<0.001$). Thus, chance does not explain the positive results seen in these studies.

Conclusion for Animal Carcinogenicity Studies

There are several general issues that pertain to all animal carcinogenicity studies. There is considerable genetic variability across animal strains both over time and space. It is difficult to compare experiments done in different laboratories even when using the same strain of animal. This is obvious when you examine the rates for hepatocellular adenomas in Wistar rats across the three studies using this strain. Thus, each study

should be considered separately with regard to the findings in that study before being compared across studies.

The use of a p-value of 0.05 as the cut off for increasing tumor incidence does not account for trends in the data across multiple studies. Three studies with marginal responses of 6-8% in a given tumor could, when pooled for analysis, lead to highly significant findings. This issue is well-recognized in epidemiology but not usually considered in toxicology because of a lack of replicate studies. This case is fairly unique because of the larger number of studies available for analysis and requires a more rigorous evaluation of the data such as the pooled analysis presented in this report.

Pooling of the data for the evaluation of replicate studies makes sense as it addresses the question "Does the data as a whole support a finding of increased cancer incidence in these studies?" Some toxicologists may argue that the studies are not replicates and hence cannot be pooled. But if they are not replicates, then they cannot be compared to see if there is consistency across the studies. This is because there may be some subtle change from one study to another that leads to a positive finding in one study but a negative finding in other studies. Thus, either the studies are not good replicates so you cannot compare across studies and you cannot pool them, or they are good replicates so you can compare across studies and you can pool them. There is no argument that would support a comparison across studies that is appropriate when pooling is inappropriate.

There were seven rat studies and five mouse studies that were of sufficient quality and with sufficient details available for inclusion in this evaluation.

Glyphosate has been demonstrated to cause cancer in two strains of rats and one strain of mice. Glyphosate causes hepatocellular adenomas in male Wistar rats and, to a lesser degree, in male Sprague-Dawley rats, mammary gland adenomas and adenocarcinomas in female Wistar rats, skin keratocanthomas in male Wistar rats, and kidney adenomas and thyroid C-cell adenomas and carcinomas in male Sprague-Dawley rats. Glyphosate causes hemangiosarcomas, kidney tumors and malignant lymphomas in male CD-1 mice and hemangiomas in female CD-1 mice and possibly causes malignant lymphomas, kidney adenomas in male Swiss albino mice and hemangiomas in female Swiss albino mice. Thus, glyphosate causes cancer in mammals.

Mechanisms Relating to Carcinogenicity

Many human carcinogens act via a variety of mechanisms causing various biological changes, taking cells through multiple stages from functioning normally to becoming invasive with little or no growth control (carcinogenic). **Hanahan and Weinberg (2011)**[112] identified morphological changes in cells as they progress though this multistage process and correlated these with genetic alterations to develop what they refer to as the "hallmarks of cancer." These hallmarks deal with the entire process of carcinogenesis and not necessarily with the reasons that cells begin this process or the early stages in the process where normal protective systems within the cells remove

potentially cancerous cells from the body. While tumors that arise from a chemical insult to the cell may be distinct from other tumors by mutational analysis, they all exhibit the hallmarks as described by **Hanahan and Weinberg (2011)**.

Systematic review of all data on the mechanisms by which a chemical causes cancer is complicated by the absence of widely accepted methods for evaluating mechanistic data to arrive at an objective conclusion on human hazards associated with carcinogenesis. Such systematic methods exist in other contexts[113], but are only now being accepted as a means of evaluating literature in toxicological evaluations[114-117].

In this portion of the report, I am focusing on the mechanisms that can cause cancer. **Smith et al. (2015)**[37] discussed the use of systematic review methods in identifying and using key information from the literature to characterize the mechanisms by which a chemical causes cancer. They identified 10 "Key Characteristics of Cancer" useful in facilitating a systematic and uniform approach to evaluating mechanistic data relevant to carcinogens. These 10 characteristics are presented in Table 16 (copied from Table 1 of **Smith et al. (2015)**[37]). While there is limited evidence on glyphosate for most of the key characteristics, genotoxicity (characteristic two) and oxidative stress (characteristic five) have sufficient evidence to warrant a full review.

Genotoxicity

Genotoxicity refers to the ability of an agent (chemical or otherwise) to damage the genetic material within a cell, thus increasing the risks for a mutation. Genotoxic substances interact with the genetic material, including DNA sequence and structure, to damage cells. DNA damage can occur in several different ways, including single- and double-strand breaks, cross-links between DNA bases and proteins, formation of micronuclei and chemical additions to the DNA.

Just because a chemical can damage DNA does not mean it will cause mutations. So, while all chemicals that cause mutations are genotoxic, all genotoxic chemicals are not necessarily mutagens. Does that mean that the genotoxicity of a chemical can be ignored if all assays used for identifying mutations in cells following exposure to a chemical are negative? The answer to that question is no and is tied to the limitations in tests for mutagenicity (the ability of a chemical to cause mutations in a cell). It is unusual to see an evaluation of the sequence of the entire genome before exposure with the same sequence after exposure to determine if the genome has been altered (mutation). There are assays that can evaluate a critical set of genes that have previously been associated with cancer outcomes (e.g. cancer oncogenes), but these are seldom applied. In general, mutagenicity tests are limited in the numbers of genes they actually screen and the manner in which these screens work.

Because screening for mutagenicity is limited in scope, any genetic damage caused by chemicals should raise concerns because of the possibility of a mutation arising from that genetic damage. In what follows, I will systematically review the scientific findings available for evaluating the genotoxic potential of glyphosate. This will be divided into six separate sources of data based on the biological source of that data: (1) data from exposed humans, (2) data from exposed human cells in a laboratory setting, (3) data

from exposed mammals (non-human), (4) data from exposed cells of mammals (non-human) in the laboratory, (5) data from non-mammalian animals and others, and (5) data from cells from non-mammalian animals and others. These six areas are based upon the priorities one would apply to the data in terms of impacts. Seeing genotoxicity in humans is more important than seeing genotoxicity in other mammals, which is more important than seeing genotoxicity in non-mammalian systems. In addition, seeing genotoxicity in whole, living organisms (*in vivo*) carries greater weight than seeing responses in cells in the laboratory (*in vitro*). Basically, the closer the findings are to real, living human beings, the more weight they should be given.

Table 16: Key characteristics of carcinogens, **Smith et al. (2016)**[37]

Characteristic	Examples of relevant evidence
1. Is electrophilic or can be metabolically activated	Parent compound or metabolite with an electrophilic structure (e.g., epoxide, quinone), formation of DNA and protein adducts
2. Is genotoxic	DNA damage (DNA strand breaks, DNA–protein cross-links, unscheduled DNA synthesis), intercalation, gene mutations, cytogenetic changes (e.g., chromosome aberrations, micronuclei)
3. Alters DNA repair or causes genomic instability	Alterations of DNA replication or repair (e.g., topoisomerase II, base-excision or double-strand break repair)
4. Induces epigenetic alterations	DNA methylation, histone modification, microRNA expression
5. Induces oxidative stress	Oxygen radicals, oxidative stress, oxidative damage to macromolecules (e.g., DNA, lipids)
6. Induces chronic inflammation	Elevated white blood cells, myeloperoxidase activity, altered cytokine and/or chemokine production
7. Is immunosuppressive	Decreased immunosurveillance, immune system dysfunction
8. Modulates receptor-mediated effects	Receptor in/activation (e.g., ER, PPAR, AhR) or modulation of endogenous ligands (including hormones)
9. Causes immortalization	Inhibition of senescence, cell transformation
10. Alters cell proliferation, cell death or nutrient supply	Increased proliferation, decreased apoptosis, changes in growth factors, energetics and signaling pathways related to cellular replication or cell cycle control, angiogenesis

Abbreviations: AhR, aryl hydrocarbon receptor; ER, estrogen receptor; PPAR, peroxisome proliferator–activated receptor. Any of the 10 characteristics in this table could interact with any other (e.g., oxidative stress, DNA damage, and chronic inflammation), which when combined provides stronger evidence for a cancer mechanism than would oxidative stress alone.

The data being included in this review come from the peer-reviewed scientific literature, the summaries of reports in regulatory documents that are proprietary and for which I have limited access to the original work, and reports from industry that are proprietary to which I have been given greater access. All of these studies are included in the overall evaluation of causation.

Genotoxicity in Humans *in-vivo*

Three studies have evaluated the potential genotoxicity of glyphosate formulations in exposed humans. **Paz-y-Miño et al. (2007)**[118] analyzed the blood of 24 exposed individuals (living within 3 kilometers of spraying) and 21 unexposed individuals (living 80 kilometers away from the spraying area) for DNA damage using the comet assay. All study subjects were from Ecuador and none of the controls or exposed individuals smoked, drank alcohol, took non-prescription drugs or had been exposed to pesticides during the course of their normal daily lives. Exposed and control individuals did some cultivating and harvesting but without pesticides or herbicides. Exposed individuals were analyzed within two months of spraying for the eradication of plants associated with illegal narcotics. An average of 200 cells per person were ranked between 0-400 depending on the amount of DNA in the comet's tail in order to calculate the mean amount of DNA damage. There was a significant difference between the mean total migration level of exposed individuals to controls ($p<0.001$). Data was given for each individual classified into five groups based upon the amount of DNA in the comet's tail. There was clearly a shift in the distribution of DNA in cells with the controls never seeing scores in the top two categories while all but three exposed had some scores in the top two categories. In essence, some of the DNA had been fragmented by the exposure.

In a second study by the same group, **Paz-y-Miño et al. (2011)**[119] evaluated the karyotypes (the chromosome count of the individuals and any alterations to the chromosomes as seen under a microscope) of 92 people living in 10 communities in northern Ecuador. Controls were from areas without spraying and both controls and exposed subjects had no history of exposure to smoking or other genotoxic compounds. This study saw no changes between controls and exposed subjects for 182 karyotypes evaluated.

Bolognesi et al. (2009)[120] studied women of reproductive age and their spouses in five areas of Colombia, four of which are subject to spraying for either narcotics control or sugar cane growing. There were 60 subjects from the Santa Marta area (organic coffee is grown without the use of pesticides), 52 from Boyaca (manual spraying for illicit drugs), 58 from Putumayo (aerial spraying for illicit drugs using a glyphosate formulation), 63 from Nariño (same exposure as Putumayo) and 28 from Valle del Cauca (aerial spraying of Roundup 747 (74.7% glyphosate) without additional adjuvant for sugar cane maturation). All subjects were interviewed with a standardized questionnaire designed to obtain information about current health status, health history, lifestyle and potential exposure to possible confounding factors (smoking, use of medicinal products, severe infections or viral diseases during the last six months, recent vaccinations, presence of known indoor/outdoor pollutants, exposure to diagnostic x-rays, and previous radio- or chemotherapy). In Santa Marta, blood samples were taken

once, during the initial interview. In Boyaca, blood samples were taken at the initial interview and 1 month later. In Nariño, Putumayo and Valle del Cauca, blood samples were taken at the initial interview, within five days after spraying and 4 months later. In lymphocytes, binucleated cells with micronuclei (BNMN) were lowest in Santa Marta and similar in the four exposed regions prior to exposure. Statistically significant increases in BMNM in Nariño, Putumayo and Valle del Cauca were seen between first and second sampling. The mean BNMN in Nariño and Putumayo was greater in respondents who self-reported direct contact with sprayed fields, but differences were not statistically significant. Multiple linear regression demonstrated statistically significant increases in BMNM in all four exposed regions post exposure when compared to pre-exposure and controlling for all other variables ($p<0.001$). The largest total change in mean BMNM values pre-exposure compared to immediate post exposure occurred in Valle del Cauca where spraying is done using Roundup with no additional adjuvant.

Kier (2015)[121] identified 16 additional studies of pesticide use that included some exposure to glyphosate. Eleven of the 16 studies demonstrated some degree of genotoxicity in the human populations studied but did not adequately attribute the exposure primarily to glyphosate so they are not included in this review.

In summary, two of the three studies in which genotoxicity endpoints were evaluated in humans in areas with exposure to glyphosate spraying showed statistically increased changes in DNA damage in blood. In the strongest study, in three areas where chromosomal damage (micronuclei) was examined in individuals pre- and post-spraying (<5 days) showed statistically significant increases. In one other area where post-exposure damage was measured one month after exposure, there was little change.

Genotoxicity In Human Cells (*in vitro*)

Studies have explored the *in vitro* genotoxicity of glyphosate using a variety of different cell types (lymphocytes, fibroblasts, and immortalized cells from cancers of the larynx, mouth, blood and liver) using several different assays for markers of genotoxicity with or without metabolic activation.

Mladinic et al. (2009)[122] induced DNA strand breaks (comet assay) from exposure to glyphosate (purity not given) in lymphocytes from three healthy human donors (questionnaire used to exclude genotoxic exposures) at concentrations of 3.5, 92.8 and 580 µg/ml with S9 activation and saw effects at only the highest doses for cells without S9 activation.

Alvarez-Moya et al. (2014)[123] conducted a similar study using lymphocytes from human volunteers (questionnaire used to exclude genotoxic exposures) and exposure to glyphosate (96% purity) at concentrations of 0.12, 1.2, 12 and 120 µg/ml. A significant increase in DNA strand breaks (comet assay) was seen for all exposure groups with a clear dose-response relationship without metabolic activation (metabolic activation was not tested).

Using human HEP-2 cells, **Manas et al. (2009)**[124] induced DNA damage (comet assay) by

glyphosate (96% pure) at all concentrations ranging from 676 µg/ml to 1270 µg/ml (no S9 activation tested). Cell viability at the highest concentration was below 80% and values at the other concentrations were not given.

Monroy et al. (2005)[125] induced significant DNA damage (comet assay) in fibroblast GM 38 cells at concentrations of glyphosate (technical grade, purity not given) ranging from 676 µg/ml to 1000 µg/ml with a clear dose-response pattern. Over this same concentration range, they also saw concentration-dependent decreases in cell viability at all doses making the comet assay results difficult to interpret. In a similar analysis in the same paper, using fibrosarcoma HT1080 cells, they also saw concentration-dependent DNA damage and loss of cell viability. Activation by S9 was not used in either experiment.

Lueken et al. (2004)[126] induced DNA damage (comet assay) in fibroblasts GM 5757 at a concentration of glyphosate (98.4% purity) of 12,680 µg/ml in combination with exposure to 40 or 50 mM H_2O_2. Activation by S9 was not used in this experiment. According to the authors, cell viability at this exposure level was above 80%.

Koller et al. (2012)[127] significantly induced DNA damage (comet assay) in human TR146 cells (buccal carcinoma cells) from exposure to glyphosate (>95% purity) in a dose-dependent fashion at concentrations of 20 and 40 µg/ml. Above 40 µg/ml, there was a significant increase in tail intensity relative to controls, but the actual amount increased did not change as the dose increased (plateau). Using Roundup (Ultra Max) the authors saw virtually the same level of DNA damage at 20 and 40 µg/ml, but the concentration response continued to increase above that exposure. These experiments did not use S9 activation. They also used the CBMN assay in the same system to evaluate the total number of micronuclei in binucleated cells (MNI), the number of binucleated cells with micronuclei (BN-MNI), the number of nuclear buds (NB) and the number of nucleoplasmic bridges (NPB) caused by glyphosate and Roundup exposure. Two endpoints (NB, NPB) had significant increases at concentrations of 10, 15 and 20 µg/ml and two (MNi, BN-MNi) were significantly elevated for concentrations of 15 and 20 µg/ml. Equivalent Roundup exposures resulted in significant increases in all four measures of DNA damage at 10, 15 and 20 µg/ml. The results for the Roundup were greater than for glyphosate alone.

Gasnier at al. (2009)[128] exposed cells from the hepatoma cell line HepG2 to glyphosate (purity not given) and four glyphosate formulations. Only one glyphosate formulation was tested for DNA damage (comet assay) and they saw significant effects at equivalent concentrations of 0.05 µ/ml to 4 µg/ml of glyphosate (p-values not given). No p-values are provided and presentation of the results does not provide a clear means to compare these results with other studies. This study will not be used in the evaluation.

Manas et al. (2009)[124] obtained human blood samples from three healthy, non-smoking women and three healthy men with no history of pesticide exposure. Lymphocytes were cultured with glyphosate (96% purity) at concentrations of 34, 203, and 1015 µg/ml with no statistically significant changes in chromatid breaks,

chromosome breaks, chromatid gaps, chromosome gaps, dicentrics, acentric fragments, or endoreduplication.

Mladinic et al. (2009)[129] used blood from three non-smoking, healthy volunteers to evaluate the formation of micronuclei, nuclear buds and nucleoplasmic bridges as a function of exposure to glyphosate (98% purity). Significant changes in micronuclei were seen following exposure to glyphosate at 92.8 and 580 µg/ml in S9 activated cells, but not those without metabolic activation. Changes in nuclear buds were seen at 580 µg/ml for both S9 activated and non-activated cells while significant changes in nucleoplasmic bridges were seen only at 580 µg/ml in S9 activated cells. This study contained a positive control (ethyl methanesulfonate at 200 µg/ml) which was also negative in all assays, many times showing effects below that seen for glyphosate.

Bolognesi et al. (1997)[130] obtained blood from two healthy female donors and exposed it to glyphosate (99.9% purity) or a Roundup formulation (30.4% glyphosate). At concentrations of 1000, 3000 and 6000 µg/ml of glyphosate and at 100 and 330 µg/ml of glyphosate formulation, significant changes in sister chromatid exchanges (SCEs) were seen. At 330 µg/ml, a non-significant increase in SCEs was seen for glyphosate alone that was approximately 20% below that seen for an equivalent glyphosate exposure from the Roundup formulation. This study did not consider S9 activation.

Lioi et al. (1998)[124, 131] obtained blood from three healthy donors and exposed it to glyphosate (>98% purity). At concentrations of 1.4, 2.9, and 8.7 µg/ml of glyphosate, significant changes in sister chromatid exchanges (SCEs) and chromosomal aberrations were seen. This study did not consider S9 activation.

Vigfusson and Vyse (1980)[132] exposed cultured human lymphocytes from two people to Roundup (% glyphosate unknown) at concentrations of 250, 2500 and 25000 µg/ml. Results for the highest concentration were not provided due to lack of cell growth in culture. SCEs were shown to be significantly increased for the remaining two concentrations in one donor and only for the lowest concentration in the other. While the relative SCE counts seen in this paper are similar to those from **Bolognesi et al. (1997),** the absolute counts in the controls are roughly three times higher in this study. This study did not consider S9 activation.

Genotoxicity in Non-Human Mammals (*in vivo*)

Bolognesi et al. (1997)[130] exposed groups of three Swiss CD-1 male mice by Intraperitoneal (IP) injection with a single dose of glyphosate (99.9% purity, 300 mg/kg) or Roundup (900 mg/kg, equivalent to 270 mg/kg glyphosate). Animals were sacrificed at four and 24 hours after injection and livers and kidney were removed to obtain crude nuclei from the adhering tissues. Both tissues demonstrated significant increases in DNA single-strand breaks ($p<0.05$) at four hours for both glyphosate and Roundup with no discernable difference between the responses. At 24 hours, the presence of strand breaks was reduced and no longer statistically significant from controls.

Peluso et al. (1998)[133] exposed groups of six (controls, lowest doses of glyphosate-salt and Roundup) or three Swiss CD-1 mice (males and females, specific numbers not

specified, liver and kidney tissues combined for analysis) to the isopropylammonium salt of glyphosate or Roundup (30.4% isopropylammonium salt of glyphosate) for 24 hours. DNA adducts (^{32}P-DNA post labeling) were not evident in mice exposed to the glyphosate-salt alone in either liver or kidney, but were present in liver and kidney at all tested doses of Roundup showing a dose-response pattern.

Rank et al. (1993)[134] exposed male and female NMRI mice (three to five per sex) to glyphosate isopropylamine salt (purity not specified) and Roundup (480 g glyphosate isopropylamine salt per liter) by intraperitoneal injection. After 24 or 48 hours (only 24 hours for Roundup), polychromatic erythrocytes from bone marrow were extracted and micronuclei counted from a sample of 1000 cells. No significant increases were seen for any concentration in glyphosate-exposed animals (100, 150 and 200 mg/kg) or Roundup-exposed animals (133 and 200 mg/kg glyphosate equivalent dose). The positive controls, while not statistically significant, showed an increase in micronuclei.

Bolognesi et al (1997)[130] exposed groups of three, four or six male Swiss CD-1 mice to glyphosate (99.9% purity) and Roundup (30.4% glyphosate) by intraperitoneal injection in two equal doses given 24 hours apart. After six or 24 hours following the last exposure, polychromatic erythrocytes from bone marrow were extracted and micronuclei counted from a sample of 1000 cells. Mice given two doses of 150 mg/kg of glyphosate showed a non-significant increase in micronuclei at 6 hours and a significant increase at 24 hours. In contrast, mice given two doses of 225 mg/kg glyphosate equivalent of Roundup showed a significant increase in micronuclei at both six and 24 hours. The relative differences in mean absolute increase (subtract mean response in controls) in micronucleii between glyphosate and Roundup at 24 hours was 3.6 whereas the relative difference in glyphosate equivalent dose was 1.5 indicating a greater effect of the glyphosate formulation.

Manas et al. (2009)[124] exposed groups of male and female Balb C mice (group size not given, tissues combined for analysis) to glyphosate (96% purity) by intraperitoneal injection in two equal doses given 24 hours apart. Twenty-four hours post exposure, polychromatic erythrocytes from bone marrow were extracted and micronuclei counted from a sample of 1000 cells. No significant increases were seen at doses of 50 mg/kg and 100 mg/kg in glyphosate-exposed animals but a significant increase was seen at 400 mg/kg. The positive controls showed a statistically significant increase in micronuclei (roughly three times the control rate).

Dimitrov et al. (2006)[135] exposed groups of eight male C57BL mice (tissues combined for analysis) to Roundup (41% glyphosate) via gavage at a dose of 1080 mg/kg. At 6, 24, 72, 96, or 120 hours post exposure, polychromatic erythrocytes from bone marrow were extracted and micronuclei counted from a sample of 4000 cells (500 per animal). No significant increases were seen. They also looked for chromosomal damage in these animals and saw no significant increases. The positive controls showed a statistically significant increase in micronuclei.

Prasad et al. (2009)[136] exposed groups of 15 male Swiss CD-1 mice to Roundup (30.4% glyphosate) by IP injection at doses of 25 and 50 mg/kg. At 24, 48 or 72 hours post

exposure, polychromatic erythrocytes from bone marrow were extracted and micronuclei counted from a sample of 2000 cells per animal, five animals per sacrifice. Micronucleii counts were significantly increased ($p<0.05$) at all doses at all times relative to controls. In addition, the number of cells with chromosomal aberrations was significantly increased for all doses at all times. The control rate of micronuclei was similar to that of **Bolognesi et al. (1997),** but about 50% greater response for a dose that was approximately 10 times smaller.

Grisolia et al. (2002)[137] exposed groups of Swiss mice (sex and sample size not given) to Roundup (480 g glyphosate isopropylamine salt per liter) by IP injection at doses of 50, 100 and 200 mg/kg Roundup in two doses separated by 24 hours. At 24 hours post exposure, polychromatic erythrocytes from bone marrow were extracted and micronuclei counted from a sample of 2000 cells per animal. Micronuclei counts were not increased at any dose. This exposure appears to be the same formulation of Roundup used in the study by **Rank et al. (1993)** which was also negative.

Coutinho do Nascimento and Grisolia (2000)[138] exposed groups of six male mice (strain not given) to Roundup (% glyphosate not given) by IP injection at doses of 50, 100 and 200 mg/kg in two doses separated by 24 hours. At 24 hours post exposure, polychromatic erythrocytes from bone marrow were extracted and micronuclei counted from a sample of 1000 cells per animal. A significant increase in micronuclei were seen at a dose of 85 mg/kg. No increase was seen at 42 or 170 mg/kg.

Cavusoglu et al. (2011)[139] exposed groups of six Swiss albino mice by IP injection with a single dose of glyphosate formulation (RoundupUltra Max, 450 g/l glyphosate, 50 mg/kg glyphosate equivalent dose). Animals were sacrificed at three days after injection. Micronuclei in normochromatic erythrocytes were counted from a sample of 1000 cells per animal. There was a significant increase in micronuclei in erythrocytes ($p<0.05$). *G. bilboa* eliminated these effects.

Chan and Mahler (1992)[140] exposed groups of 10 male and female $B6C3F_1$ mice to glyphosate (98.6% purity) in feed at doses of 0, 507, 1065, 2273, 4776, and 10780 mg/kg in males and 0, 753, 1411, 2707, 5846, and 11977 mg/kg in females for 13 weeks. At sacrifice, polychromatic erythrocytes from peripheral blood were extracted and micronuclei counted from a sample of 10,000 cells. No significant increases were seen at any of the tested doses.

Li and Long (1988)[141] exposed groups of 18 male and female Sprague-Dawley rats to glyphosate (98% purity) by IP injection at a dose of 1000 mg/kg. At 6, 12 and 24 hours post treatment, 6 animals of each sex were sacrificed and polychromatic erythrocytes from bone marrow were extracted and micronuclei counted from a sample of 50 cells per animal. The percentage of cells with chromosomal aberrations was not increased at any time point following exposure.

Genotoxicity in Non-Human Mammalian Cells (*in vitro*)

Li and Long (1988)[141] incubated Chinese hamster ovary cells (CHO-K1BH4) with glyphosate (98% purity) for three hours at concentrations of 5, 10, 50 and 100 mg/ml.

Cells were then plated using 200 cells per sample in triplicate and incubated for 8-12 days. Colonies were then counted and results expressed as mutant frequency. No positive results were seen in any experimental group with or without S9 activation. It is not clear why there is such a large difference in the incubation times in the various groups in this experiment, nor is it clear which groups incubated longer. In a second study in the same publication, non-induced primary rat hepatocytes (Fischer 344) were incubated with seven concentrations of glyphosate (12.5 ng/ml to 125 µg/ml) for 18-20 hours. No significant increases were seen for net grains per nucleus at any exposure concentration. There was a four-fold increase in the lowest exposure groups relative to controls and then every other treated group was below the control response. This is a very unusual finding and could be due to the way in which the data is adjusted for net grains in cytoplasm. The authors calculated net grains per nucleus by subtracting the highest cytoplasmic count from the nuclear count; if cytoplasmic count is increased by glyphosate this could bias the findings making any increase in nuclear count disappear. No data is provided to resolve this issue.

Roustan et al. (2014)[142] incubated Chinese hamster ovary cells (CHO-K1) with glyphosate (purity not provided) for three hours at concentrations of 2, 5, 10, 15, 17.5, 20, and 22.5 mg/ml. Cells were then plated using 200 cells per sample in triplicate and incubated for 24 hours. For each exposure concentration, 2000 bi-nucleated cells were examined for micronuclei. No positive results were seen in any experimental group without S9 activation but the four highest exposure groups were significant with a clear concentration-response pattern when S9 activation was present.

Lioi et al. (1998)[131] exposed lymphocytes from three unrelated healthy cows to glyphosate (>98% purity) for 72 hours to concentrations of 3, 14.4 and 28.7 µg/ml without S9 activation. Chromosomal aberrations scored from 150 cells were significantly increased ($P<0.05$) for all exposure concentrations of glyphosate with a clear concentration-response pattern. Similarly, SCEs per cell were increased at all concentrations ($p<0.05$) but no concentration response pattern was evident.

Sivikova and Dianovsky (2006)[143] exposed lymphocytes from two healthy young bovine bulls to glyphosate formulation (62% glyphosate) for 2, 24 and 48 hours using concentrations of 4.7, 9.5, 23.6, 47.3, 94.6 and 190 µg/ml without S9 activation. Chromosomal aberrations scored from 100 cells were not significantly increased ($P<0.05$) without S9 activation for any 24-hour exposure concentration of glyphosate (2- and 48-hours exposures were not done). SCEs per cell were increased at all 24-hour exposure concentrations ($p<0.05$) except the lowest concentration. At 48-hours, significant increases of SCEs per cell were seen at concentrations at or above 47.3 µg/ml (2-hour exposures were not done). Finally, after two hours of exposure with S9 activation, significant effects were seen at 5 and 10 µg/ml but not at 15 µg/ml (24- and 48-hour exposures were not done for S9 activation).

Holeckova (2006)[144] exposed lymphocytes from two healthy young bovine bulls to glyphosate formulation (62% glyphosate) for 24 hours to concentrations ranging from 28 to 1120 µmol/L without S9 activation. A significant increase in polyploidy was observed at 56 µmol/L, all other comparisons were without significance. However, this

one finding cannot be easily dismissed because all exposure groups above this concentration had too few cells for evaluation. This study did not consider S9 activation.

Genotoxicity in Non-Human Systems (*in vivo* and *in vitro*)

Four studies[123, 145-147] in fish have seen positive results for genotoxicity (DNA strand breaks, different assays) following exposure to glyphosate. In addition, one study[148] in oyster sperm and embryos exposed to glyphosate saw no increase in DNA damage (comet assay) and one study[149] in two strains of Drosophila melanogaster showed an increase in mutations (wing spot test) at the higher doses of exposure.

Fourteen studies[137, 145, 147, 150-160] in multiple fish species evaluated the relationship between various glyphosate formulations and genotoxicity with all studies showing positive results for various endpoints (DNA strand breaks, micronucleus formation, and chromosomal aberrations). Two of the studies[150, 152] were negative for micronucleus formation after exposure to glyphosate formulations and one of these[150] was also negative for chromosomal aberrations but both were positive in other markers of genotoxicity. Two studies[161, 162] demonstrated genotoxicity (DNA strand breaks, micronuclei) in caiman from *in-vivo* exposure to a glyphosate formulation. Three studies[163-165] demonstrated genotoxicity (DNA strand breaks, micronucleus formation) in frogs or tadpoles from exposure to glyphosate formulations. One study[148] in oyster sperm and embryos, one study[166] in clams and one study[167] in mussels exposed to a glyphosate formulation saw no increase in DNA damage (comet assay). One study[168] in snails saw increased DNA damage (comet assay) following exposure to a glyphosate formulation. Two studies[169, 170] in worms saw mixed results for DNA damage (comet assay) with one of these studies[169] showing a positive result for micronucleus formation. One study[171] in Drosophila melanogaster showed an increase in sex-linked recessive lethal mutations.

In the published literature, five studies evaluated the impact of glyphosate in *in vitro* systems. Two of these studies[172, 173] looked at genotoxicity of glyphosate in combination with UVB radiation and saw significant increases in DNA strand breaks (FADU assay) in bacteria without metabolic activation. One study[174] in eukaryote fish saw a significant increase in DNA strand breaks (comet assay) without S9 activation. Another study[141] showed no increase in reverse mutations in two strains of bacteria with and without S9 activation.

Williams et al. (2000)[175] summarized the literature regarding the use of reverse mutation assays in *S. typhimurium* (Ames Test). Four studies using glyphosate and five studies of glyphosate formulations were all negative. They cited one study[134] of a glyphosate formulation that was positive with S9 activation and negative without S9 activation. However, this study was positive with S9 activation in TA100 cells, negative with S9 activation in TA98 cells, negative without S9 activation for TA100 cells and positive without activation for TA98 cells. They also summarized two studies of glyphosate in *e. coli* that were negative with and without activation.

Two additional studies[141, 176] of glyphosate using reverse mutation assays are available

from the scientific literature, both of which are negative.

Regulatory Studies

EFSA[89] cited 14 reverse mutation assays in *S. typhimurium* (Ames Test), most of which were tested in strains TA 98, 100, 1535, 1537 (Table B.6.4-1). All 14 studies are listed as negative by EFSA. Actual data is provided for only one of the 14 studies and this study is clearly negative. **EPA**[61] cited 27 reverse mutation assays in *S. typhimurium* (Ames Test), most of which were tested in strains TA 98, 100, 1535, 1537 (EPA Table 5.1). All 27 studies are listed as negative. No data is provided for any of the studies. **Kier and Kirkland (2013)**[177] cited results from 18 bacterial reverse mutation assays of glyphosate and 16 of glyphosate formulations. Tabulated results and background information were provided for all 34 studies. Six studies of glyphosate alone demonstrated positive findings in one or more groups.

EFSA[89] cites three studies of gene mutations in mammalian cells, all of which are listed as negative (EFSA Table B.6.4-5), two use the mouse lymphoma assay, and one uses the Chinese hamster ovary cell/hypoxanthine-guanine phosphoribosyl transferase (CHO/HGPRT) mutation assay. **EPA**[61] cites four studies, three of which appear to be the same as those cited by EFSA (EPA Table 5.2) and the fourth is another mouse lymphoma assay. All four are listed as negative. **Kier and Kirkland (2013)**[177] cite two of the mouse lymphoma studies and provide tabulated data. Neither study shows any indication of a statistically significant increase in mutation frequency at the thymidine kinase locus of L5178 mouse lymphoma tk(+/-) cells.

EFSA[89] cites one *in vitro* study of DNA damage and repair in mammalian cells which is listed as negative (EFSA Table B.6.4-6). This study is of unscheduled DNA synthesis (UDS assay) in primary rat lymphocytes. They also list five studies of chromosome aberrations (EFSA Table B.6.4-8), which are characterized as negative. Two studies are in human lymphocytes and two are in Chinese hamster lung (CHL) cells. Data for one of the studies in CHL is provided in tabular form and is clearly negative. **EPA**[61] cites eight *in vitro* studies of chromosome aberrations in mammalian cells (EPA Table 5.3); two of these studies match studies in the EFSA report. Four of the studies are from the literature[124, 131, 143, 178] and are reviewed above. Surprisingly, EPA refers to the study by **Manas et al. (2009)**[124] as negative although it was clearly positive in the comet assay., Additionally, EPA refers to the study by **Sivikova and Dainovsky (2006)**[143] as negative even though they saw clear effects of glyphosate on SCEs. Basically, all four of the literature studies cited by EPA are positive yet EPA lists only two of the four as positive. The remaining four studies are noted as negative; however, no data is supplied for these studies. **Kier and Kirkland (2013)**[177] cites eight literature studies (all reviewed above) and three regulatory studies with glyphosate exposure. The three regulatory studies are listed as negative, and the data are available as a table in the supplement material to **Kier and Kirkland (2013)**; these studies are negative at all tested concentrations in CHL cells; one matches the study data provided by EFSA[89].

EFSA[89] cites nine micronucleus assays, three in Swiss Albino mice, two in NMRI mice, two in CD-1 mice, one in Sprague-Dawley rats, and one in CD rats (EFSA Table B.6.4-12). They list one study in Swiss Albino mice as weakly positive in males, one study in CD-1 mice as positive at the highest dose (data for this study is provided) and all other studies as negative. They discard one study with low doses in male Swiss mice, but the tables provided for this study show a clearly significant result at the highest dose used (30 mg/kg) and clear dose-response. They provide data for two of the negative studies which indicate these studies were indeed negative. **EPA**[61] (EPA Table 5.5) cites 20 micronucleus assays, four are available in the scientific literature and three are reviewed above (the fourth reference[179] was unavailable to me at the time of preparation of this report). The remaining 16 studies include six studies in Swiss Albino mice, four studies in CD-1 mice, three studies in NMRI mice, two studies in Sprague-Dawley rats and one study in Wistar rats. Since EFSA does not provide names associated with their micronucleus studies, I cannot determine if any of the studies cited by the EPA are the same as those cited by EFSA. EPA lists two of the literature studies as positive and two as negative (matching my reviews for the three studies I have access to) and all but one of the regulatory studies as negative (the one positive study was in Swiss-Albino mice). **Kier and Kirkland (2013)**[177] cite 12 regulatory micronucleus assays of glyphosate and provide data tables for all 12. All 12 of these studies are cited by EPA. **Kier and Kirkland (2013)** list 11 studies as negative and one as inconclusive. However, four of the studies show positive effects in at least one sex-by-treatment group. One of these four studies they list as inconclusive and the remaining three studies are determined to be negative because the response is within the range of the historical controls. As was discussed for the animal carcinogenicity studies, the correct group to use is the concurrent control. **Kier and Kirkland (2013)**[177] also cite 12 regulatory studies and three literature studies where animals are exposed to a glyphosate formulation. Two of the literature studies are reviewed above and the remaining study[179] was unavailable. Data for the 12 regulatory studies are all provided in tables by **Kier and Kirkland (2013)** and show two positive studies in CD-1 mice and negative studies for the remaining 10.

Summary for Genotoxicity

This is a complicated area from which to draw a conclusion due to the diversity of the studies available (there are multiple species, multiple strains within a species, multiple cell types from multiple species, differing lengths of exposure, differing times of evaluation after exposure, differing exposures, numerous markers of genotoxicity, and finally both glyphosate and multiple different glyphosate formulations). There are three studies that evaluate the genotoxicity of glyphosate in humans directly, 36 experiments in eight strains of mice, three studies in rats, nine studies in human lymphocytes and four studies in other human cells, 12 studies in non-human mammalian cell lines (two using mouse cells, five using hamster cells, two using rat cells and three using cells from cows), a large number of studies in a wide variety of non-mammalian species, and a plethora of studies, mostly identical, in bacteria.

Some conclusions are straightforward"; glyphosate does not appear to cause reverse mutations for histidine synthesis in *Salmonella typhimurium*, regardless of whether

these reverse mutations are due to frameshift mutations or point mutations. I am cautious in this determination because there were several studies with positive results, but no clear pattern is evident. There is ample evidence supporting the conclusion that glyphosate formulations and glyphosate can cause genotoxicity in non-mammalian animal species. This clearly indicates that both glyphosate and the formulations are able to cause injury to DNA. So while findings of genotoxicity in these species do not speak directly to the hazard potential in humans, they do support a cause for concern.

The more important studies are those that have been done using mammalian systems, human cells and direct human contact. Table 16 summarizes these studies in a simple framework that allows all of the experimental data to be seen in one glance. This table does not address the subtlety needed to interpret any one study, but simply demonstrates when a study produced positive versus negative results.

Clearly, for *in vitro* evaluations in human cells, the majority of the studies have produced positive results. There was only one regulatory study evaluating glyphosate genotoxicity in human lymphocytes from healthy volunteers and that study was negative. The study was not significantly different from the other six studies in this category, five of which produced positive results. The majority of these studies used either the comet assay (a simple way for measuring any type of DNA strand break) or methods that counted specific types of strand breaks in the cells (e.g. SCEs, micronuclei, nuclear buds and nucleoplasmic bridges). From these assays, we can conclude there is DNA damage. For glyphosate formulations, there are only three studies in humans *in vivo*, two of which were positive.

The magnitude of the concentrations used in these studies could potentially lead to false positives if the glyphosate is causing cytotoxicity in the cells. All six studies using the comet assay were positive with no study showing a negative response below 10 µg/ml and mixed results below that with positive results at 0.12 and 3.5 µg/ml and negative results at 2.91 and 10 µg/ml. In general, the comet assays provide strong support for genotoxicity.

The four studies that directly addressed specific types of strand breaks in cells following exposure to glyphosate showed markedly different responses across the various concentrations used. **Manas et al. (2009)** saw no changes in chromatid breaks, chromosome breaks, chromatid gaps, chromosome gaps, dicentrics, acentric fragments or endoreduplication over the range of concentrations 3.4-1015 µg/ml. In contrast, **Lioi et al. (1998)** saw changes in SCEs over concentrations ranging from 1.4 to 8.7 µg/ml. Both studies were done in lymphocytes from volunteers. **Mladinic et al. (2009)** saw significant changes in micronuclei above 92.8 µg/ml and **Bolognesi et al. (1997)** saw positive changes in SCEs above 1000 µg/ml but not at 330 µg/ml. While changes have been seen in three of the four studies, the actual concentrations in which the changes are seen is not consistent across studies. I conclude that glyphosate causes DNA strand breaks, which is indicative of genotoxicity.

The micronucleus assays in rodents examining glyphosate genotoxicity are either all positive in one strain or all negative in one strain with the exception of the three studies

in CD-1 mice and four studies in Swiss Albino mice. For the positive studies, we can ask the question of whether, in this strain, the actual number of micronuclei are consistent.

Table 17: Summary of *in vivo* and *in vitro* genotoxicity studies of glyphosate and glyphosate formulations in mammals[1]

In vivo or *in vitro*	Species	Cell type or tissue	Glyphosate[2]		Glyphosate Formulations	
			Number Positive	Number Negative	Number Positive	Number Negative
In vivo	Humans	Peripheral blood			2	1
in vitro	Humans	lymphocytes	5	2(1)	2	
		Hep 2	1			
		GM 38 HT1080	1			
		GM 5757	1			
		TR146	1		1	
In vivo	Swiss CD-1 Mouse	Liver/Kidney	1	1	2	
In vivo (micro-nucleus assay)	NMRI mouse	Erythrocytes		4(3)		2(1)
	Swiss CD-1 mouse		1		2	
	Balb C mouse		1			
	B6C3F$_1$ mouse			1		
	Swiss mouse		1(1)			3(2)
	CD-1 mouse		2(2)	1(1)	2 (2)	6 (6)
	Swiss albino mouse		1(1)	3(3)	1	
	C57BL mouse					1
	Mouse (not specified)				1	
	Rats (all)			2(1)		1(1)
In vitro	Mouse	L5178 lymphoma		2(2)		
	Chinese hamster	Lung		3(3)		
	Chinese hamster	ovary	1	1		
	Fischer rat	liver		1		

	Rat	Lymphocytes		1(1)		
	Bovine	Lymphocytes	1		2	

[1]each entry in the table corresponds to a single study where a study is positive if at least one valid positive finding emerged from the study $p<0.05$; entries in the table are only for studies where data was available to review including data from EFSA[89] and Kier and Kirkland (2000)[177]; [2]numbers are the total number of studies in this category, numbers in parentheses are the subset of studies that are regulatory studies

In Swiss Albino mice, all four studies were done with males and females. Exposures were by oral gavage for the positive study (in female mice) and IP injection by the negative studies. The positive study was at 5000 mg/kg and the highest dose in any of the negative studies was 3024 mg/kg. Finally, the control response in the positive study was 6.7 micronucleated PCE per 1000 PCE whereas the controls in the three negative studies were between 0 and 0.6 micronucleated PCE per 1000 PCE. Any of these differences could easily explain the differences in response so the positive result in Swiss Albino mice should be accepted.

For CD-1 mice, the one negative micronucleus study was by oral gavage in males and females at a single dose of 5000 mg/kg. One of the positive studies was also by oral gavage in males at a single dose of 2000 mg/kg. Because of the nature of statistical noise, these two studies could both occur whether there is a true effect or not. For the other positive study, the dose was by IP injection in male mice with a positive response at 600 mg/kg that was more than double the response of the controls. These data support the finding that glyphosate can cause micronuclei in male CD-1 mice, which is indicative of genotoxicity.

The remaining *in vitro* assays in mammalian cells exposed to glyphosate show mixed results. The mouse lymphoma assay and the Chinese hamster ovary assays are looking for specific mutations that will allow these cells to grow in culture. The Chinese hamster lung, the two rat assays and the assay in bovine lymphocytes are measuring DNA damage and provide mixed results. In general, these responses appear to be negative with the exception of those seen in bovine lymphocytes that appear to show a positive increase in SCEs following exposure to glyphosate.

For glyphosate formulations, the main difference between the findings for glyphosate and those for the glyphosate formulations is the direct evidence for genotoxicity in humans and the micronucleus assays in Swiss mice. The observation of genotoxicity in humans following exposure to glyphosate formulations must carry the greatest weight in the overall analysis and two of the three studies were positive with the strongest study by **Bolognesi et al. (2009)**[120] showing the strongest response.

For the Swiss mouse studies of micronuclei, the fact that all three studies are negative for glyphosate formulations while one study is positive for glyphosate creates a clear disagreement. The positive study is an oral gavage study with an effect seen in male mice at 30 mg/kg/day. The two negative regulatory studies for glyphosate formulations were done at 2000 mg/kg (about 500 mg/kg glyphosate equivalent), were also oral

gavage studies and were replicates done in the same laboratory at different times. The remaining negative study used glyphosate formulation doses of 50-200 mg/kg (25-100 mg/kg glyphosate equivalent) but was done by intraperitoneal injection. With the exception of the different routes of exposure, the differences between these studies cannot be resolved.

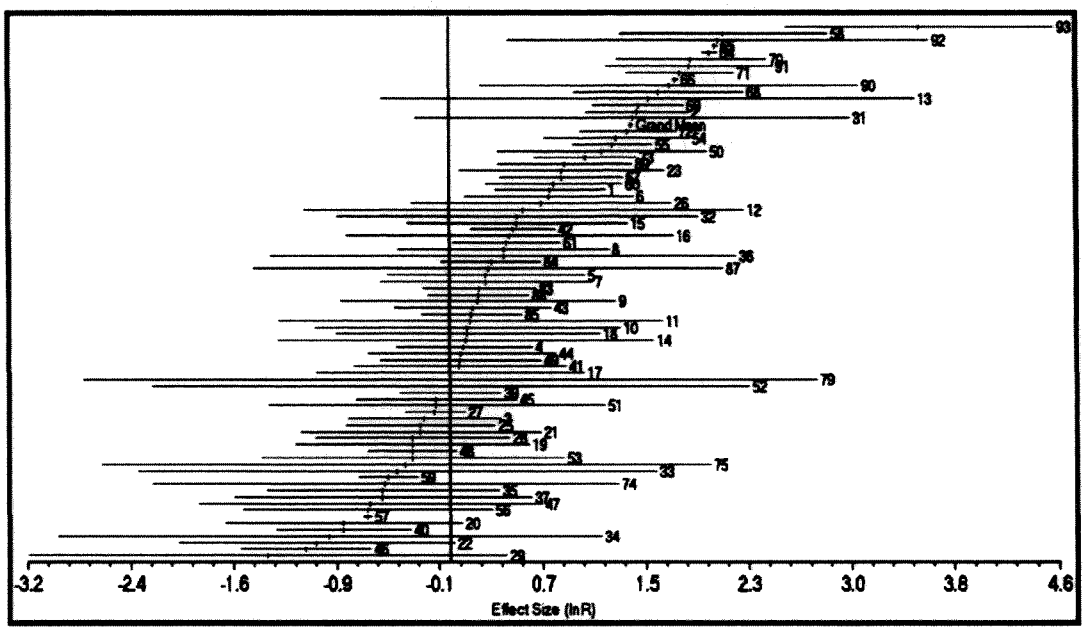
In this case, a pooled analysis of the data is not possible because in almost every case, no one study is a clear replicate of another. Instead, the appropriate approach would be to do a meta-analysis and evaluate which aspects of the experimental designs are important to producing positive findings of genotoxicity. The studies with the most data for this type of analysis are the various *in vivo* assays of micronucleus formation. **Ghisi et al. (2016)**[180] did a systematic search to identify all published studies evaluating the ability of glyphosate or glyphosate formulations to induce micronuclei *in vivo*. The authors also used the data from **Kier and Kirkland (2013)**[177] summarized above. An experiment, in their evaluation, was defined by sex/species/route/form of glyphosate so that some studies doing both sexes using glyphosate and a glyphosate formulation will enter multiple times into the analysis. They identified 93 experiments from which it was possible to do a meta-analysis. Data were extracted for each study and the log ratio of the mean of each experimental group to the mean control response (E+) was used to evaluate effect sizes in the meta-analysis. For this meta-analytic mean, a value below zero suggests no genotoxicity while a value above zero suggests increased genotoxicity. A test of heterogeneity (Cochran's Q statistic discussed earlier for the epidemiological data) was also evaluated.

Figure 2 is a reprint of Figure 1 from the study by **Ghisi et al. (2016)**[180] and is a forest plot from all studies they evaluated for glyphosate and glyphosate formulations. It is clear from this plot that the predominant response is positive in these data with an overall grand mean response across all studies of E+=1.37 and a 95% confidence interval of (1.356-1.381) (this is highly statistically significant with a $p<0.0001$). The Qt value for the grand mean was also statistically significant suggesting there are other explanatory variables in the data that would help to explain the overall variance.

Categorical variables were then used to make comparisons across the various strata in the data to identify which experimental conditions show the largest impacts on the mean response. Mammalian species presented a higher mean effect (E+=1.379; 1.366-1.391) than non-mammalian species (E+=0.740; 0.641-0.840). Glyphosate formulations showed a greater mean response (E+=1.388; 1.375-1.400) than did glyphosate (E+=0.121; 0.021-0.221), but both were significantly greater than zero. The mean response in studies using only male animals (E+=1.833; 1.819-1.847) was significantly different from zero as were studies using both males and females (E+=0.674;0.523-0.825) whereas the mean response in studies using only females (E+0.088; -0.153-0.328) was not. Peer-reviewed studies had higher mean response (E+=1.394; 1.381-1.407) compared to regulatory studies (E+=0.114; 0.027-0.202), but both means were significantly greater than zero, indicating an overall genotoxic effect. Other variables were examined such as length of exposure and magnitude of exposure that had very little impact on the overall findings.

The meta-analysis by **Ghisi et al. (2016)**[180] provides strong support for the hypothesis that exposure to glyphosate and glyphosate formulations increases the formation of micronuclei *in vivo*. This means that glyphosate and glyphosate formulations are damaging DNA in living, functioning organisms with intact DNA repair capacity strengthening the finding that glyphosate is genotoxic to humans.

Figure 2: Forest plot of studies evaluating micronucleus frequency in glyphosate exposure, arranged by effects size. The plot shows the estimate of the response ratio and 95% confidence interval (CI) of each experiment included in the meta-analysis. The number beside the bars represents the reference number of each experiment as in Table 1 of Ghisi et al. (2016)[180]. Grand Mean is the overall mean effects size of all studies. [Reprinted from Ghisi et al. (2016)[180]]



From a simply statistical perspective, there is another way in which one can decide if the positive findings in the micronucleus assays in the mice are due to chance. For the glyphosate studies, if one adds up all of the individual experimental groups, there are 79 total groups which correspond to 79 statistical tests. Assuming the critical testing level is 0.05 for all of the tests, one would expect to see just under four positive findings, yet six are observed. For the glyphosate formulations, there were 70 experimental groups so one expects 3.5 positive findings yet 12 are observed ($p<0.01$). Overall, there were a total of 149 experimental groups examined in mice for micronucleus formation and we observed 18 (7.5 expected, $p<0.01$). Repeating this analysis on the basis of studies instead of experimental groups, there were 15 studies for glyphosate (expected number is 0.75 positive) yet six positive were observed ($p<0.01$). For the glyphosate formulations, there were 18 studies (expected number is 0.9 positive) yet six positive

are observed (p<0.01). Now expanding to all 69 studies presented in Table 17, there were 33 positive studies, but the expectation is a mere 3.5 (p<0.01).

It is clear that both glyphosate and glyphosate formulations have genotoxic potential. But which is worse? Of the 69 experiments in Table 17, there were eight experiments from five research publications that addressed both glyphosate and a glyphosate formulation in the same laboratory. Of these, two were negative for both glyphosate and the formulation and do not contribute to a discussion of relative potency. The remaining six can provide some guidance on the relative potency of glyphosate to glyphosate formulations. In **Koller et al. (2007)**[127], tail intensity for the comet assay were virtually identical when the amount of glyphosate in the formulation was compared to the results using glyphosate alone. In the same paper, micronuclei and related biomarkers were consistently higher in the glyphosate formulation by 10-20%. In **Bolognesi et al. (1997),** DNA strand breaks in liver and kidney in Swiss CD-1 mice were virtually identical under equivalent doses of glyphosate and glyphosate formulations. In their micronucleus assay, the glyphosate formulation was approximately 50% more potent. Finally, **Bolognesi et al. (1997),** in their analysis of SCEs in human lymphocytes, the glyphosate formulation was approximately twice as effective as glyphosate alone. In **Peluso et al. (1988)**[133], DNA adducts in livers and kidneys were only seen in mice treated with the glyphosate formulation, so these findings are not likely to be due to glyphosate. The data suggest a small increase in the potential for genotoxicity for glyphosate formulations relative to the genotoxicity one would see with glyphosate alone.

In summary, the data support a conclusion that both glyphosate and glyphosate formulations are genotoxic. Thus, there is a reasonable mechanism supporting the increases in tumors caused by glyphosate and glyphosate formulations in humans and animals.

Oxidative Stress

Oxidative stress refers to an imbalance between the production of reactive oxygen species (free radicals) in a cell and the antioxidant defenses the cell has in place to prevent this. Oxidative stress has been linked to both the causes and consequences of several diseases[181-186] including cancer[37, 187-191]. Multiple biomarkers exist for oxidative stress; the most common being the increased antioxidant enzyme activity, depletion of glutathione or increases in lipid peroxidation. In addition, many studies evaluating oxidative stress used antioxidants following exposure to glyphosate to demonstrate that the effect of the oxidative stress can be diminished.

Oxidative Stress in Human Cells (*in vitro*)

Mladinic et al. (2009)[122] examined the induction of oxidative stress from exposure to glyphosate (98% purity) in lymphocytes from three healthy human donors (questionnaires were used to exclude other genotoxic exposures) at concentrations of 0.5, 2.91, 3.5, 92.8 and 580 µg/ml. Cells with and without S9 activation saw increases in total antioxidant capacity at only the highest dose for cells without S9 activation although a clear concentration response pattern was seen with S9 activation.

Kwiatkowska et al. (2014)[192] examined the induction of oxidative stress from exposure to glyphosate (purity not given) in erythrocytes obtained from healthy donors in the Blood Bank of Lodz, Poland. Erythrocytes were exposed to concentrations of 1.7, 8.4, 17, 42.3, 85 and 845 µg/ml and incubated for 1 hour. Oxidative stress (oxidation of dihydrorhodamine 123) was significantly increased at 42.3, 85 and 845 µg/l with a clear concentration-response pattern.

Chaufan et al. (2014)[193] examined the induction of oxidative stress from exposure to glyphosate (95% purity) and Roundup UltraMax (74.7% glyphosate) in HepG2 cells (human hepatoma cell line). Exposure concentrations were 900 µg/ml for glyphosate and 40 µg/ml for the glyphosate formulation. After incubation for 24 hours, oxidative stress (expressed as the activity of superoxide dismutase (SOD), catalase (CAT), glutathione (GSH) and glutathione-S-transferase (GST)) was significantly increased ($p<0.0-5$) for the glyphosate formulation (increased SOD activity) but not for glyphosate alone.

Coalova et al. (2014)[194] examined the induction of oxidative stress from exposure to a glyphosate formulation (Atanor, 48% glyphosate) or with a surfactant (Impacto) in Hep-2 cells (human epithelial cell line). Exposure concentrations were 376.4 µg/ml for Atanor, 12.1 µg/ml for Impacto and 180.2 µg/ml for a mixture of the two. After incubation for 24 hours, oxidative stress (measured as activity of SOD, CAT, GSH, and GST) was significantly increased for Impacto, Atanor and the mixture (CAT and GSH only, $p<0.05$ or $p<0.01$).

Gehin et al. (2005)[195] examined the induction of oxidative stress from exposure to glyphosate (purity unknown) and a glyphosate formulation (Roundup 3 plus, 21% glyphosate) in HaCaT cells (human keratinocyte cell line). Glyphosate induced cytotoxicity in the cells which was reduced or eliminated by antioxidants. The authors attributed the cytotoxicity to oxidative stress.

Elie-Caille et al. (2010)[196] examined the induction of oxidative stress from exposure to glyphosate (purity unknown) in HaCaT cells (human keratinocyte cell line). Exposure concentrations ranged from 1700 µg/l to almost 12,000 µg/ml. Glyphosate induced cytotoxicity in the cells and increased hydrogen peroxide H_2O_2 (dichlorodihydrofluorescein diacetate assay). This study used exceptionally high concentrations that may be inducing cytotoxicity by means that are independent of the oxidative stress observed. Measuring oxidative stress using the dichlorodihydrofluorescein diacetate assay has limitations[197, 198].

George and Shukla (2013)[199] examined the induction of oxidative stress from exposure to a glyphosate formulation (Roundup Original, 41% glyphosate) in HaCaT cells (human keratinocyte cell line). Exposure concentration ranged from 1.7 µg/ml to 17,000 µg/ml and exposure was for 24 hours. Glyphosate significantly induced the formation of reactive oxygen species (dichlorodihydrofluorescein diacetate assay) at all exposures in a concentration-dependent fashion. Prior treatment of the cells with N-Acetylcysteine reduced the impact of glyphosate, but did not eliminate it. Measuring oxidative stress using dichlorodihydrofluorescein diacetate has limitations[197, 198] that affect the clear

interpretation of these results.

Oxidative Stress in Non-Human Mammals (*in vivo*)

Bolognesi et al. (1997)[130] exposed groups of three Swiss CD-1 male mice by IP injection with a single dose of glyphosate (99.9% purity, 300 mg/kg) or Roundup (900 mg/kg, equivalent to 270 mg/kg glyphosate). Animals were sacrificed at eight and 24 hours after injection and livers and kidney were removed to obtain crude nuclei from the adhering tissues. Samples of liver and kidneys from these mice were evaluated for levels of 8-hydroxy-2'-deoxyguanosine (8-OHdG) which is a biomarker of oxidative stress[200]. There was a significant increase in the liver of 8-OHdG at 24 hours following glyphosate exposure, but not at eight hours and not in the kidney. At both eight hours and 24 hours, Roundup increased 8-OHdG in the kidneys, but the mild increase seen in the liver at 24 hours was not significant.

Cavusoglu et al. (2011)[139] exposed groups of six Swiss albino mice by IP injection of a glyphosate formulation (RoundupUltra Max, 450 g/l glyphosate, 50 mg/kg formulation). At the end of dosing, animals were fasted overnight then sacrificed. There was a significant increase in malondialdehyde in both liver and kidney and a significant decrease in GSH in liver and kidney from exposure to the glyphosate formulation. *G. bilboa* eliminated these effects.

Jasper et al. (2012)[201] exposed groups of 10 male and 10 female Swiss albino mice via oral gavage for 15 days to a glyphosate formulation (Roundup Original, 41% glyphosate, 50 mg/kg glyphosate equivalent dose). Animals were sacrificed at three days after injection. There was a significant increase in thiobarbituric acid-reactive substances (TBARS) in the liver for both male and female mice at both doses ($p<0.05$). The concentration of non-protein thiols was elevated in both dose groups for males and for the high dose only in females (no dose-response was seen for this endpoint).

Astiz et al. (2009)[202] exposed groups of four male Wistar rats by IP injection to a single dose of glyphosate (purity unknown, 10 mg/kg). Animals were injected three times per week for five weeks and then sacrificed. Thiobarbituric acid-reactive substances (TBARS assay), protein carbonyls (PCOSs), total glutathione levels, individual glutathione levels, SOD and CAT were all measured as biomarkers for oxidative stress in plasma, brain, liver and kidney. Glyphosate significantly increased TBARS in all tissues ($p<0.01$), total glutathione in brain ($p<0.01$), SOD in liver and brain ($p<0.01$) and CAT in brain. In a follow-up report[203], they demonstrate that lipoic acid eliminates or severely reduces the impacts of glyphosate on the brain.

Cattani et al. (2014)[204] exposed groups of four pregnant Wistar rats to glyphosate formulation (Roundup Original, 360 g/L glyphosate) in drinking water from gestational days 5-15 at a dose of 71.4mg/kg. Fifteen day-old pups (2 per dam) were examined for oxidative stress markers in the hippocampus. Pups had a significant increase in TBARS ($p<0.05$) and a significant decrease in GSH ($p<0.01$).

George et al. (2010)[82] exposed groups of four Swiss albino mice to a glyphosate formulation (Roundup Original, 36g/L glyphosate) at a dose of 50 mg/kg (glyphosate

equivalent dose) via a single topical application. Proteomic analysis of skin from the treated animals saw alterations in SOD1, CA III and PRX II, proteins known to play a role in the management of oxidative stress.

Oxidative Stress in Non-Mammalian Systems

As for genotoxicity, oxidative stress from exposure to glyphosate and glyphosate formulations have been studied in various aquatic organisms; reviewed in **Slaninova et al. (2009)**[205]. Many of the studies reviewed by **Slaninova et al. (2009)** showed associations with glyphosate and oxidative stress in various organs. Since that review, additional studies have been completed that also demonstrate a positive association between glyphosate and oxidative stress[147, 156-159, 206-217].

Summary for Oxidative Stress

Seven studies addressed oxidative stress in human cells and another six studies addressed it in mammalian systems. In lymphocytes and erythrocytes from healthy donors, oxidative stress was detected as low as 580 µg/ml in lymphocytes and at 42.3 µg/ml in erythrocytes. In Hep-G2 cells, no increased oxidative stress was seen for a single concentration of 900 µg/l. In two studies in HaCat cells, glyphosate induced oxidative stress in a continuous model fit to the results in one study and at the lowest concentration (1700 µg/ml) in the other. The most convincing studies in human cells for oxidative stress are the two studies in human blood.

In Swiss CD-1 male mice, increased oxidative stress was seen in the liver at 24 hours, but not at four hours after injection of 300 mg/kg glyphosate. No increase was seen in the kidney. In Wistar rats, repeated IP dosing with glyphosate lead to increased oxidative stress in multiple organs using multiple biomarkers. Thus, all of the laboratory studies demonstrated oxidative stress with a significant finding in the rat study.

In Hep-G2 cells, a glyphosate formulation demonstrated a robust increase in oxidative stress at 40 µg/ml. Given the negative response in this cell line for glyphosate alone, it must be concluded that this response is not due to glyphosate. In HEP-2 cells, a glyphosate formulation demonstrated a robust increase in oxidative stress via multiple biomarkers at 376 µg/ml and when a surfactant is added, at 180.2 µg/ml. In HaCaT cells, a glyphosate formulation demonstrated significant increases in oxidative stress from doses starting as low as 1.7 µg/ml in a concentration-dependent fashion. No studies were available in human lymphocytes.

In Swiss CD-1 mice, a glyphosate formulation significantly increased oxidative stress in the kidney but only demonstrated a mild (non-significant) increase in the liver. This study evaluated oxidative stress at two different time points following exposure and saw responses that differed over time. The strong increase in the liver for glyphosate but not glyphosate formulation, suggests a complicated response pattern for pure glyphosate versus the formulation that could be linked to the time since exposure. In Swiss Albino mice, a glyphosate formulation demonstrated increased oxidative stress by two separate biomarkers in both the liver and the kidney. In a second study in Swiss albino mice using a different biomarker but a similar dose, increased oxidative stress

was seen in both the liver and the kidney. In Wistar rat pups exposed in utero, an increase in oxidative stress was seen in the hippocampus. In Swiss albino mice, topical application of a glyphosate formulation to the skin resulted in a proteomic fingerprint suggesting oxidative stress was increased.

Though there are fewer studies for oxidative stress than there are for genotoxicity, the robust response seen here in human cells and in rodent studies clearly supports a role for both glyphosate and glyphosate formulations in inducing oxidative stress. Thus, there is a second reasonable mechanism through which the tumors seen in humans and those seen in animals can be caused by glyphosate and glyphosate formulations.

Summary for Biological Plausibility

In the evaluation of causality, the evidence for biological plausibility is overwhelming. Glyphosate clearly causes multiple cancers in mice, two cancers in the hematopoietic system similar to what is seen in humans, causes cancer in rats, is genotoxic and induces oxidative stress. The findings are clear for both glyphosate alone and for glyphosate formulations. **There is strong support for biological plausibility in support of a causal association of glyphosate and glyphosate formulations with NHL.**

Biological Gradient

Only three of the epidemiological studies provided information on biological gradients in their publications.

Eriksson et al. (2008)[46] divided their cases and controls into those with ≤10 days per year of exposure and those with >10 days per year of exposure. The ORs were calculated using a multivariate analysis that included agents with statistically significant increased OR, or with an OR > 1.50 and at least 10 exposed subjects. ORs for glyphosate were 1.69 (0.70-4.07) for ≤10 days per year and 2.36 (1.04-5.37) for >10 days per year. In their multivariate analysis, latency periods of 1-10 years showed an OR of 1.11 (0.24-5.08) and >10 years had an OR of 2.26 (1.16-4.40). Thus, they show an increase with intensity of exposure and with latency.

McDuffie et al. (2001)[50], using a conditional logistic regression analysis controlling for major chemical classes of pesticides and all other covariates with $p<0.05$, the OR for ≤2 days per year of exposure was 1.0 (0.63-1.57) and for >2 days per year, the OR was 2.12 (1.20-3.73). Thus, they show an increase with intensity of exposure.

De Roos et al. (2005)[45] used three exposure metrics in their analyses: a) ever personally mixed or applied pesticides containing glyphosate; b) cumulative exposure days of use of glyphosate (years of use times days per year); and c) intensity weighted cumulative exposure days (years of use times days per year times intensity of use). For exposure measurements b and c, they divided the respondents into tertiles chosen *a priori* to avoid having sparse data when dealing with rare tumors. For cumulative exposure days and using the lowest exposed tertile as the reference group, the RRs drop with values of 0.7 (0.4-1.4) and 0.9 (0.5-1.6) for tertiles 2 and 3 respectively adjusted for demographic and lifestyle factors and other pesticides (30,699 subjects). When

intensity-weighted exposure days are examined, the RRs drop with values of 0.6 (0.3-1.1) and 0.8 (0.5-1.4) for tertiles 2 and 3, respectively adjusted for demographic and lifestyle factors and other pesticides (30,699 subjects). Thus, they do not see a biological gradient in their responses. However, the high frequency of exposure to many pesticides (e.g. 73.8% were exposed to 2,4-D) means subjects with low exposure to glyphosate were likely to be exposed to other agents that may also induce NHL; this could reduce the RRs in the higher exposure classes because it would inflate the RR in the low-exposure referent group.

Eriksson et al. (2008)[46] and **McDuffie et al. (2001)**[50] had consistent results for intensity of exposure per year (≤2 days per year, OR=1.0; ≤10 days per year, OR=1.69; >2 days per year, OR=2.12; >10 days per year, OR=2.26). It is not possible to resolve the remaining differences between these three studies nor is it easy to argue that one study has more weight on this question than any other. The studies use different measures of exposure or time since exposure, are done on different populations and have different statistical power to detect a trend.

In rodent carcinogenicity studies, there is clear evidence of a biological gradient.

In general, there is support that a biological gradient exists for the epidemiological data and thus support from this aspect of the Bradford-Hill evaluation.

Temporal Relationship

Exposure must come before the cancers occur otherwise the epidemiology studies are useless. In this case, it is clear that exposure came before the onset of NHL. **The need for a temporal relationship in the data supporting a causal association between glyphosate and NHL is satisfied.**

Specificity

There are other causes of NHL[218-221] so this group of cancers is not specific to glyphosate. **There is little support for specificity**.

Coherence

Humans, coming into contact with glyphosate, can absorb the compound into their bodies where it has been measured in blood and in urine[56, 222-226]. In laboratory animals, absorption, distribution and elimination of glyphosate and glyphosate compounds have been studied[140, 227] and show that glyphosate gets into the animal's bodies, distributes to numerous organs and is eliminated in urine. The animal cancer studies clearly demonstrate that glyphosate in mammals can have toxic effects.

Mouse models have long served as surrogates for humans in understanding and developing treatments for many diseases. The same holds true for lymphoid tumors seen in humans. For over 30 years, mouse models have been studied and evaluated as surrogates for NHL[228-232]. These publications and the associated classification systems for humans and mice indicate a close linkage between the diseases in humans and mice.

Thus, coherence is supported by the increased risk of malignant lymphomas in CD-1 mice, the marginal increase in these tumors in Swiss mice and the strong similarity between malignant lymphomas in mice and NHL in humans.

There is strong support for coherence in the data supporting a causal association of glyphosate and glyphosate formulations with NHL.

Experimental Evidence in Humans

There is no experimental evidence in humans since purposely exposing humans to a pesticide, especially one that is probably carcinogenic, is not ethical and would never pass review by a human subject's advisory board.

Analogy

I am unaware of any analogous compounds from the scientific literature. This, however, is not an area where I have sufficient background to express an opinion.

Summary

Table 18 summarizes the information for each of Hill's aspects of causality. For these data, causality is strengthened because the available epidemiological studies show a consistent positive association between cancer and the exposure. The studies do not show different responses with some studies being positive and others negative, nor do they show any heterogeneity when analyzed together. And, in answer to Hill's question, the relationship between NHL and glyphosate exposure has been observed by different persons, in different places, circumstances, and times.

Causality is strengthened for these data because the strength of the observed associations, when evaluated simultaneously, are statistically significant, the findings are uni-directional and the results are unlikely to be due to chance. Even though none of the individual studies provide relative risks or odds ratios that are large and precise, the meta-analysis has objectively shown that the observed association across these studies is significant and supports a positive association between NHL and glyphosate.

Biological plausibility is strongly supported by the animal carcinogenicity data and the mechanistic data on genotoxicity and oxidative stress. When addressing biological plausibility, the first question generally asked is "Can you show that glyphosate causes cancers in experimental animals?" In this case, the answer to that question is clearly yes. Glyphosate has been demonstrated to cause cancer in two strains of rats and one strain of mice. Glyphosate has been demonstrated to cause cancer in two strains of rats and one strain of mice. Glyphosate causes hepatocellular adenomas in male Wistar rats and, to a lesser degree, in male Sprague-Dawley rats, mammary gland adenomas and adenocarcinomas in female Wistar rats, skin keratocanthomas in male Wistar rats, and kidney adenomas and thyroid C-cell adenomas and carcinomas in male Sprague-Dawley rats. Glyphosate causes hemangiosarcomas, kidney tumors and malignant lymphomas in male CD-1 mice and hemangiomas in female CD-1 mice and possibly

causes malignant lymphomas, kidney adenomas in male Swiss albino mice and hemangiomas in female Swiss albino mice. Thus, glyphosate causes cancer in mammals. Thus, it is biologically plausible that glyphosate alone can cause cancer in mammals.

The next question generally asked is "Does the mechanism by which glyphosate causes cancer in experimental animals also work in humans?" The best understood mechanism by which chemicals cause cancer in both humans and animals is through damaging DNA that leads to mutations in cells that then leads to uncontrolled cellular replication and eventually cancer. It is absolutely clear from the available scientific data that both glyphosate and glyphosate formulations are genotoxic. This has been amply demonstrated in humans that were exposed to glyphosate, in human cells *in vitro*, in experimental animal models and their cells *in vitro* and *in vivo*, and in wildlife. One way in which DNA can be damaged is through the presence of free oxygen radicals that overwhelm a cell's antioxidant defenses. Glyphosate induces this type of oxidative stress, providing additional support for a biological mechanism that works in humans.

Table 18: Summary conclusions for Hill's nine aspects of epidemiological data and related science

Aspect	Conclusion	Reason
Consistency of the observed association	Strong	Multiple studies, all are positive, meta-analysis shows little heterogeneity, different research teams, different continents, different questionnaires, no obvious bias or confounding
Strength of the observed association	Strong	Six core epidemiology studies all show the same modest increase, significant meta-analyses
Biological plausibility	Very Strong	Multiple cancers in multiple species, not due to chance, increased risk of rare tumors, convincing evidence for genotoxicity and oxidative stress
Biological gradient	Moderate	Clearly seen in the two case-control studies that evaluated it, not seen in the cohort study
Temporal relationship of the observed association	Satisfied	Exposure clearly came before cancers
Specificity of the observed association	Not needed	NHL has other causes, this does not subtract from the causal argument
Coherence	Strong	Glyphosate is absorbed, distributed and excreted from the body, cancers seen in the mice have strong similarity to human NHL
Evidence from human experimentation	No data	No studies are available
Analogy	No data	No studies available in the literature

In general, there is support that a biological gradient exists for the epidemiological data and thus support from this aspect of the Bradford-Hill evaluation. Glyphosate ORs increased with time since first exposure and with intensity of use per year in the two case-control studies that evaluated at least one of these issues.

There is clearly the proper temporal relationship with the exposure coming before the cancers.

The human evidence is coherent. The basic findings in humans agree with the animal evidence for absorption, distribution and elimination of glyphosate. Also, one of the tumors seen in mice has almost the same etiology as NHL.

NHL is not specific to glyphosate exposure. There is no experimental evidence in humans and I did not find any references where researchers looked for analogous compounds with similar toxicity.

Hill (1965)[36] asks *"is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?"* There is no better way of explaining the scientific evidence relating glyphosate to an increase in NHL in humans than cause and effect.

In my opinion, glyphosate probably causes NHL and, given the human, animal and experimental evidence, I assert that, to a reasonable degree of scientific certainty, the probability that glyphosate causes NHL is high.

The IARC Assessment of Glyphosate

In March 2015, the International Agency for Research on Cancer (an agency of the World Health Organization) brought together seventeen scientists (the Working Group) to evaluate the scientific evidence on whether glyphosate can cause cancer in humans. This group also contained one invited specialist (myself) to aid the Working Group (WG) in going through the science but who was not allowed to join discussions on the final conclusion or write any part of the document. The Working Group concluded that glyphosate falls in the category *"probably carcinogenic to humans (Group 2A)"*[56].

The IARC preamble[30] guides Working Groups on how to evaluate scientific literature to determine if something is a hazard. All Working Groups follow these guidelines and this process is accepted worldwide as a proper way to evaluate the literature for a hazard (e.g., the European Chemical Agency cites the IARC review process as guidance and then uses the exact same wording as IARC does to guide their own hazard evaluation process[34]).

The WG examined the epidemiological data and classified it as *"limited evidence of carcinogenicity,"* which is defined to mean *"a positive association has been observed between exposure to the agent and cancer for which a causal interpretation is considered to be credible, but chance, bias or confounding could not be ruled out with*

reasonable confidence." This is a precise and clear description of the strength of the evidence from the epidemiological studies.

The WG examined the evidence from animal carcinogenicity studies and classified it as "*sufficient evidence of carcinogenicity,*" which IARC defines as: "*a causal relationship has been established between the agent and an increased incidence of malignant neoplasms or of an appropriate combination of benign and malignant neoplasms in (a) two or more species of animals or (b) two or more independent studies in one species carried out at different times or in different laboratories or under different protocols. A single study in one species and sex might be considered to provide sufficient evidence of carcinogenicity when malignant neoplasms occur to an unusual degree with regard to incidence, site, type of tumour or age at onset, or when there are strong findings of tumours at multiple sites*." Based on the data available to IARC at the time of their review and the restrictions placed on the studies they can review by the Preamble, this conclusion is justified and correct.

One of the major criticisms of the WG review was that the WG did not review all of the animal carcinogenicity data that was available to the regulatory bodies and thus came to the wrong conclusions on the animal cancer data. In this review, I evaluated all 19 animal carcinogenicity experiments that have been collectively mentioned by any agency that reviews glyphosate. Where possible, I have analyzed the original data and used sound statistical methods to test for significant increases in cancer incidence in animals exposed to glyphosate. My conclusion is that the WG would have called this data "*sufficient evidence*" to support their findings despite not reviewing the additional studies analyzed herein. Despite the fact the industry kept these studies confidential, nothing contained in the withheld studies would have changed the WG conclusion.

On the mechanistic data, the IARC Working Group reviewed the same data that I reviewed, but I also evaluated, where possible, the proprietary data supporting the regulatory decisions. Where possible, I reanalyzed that data to be certain the results being presented were accurate. The IARC Working Group, using the guidelines set forth in their Preamble, declared strong support for the biological mechanisms of genotoxicity and oxidative stress. As I have shown here, there is strong support for these two mechanisms, even with the proprietary evidence from the industry studies. Thus, the IARC Working Group reached the correct conclusion.

To decide on a final classification for a compound, the IARC Preamble provides guidance on how the classification of the three areas are to be used. If the data in humans is "*limited*" and the data from animal carcinogenicity studies is "*sufficient*," the discussions should begin with Class 2A, "*the agent is probably carcinogenic to humans*." Then, given the overall quality of the data set, the strength of the evidence from the mechanistic studies and any additional scientific issues that need to be considered, the Working Group will determine whether the data justifies a different category. In this case, the Working Group concluded 2A was the right category and I still believe the evidence supports that finding.

The EPA Assessment of Glyphosate

Like IARC, the EPA has guidelines that are to be followed when evaluating scientific literature and making a determination about the carcinogenic potential of a chemical. Those guidelines have been developed over many years and are based on sound scientific guidance that myself and many other scientists have provided to the Agency. For their evaluation of glyphosate, the Agency did not follow their own guidelines, nor did they follow sound scientific practice. This opinion is consistent with the review done by the **EPA FIFRA Scientific Advisory Panel**[54]. In addition, the Agency failed to find all of the relevant animal cancer studies and misinterpreted several of them. The major problems with the Agency evaluation are:

- Misinterpretation of the epidemiological evidence, confusing the potential for bias and potential for confounding with real bias and real confounding, allowing them to give almost no weight to the case-control studies in favor of the one cohort study;
- Misinterpretation of the findings in the meta-analysis;
- Failure to properly use historical controls in the analysis of the animal carcinogenicity studies; declaring a significant finding as not due to the compound if it is in the range of the historical controls;
- Failure to analyze all tumors in all studies relying upon the industry submissions to have done this correctly;
- Failure to follow their guidelines on what constitutes a positive finding, disregarding significant trend tests when no corresponding pairwise comparisons are also significant;
- Disregarding positive findings in doses that are clearly not above the maximum dose the animals could be given with compromising the integrity of the study;
- Using unreasonable arguments about the overall false positive rates in the study without actually doing an analysis of this issue;
- Failing to recognize the similar findings in similar studies and to do a pooled analysis to determine if the negative effects in one study cancel out the positive effects in another;
- Giving very little weight to studies from the literature and relying almost entirely on studies provided by industry that have not undergone peer review for both quality and, more importantly in some cases, interpretation of the findings; and
- Comparing results across different species and strains for the animal cancer studies and the mechanistic studies with little regard for unique findings in any one study and consistent findings across multiple studies.

Similar comments apply to the evaluation done by the **European Food Safety Authority**[89] and the **European Chemical Agency**[233]. My detailed comments to these

agencies on their risk assessments are attached. There were comments to my comments to EPA by other scientists and I also responded to those comments in the EPA docket for glyphosate. These are also included in the attached Appendices.

Dr. Christopher J. Portier

Cited References

1. Portier, C. and D. Hoel, *Optimal design of the chronic animal bioassay.* J Toxicol Environ Health, 1983. **12**(1): p. 1-19.
2. Portier, C.J. and D.G. Hoel, *Design of the Chronic Animal Bioassay for Goodness of Fit to Multistage Models.* Biometrics, 1983. **39**(3): p. 809-809.
3. Bailer, A.J. and C.J. Portier, *Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples.* Biometrics, 1988. **44**(2): p. 417-31.
4. Portier, C.J. and A.J. Bailer, *Testing for increased carcinogenicity using a survival-adjusted quantal response test.* Fundam Appl Toxicol, 1989. **12**(4): p. 731-7.
5. Portier, C.J., J.C. Hedges, and D.G. Hoel, *Age-specific models of mortality and tumor onset for historical control animals in the National Toxicology Program's carcinogenicity experiments.* Cancer Res, 1986. **46**(9): p. 4372-8.
6. Portier, C.J. and A.J. Bailer, *2-Stage Models of Tumor-Incidence for Historical Control Animals in the National Toxicology Programs Carcinogenicity Experiments.* Journal of Toxicology and Environmental Health, 1989. **27**(1): p. 21-45.
7. Portier, C.J. and L. Edler, *Two-stage models of carcinogenesis, classification of agents, and design of experiments.* Fundam Appl Toxicol, 1990. **14**(3): p. 444-60.
8. Portier, C.J., et al., *Biologically based models for risk assessment.* IARC Sci Publ, 1990(104): p. 20-8.
9. Kopp-Schneider, A. and C.J. Portier, *Distinguishing between models of carcinogenesis: the role of clonal expansion.* Fundam Appl Toxicol, 1991. **17**(3): p. 601-13.
10. Kopp-Schneider, A., C.J. Portier, and F. Rippmann, *The application of a multistage model that incorporates DNA damage and repair to the analysis of initiation/promotion experiments.* Math Biosci, 1991. **105**(2): p. 139-66.
11. Kopp-Schneider, A., C.J. Portier, and F. Rippmann, *The Application of a Multistage Model That Incorporates DNA Damage and Repair to the Analysis of Initiation Promotion Experiments.* Mathematical Biosciences, 1991. **105**(2): p. 139-166.
12. Kopp-Schneider, A. and C.J. Portier, *Birth and Death Differentiation Rates of Papillomas in Mouse Skin.* Carcinogenesis, 1992. **13**(6): p. 973-978.

13. Portier, C.J., A. Kopp-Schneider, and C.D. Sherman, *Using Cell Replication Data in Mathematical-Modeling in Carcinogenesis.* Environmental Health Perspectives, 1993. **101**: p. 79-86.
14. Kopp-Schneider, A., C.J. Portier, and C.D. Sherman, *The Exact Formula for Tumor-Incidence in the 2-Stage Model.* Risk Analysis, 1994. **14**(6): p. 1079-1080.
15. Portier, C., et al., *Modeling the number and size of hepatic focal lesions following exposure to 2378-TCDD.* Organohalogen Compounds, 1994. **21**: p. 393-397.
16. Kopp-Schneider, A. and C.J. Portier, *Carcinoma formation in NMRI mouse skin painting studies is a process suggesting greater than two stages.* Carcinogenesis, 1995. **16**(1): p. 53-9.
17. Portier, C.J., A. Kopp-Schneider, and C.D. Sherman, *Calculating tumor incidence rates in stochastic models of carcinogenesis.* Mathematical Biosciences, 1996. **135**(2): p. 129-146.
18. Portier, C.J., et al., *Modeling the number and size of hepatic focal lesions following exposure to 2,3,7,8-TCDD.* Toxicology and Applied Pharmacology, 1996. **138**(1): p. 20-30.
19. Kopp-Schneider, A., C. Portier, and P. Bannasch, *A model for hepatocarcinogenesis treating phenotypical changes in focal hepatocellular lesions as epigenetic events.* Math Biosci, 1998. **148**(2): p. 181-204.
20. Toyoshiba, H., et al., *Gene interaction network analysis suggests differences between high and low doses of acetaminophen.* Toxicol Appl Pharmacol, 2006. **215**(3): p. 306-16.
21. Gohlke, J.M. and C.J. Portier, *The forest for the trees: a systems approach to human health research.* Environ Health Perspect, 2007. **115**(9): p. 1261-3.
22. Gohlke, J.M., et al., *Genetic and environmental pathways to complex diseases.* BMC Syst Biol, 2009. **3**: p. 46.
23. Gohlke, J.M., et al., *Estimating the global public health implications of electricity and coal consumption.* Environ Health Perspect, 2011. **119**(6): p. 821-6.
24. Aylward, L.L., et al., *Evaluation of biomonitoring data from the CDC National Exposure Report in a risk assessment context: perspectives across chemicals.* Environ Health Perspect, 2013. **121**(3): p. 287-94.
25. Trong, I. and C. Portier. *Proceedings of the Viet Nam – United States Scientific Conferences on Human Health and Environmental Effects of Agent Orange/Dioxin, Part 1 and 2.* 2002. Ha Noi, Vietnam: US National Institute of Environmental Sciences and the Government of Vietnam.
26. Portier, C.J. and M.S. Wolfe, eds. *EMF Science Review Symposium Breakout Group Report for Epidemiology Research Findings*. 1998, National Institute of Environmental Health Sciences: Research Triangle Park, North Carolina.
27. Portier, C.J. and M.S. Wolfe, eds. *EMF Science Review Symposium Breakout Group Report for Clinical and In Vivo Laboratory Findings*. NIH Publication Number 98-4400. 1998, National Institute of Environmental Health Sciences: Research Triangle Park, North Carolina.
28. Portier, C.J. and M.S. Wolfe, eds. *Assessment of Health Effects from Exposure to Power-Line Frequency Electric and Magnetic Fields*. NIH Publication Number 98-

3981. 1998, National Institute of Environmental Health Sciences: Research Triangle Park, North Carolina. 508.

29. Portier, C.J., et al., *A Human Health Perspective on Climate Change*, N.I.o.E.H.S. Health and Human Services, Editor. 2010, Environmental Health Perspectives: Research Triangle Park, NC72.
30. IARC. *Preamble to the IARC Monographs*. 2006 [cited 2017 3/8/17]; Available from: http://monographs.iarc.fr/ENG/Preamble/CurrentPreamble.pdf.
31. National Research Council. Science and Judgment in Risk Assessment. 1994, Washington (DC): National Academy of Sciences Press.
32. National Research Council. Science and Decisions: Advancing Risk Assessment. 2009, Washington (DC): National Academies Press.
33. USEPA, *Guidelines for Carcinogen Risk Assessment*, U.E.P. Agency, Editor. 2005: Washington DC166.
34. European Chemicals Agency, *Guidance on the Application of the CLP Criteria: Guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures*. 2015, European Chemicals Agency: Helsinki, Finland.
35. National Toxicology Program. *Handbook for Preparing Report on Carcinogens Monographs*. 2015 [cited 2017 3/8/17]; Available from: https://ntp.niehs.nih.gov/ntp/roc/handbook/roc_handbook_508.pdf.
36. Hill, A.B., *The Environment and Disease: Association or Causation?* Proc R Soc Med, 1965. **58**: p. 295-300.
37. Smith, M.T., et al., *Key Characteristics of Carcinogens as a Basis for Organizing Data on Mechanisms of Carcinogenesis.* Environ Health Perspect, 2016. **124**(6): p. 713-21.
38. Chang, E.T. and E. Delzell, *Systematic review and meta-analysis of glyphosate exposure and risk of lymphohematopoietic cancers.* J Environ Sci Health B, 2016: p. 1-27.
39. Cantor, K.P., et al., *Pesticides and other agricultural risk factors for non-Hodgkin's lymphoma among men in Iowa and Minnesota.* Cancer Res, 1992. **52**(9): p. 2447-2455.
40. Nordstrom, M., et al., *Occupational exposures, animal exposure and smoking as risk factors for hairy cell leukaemia evaluated in a case-control study.* Br.J Cancer, 1998. **77**(11): p. 2048-2052.
41. Hardell, L. and M. Eriksson, *A case-control study of non-Hodgkin lymphoma and exposure to pesticides.* Cancer, 1999. **85**(6): p. 1353-1360.
42. Hardell, L., M. Eriksson, and M. Nordstrom, *Exposure to pesticides as risk factor for non-Hodgkin's lymphoma and hairy cell leukemia: pooled analysis of two Swedish case-control studies.* Leuk.Lymphoma, 2002. **43**(5): p. 1043-1049.
43. De Roos, A.J., et al., *Integrative assessment of multiple pesticides as risk factors for non-Hodgkin's lymphoma among men.* Occup Environ Med, 2003. **60**(9): p. E11.
44. Lee, W.J., et al., *Non-Hodgkin's lymphoma among asthmatics exposed to pesticides.* Int J Cancer, 2004. **111**(2): p. 298-302.

45. De Roos, A.J., et al., *Cancer incidence among glyphosate-exposed pesticide applicators in the Agricultural Health Study.* Environ Health Perspect, 2005. **113**(1): p. 49-54.
46. Eriksson, M., et al., *Pesticide exposure as risk factor for non-Hodgkin lymphoma including histopathological subgroup analysis.* Int J Cancer, 2008. **123**(7): p. 1657-63.
47. Orsi, L., et al., *Occupational exposure to pesticides and lymphoid neoplasms among men: results of a French case-control study.* Occup Environ Med, 2009. **66**(5): p. 291-8.
48. Hohenadel, K., et al., *Exposure to multiple pesticides and risk of non-Hodgkin lymphoma in men from six Canadian provinces.* Int J Environ Res Public Health, 2011. **8**(6): p. 2320-30.
49. Cocco, P., et al., *Lymphoma risk and occupational exposure to pesticides: results of the Epilymph study.* Occup Environ Med, 2013. **70**(2): p. 91-8.
50. McDuffie, H.H., et al., *Non-Hodgkin's lymphoma and specific pesticide exposures in men: cross-Canada study of pesticides and health.* Cancer Epidemiol Biomarkers Prev, 2001. **10**(11): p. 1155-1163.
51. Zahm, S.H., et al., *A case-control study of non-Hodgkin's lymphoma and the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) in eastern Nebraska.* Epidemiology, 1990. **1**(5): p. 349-356.
52. Hoar, S.K., et al., *Agricultural herbicide use and risk of lymphoma and soft-tissue sarcoma.* JAMA, 1986. **256**(9): p. 1141-1147.
53. Blair, A., et al., *Reliability of reporting on life-style and agricultural factors by a sample of participants in the Agricultural Health Study from iowa.* Epidemiology, 2002. **13**(1): p. 94-9.
54. FIFRA Scientific Advisory Panel Meeting Minutes, *Meeting Minutes and Final Report of the December 13-16, 2016 FIFRA SAP Meeting Held to Consider and Review Scientific Issues Associated with EPA's Evaluation of the Carcinogenic Potential of Glyphasate*, O.o.P. Programs, Editor. 2017, US Environmental Protection Agency: Washington, DC101.
55. Portier, C.J., et al., *Differences in the carcinogenic evaluation of glyphosate between the International Agency for Research on Cancer (IARC) and the European Food Safety Authority (EFSA).* J Epidemiol Community Health, 2016. **70**(8): p. 741-5.
56. IARC Working Group, *Glyphosate*, In: *Some Organophosphate Insecticides and Herbicides: Diazinon, Glyphosate, Malathion, Parathion, and Tetrachlorvinphos.* 2015, IARC Monogr Prog. V. 112. p. 1-92.
57. Schinasi, L. and M.E. Leon, *Non-Hodgkin lymphoma and occupational exposure to agricultural pesticide chemical groups and active ingredients: a systematic review and meta-analysis.* Int J Environ Res Public Health, 2014. **11**(4): p. 4449-527.
58. Ioannidis, J.P., *Interpretation of tests of heterogeneity and bias in meta-analysis.* J Eval Clin Pract, 2008. **14**(5): p. 951-7.
59. OECD, *Guidance Document 116 on the Conduct and Design of Chronic Toxicity and Carcinogenicity Studies*, H.a.S.P. Environment, Editor. 2012, OECD: Paris.

60. Haseman, J.K., *Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies.* Environ Health Perspect, 1984. **58**: p. 385-92.
61. Environmental Protection Agency, *Glyphosate Issue Paper: Evaluation of Carcinogenic Potential*. 2016: Washington, DC.
62. NRC Committee to Review the Styrene Assessment in the National Toxicology Program 12th Report on Carcinogens, in *Review of the Styrene Assessment in the National Toxicology Program 12th Report on Carcinogens: Workshop Summary*. 2014, National Academies Press: Washington (DC).
63. Bieler, G.S. and R.L. Williams, *Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity.* Biometrics, 1993. **49**(3): p. 793-801.
64. Piegorsch, W.W. and A.J. Bailer, *Analyzing environmental data*. 2005, Chichester, West Sussex, England ; Hoboken, NJ: Wiley. xv, 496 p.
65. Haseman, J.K., G.A. Boorman, and J. Huff, *Value of historical control data and other issues related to the evaluation of long-term rodent carcinogenicity studies.* Toxicol Pathol, 1997. **25**(5): p. 524-7.
66. Haseman, J.K., J. Huff, and G.A. Boorman, *Use of historical control data in carcinogenicity studies in rodents.* Toxicol Pathol, 1984. **12**(2): p. 126-35.
67. OECD, *Combined Chronic Toxicity\Carcinogenicity Studies, OECD Guidelines for Testing of Chemicals, No. 453*, OECD, Editor. 2009: Paris.
68. Atkinson, C., Strutt, A., Henderson, W., et al. , *104-Week Chronic Feeding/ Oncogenicity study in rats with 52-week interim kill.* 1993.
69. Brammer., *Glyphosate Acid: Two Year Dietary Toxicity and Oncogenicity Study in Wistar Rats.* 2001: Central Toxicology Laboratory, Alderley Park Macclesfield, Cheshire, UK.
70. Burnett, P., Borders, J.;,Kush, J. , *Report to Monsanto Company: Two Year Chronic Oral Toxicity Study with CP- 76100 in Albino Rats*. 1979, Industrial Bio-Test Laboratories, Inc.: on behalf of Monsanto Co.
71. Chruscielska, K.B., J.; Kita, K, et al., *Glyphosate: Evaluation of chronic activity and possible far - reaching effects - Part 1. Studies on chronic toxicity.* Pestycydy, 2000. **3-4**: p. 10.
72. Enemoto, K., *24-Month Oral Chronic Toxicity and Oncogenicity Study in Rats, Vol. 1*. 1997: The Institute of Environmental Toxicology, Kodaira-shi, Tokyo, Japan.
73. Excel, *Combined chronic toxicity/carcinogenicity study of glyphosate technical in Sprague Dawley rats.* . 1997: Indian Institute of Toxicology, Pune, India.
74. Lankas, G., P., *A Lifetime Study of Glyphosate in Rats*. 1981, Monsanto: Report No. 77-2062 prepared by Bio Dynamics, Inc.
75. Pavkov, K.L., Wyand, S. , *Two-Year Chronic Toxicity and Oncogenecity Dietary Study with SC-0224 in Rats.* 1987: Stauffer Chemical Company.
76. Reyna, M.S., Gordon, D.E. , *Two- Year Chronic Oral Toxicity Study with CP67573 in Albino Rats: Monsanto*. 1974.
77. Seralini, G.E., et al., *Republished study: long-term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize.* Environmental Sciences Europe, 2014. **26**(1): p. 14.

78. Stout, L.D.a.R., P.A., *Chronic Study of Glyphosate Administered in Feed to Albino Rats*. 1990: Monsanto.
79. Suresh, T.P., *Combined Chronic Toxicity and Carcinogenicity Study with Glyphosate Technical in Wistar Rats*. 1996, Syngenta: Toxicology Department Rallis Research Centre, Rallis India Limited.
80. Wood, E., Dunster, J., Watson, P., and Brooks, P. , *Glyphosate Technical: Dietary Combined Chronic Toxicity/Carcinogenicity Study in the Rat*. 2009: Harlan Laboratories Limited, Shardlow Business Park, Shardlow, Derbyshire DE72 2GD, UK. Study No. 2060-012. April, 23, 2009.
81. Atkinson, C., Martin, T., Hudson, P., and Robb, D., *Glyphosate: 104 week dietary carcinogenicity study in mice*. 1993: Inveresk Research International, Tranent, EH33 2NE, Scotland. IRI Project No. 438618. April 7, 1993.
82. George, J., et al., *Studies on glyphosate-induced carcinogenicity in mouse skin: a proteomic approach*. J Proteomics, 2010. **73**(5): p. 951-64.
83. Knezevich, A.L. and G.K. Hogan, *A chronic feeding study of glyphosate in mice: Monsanto*. 1983: Bio/Dynamic Inc., dated July 21, 1983. Report No. 77-2011. .
84. Kumar, D.P.S., *Carcinogenicity Study with Glyphosate Technical in Swiss Albino Mice*. 2001: Toxicology Department Rallis Research Centre, Rallis India Limited. Study No. TOXI: 1559.CARCI-M.
85. Pavkov, K.L., Turnier, J.C. , *Two-Year Chronic Toxicity and Oncogenecity Dietary Study with SC-0224 in Mice*. 1987: Stauffer Chemical Company.
86. Reyna, M.S. and D.E. Gordon, *18-Month Carcinogenicity Study with CP67573 in Swiss White Mice: Monsanto*. 1973: Industrial Bio-Test Laboratories, Inc.
87. Sugimoto, K., *18-Month Oral Oncogenicity Study in Mice, Vol. 1 and 2*. 1997: The Institute of Environmental Toxicology, 2-772, Suzuki-cho, Kodaira-shi, Tokyo, 187, Japan, Study No.:IET 94-0151.
88. Wood, E., Dunster, J., Watson, P., and Brooks, P. , *Glyphosate Technical: Dietary Carcinogenicity Study in the Mouse*. 2009: Harlan Laboratories Limited, Shardlow Business Park, Shardlow, Derbyshire DE72 2GD, UK. Study No. 2060-011. April, 22, 2009.
89. (BfR)., G.F.I.f.R.A., *Glyphosate Renewal Assessment Report: Toxicology and Metabolism*. 2013. V. 3. p.
90. (BfR)., G.F.I.f.R.A., *Glyphosate Renewal Assessment Report: Assessment of IARC Monographies Volume 112 (2015)*. 2015. V. 3Glyphosate Addendum I to RAR. p.
91. Greim, H., et al., *Evaluation of carcinogenic potential of the herbicide glyphosate, drawing on tumor incidence data from fourteen chronic/carcinogenicity rodent studies*. Crit Rev Toxicol, 2015. **45**(3): p. 185-208.
92. Blettner, M., et al., *Traditional reviews, meta-analyses and pooled analyses in epidemiology*. Int J Epidemiol, 1999. **28**(1): p. 1-9.
93. Friedenreich, C.M., *Methods for pooled analyses of epidemiologic studies*. Epidemiology, 1993. **4**(4): p. 295-302.
94. Portier, C.J., *Comments on Glyphosate Issue Paper (EPA-HQ-OPP-2016-0385-0094)*. 2016, US EPA Docket Number (EPA-HQ-OPP-2016-0385-0094): Washington DC.

95. Hogan, g.k., *A chronic Feeding Study of Glyphosate (Roundup Technical) in Rats*. 1981, Monsanto: Bio/dynamics Inc., E. Millstone, NJ.
96. EPA, *EPA Memo Stout and Ruecker*, I. William Dykstra.Toxicology Branch, Editor. 1991. V. MRID 416438-01 Tox review 008897. p.
97. Uno, H., et al., *A versatile test for equality of two survival functions based on weighted differences of Kaplan-Meier curves*. Stat Med, 2015. **34**(28): p. 3680-95.
98. Giknis, M. and C. Clifford. *Neoplastic and Non-Neoplastic Lesions in the Charles River Wistar Hannover [Crl:WI(Han)] Rat*. 2011; Available from: http://www.criver.com/files/pdfs/rms/wistarhan/rm_rm_r_wistar_han_tox_data_2011.aspx.
99. Acquavella, J., et al., *Glyphosate epidemiology expert panel review: a weight of evidence systematic review of the relationship between glyphosate exposure and non-Hodgkin's lymphoma or multiple myeloma*. Crit Rev Toxicol, 2016. **46**(sup1): p. 28-43.
100. Lacayo, H., *Memorandum: Use of historical data in determiningthe weight of evidence from kidney tumor incidence in the Glyphosate two-year feeding study and some remarks on false positives*, S.M.S. Branch, Editor. 1985, US EPA: Washington, DCC.
101. Chandra, M. and C.H. Frith, *Spontaneous neoplasms in aged CD-1 mice*. Toxicol Lett, 1992. **61**(1): p. 67-74.
102. Giknis, M. and C. Clifford, *Spontaneous Neoplastic Lesions in the Crl:CD-1(ICR)BR Mouse*. 2000, Charles River Laboratories.
103. Giknis, M. and C. Clifford, *Spontaneous Neoplastic Lesions in the Crl:CD-1(ICR)BR Mouse in Control Groups from 18 Month to 2 year Studies*. 2005, Charles River Laboratories.
104. Eustis, S.L., et al., *The utility of multiple-section sampling in the histopathological evaluation of the kidney for carcinogenicity studies*. Toxicol Pathol, 1994. **22**(5): p. 457-72.
105. Kasza, L., *Memorandum: Glyphosate - Evaluation of Kidney Tumors in Male Mice. Chronic Feeding Study*, T.B. Branch, Editor. 1985, US EPA: Washington, DCC.
106. de Vivar Chevez, A.R., J. Finke, and R. Bukowski, *The role of inflammation in kidney cancer*. Adv Exp Med Biol, 2014. **816**: p. 197-234.
107. Dixon, D., et al., *Summary of chemically induced pulmonary lesions in the National Toxicology Program (NTP) toxicology and carcinogenesis studies*. Toxicol Pathol, 2008. **36**(3): p. 428-39.
108. Son, W.C. and C. Gopinath, *Early occurrence of spontaneous tumors in CD-1 mice and Sprague-Dawley rats*. Toxicol Pathol, 2004. **32**(4): p. 371-4.
109. OECD, *Carcinogenicity Studies, OECD Guideline for the Testing of Chemicals, No. 451*, OECD, Editor. 2009: Paris.
110. Martens, M., *Memorandum to Joe Haseman, March 2, 2017: Analysis tumor data RAC 40 meeting*, G.T. Force, Editor. 2017: Helsinki, Finland.
111. Haseman, J.K., *Comments on the laboratory animal carcinogenicity studies evaluated in EPA's Glyphosate Issue Paper: Evaluation of Carcinogenic Potential,*

and an assessment of the comments by Dr. Christopher Portier on these studies. 2016, US EPA Docket Number (EPA-HQ-OPP-2016-0385-0094): Washington DC.

112. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
113. Higgins, J.P.T. and S. Green, *Cochrane Handbook for Systematic Reviews of Interventions*. 2011, The Cochrane Collaboration.
114. Birnbaum, L.S., et al., *Implementing systematic review at the National Toxicology Program: status and next steps.* Environ Health Perspect, 2013. **121**(4): p. A108-9.
115. Murray, H.E. and K.A. Thayer, *Implementing systematic review in toxicological profiles: ATSDR and NIEHS/NTP collaboration.* J Environ Health, 2014. **76**(8): p. 34-5.
116. Rooney, A.A., et al., *Systematic review and evidence integration for literature-based environmental health science assessments.* Environ Health Perspect, 2014. **122**(7): p. 711-8.
117. Vandenberg, L.N., et al., *A proposed framework for the systematic review and integrated assessment (SYRINA) of endocrine disrupting chemicals.* Environ Health, 2016. **15**(1): p. 74.
118. Paz-y-Miño, C., et al., *Evaluation of DNA damage in an Ecuadorian population exposed to glyphosate.* Genetics and Molecular Biology, 2007. **30**(2).
119. Paz-y-Miño, C., et al., *Baseline determination in social, health, and genetic areas in communities affected by glyphosate aerial spraying on the northeastern Ecuadorian border.* Reviews on Environmental Health, 2011. **26**(1).
120. Bolognesi, C., et al., *Biomonitoring of genotoxic risk in agricultural workers from five colombian regions: association to occupational exposure to glyphosate.* J Toxicol Environ Health A, 2009. **72**(15-16): p. 986-97.
121. Kier, L.D., *Review of genotoxicity biomonitoring studies of glyphosate-based formulations.* Crit Rev Toxicol, 2015. **45**(3): p. 209-18.
122. Mladinic, M., et al., *Evaluation of genome damage and its relation to oxidative stress induced by glyphosate in human lymphocytes in vitro.* Environ Mol Mutagen, 2009. **50**(9): p. 800-7.
123. Alvarez-Moya, C., et al., *Comparison of the in vivo and in vitro genotoxicity of glyphosate isopropylamine salt in three different organisms.* Genet.Mol Biol, 2014. **37**(1): p. 105-110.
124. Manas, F., et al., *Genotoxicity of glyphosate assessed by the comet assay and cytogenetic tests.* Environ Toxicol Pharmacol, 2009. **28**(1): p. 37-41.
125. Monroy, C.M., et al., *Citotoxicidad y genotoxicidad en células humanas expuestas in vitro a glifosato.* Biomédica, 2005. **25**: p. 335-45.
126. Lueken, A., et al., *Synergistic DNA damage by oxidative stress (induced by H2O2) and nongenotoxic environmental chemicals in human fibroblasts.* Toxicology Letters, 2004. **147**(1): p. 35-43.
127. Koller, V.J., et al., *Cytotoxic and DNA-damaging properties of glyphosate and Roundup in human-derived buccal epithelial cells.* Arch Toxicol, 2012. **86**(5): p. 805-13.

128. Gasnier, C., et al., *Glyphosate-based herbicides are toxic and endocrine disruptors in human cell lines.* Toxicology, 2009. **262**(3): p. 184-91.
129. Mladinic, M., P. Perkovic, and D. Zeljezic, *Characterization of chromatin instabilities induced by glyphosate, terbuthylazine and carbofuran using cytome FISH assay.* Toxicol Lett, 2009. **189**(2): p. 130-7.
130. Bolognesi, C., et al., *Genotoxic Activity of Glyphosate and Its Technical Formulation Roundup.* J. Agric. Food Chem., 1997. **45**: p. 1957-1962.
131. Lioi, M.B., et al., *Genotoxicity and oxidative stress induced by pesticide exposure in bovine lymphocyte cultures in vitro.* Mutat Res, 1998. **403**(1-2): p. 13-20.
132. Vigfusson, N.V. and E.R. Vyse, *The effect of the pesticides, Dexon, Captan and Roundup, on sister-chromatid exchanges in human lymphocytes in vitro.* Mutat Res, 1980. **79**(1): p. 53-57.
133. Peluso, M., et al., *32P-postlabeling detection of DNA adducts in mice treated with the herbicide Roundup.* Environ Mol Mutagen, 1998. **31**(1): p. 55-59.
134. Rank, J., et al., *Genotoxicity testing of the herbicide Roundup and its active ingredient glyphosate isopropylamine using the mouse bone marrow micronucleus test, Salmonella mutagenicity test, and Allium anaphase-telophase test.* Mutat Res, 1993. **300**(1): p. 29-36.
135. Dimitrov, B.D., et al., *Comparative genotoxicity of the herbicides Roundup, Stomp and Reglone in plant and mammalian test systems.* Mutagenesis, 2006. **21**(6): p. 375-82.
136. Prasad, S., et al., *Clastogenic effects of glyphosate in bone marrow cells of swiss albino mice.* J Toxicol, 2009. **2009**: p. 308985.
137. Grisolia, C.K., *A comparison between mouse and fish micronucleus test using cyclophosphamide, mitomycin C and various pesticides.* Mutat Res, 2002. **518**(2): p. 145-150.
138. Coutinho do Nascimento A and G. C., *Comparative analysis between micronuclei tests in mice and in peripheral erythrocytes of Oreochromis niloticus in evaluation of mutagenic potential of the agrotoxins deltamethrin, dicofol, glyphosate, and imazapyr.* Pesticides: R Ecotoxicol E Meio Ambiente, Curitiba, 2000. **10**: p. 8.
139. Cavusoglu, K., et al., *Protective effect of Ginkgo biloba L. leaf extract against glyphosate toxicity in Swiss albino mice.* J Med Food, 2011. **14**(10): p. 1263-72.
140. Chan, P. and J. Mahler, *Toxicity Studies of Glyphosate (CAS No. 1071836) Administered in Dosed Feed to F344/N Rats and B6C3F1 Mice*, N.T. Program, Editor. 1992: Research Triangle Park, NC.
141. Li, A.P. and T.J. Long, *An evaluation of the genotoxic potential of glyphosate.* Fundam.Appl Toxicol, 1988. **10**(3): p. 537-546.
142. Roustan, A., et al., *Genotoxicity of mixtures of glyphosate and atrazine and their environmental transformation products before and after photoactivation.* Chemosphere, 2014. **108**: p. 93-100.
143. Sivikova, K. and J. Dianovsky, *Cytogenetic effect of technical glyphosate on cultivated bovine peripheral lymphocytes.* Int J Hyg Environ Health, 2006. **209**(1): p. 15-20.

144. HOLEČKOVÁ, B., *EVALUATION OF THE IN VITRO EFFECT OF GLYPHOSATE-BASED HERBICIDE ON BOVINE LYMPHOCYTES USING CHROMOSOME PAINTING.* Bull Vet Inst Pulawy, 2006. **50**: p. 533-536.
145. Moreno, N.C., S.H. Sofia, and C.B. Martinez, *Genotoxic effects of the herbicide Roundup Transorb and its active ingredient glyphosate on the fish Prochilodus lineatus.* Environ Toxicol Pharmacol, 2014. **37**(1): p. 448-54.
146. Lopes, F.M., et al., *Effect of glyphosate on the sperm quality of zebrafish Danio rerio.* Aquat Toxicol, 2014. **155**: p. 322-6.
147. Guilherme, S., et al., *Are DNA-damaging effects induced by herbicide formulations (Roundup(R) and Garlon(R)) in fish transient and reversible upon cessation of exposure?* Aquat Toxicol, 2014. **155**: p. 213-21.
148. Akcha, F., C. Spagnol, and J. Rouxel, *Genotoxicity of diuron and glyphosate in oyster spermatozoa and embryos.* Aquat Toxicol, 2012. **106-107**: p. 104-13.
149. Kaya, B., et al., *Use of the Drosophila wing spot test in the genotoxicity testing of different herbicides.* Environ Mol Mutagen, 2000. **36**(1): p. 40-6.
150. Cavalcante, D.G., C.B. Martinez, and S.H. Sofia, *Genotoxic effects of Roundup on the fish Prochilodus lineatus.* Mutat Res, 2008. **655**(1-2): p. 41-6.
151. Cavas, T. and S. Konen, *Detection of cytogenetic and DNA damage in peripheral erythrocytes of goldfish (Carassius auratus) exposed to a glyphosate formulation using the micronucleus test and the comet assay.* Mutagenesis, 2007. **22**(4): p. 263-8.
152. de Castilhos Ghisi, N. and M.M. Cestari, *Genotoxic effects of the herbicide Roundup((R)) in the fish Corydoras paleatus (Jenyns 1842) after short-term, environmentally low concentration exposure.* Environ Monit Assess, 2013. **185**(4): p. 3201-7.
153. De Souza Filho, J., et al., *Mutagenicity and genotoxicity in gill erythrocyte cells of Poecilia reticulata exposed to a glyphosate formulation.* Bull Environ Contam Toxicol, 2013. **91**(5): p. 583-7.
154. Gholami-Seyedkolaei, S.J., et al., *Optimization of recovery patterns in common carp exposed to roundup using response surface methodology: evaluation of neurotoxicity and genotoxicity effects and biochemical parameters.* Ecotoxicol Environ Saf, 2013. **98**: p. 152-61.
155. Guilherme, S., et al., *European eel (Anguilla anguilla) genotoxic and pro-oxidant responses following short-term exposure to Roundup--a glyphosate-based herbicide.* Mutagenesis, 2010. **25**(5): p. 523-30.
156. Guilherme, S., et al., *Differential genotoxicity of Roundup((R)) formulation and its constituents in blood cells of fish (Anguilla anguilla): considerations on chemical interactions and DNA damaging mechanisms.* Ecotoxicology, 2012. **21**(5): p. 1381-90.
157. Marques, A., et al., *Progression of DNA damage induced by a glyphosate-based herbicide in fish (Anguilla anguilla) upon exposure and post-exposure periods--insights into the mechanisms of genotoxicity and DNA repair.* Comp Biochem Physiol C Toxicol Pharmacol, 2014. **166**: p. 126-33.

158. Marques, A., et al., *Erratum To: Progression of DNA damage induced by a glyphosate-based herbicide in fish (Anguilla anguilla) upon exposure and post-exposure periods--insights into the mechanisms of genotoxicity and DNA repair.* Comp Biochem Physiol C Toxicol Pharmacol, 2015. **168**: p. 1.
159. Nwani, C.D., et al., *DNA damage and oxidative stress modulatory effects of glyphosate-based herbicide in freshwater fish, Channa punctatus.* Environ Toxicol Pharmacol, 2013. **36**(2): p. 539-47.
160. Vera-Candioti, J., S. Soloneski, and M.L. Larramendy, *Evaluation of the genotoxic and cytotoxic effects of glyphosate-based herbicides in the ten spotted live-bearer fish Cnesterodon decemmaculatus (Jenyns, 1842).* Ecotoxicol Environ Saf, 2013. **89**: p. 166-73.
161. Poletta, G.L., et al., *Genetic, enzymatic and developmental alterations observed in Caiman latirostris exposed in ovo to pesticide formulations and mixtures in an experiment simulating environmental exposure.* Ecotoxicol Environ Saf, 2011. **74**(4): p. 852-9.
162. Poletta, G.L., et al., *Genotoxicity of the herbicide formulation Roundup (glyphosate) in broad-snouted caiman (Caiman latirostris) evidenced by the Comet assay and the Micronucleus test.* Mutat Res, 2009. **672**(2): p. 95-102.
163. Clements, C., S. Ralph, and M. Petras, *Genotoxicity of select herbicides in Rana catesbeiana tadpoles using the alkaline single-cell gel DNA electrophoresis (comet) assay.* Environ Mol Mutagen, 1997. **29**(3): p. 277-288.
164. Meza-Joya, F.L., M.P. Ramirez-Pinilla, and J.L. Fuentes-Lorenzo, *Toxic, cytotoxic, and genotoxic effects of a glyphosate formulation (Roundup(R)SL-Cosmoflux(R)411F) in the direct-developing frog Eleutherodactylus johnstonei.* Environ Mol Mutagen, 2013. **54**(5): p. 362-73.
165. Yadav, S.S., et al., *Toxic and genotoxic effects of Roundup on tadpoles of the Indian skittering frog (Euflictis cyanophlyctis) in the presence and absence of predator stress.* Aquat Toxicol, 2013. **132-133**: p. 1-8.
166. dos Santos, K.C. and C.B. Martinez, *Genotoxic and biochemical effects of atrazine and Roundup((R)), alone and in combination, on the Asian clam Corbicula fluminea.* Ecotoxicol Environ Saf, 2014. **100**: p. 7-14.
167. Conners, D.E. and M.C. Black, *Evaluation of Lethality and Genotoxicity in the Freshwater Mussel Utterbackia imbecillis (Bivalvia: Unionidae) Exposed Singly and in Combination to Chemicals Used in Lawn Care.* Archives of Environmental Contamination and Toxicology, 2004. **46**(3).
168. Mohamed, A.H., *Sublethal toxicity of Roundup to immunological and molecular aspects of Biomphalaria alexandrina to Schistosoma mansoni infection.* Ecotoxicol Environ Saf, 2011. **74**(4): p. 754-60.
169. Muangphra, P., W. Kwankua, and R. Gooneratne, *Genotoxic effects of glyphosate or paraquat on earthworm coelomocytes.* Environ Toxicol, 2014. **29**(6): p. 612-20.
170. Piola, L., et al., *Comparative toxicity of two glyphosate-based formulations to Eisenia andrei under laboratory conditions.* Chemosphere, 2013. **91**(4): p. 545-51.
171. Kale, P.G., et al., *Mutagenicity testing of nine herbicides and pesticides currently used in agriculture.* Environ Mol Mutagen, 1995. **25**(2): p. 148-153.

172. Chen, L., et al., *The combined effects of UV-B radiation and herbicides on photosynthesis, antioxidant enzymes and DNA damage in two bloom-forming cyanobacteria.* Ecotoxicol Environ Saf, 2012. **80**: p. 224-30.
173. Wang, G., et al., *Damage to DNA caused by UV-B radiation in the desert cyanobacterium Scytonema javanicum and the effects of exogenous chemicals on the process.* Chemosphere, 2012. **88**(4): p. 413-7.
174. Alvarez-Moya, C., et al., *Evaluation of genetic damage induced by glyphosate isopropylamine salt using Tradescantia bioassays.* Genet.Mol Biol, 2011. **34**(1): p. 127-130.
175. Williams, G.M., R. Kroes, and I.C. Munro, *Safety evaluation and risk assessment of the herbicide Roundup and its active ingredient, glyphosate, for humans.* Regul Toxicol Pharmacol, 2000. **31**(2 Pt 1): p. 117-65.
176. Moriya, M., et al., *Further mutagenicity studies on pesticides in bacterial reversion assay systems.* Mutat Res, 1983. **116**(3-4): p. 185-216.
177. Kier, L.D. and D.J. Kirkland, *Review of genotoxicity studies of glyphosate and glyphosate-based formulations.* Crit Rev Toxicol, 2013. **43**(4): p. 283-315.
178. Lioi, M.B., et al., *Cytogenetic damage and induction of pro-oxidant state in human lymphocytes exposed in vitro to gliphosate, vinclozolin, atrazine, and DPX-E9636.* Environ Mol Mutagen, 1998. **32**(1): p. 39-46.
179. Chruscielska, K.B., J.; Kita, K, et al., *Glyphosate: Evaluation of chronic activity and possible far - reaching effects - Part 2. Studies on Mutagenic Activity.* Pestycydy, 2000. **3-4**: p. 5.
180. Ghisi Nde, C., E.C. de Oliveira, and A.J. Prioli, *Does exposure to glyphosate lead to an increase in the micronuclei frequency? A systematic and meta-analytic review.* Chemosphere, 2016. **145**: p. 42-54.
181. Aghadavod, E., et al., *Role of Oxidative Stress and Inflammatory Factors in Diabetic Kidney Disease.* Iran J Kidney Dis, 2016. **10**(6): p. 337-343.
182. Kamceva, G., et al., *Cigarette Smoking and Oxidative Stress in Patients with Coronary Artery Disease.* Open Access Maced J Med Sci, 2016. **4**(4): p. 636-640.
183. Qureshi, M.A., Y.O. Kim, and D. Schuppan, *Hepatocellular carcinoma in nonalcoholic fatty liver disease: A link between oxidative stress and T-cell suppression.* Hepatology, 2016. **64**(5): p. 1794-1797.
184. Sayanthooran, S., et al., *Upregulation of Oxidative Stress Related Genes in a Chronic Kidney Disease Attributed to Specific Geographical Locations of Sri Lanka.* Biomed Res Int, 2016. **2016**: p. 7546265.
185. Turkmen, K., *Inflammation, oxidative stress, apoptosis, and autophagy in diabetes mellitus and diabetic kidney disease: the Four Horsemen of the Apocalypse.* Int Urol Nephrol, 2016.
186. Vakonaki, E., et al., *Complex interplay of DNA damage, DNA repair genes, and oxidative stress in coronary artery disease.* Anatol J Cardiol, 2016. **16**(12): p. 939.
187. Hecht, F., et al., *The role of oxidative stress on breast cancer development and therapy.* Tumour Biol, 2016. **37**(4): p. 4281-91.
188. Li, L. and F. Chen, *Oxidative stress, epigenetics, and cancer stem cells in arsenic carcinogenesis and prevention.* Curr Pharmacol Rep, 2016. **2**(2): p. 57-63.

189. Perse, M., *Oxidative stress in the pathogenesis of colorectal cancer: cause or consequence?* Biomed Res Int, 2013. **2013**: p. 725710.
190. Prasad, S., et al., *Oxidative Stress and Cancer: Advances and Challenges.* Oxid Med Cell Longev, 2016. **2016**: p. 5010423.
191. Toyokuni, S., *Oxidative stress as an iceberg in carcinogenesis and cancer biology.* Arch Biochem Biophys, 2016. **595**: p. 46-9.
192. Kwiatkowska, M., B. Huras, and B. Bukowska, *The effect of metabolites and impurities of glyphosate on human erythrocytes (in vitro).* Pestic Biochem Physiol, 2014. **109**: p. 34-43.
193. Chaufan, G., I. Coalova, and C. Rios de Molina Mdel, *Glyphosate commercial formulation causes cytotoxicity, oxidative effects, and apoptosis on human cells: differences with its active ingredient.* Int J Toxicol, 2014. **33**(1): p. 29-38.
194. Coalova, I., C. Rios de Molina Mdel, and G. Chaufan, *Influence of the spray adjuvant on the toxicity effects of a glyphosate formulation.* Toxicol In Vitro, 2014. **28**(7): p. 1306-11.
195. Gehin, A., et al., *Vitamins C and E reverse effect of herbicide-induced toxicity on human epidermal cells HaCaT: a biochemometric approach.* Int J Pharm, 2005. **288**(2): p. 219-26.
196. Elie-Caille, C., et al., *Morphological damages of a glyphosate-treated human keratinocyte cell line revealed by a micro- to nanoscale microscopic investigation.* Cell Biol Toxicol, 2010. **26**(4): p. 331-9.
197. Bonini, M.G., et al., *The oxidation of 2',7'-dichlorofluorescin to reactive oxygen species: a self-fulfilling prophesy?* Free Radic Biol Med, 2006. **40**(6): p. 968-75.
198. Kalyanaraman, B., et al., *Measuring reactive oxygen and nitrogen species with fluorescent probes: challenges and limitations.* Free Radic Biol Med, 2012. **52**(1): p. 1-6.
199. George, J. and Y. Shukla, *Emptying of Intracellular Calcium Pool and Oxidative Stress Imbalance Are Associated with the Glyphosate-Induced Proliferation in Human Skin Keratinocytes HaCaT Cells.* ISRN Dermatol, 2013. **2013**: p. 825180.
200. Valavanidis, A., T. Vlachogianni, and C. Fiotakis, *8-hydroxy-2' -deoxyguanosine (8-OHdG): A critical biomarker of oxidative stress and carcinogenesis.* J Environ Sci Health C Environ Carcinog Ecotoxicol Rev, 2009. **27**(2): p. 120-39.
201. Jasper, R., et al., *Evaluation of biochemical, hematological and oxidative parameters in mice exposed to the herbicide glyphosate-Roundup((R)).* Interdiscip Toxicol, 2012. **5**(3): p. 133-40.
202. Astiz, M., M.J. de Alaniz, and C.A. Marra, *Antioxidant defense system in rats simultaneously intoxicated with agrochemicals.* Environ Toxicol Pharmacol, 2009. **28**(3): p. 465-73.
203. Astiz, M., M.J. de Alaniz, and C.A. Marra, *The oxidative damage and inflammation caused by pesticides are reverted by lipoic acid in rat brain.* Neurochem Int, 2012. **61**(7): p. 1231-41.
204. Cattani, D., et al., *Mechanisms underlying the neurotoxicity induced by glyphosate-based herbicide in immature rat hippocampus: involvement of glutamate excitotoxicity.* Toxicology, 2014. **320**: p. 34-45.

205. Slaninova, A., et al., *A review: oxidative stress in fish induced by pesticides.* Neuro.Endocrinol Lett, 2009. **30 Suppl 1**: p. 2-12.
206. Cattaneo, R., et al., *Toxicological responses of Cyprinus carpio exposed to a commercial formulation containing glyphosate.* Bull Environ Contam Toxicol, 2011. **87**(6): p. 597-602.
207. Costa, M.J., et al., *Oxidative stress biomarkers and heart function in bullfrog tadpoles exposed to Roundup Original.* Ecotoxicology, 2008. **17**(3): p. 153-63.
208. de Menezes, C.C., et al., *Roundup effects on oxidative stress parameters and recovery pattern of Rhamdia quelen.* Arch Environ Contam Toxicol, 2011. **60**(4): p. 665-71.
209. Ferreira, D., et al., *Assessment of oxidative stress in Rhamdia quelen exposed to agrichemicals.* Chemosphere, 2010. **79**(9): p. 914-21.
210. Geret, F., et al., *Effects of low-dose exposure to pesticide mixture on physiological responses of the Pacific oyster, Crassostrea gigas.* Environ Toxicol, 2013. **28**(12): p. 689-99.
211. Glusczak, L., et al., *Acute exposure to glyphosate herbicide affects oxidative parameters in piava (Leporinus obtusidens).* Arch Environ Contam Toxicol, 2011. **61**(4): p. 624-30.
212. Guilherme, S., et al., *DNA damage in fish (Anguilla anguilla) exposed to a glyphosate-based herbicide -- elucidation of organ-specificity and the role of oxidative stress.* Mutat Res, 2012. **743**(1-2): p. 1-9.
213. Lushchak, O.V., et al., *Low toxic herbicide Roundup induces mild oxidative stress in goldfish tissues.* Chemosphere, 2009. **76**(7): p. 932-7.
214. Modesto, K.A. and C.B. Martinez, *Roundup causes oxidative stress in liver and inhibits acetylcholinesterase in muscle and brain of the fish Prochilodus lineatus.* Chemosphere, 2010. **78**(3): p. 294-9.
215. Ortiz-Ordonez, E., et al., *Effect of Yerbimat herbicide on lipid peroxidation, catalase activity, and histological damage in gills and liver of the freshwater fish Goodea atripinnis.* Arch Environ Contam Toxicol, 2011. **61**(3): p. 443-52.
216. Sinhorin, V.D., et al., *Effects of the acute exposition to glyphosate-based herbicide on oxidative stress parameters and antioxidant responses in a hybrid Amazon fish surubim (Pseudoplatystoma sp).* Ecotoxicol Environ Saf, 2014. **106**: p. 181-7.
217. Uren Webster, T.M., et al., *Effects of glyphosate and its formulation, roundup, on reproduction in zebrafish (Danio rerio).* Environ Sci Technol, 2014. **48**(2): p. 1271-9.
218. Adamson, P., et al., *Time trends in the registration of Hodgkin and non-Hodgkin lymphomas in Europe.* Eur J Cancer, 2007. **43**(2): p. 391-401.
219. Engels, E.A., *Infectious agents as causes of non-Hodgkin lymphoma.* Cancer Epidemiol Biomarkers Prev, 2007. **16**(3): p. 401-4.
220. Hennessy, B.T., E.O. Hanrahan, and P.A. Daly, *Non-Hodgkin lymphoma: an update.* Lancet Oncol, 2004. **5**(6): p. 341-53.
221. Muller, A.M., et al., *Epidemiology of non-Hodgkin's lymphoma (NHL): trends, geographic distribution, and etiology.* Ann Hematol, 2005. **84**(1): p. 1-12.

222. Acquavella, J.F., et al., *Glyphosate Biomonitoring for Farmers and Their Families: Results from the Farm Family Exposure Study.* Environmental Health Perspectives, 2003. **112**(3): p. 321-326.
223. Curwin, B.D., et al., *Urinary pesticide concentrations among children, mothers and fathers living in farm and non-farm households in iowa.* Ann Occup Hyg, 2007. **51**(1): p. 53-65.
224. Wester, R.C., et al., *Glyphosate skin binding, absorption, residual tissue distribution, and skin decontamination.* Fundam.Appl Toxicol, 1991. **16**(4): p. 725-732.
225. Wester, R.C., D. Quan, and H.I. Maibach, *In vitro percutaneous absorption of model compounds glyphosate and malathion from cotton fabric into and through human skin.* Food Chem Toxicol, 1996. **34**(8): p. 731-735.
226. Zouaoui, K., et al., *Determination of glyphosate and AMPA in blood and urine from humans: about 13 cases of acute intoxication.* Forensic Sci Int, 2013. **226**(1-3): p. e20-5.
227. Brewster, D.W., J. Warren, and W.E. Hopkins, *Metabolism of glyphosate in Sprague-Dawley rats: tissue distribution, identification, and quantitation of glyphosate-derived materials following a single oral dose.* Fundam.Appl Toxicol, 1991. **17**(1): p. 43-51.
228. Begley, D.A., et al., *Finding mouse models of human lymphomas and leukemia's using the Jackson laboratory mouse tumor biology database.* Exp Mol Pathol, 2015. **99**(3): p. 533-6.
229. Hori, M., et al., *Non-Hodgkin lymphomas of mice.* Blood Cells Mol Dis, 2001. **27**(1): p. 217-22.
230. Morse, H.C., 3rd, J.M. Ward, and M.A. Teitell, *Mouse models of human B lymphoid neoplasms*, in *The Lymphoid Neoplasms*, I.T. Magrath, Editor. 2010, CRC Press: Boca Ratan, FL.
231. Pattengale, P.K. and C.R. Taylor, *Experimental models of lymphoproliferative disease. The mouse as a model for human non-Hodgkin's lymphomas and related leukemias.* Am J Pathol, 1983. **113**(2): p. 237-65.
232. Ward, J.M., *Lymphomas and leukemias in mice.* Exp Toxicol Pathol, 2006. **57**(5-6): p. 377-81.
233. German Federal Institute for Occupational Safety and Health, *Proposal for Harmonized Classification and Labeling: Glyphosate*, F.O.f. Chemicals, Editor. 2016: Dortmund, Germany.