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.

<u>Human Evidence Findings</u>

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

Year	Strain	Length[1]	Top Dose[2]	Renal Tumors	Hemangio-sarcomas	Malignant Lymphoma
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

Year	Strain	Length	Doses (mg/kg/d)	Response	p-Trend (p-poly3)
1983	Crl:CD-1	24	157, 814, 4841	1/50, 0/49, 1/50, 3/50	0.03 (0.03)
1993	?:CD-1	24	100, 300, 1000	2/50, 2/50, 0/50, 0/50	0.94 (0.94)
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	0/49, 0/49, 1/50, 2/50	0.04 (0.04)
2009	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

Year	Strain	p-Trend (p-poly3)
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

Year	Strain	Length	Doses (mg/kg/d)	Response	p-Trend (p-poly3)
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

Year	Strain	p-Trend (p-poly3)
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

Year	Strain	Length	Doses (mg/kg/d)	Response	p-Trend (p-poly3)
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

Year	Strain	p-Trend (p-poly3)
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.