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

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

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