PAN Germany
Pestizid Aktions-Netzwerk e.V.

Pesticide Action Network Europe

Repeating wrong statements does not make them more correct

A comment on the answer of EFSA's Chief Executive, Bernhard Url, to the open letter of Prof. Christopher Portier

by Peter Clausing, PAN Germany

Hamburg - January 2016

Introduction

In a three-page letter with a 14-page Annex (1) published on EFSA's website EFSA's Chief Executive responds to the open letter (2) to EU Commissioner Vytenis Andriukaitis by Professor Christopher Portier, signed by 96 scientists, subsequently designated *Response*.

As in the previously published analyses (3, 4) the following comments concentrate on the questions around the mouse carcinogenicity studies, because according to Regulation on classification, labelling and packaging ([CLP] 1272/2008, Annex I; 3.6.2.1) an active ingredient of pesticides is to be classified as a carcinogen 1B ("presumed human carcinogen"), if there is "sufficient evidence" from experiments "to demonstrate animal carcinogenicity". The term 'sufficient' has been adopted from the IARC (cf. CLP Regulation 1272/2008, Annex I; 3.6.2.2.3) and is defined 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".

In the *Response* it is claimed that the EFSA assessment "is based on weight of evidence fully in line with the CLP criteria and the ECHA guidance (ECHA, 2013; 2015), regarding the biological relevance of observed incidences for the assessment of the carcinogenicity potential of glyphosate."

These unfunded claims refer to:

- distorted arguments on the chosen method of statistical analysis;
- alleged inconsistencies within and between studies (dose-response-effects); and
- a repetition of previously criticised "lack of biological relevance"-arguments (excessive toxicity, historical control data, alleged viral infections in one study).

Statistical Analysis

A key argument in the *Response* was that for the statistical analysis of the five mouse studies which were evaluated by BfR/EFSA (5) the use of pair-wise comparisons was appropriate whereas the Cochran-Armitage-Trend Test (used by the IARC) was not. In the *Repsonse* it is argued "that the planning of a study before the initiation of the experimentation as established in the respective protocol – which includes the planned statistical analysis – is a key element in assessing the quality of the study; therefore deviations from the statistical analysis used by the study authors should be limited and properly justified."

Then, the *Response* cites from OECD Guidance (6) which states that "the statistical methods most appropriate for the analysis of the data collected should be established at the time of designing the experiment and before the study starts." And continues to claim that "(t)he studies under consideration were designed for pair-wise comparisons".

The problem with EFSA's argumentation is threefold.

First, citing from the OECD guidance it neglects that this guidance uses the plural ("statistical method**s** most appropriate") for good reasons. As explained in the next paragraph a good study plan should describe several statistical methods to be applied depending on the conditions when to use which method. To **design** studies for pair-wise comparisons is simply not possible. Therefore, this is a false statement. EFSA would be in trouble if it should explain how such a design in distinction from a "design" for trend tests would look like.

The type of data generated may require different statistical approaches, e.g. depending whether the survival rate of the animals was affected by treatment or not. This, however, cannot be foreseen when planning the study. That is why a good study plan should consider different options of statistical analysis and the circumstances when to use which option. In the past, this however has rarely been done in the day-to-day practice of regulatory toxicology. Transparency is repeatedly emphasized in the *Response*. Therefore, the EFSA should disclose the study plans of the five mouse studies or at least the sections on statistics of these studies plans to prove to the public that "the statistical methods most appropriate" had been established at the time of designing the experiment.

Second, in the *Response* it is stated, that "deviations from the statistical analysis used by the study authors should be limited and properly justified" (emphasis added). This should not preclude the BfR and the EFSA from a post-hoc application of the most appropriate statistical method. While the OECD Guidance of 2012 (6) discusses both approaches for statistical assessment, i.e. pair-wise comparisons and trend tests, it clearly points to the use of trend tests as the method of choice in its flow chart on page 123. In addition, this guidance states: "In general, testing a trend which is a more specific hypothesis has greater power than a pair-wise comparison" (6, p.118).

Third, the EFSA while insisting on the use of the statistical method with less power states in its *Response*: "It should also be noted that there are no valid studies with statistical significant effects confirmed by both statistical approaches." This statement gives the impression that a confirmation by both statistical approaches would be a pre-condition for considering differences as significant. This statement ignores the identical clear wording which can be found in both the old (10, p. 62) and the new (7, p. 116) OECD Guidance: "Significance in either kind of test is sufficient to reject the hypothesis that chance accounts for the result." (emphasis added).

In summary, the use of pair-wise testing claimed in the *Response* as appropriate has no scientific basis. Instead, the arguments used in the *Response* are either based on invalidated contentions (studies were "designed" for pair-wise comparison) or on out-of-context-reference to applicable OECD guidance.

Biological relevance

For a start, it should be noted that originally BfR's argumentation focused heavily on the (lack of) statistical significance when discussing the increased incidence of malignant lymphoma in the 31 March version of the RAR. For instance in Volume 1 of this document it was argued: "Taking all this information together, a treatment-related effect in the study by (2001, ASB2012-11491) in Swiss albino mice cannot be completely excluded. However, the weak increase in malignant lymphoma even over the historical control of the performing laboratory was clearly confined to this single study and strain since it was not reproducible in four other valid long-term studies" (p. 65, emphasis added). After the publication of the IARC monograph on glyphosate, the BfR re-assessed the studies using the Cochran-Armitage-Trend Test and found out that statistical significance was observed in all five studies. At this point statistical significance was declared unimportant: "It should be avoided to base any conclusion only on the statistical significance of an increased tumour incidence identified in a single study without consideration of the biological significance of the finding" (Addendum 1 to RAR, page iii, emphasis added). Also it should be noted that the BfR in its addendum argues that statistical significance **in a single study** should not be the basis for concluding

carcinogenicity, while there is a significant increase in the incidence in three different tumour types in a total of five different studies.

The first claim concerning the alleged lack of biological relevance offered in the *Response* refers to inconsistent effects, "both within (lack of dose response) and between studies (inconsistencies between results observed at the same dose in different equivalent studies)".

As shown below, a dose-response was visible in most of the studies, in particular in two studies for renal adenoma and malignant lymphoma.

Number of affected males (49 to 50 per group). Data from RAR.

Tumour type	Year of Report	Control	Low Dose	Mid Dose	High Dose
Renal adenoma	1983	0	0	1	3
	1997	0	0	0	2
	2001	0	0	1	2
Heamangiosarcoma	1993	0	0	0	4
	1997	0	0	0	2
Malignant Lymphoma	1993	4	2	1	6
	1997	2	2	0	6
	2001	10	15	16	19
	2009	0	1	2	5

It is common in biological systems like mouse carcinogenicity studies that effects cannot be reproduced in a "congruent" manner, because of biological variability. It can be argued that a significant increase observed in three out of five studies as in case of renal adenoma and malignant lymphoma should be considered as consistency between studies. These effects were seen in only male animals. A careful consideration of the mode of action is considered necessary "to see if the response is consistent with the postulated mode of action" (6, p.377). It would be the responsibility of the BfR/EFSA to perform or to require this. But by refusing to acknowledge carcinogenic effects *per se* these institutions dispose of the necessity of such an assessment or the requirement of targeted mechanistic studies from industry.

The second claim in the *Response* with regard to the alleged lack of biological relevance refers to "excessive toxicity" as a confounding factor. The EFSA rightfully states that "(s)uch toxicity **can** cause effects such as cell death (necrosis) with associated regenerative hyperplasia, which in turn can lead to tumour development as a secondary consequence". However such a statement should be accompanied by facts describing that tissue necrosis (at the site of tumour) really occurred. Once again this gives reason to demand full disclosure of the study reports or – for the sake of transparency – at least those parts that support such a claim.

Another aspect of excessive toxicity is the "Maximum Tolerated Dose" (MTD) which is "conventionally defined as the highest dose to produce toxic effects without causing death and to decrease body weight gain by no more than 10% relative to controls" (7, p. 53). However, when arguing that the MTD was exceeded, as it is done in the *Response* with

regard to the mouse carcinogenicity studies, two things need to be taken into consideration: First, it should be noted that one concern of reduced body weight gain of too high doses in carcinogenicity studies is that a decreased body weight gain could mask carcinogenic effects rather than exaggerating them: “It is now recognised that there is a positive correlation between body weight and the occurrence of certain tumours in rodent species and strains used in safety assessment or for hazard identification; … Moreover, the lower the body weight, the less sensitive the animal may be to agent-induced toxicity, including cancer.” (7, p. 64). Moreover, it should be noted, that in the study where the data were available in the RAR (study of 1997, Volume 3, Annex B.6., p. 522) the decreased body weight gain was associated with a lower food consumption (most likely due to palatability) casting further doubt on the “excessive-toxicity”-argument.

Finally, as discussed in earlier analyses (4, 5) two studies with significant increases in th tumour incidence had top doses at or below 1.000 mg/kg. In addition, as pointed out in the OECD guidance (7, p. 53) in particular toxicokinetics is a criterion to be taken into consideration when selecting an adequate top dose. Glyphosate has an absorption rate of only 20-30% after oral administration. Taking this into account none of the top doses used in any of the mouse studies seem to be exaggerated.

The third claim in the Response with regard to biological relevance concerns the lack of preneoplastic lesions changes in organs where tumours occurred which would show a histopathological continuum in the development of tumours. Besides that a lack of the observation of preneoplastic changes would not invalidate a neoplastic tumor finding,

this can only be judged by re-evaluating the histopathological slides which would be desirable in the light of the many inconsistencies presented by the EFSA. In addition, it is worth noting that for the study of 1993 the incidence of malignant lymphoma listed in Table 2-6.9 carries the footnote “based on histopathological examination of lymph nodes with macroscopic changes” (RAR, Volume 1, p.64, emphasis added). When tumour incidences are determined only in organs with macroscopic changes, how is it possible to assess preneoplastic changes at all? Here, a re-evaluation of the histopathological slides seems inevitable. EFSA needs to clarify the situation.

The fourth claim in the *Response* with regard to biological relevance concerns the use of historical controls. Here, the *Response* states without any additional explanation “that the letter signatories have misinterpreted the efforts made by the German RMS to get supportive information for those studies with no valid historical controls” and that the EFSA “only considered valid the historical control data from the performing laboratory”. This is a bold distortion of facts. First of all, by claiming that historical control data speak against the biological relevance EFSA and BfR conceal that there are two studies where the (valid!) historical controls (the studies of 1997 and 2001) support the finding of an increased incidence of malignant lymphoma. For the study of 2001 it is stated “The incidence was statistically significantly elevated as compared to the actual control groups in this study, was above the mean values of the (relatively small) historical control and, for males, outside the historical control range” (RAR Volume 3, Annex B.6., p. 510). For the study of 1997 it is stated “8 of 9 studies had a control incidence below 12 % (6 % or lower) as observed now at the top dose level” (RAR Volume 3, Annex B.6., p. 510). Turning these facts upside down, in the RAR it says (referring to the studies of 2001 and 2009): “The slightly higher incidences in top dose males in the two studies in CD-1 mice were not statistically significant and fully

covered by historical control data" (RAR Volume 1, p.65). For the 2001 study, this is repeated in the Addendum: "Also in the study with Swiss mice, which have considerably higher background incidences for malignant lymphomas, the observed incidences were within the historical control range." (Addendum 1 to RAR, p.92, emphasis added). Furthermore it needs to be emphasized that for the study of 2009 where the German RMS "made efforts" to get supportive information, it was admitted that the historical control data supplied by this laboratory were unusable: "However, the quality and regulatory value of the historical control data is very much compromised by the fact that the sexes were not considered separately" (RAR, Volume 3, Annex B.6., p. 517). This did not stop the RMS to claim in Volume 1 of the RAR that the "higher incidences in top dose males … was fully covered by historical control data" (RAR Volume 1, p.65).

The other "efforts made by the German RMS to get supportive information" as stated in the Response, consisted in dismissing important rules for the use of historical control data as established by the OECD (see 3, p. 4-6 and 4, p. 7 for a detailed discussion), after this dismissal historical controls largely disconnected from the original studies are abusively used in an attempt to invalidate the findings of carcinogenic effects of glyphosate.

The issue of presumed viral infections

To grasp the spirit in which the *Response* is written, it is worth looking at the section on "Additional considerations of the tumours reported in the IARC monograph". The evolution of arguments highlights the approach the EFSA is using:

1. In the RAR itself the issue of viral infections in Swiss albino mice as a contributing factor to the occurrence of malignant lymphoma is as follows: "According to a more recent article (Taddesse-Heath et al., 2000, ASB2015-2535), a much higher incidence of hematopoietic neoplasia of 58% was observed in a colony of CFW Swiss mice in the USA. … The authors ascribed these tumours mainly to 'infectious expression of murine leukemia viruses'. <u>It is not known to which extent such a latent infection might have contributed to lymphoma incidences reported earlier or even in the studies described in this RAR</u>" (RAR Volume 3 Annex B.6, p. 511, empahsis added).
2. In the EFSA conclusion this reads as follows: "The study was re-considered during the second experts' teleconference (TC 117) as not acceptable due to viral infections that could influence survival as well as tumour incidence – especially lymphomas" (11, p. 10). At this point it should be noted that the survival rate of the mice in this study was within the historical range (see Table B.6.5-41, RAR Volume 3, Annex B.6., p. 504).
3. When discussing <u>renal tumours</u> reported in mice, it is stated in the Response that "(a) fifth study performed on Swiss albino mice (2001) was concluded to be unreliable since the health of the animals in the study was <u>clearly compromised due to viral infections in all groups including concurrent control</u>" (emphasis added).

This is a remarkable twist in the argumentation. While in the RAR itself with reference to a single publication it is admitted that "it is not known to which extent such a latent infection might have contributed to lymphoma incidences reported … in the studies described in this RAR", a later teleconference declares the study inacceptable "due to viral infections" which then "clearly compromised" the health of the animals in this study according to the *Response*.

Conclusion

In the Response letter of EFSA's Chief Executive, a number of invalid arguments which had been criticised earlier were repeated. Furthermore in the Response it is claimed that trend tests cannot be used for the statistical analysis in the tumour studies presented in the RAR. As shown above, this claim has neither a formal nor a scientific basis. Furthermore, a bold distortion of facts has been identified in the Response concerning the use of historical control data. Other aspects, such as the "invalidation" of one study based on an alleged virus infection represent a distortion of the facts contained in the RAR.

The Commissioner Vyteni Andriukaitis should require a correction of these distortions and inconsistencies before making a decision about the future fate of glyphosate as an active ingredient of pesticides in the European Union.

In a more general sense it is urgent to correct the way such issues are handled to re-establish confidence the authorities.

Notes

1. http://www.efsa.europa.eu/sites/default/files/EFSA_response_Prof_Portier.pdf
2. http://www.efsa.europa.eu/sites/default/files/Prof_Portier_letter.pdf
3. http://www.pan-germany.org/download/PAN_Germany_Addendum_analysis_09112015.pdf
4. http://www.pan-germany.org/download/Analysis_EFSA-Conclusion_151201.pdf
5. BfR = Bundesinstitut für Risikobewertung (Germany's Federal Institute for Risk Assessment)
6. https://echa.europa.eu/documents/10162/13562/clp_en.pdf
7. http://www.oecd.org/officialdocuments/displaydocument/?cote=ENV/JM/MONO(2011)47&doclanguage=en
8. RMS = Reporting Member State
9. RAR = Renewal Assessment Report
10. http://www.oecd.org/officialdocuments/displaydocument/?cote=env/jm/mono(2002)19&doclanguage=en
11. EFSA (2015): Conclusion on the peer review of the pesticide risk assessment of the active substance glyphosate. EFSA Journal 2015;13 (11):4302.

Imprint

Pestizid Aktions-Netzwerk (PAN) e.V. Nernstweg 32 D-22765 Hamburg Phone: +49 (0)40-399 19 10-0 Email: info@pan-germany.org	**PAN Europe** Rue de la Pépinière 1 B-1000 Brussels Phone: + 32 2503 0837

Hamburg, 20 January 2016

Author: Dr. Peter Clausing,
Phone:, +49-176 7801 2705, Email: peter.clausing@pan-germany.org,