Appendix A

Review of Meta-Analyses on Glyphosate

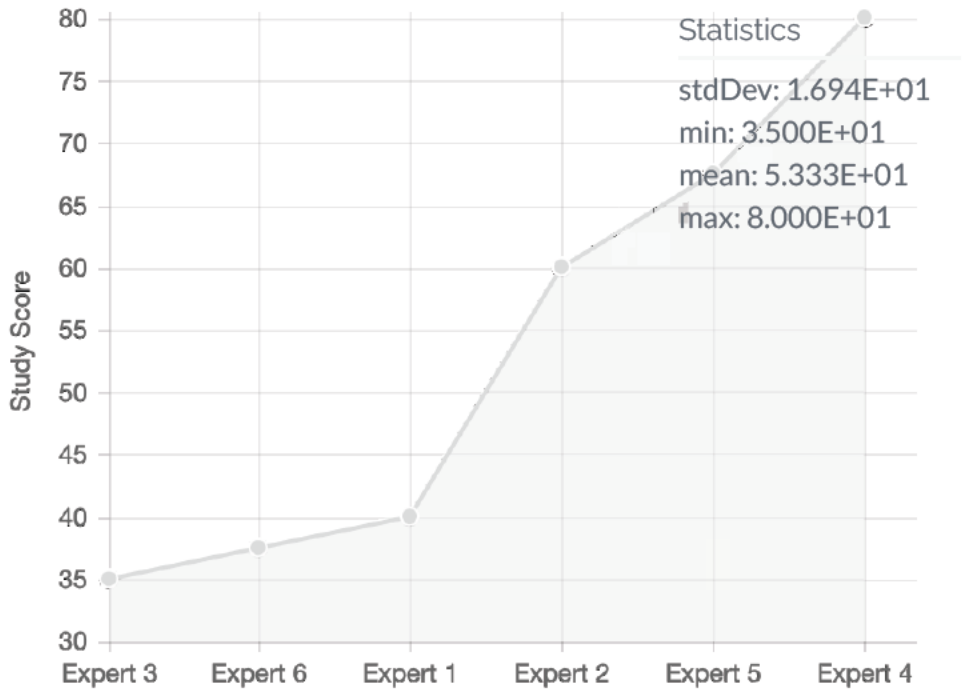
Scoring meta-analyses studies based on their strengths and limitations

Generated: 2021 08 06 20:42:50 +0000
URL: https://app.scipinion.com/scipis/352/report

1) META-ANALYSIS SCORING

Result 1.1 *(ID: 5172)*
Question 1.1 *(ID: 4156)*

On a scale of 1-10 (1=lowest, 10=highest) what is your degree of confidence in the following factors for study A: Shinasi and Leon (2014). Please provide an explanation for EACH rating in the space below.



Answer Explanations

40.0 Expert 1

Factor	Factor Rating::1
Methods::25	4
Results::25	4
Discussion/Conclusions::25	4
Application to Risk Assessment Decision-Making::25	4

This was the first systematic review and meta-analysis to assess the link between glyphosate and non-Hodgkin's lymphoma. This is a comprehensive systematic review and meta-analysis of a large number of pesticide chemical groups (n=21) and 80 active ingredients. Only two electronic databases were searched (PubMed, Web of Science) as part of the literature search strategy. Only studies in English were searched. They did not follow the PRISMA Guidelines. They did not undertake a quality assessment of the studies. There was no assessment of publication bias. Sub-group analysis was conducted.

37.5 Expert 6

Factor	Factor Rating::1
Methods::25	5
Results::25	4
Discussion/Conclusions::25	4
Application to Risk Assessment Decision-Making::25	2

Only ever exposure was considered. Only NHL was explored, although acknowledging the heterogeneity of lymphoma subtypes. Glyphosate results are presented as significantly increased, and a note in the results section mentioned that an analysis of exposures starting from 1975, the year when glyphosate was introduced in the market, showed significantly increased risk. There is no mention of such findings in the Discussion.

67.5 Expert 5

Factor	Factor Rating::1
Methods::25	7
Results::25	7
Discussion/Conclusions::25	7
Application to Risk Assessment Decision-Making::25	6

Methods: In the articles used for this meta-analysis a variety of exposure measurements were used. The Schinasi et al authors decided to use the results provided for the dichotomously defined exposure with the greatest number of cases. Results: It is not clear how the exposure information from the underlying studies were combined in the final meta-analysis. It seems that a lot of important confounding variables are controlled in some studies (ie., Eriksson et al, 2008) while in others not (i.e, Hardell et al, 2002). Discussion/Conclusions: Due to the decision to use dichotomous exposure reported in the studies, several aspects of exposure, such as exposure lags and duration of exposure were ignored. Application to Risk assessment decision-making: The results of this review should be interpreted and used with caution. Although it has combined evidence from large studies there is a lack of detail in how exposure categories were treated. Eriksson et al, 2008 showed a dose-response relationship when >10 days/y exposed while De Roos et al 2005 using categories for multiple exposure (measured in cumulative days) did not find a relationship.

60.0 Expert 2

Factor	Factor Rating::1
Methods::25	6
Results::25	7
Discussion/Conclusions::25	6
Application to Risk Assessment Decision-Making::25	5

For the most part, this was a well-conducted systematic review and meta-analyses. The search strategy was clearly articulated. The paper considered several other pesticides other than glyphosate, and therefore, only a portion of the papers was directly relevant to associations between glyphosate

and NHL. I have some concerns about the methods used to identify the relevant measure of association. These are described in more detail below. While the search strategy was well defined, and formally presented in the supplementary materials for the paper, the authors did not give sufficient details about how the review of the papers was carried out by the team. Generally speaking, multiple reviewers carry out the review of abstracts for relevance, and processes are put in place to deal with inconsistencies between reviewers but these methods are not described in the paper. The authors did search multiple databases (Pubmed and Web of Science) for relevant papers which is important to do in a systematic review. It is unclear whether they searched the grey literature for possible theses and governmental reports to inform the analyses. However, the fact that subsequent meta-analyses did not yield publications in this area provides some re-assurance, and it is my view that the appropriate papers to conduct the meta-analyses were identified. As with other meta-analyses of this research question, meta-analytic methods are severely limited due to the small number of studies. To their credit, the authors acknowledge this. Standard meta-analysis procedures, such as assessing heterogeneity and publication bias cannot be carried out due to the small numbers of studies. Not only are there relatively few studies that were identified to be included in the meta-analyses, but the precision of the measures of association (odds ratios) in these individual studies are poor as they are based on small numbers of cases (and controls). This is particularly problematic for assessing causality for subtypes of leukemia. To illustrate, the odds ratio for B cell lymphoma from the Coco study is based on 4 cases, and the odds for those exposed relative to those not exposed is 3.1 (0.6-17.1) lacks precision. Similarly, the odds ratio and 95% confidence interval for lymphocytic leukemia was 0.9 and 0.6 to 5.8. The summary meta-analysis risk estimate for glyphosate (exposed compared to non-exposed was 1.5 (95% CI: 1.1-2.0). This is similar to the summary estimate from the Chang and Delzell paper (1.3, 95% CI=1.0-1.6). On closer inspection these two results differed due to the decision of Shinasi and Leon to use the higher odds ratio presented in the Hardell study (OR=3.0, 95% CI=1.1-8.5) rather than the adjusted OR=1.85 (95% CI=0.55-6.20) – these two risk estimates are presented in Table VII of the Hardell et al case-control study. While I can appreciate the choice to use the higher risk estimate (OR=3.0) the other estimate may be preferred because it is an adjusted risk estimate, and the authors themselves indicated an intent to use adjusted measures of associations in their methodology. Sinasi and Leon chose to use other adjusted risk estimates for the other studies to it is not clear why they opted not to for the Hardell study. In addition, they also selected the higher risk estimate from the De Roos et al 2003 study (logistic OR=2.1, 95% CI=1.1-4.0) rather than the estimate from hierarchical regression (OR=1.6, 95% CI=0.9-2.8) which was substantially lower and not statistically significant. Finally, they selected the higher summary risk estimate from the Eriksson et al paper (OR=2.0, 95% CI=1.1-3.7 – Table 4, page 4505) whereas Eriksson also published an adjusted risk estimate of 1.51 (95% CI=0.77 to 2.94 in their paper – Table VII, page 1661). The preferred approach would have been to conducted separate meta-analyses to estimate the summary overall measures of association using these two different measures from De Roos as was done by Chang and Delzell. In my view, the Chang and Delzell is a superior paper. In addition, I sometimes have concerns with the measures of association reported by Hardell and colleagues in his publications that span across several topics. His measures of associations tend to be outliers, and usually substantially higher than estimates from other studies. This is shown with his work on cell phones and glioma (https://pubmed.ncbi.nlm.nih.gov/25466607/), as well as pentachlorophenol (see slide 12 of this presentation: http://www.occupationalcancer.ca/wp-content/uploads/2013/05/Demers_PCP-Users_NTP_May2013.pdf). My comment is not intended to criticisz this research group, but I am just making a general observation that publications by Hardell et al tend to be much higher than other

research papers across many exposures. A minor comment, in the Shnasi and Leon paper (2014) the forest plot figure S1 (D) seemingly includes an incorrect confidence interval for the Hardell study. On the figure, the 95% CI goes from 1.08 to 8.54 while the 95% confidence interval as presented in Hardell's paper is 1.08 to 8.52. This is a very minor error, and unlikely to substantially change summary risk estimate presented in the Shinasi and Leon paper. Shinasi and Leon's discussion section paper appropriately acknowledged limited ability to assess exposure response relationships between exposure and outcome and evaluate the relevance of exposure lags, and duration of exposure. That said, I take the view that the validity of the risk estimates for each individual study needs to be assessed when evaluating the validity of the summary risk estimate derived from a meta-analysis. Shinasi and Leon's paper would be be better had it discussed some of these limitations. The summary risk estimate (OR=2.0, 95% CI=1.1-3.6) for B cell lymphoma (page 4513) is based on only 2 studies, and one of these may be influenced by substantial bias (Coco et al, see below). As a result, the validity of this summary measure is questionable. The validity of a meta-analysis is dependent on the quality of the individual studies, and therefore authors should strive to identify important sources of bias in the individual studies. In my view, the relevance of participation bias was not given sufficient description. For example, one of the studies looked at (Coco et al) was only able to achieve a response rate of 52% in the control series. This could introduce an important source of bias if agricultural workers were underrepresented in the control series. The impact of this bias would be an upward bias in risk due to underrepresentation of agricultural workers in the control series. With respect to the criteria of consistency, there was substantial variability in the odds ratios with 3 of the 6 studies (Table 4 p page 4505) essentially null findings. Nonetheless, if one were to use the adjusted risk estimate from the Hardell study (OR=1.85) you would have three studies showing essentially a double in risk of NHL among those exposed, with the remaining studies null. Altogether I would conclude that the evidence in terms of consistency was 'weak', and there is the potential for bias. Similarly, there is potential for bias in almost any control series and the paper does not explore this source of bias. The paper was published in the journal International J Environmental Research in Public Health. I have also published there, but do recognize that some MDPI journals have been considered as predatory having made Jeffrey Beall's list. Moreover, the MDPI journal, Entropy, published a review paper by Samsel et al claiming glyphosate may be the most important factor in the development of a number of health outcomes The paper itself does not contain any primary research results and was subsequently criticized (Kloor, Keith. "When Media Uncritically Cover Pseudoscience". Discover Magazine. 2014.) For me, there are no apparent conflicts of interest by the authors, and the work was funded through government funding in the EU. Overall, the study provides very weak evidence to support a causal association between glyphosate and NHL. This is due primarily to limited number of studies, crude characterization of exposure (ever vs never). The summary risk estimates are more elevated that what I should be due to decisions of the authors to consistently incorporate the higher risk estimates that were presented in the individual level studies into their meta-analyses. This may well be traced back to their stated decision to include results from the 'most complete and updated analyses with the greatest number of participants'. As a result they have opted at the onset to place more validity on measures of association derived with a greater number of participants, than to account for the potential confounding influence of other relevant exposures. However, there criteria (3) stated in Section 2.5 indicates that 'in an effort to use the most unbiased estimate, they extracted the most adjusted risk estimate'. From my perspective, they did not follow their stated methods given they frequently selected the stronger unadjusted risk estimate in at least 3 studies.

35.0 Expert 3

Factor	Factor Rating::1
Methods::25	4
Results::25	4
Discussion/Conclusions::25	3
Application to Risk Assessment Decision-Making::25	3

Methods – The authors adequately describe their methods but don't adhere to their plan. Results – The results for univariate analyses were selected for 2 studies instead of the lower multivariate. Discussion – since this paper was not focused on glyphosate alone, the discussion was much less specific to the topic at hand.

80.0 Expert 4

Factor	Factor Rating::1
Methods::25	7
Results::25	9
Discussion/Conclusions::25	10
Application to Risk Assessment Decision-Making::25	6

The study carefully integrated quantitative evidence from meta-regressions and qualitative evidence from exposure-response modeling, as well as conducted extensive sensitivity analyses of some key features of work that can bias the results. Sensitivity analysis missed consideration of quality of exposure assessment, selection bias, and degree to which crude and adjusted effect estimates changes (a measure of residual confounding). Although the analysis did not reveal a lot of heterogeneity in meta-estimates (33%), it still would have been good to assess which factors contributed to heterogeneity, given that other similar work produced different estimates of heterogeneity. This may have been difficult due to limited number of informative studies (six). The analysis has limited value to risk assessment because it did not entail quantitative bias and uncertainty analysis. The reported effect estimates for glyphosate and NHL are imprecise, signaling that true uncertainty is even greater, as such, any qualitative decision of whether association of glyphosate and NHL is causal has weak support (certainly far less than 95% certainty). This paper analyses the same material as Chang and Delzell (2016) but makes somewhat different analytical decisions; such difference in approach do not appear to impact the conclusions that an independent reader can draw from the two articles, which place summary RR somewhere between 1 and 2 with 95% confidence.

Comments 4

SCORE 1

Expert 2 *07/16/2021 11:17*

Important comment by Expert 5/6 who indicated "Only NHL was explored, although acknowledging the heterogeneity of lymphoma subtypes". Related to this point, I think important to recognize that Leon, an author of this paper, and lead author on subsequent paper published much later in 2019 which relied only on cohort

data, and highlighted subtype relevance of NHL subtypes. Unfortunate that the latter publication never attempted to compare/contrast estimates between the two studies she was an author on. The fact that the authors made several decisions that were not described in the paper, and they did not follow their stated plan is a concern. One of the major concerns I have about all the case control studies is participation bias, and perhaps greater participation from population controls who would be unlikely to be exposed to glyphosate - one reason I think the cohort studies have more validity Also, I can appreciate the concern about looking at ever exposure (expert #6) - I don't see a major issue with this given what I see as impracticalities of standardizing exposure concentrations across studies- and if the goal to determine whether there is simply an increased risk of NHL with exposure to glyphosate, use of EVER vs NEVER is reasonable to answer this question if levels of EVER exposure is meaningfully above background levels.

SCORE
0

Expert 5

07/19/2021 01:02

I find the answer of reviewer 2 about the validity of the summary estimate very important. From this meta-analysis, it is not justified why the authors have decided to include the risk estimates they did for each study, and, where available, ignoring the adjusted estimates.

SCORE
-2

Expert 6

07/19/2021 07:25

As Olav Axelson pointed out, misclassification of diseases is as bad as misclassification of exposure. The dilution of the relevant association with a specific lymphoma subtype (if any) within the NHL group, matched with the dilution of the severely exposed within the ever exposed category, subtracts credibility to the NHL findings. This comment refers to the meta-analyses that did not explore subtypes (Donato et al, Kabat et al.). For this reason, I found of interest the meta-analysis by Kabat et al., although I would agree that results from different exposure metrics should not be combined in summary estimates, but rather looked at in the overall picture. Nonetheless, the authors tried to prevent dilution of the association (if true) by identifying the most exposed subjects within each study. Schinasi & Leon also presented results for the major B cell lymphoma subtypes, but only used the ever/never dichotomous classification of exposure, and, consistently with the other meta-analyses, picked up risks with different levels of adjustment, which is a further reason for concern. I am much less concerned about participation bias in case-control studies. Most of these studies were on the field much before IARC examined glyphosate; different from self-reports, expert assessments (Cocco P et al.) are not affected by recall bias; it is unclear why farmers should have

accepted participation less frequently, and, anyway, about 10% glyphosate use occurs in non-agricultural sectors; interviews were usually conducted blind to the case-control status. On the other hand, cohort studies only assess exposure at one point in time, and average level but not cumulative exposure was considered in the AHS study. I agree with expert 2 about the fact that "meta-analytic methods are severely limited due to the small number of studies". Currently, we are considering 6 original case-control studies, 1 follow-up of cancer incidence and 2 updates, and 1 pooled analysis of case-control studies, that were examined in 6 meta-analyses. Although exploring from different perspectives, seems to me that too many meta-analyses are stretching results from very few original studies. It is true that the measures of association in the individual studies are imprecise, because based on small numbers of cases (and controls), but this is exactly why meta-analyses are needed. A new paper on glyphosate has just been published (Meloni F, et al. Environ Health. 2021;20(1):49), and I am aware of a large new analysis in progress by the InterLymph Consortium. For this reason, it is my opinion that new meta-analyses should be put on hold until new, more powerful studies with a careful exposure assessment will be published.

SCORE
0

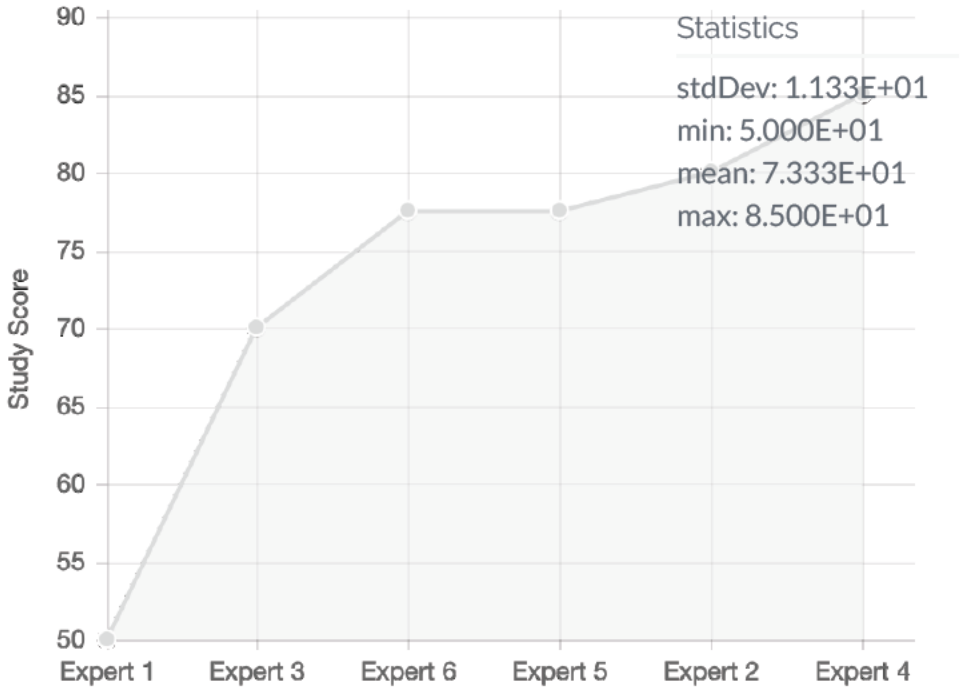
Expert 1

07/21/2021 21:27

I appreciate all the comments by experts #2, #5, and #6. This is one of the smaller Powered meta-analyses with 207 cases and 244 controls. The results are reported as relative risks, even though the majority of studies were reported using odds ratios, did the authors convert OR to RR, if so how, not mentioned in the paper. The point estimates used in this meta-analysis do not match those of the original studies. Bias associated with selection of controls is obviously an important factor in any study, but as long as the control subjects were sampled from the same population, I don't think this should be a major concern.

Result 1.2 *(ID: 5173)*
Question 1.2 *(ID: 4157)*

On a scale of 1-10 (1=lowest, 10=highest) what is your degree of confidence in the following factors for study B: Chang and Delzell (2016). Please provide an explanation for EACH rating in the space below.



Answer Explanations

70.0 — Expert 3

Factors	Factor Rating::1
Methods::25	8
Results::25	8
Discussion/Conclusions::25	9
Application to Risk Assessment Decision-Making::25	3

The methods section is strong with rules for selecting the RR to use. Chang et al. included comprehensive information on the studies included in their meta-analysis. This increases transparency for reader to see the many analyses available. The authors comment on the role of MA: "meta-analyses are not intended to identify, validate, or dispute causal relationships. .useful in proving a summary measure of association." This view is evident in their extensive discussion on components of exposure response and sources of bias. As noted by the authors …no two studies used the same set of 3 or more categories to classify glyphosate exposure (dose-response). These data really aren't well suited for risk assessment when relying on the MA results. Results – comprehensive in reporting on each study design, exposure assessment and results for each analysis. These tables were a good resource when reading the other MA>

85.0 — Expert 4

Factors	Factor Rating::1
Methods::25	9
Results::25	9
Discussion/Conclusions::25	10
Application to Risk Assessment Decision-Making::25	6

The authors analyzed the same papers as Shinasi and Leon (2014). They made somewhat different decisions about inclusion of specific RR from the six studies into meta-regression and defended their views on which RR was "most adjusted" (both articles claim to focus on the "most adjusted" effect estimates, so that decisions seems somewhat subjective). In the end though, the meta-regression RR from the two analysis are impossible to tell apart: they have 95%CI that round up to 1 to 2. The point estimates oscillate from just below 1.5 to just above but that is not the most salient feature of the result. It would be incorrect to focus only on point estimate and give next to no weight to estimated distribution of the estimates of the meta-RR. The authors conducted extensive analysis of publication bias (finding no evidence of it) as well as reasonable sensitivity analyses. Discussion of sources of bias is systematic and nuanced, giving appropriate context for interpreting the literature. The analysis has limited value to risk assessment because it did not entail quantitative bias and uncertainty analysis, e.g. it did not use cited participation rates and reliability/validity estimates of exposure assessment to evaluate what the resulting bias and cost to certainty are. (Such analyses are rare, and yet are sorely missed if epidemiology is to support risk assessment; there are precedents of such analyses in addressing questions of cancer risk of glyphosate by Prof Timothy Lash). The reported effect estimates for glyphosate are imprecise, signaling that true uncertainty is even greater, as such, any qualitative decision of whether association of glyphosate and NHL is causal has weak support (certainly far less than 95% certainty).

50.0 Expert 1

Factors	Factor Rating::1
Methods::25	5
Results::25	5
Discussion/Conclusions::25	5
Application to Risk Assessment Decision-Making::25	5

This systematic review and meta-analysis looks at glyphosate and lymphohematopoitic cancers. Three electronic databases were searched (PubMed, Web of Science, Google Scholar) as part of the literature search strategy. Two authors independently selected the studies. There was no mention as to any language restrictions. They did not follow the PRISMA Guidelines. There was a quality assessment using Austin Bradford Hill's 'viewpoints". There was no assessment of publication bias. Sub-group analysis was conducted.

77.5 Expert 6

Factors	Factor Rating::1
Methods::25	7
Results::25	8
Discussion/Conclusions::25	9

Application to Risk Assessment Decision-Making::25	7

Very well conducted, although metaestimates keep mixing case-control and cohort studies. Clear and detailed evaluation of strengths and weaknesses of each paper. Well conducted analysis of results on subtypes, trends and time since first use. Thorough and careful interpretation and discussion of results

77.5

Expert 5

Factors	Factor Rating::1
Methods::25	8
Results::25	7
Discussion/Conclusions::25	8
Application to Risk Assessment Decision-Making::25	8

Methods: The authors describe in detail and adequately the methodology and performing several sensitivity analyses and small meta-analyses based on specific assumptions. Results: The results are adequately presented via various models used in secondary analysis. No heterogeneity among the studies was observed and the authors report a marginally significant positive meta RRs for the association between glyphosate use and risk of NHL. Discussion/Conclusion: the authors suggest that the positive results should be considered with caution in light of methodological issues, such as study design, selection bias and exposure misclassification. Application to Risk Assessment Decision-Making: the study is well performed taking into consideration the methodological weaknesses of the case-control studies included. The authors have performed sensitivity analyses on several factors of heterogeneity.

80.0

Expert 2

Factors	Factor Rating::1
Methods::25	9
Results::25	9
Discussion/Conclusions::25	7
Application to Risk Assessment Decision-Making::25	7

This paper was published (2016) subsequent to the Shinasi and Leon report, though the meta-analyses was really based on the same studies. This is a well written paper, and the background provided good details on assessments by the European Food Safety Authority that concluded ‘no carcinogenic risk to humans us to be expected from glyphosate if it is used in the intended manner for the intended purpose’. The tone of the paper is more measured than the Shinasi and Leon report, which from my reading of it seems to believe apriori of an association given ‘striking increases’ in NHL and vast toxicological and immunological effects of pesticides. Chang and Delzell also draw attention (as I did above) to the fact that Dr. Shinasi et al also chose some of the higher and unadjusted risk estimates. The authors share the same philosophy that I have that meta-analyses should consider the quality of the individual studies, and their paper cited the qualitative review done by Mink et al. on this topic. The review of papers was done independently by two reviewers, and this is standard practice when conducting systematic reviews. They searched three different databases (MEDLINE, Web of

Science, Google Scholar). Importantly, they assessed the methodological qualities of each study, in particular, they considered potential for selection bias, information bias, exposure misclassification, confounding reporting bias, as well as other issues that could impact the validity of the study. The meta-analyses used the same 6 studies used by Shinasi and Leon, however the specific risk estimates differed somewhat as Chang and Delzell chose the most fully adjusted risk estimates. In my view, they selected the more pertinent measures of association. They acknowledged they could not combine multiple studies using 3 or more exposure categories as these differed across studies. Therefore, these analyses had limited ability to look at exposure response relationships. The authors assessed publication bias using the Egger test, however, the small number of studies severely limit ability to assess this form of bias. They assessed heterogeneity of risk estimates across studies and fit both fixed and random effects models in the meta-analysis. As outlined in the commonly used text (Borenstein et al, Introduction to Meta-Analyses) random effects modelling is almost always applied as it can account for differences in strength of associations that come from having slightly different study populations, designs or approaches to exposure characterization. The random effects model does not assume that the measure of association is fixed, but can vary across a distribution of values. In my view, the summary measure derived from a random effects model is preferred, and the small numbers of studies (n=6) seriously impair any ability to look at heterogeneity. So, while I think the authors followed the most appropriate meta-analytic methodology, these methods really don't perform well for such a small number of studies. In my view, the authors were thorough by conducting multiple meta-analyses of NHL and glyphosate that accounted (or assessed) the impact of using different measures of association from a particular study with respect to deriving the overall summary measure of association. For example, they modelled separately the risk estimates from De Roos (logististic regression and hierarchical regression). The summary risk estimates were very consistent (RR=1.3-1.4 across 4 different models). Their summary risk estimate for the B cell lymphoma (RR=2.0) was the same as that derived by Shinasi and Leon – however, as mentioned earlier this is only based on two studies, and I have concerns about possible participation bias in the Cocco study. In my view there is too much uncertainty with this estimate, and it is impossible to validly contrast risk estimates between glyphosate and subtype of NHL given such a small number of studies, small sample sizes, and even more so the case when there is heterogeneity in effect sizes for all NHL (combined). Two studies provided some support for an exposure response relationship (Eriksson, 2008 and McDuffie), however these findings could not be combined as it was not possible to standardize the two measures (>10 vs ≤10 days of lifetime glyphosate exposure; and >2 days of glyphosate exposure per year versus ≤2 days of exposure per year). Findings by De Roos et al found no exposure-response relationship in the cohort analyses of the Agricultural Health Study. The authors provided a nice discussion about the limitations of the studies, including participation biases, and missing data in the AHS. As such, they acknowledge many of the uncertainties, and appropriately drew attention to limitations of their analyses. I generally agree with their conclusions that there is uncertainty in the validity of the risk estimates due to possible biases, and confounding. I disagree with their characterize of the strength of the association ("marginally significant") - to me, an RR of 1.3 or 1.4 represents a >30% increase in risk for NHL, and given the number of workers exposed, if these risks were accurately, this could account large number of NHL cases. They also drew attention to important limitations of timing of exposures, and limitations particularly for case-control studies to accurately obtain information about the timing of the exposure relative to onset of disease. Finally, they acknowledged the stronger associations in the Swedish case control studies, and suggested this could be due to the decision to construct a reference group that did not use any pesticides at all (whereas others studies defined exposure group

as those that did not use glyphosate. From my perspective, there would be a need to explore these associations in a lower income country where perhaps exposure profile of glyphosate may be higher, and fewer safety precautions available to workers. With respect to financial interests, it is important to note that this study was funded by Monsanto (manufacturer of glyphosate) and therefore raises some potential for bias. The disclosure statements indicates that the authors retained sole control of the manuscript content and the findings. It does not appear as though the search was pre-registered The overall quality of the journal (Journal of Environmental Science and Health, Part B) is somewhat modest and it does have a very low impact factor (0.831 in 2019)

Comments 3

SCORE
1

Expert 2

07/16/2021 11:28

Most comments seem to suggest that this study was well carried out. I really like the comment by Expert 4 "the analysis has limited value to risk assessment because it did not entail quantitative bias and uncertainty analysis, e.g. it did not use cited participation rates and reliability/validity estimates of exposure assessment to evaluate what the resulting bias and cost to certainty are." I agree some uncertainty analysis would be useful, but these aspects are difficult to characterize, and I have concerns about doing uncertainty analysis given such a small number of study (in total). Even the small number of studies preclude a reasonable assessment of publication bias. Expert number 6 draws attention to deriving summary measure of association by "mixing of case-control and cohort studies". I do not see a major issue with this, as an odds ratio from well designed case-controls can provide a reasonable estimate of relative risk (from a cohort study) - and this 'mixing' is typically done in meta-analysis. Now, I believe the cohorts tend to provide a better estimate of the risk, and heterogeneity between cohorts and case-control studies could be explored if there were a sufficient number of studies (which I don't think there are) While Expert 1 draws attention to the fact they did not follow Prisma guidelines or discuss language, I believe they did a far superior job following the standard approaches in doing a systematic review and meta-analysis than the paper by Shinasi and Leon. Ultimately, they look at the same studies, but I think they made better decisions about what risk estimates to select from these individual studies than Schinasi and Leon did

SCORE
1

Expert 6

07/19/2021 07:22

Chang & Delzell meta-analysis presented also meta-estimates for B cell lymphoma and its major subtypes. As expert 2 pointed out the validity of a small number studies of B-cell lymphoma subtypes, with a small sample size, is questionable. Still, meta-estimates were significantly increased for B-cell lymphoma overall (two studies only), and multiple myeloma (six studies), and excesses were observed for CLL, follicular

lymphoma and hairy cell leukaemia (two studies each). The various models yielded NHL risks ranging 1.3 – 1.4 and MM risks ranging 1.4 – 1.5. Another merit of Chang & Delzell study is related to the discussion of dose-response trends and their note about the fact that "…All [the studies that evaluated dose-response trends] … relied wholly or in part … days of glyphosate use …; however, this metric has been shown to be a poor indicator of actual glyphosate received dose" [Chang & Delzell page 418, left column, subheading Exposure-response trends, NHL subtypes). The authors also give details about the exposure assessment conducted by De Roos et al 2005 in the AHS cohort: cumulative exposure was based solely on days of use, and intensity weighted cumulative days of use was what is generally defined as cumulative exposure (years*days/year*intensity). Numbers were small, but, apart from the negative finding on NHL, there was a suggestive upward trend in MM risk (which, by the way, disappeared in the Andreotti 2018 analysis of the AHS data, with an extended the period of follow-up and a follow-up interview 5 years after enrolment). I expressed concern about combining in meta-estimates results from case-control and cohort studies. expert 2 does no see any problem with that, but in my view the estimates of association in these two study designs differ, in the sense that case-control studies collect lifetime occupational histories and exposure information, while cohort studies usually rely on exposure at one point in time. I see it basically as combining results from different exposure estimates. It is true, however, that the AHS study did collect detailed information on the exposures prior to interview; in that sense, the occupational information is equivalent to that collected in the best quality case-control studies. Interestingly, Chang & Delzell discard the positive results raising methodological issues, such as study design, selection bias and exposure misclassification. As noted by expert 2, this study was funded by Monsanto, therefore some kind of conflict of interest in interpreting the results might have shifted the balance in interpreting uncertainty, by stressing weaknesses in positive studies and overlooking those of negative findings. Concerns for participation bias are also raised by expert 2 for the Cocco et al. study. This was part of a multicentre case-control study, conducted in six European countries: controls in Germany and Italy were a random sample of the respective study areas, while they were hospital controls in the remaining countries. Unless one raises such concern for case-control studies in general, it is difficult to believe that participation rate affected specifically glyphosate users in all participating countries. The main weakness of the reported finding is the small number of exposed, and the most likely explanation would be chance. This would be exactly the reason why meta-analyses are needed. Timing of exposure is an important issue. Chang & Delzell discuss it in relation to the De Roos 2003 (including the Brown study) results referred to cases diagnosed between 1979 and 1984, perhaps too close to the year glyphosate was first marketed. It is true that not many case-control studies considered

carefully this potential problem. Their report was very detailed, very well conducted, and 27 pages long. Very few high IF journals would have accepted it. So, I would not care that much about the IF of the journal that published it

SCORE
0

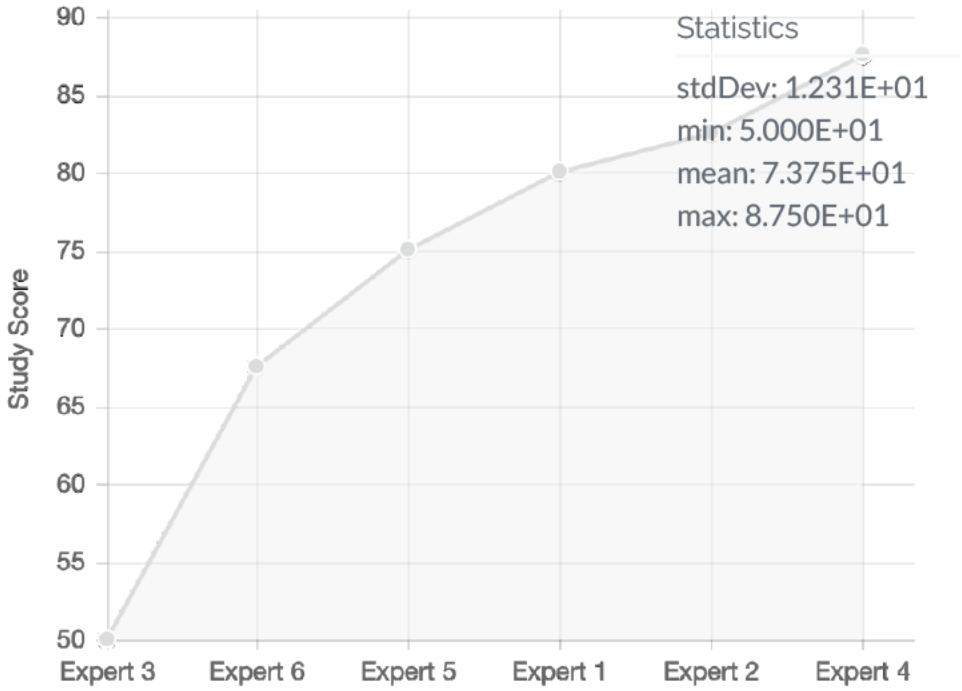
Expert 1

07/21/2021 21:49

I agree with my colleagues that this meta-analysis is an improvement on the previous one conducted by Schinasi & Leon 2014. This analysis used 7 studies with 257 cases and 377 controls. Again, the results are reported as relative risks, even though the majority of studies were reported using odds ratios, did the authors convert OR to RR, if so how, not mentioned in the paper. The point estimates used in this meta-analysis do not match those of the original studies. An improvement in the analysis by comparing different studies highlights the lack of variation between the studies and the consistency of the overall polled estimate.

Result 1.3 *(ID: 5174)*
Question 1.3 *(ID: 4144)*

On a scale of 1-10 (1=lowest, 10=highest) what is your degree of confidence in the following factors for study C: Leon et al. (2019). Please provide an explanation for EACH rating in the space below.



Answer Explanations

50.0 Expert 3

Factor	Factor Rating::1
Methods::25	7
Results::25	7
Discussion/Conclusions::25	3
Application to Risk Assessment Decision-Making::25	3

The methods were appropriate, and the results were presented adequately. One limitation was not presenting the results of the 3 studies individually. Note, this isn't really a meta-analysis, since the authors had access to the raw data for all three studies. For this project, the discussion was lacking. The paper evaluated 14 chemical groups and 33 active ingredients with 5 outcomes. This publication was not specific to glyphosate in its focus of analyses or discussion. In this context, random associations would be expected. The association of glyphosate and DLBCL in the AGRICOH analyses seems to be restricted to the CNAP cohort and not elevated in the French (AGRICAN) or US (AHS) cohorts. The authors provide no hypothesis or discussion. Is this a random observation or are the exposures higher in the CNAP cohort? The lack of an exposure-response within the AHS suggests the former (a random observation). Similarly, Leon et al is poorly suited for risk assessment. First, the exposure assessments are limited by self report of application and crop cultivation. The limitations of the crop exposure matrix used by the AGRICOH were highlighted in a letter to the editor (Tomenson

2017). In their review of glyphosate and cancer, Mink et al (2012) cautioned that the validity of the risk estimates depends on the validity of the exposure assessment(s) in agricultural epidemiologic studies. Second, the analyses are limited to ever/never with no additional analyses to approximate more intense, longer duration, or higher exposures.

75.0 Expert 5

Factor	Factor Rating::1
Methods::25	8
Results::25	8
Discussion/Conclusions::25	7
Application to Risk Assessment Decision-Making::25	7

Methods: the study population and demographic variables were different for each cohort. Through cohort-specific exclusions the resulting study population included in the combined study of meta-analysis was homogeneous. The harmonisation of the exposure data allowed for the increase of the sample size and therefore the ability to detect associations if present. Results: No significant association between risk of NHL overall and ever use of glyphosate. However, the heterogeneity was, relatively, high and perhaps due to some innate differences between the three cohorts. Discussion/Conclusions: An advantage of the study is the large number of exposed cases and the inclusion of prospective cohorts which minimised a differential recall bias. However, differences were observed in the exposure information used in the meta-analysis. Only one of the three cohorts used actual exposure information based on self-report. In addition, the three cohorts differed in other variables, such as age and gender while each of the cohorts adjusted for different factors, depending on the availability of the information in each cohort. Therefore, it is uncertain whether the correct adjustments were made to account for the observed differences among the cohorts. Application to Risk Assessment Decision-Making: Although the were innate differences among the cohorts, a relatively good harmonisation of the data was described. The authors mention that a future in-depth analysis will improve the specificity of the exposure information by adding parameters such as duration, frequency, and intensity of use.

80.0 Expert 1

Factor	Factor Rating::1
Methods::25	8
Results::25	8
Discussion/Conclusions::25	8
Application to Risk Assessment Decision-Making::25	8

This is a pooled analysis of three cohort studies (Agriculture and Cancer (AGRICAN), Cancer in the Norwegian Agricultural Population (CNAP), and the Agricultural Health Study (AHS), and is not a systematic review and meta-analysis of the published literature. This analysis assessed 14 pesticides and 33 individual active chemicals. Therefore, no electronic databases were searched. There was no quality assessment. There was no assessment of publication bias. Sub-group analysis was conducted. The analysis involved a Cox-proportional Hazards model. The sample size of the cohorts is large (n=316,270).

67.5

Expert 6

Factor	Factor Rating::1
Methods::25	7
Results::25	6
Discussion/Conclusions::25	7
Application to Risk Assessment Decision-Making::25	7

Limited to cohort study. Used country specific crop exposure matrices to assess exposure in 2/3 cohorts. The analysis covered also B cell lymphoma subtypes. Only ever exposure was considered in the metanalysis. In the AHS cohort study exposure level was based on preparation of the herbicide mix, method of spraying, machinery repair, and use of personal protective equipment. Other information was available, such as crop size, type of tractor (with or without close cabine (if any) and re-entry, but it did not contribute to the exposure level estimate. Andreotti evaluated trends by days of use only, and intensity weighted lifetime days of use (cumulative exposure). However, the way cumulative exposure was calculated underscores highest exposures.

82.5

Expert 2

Factor	Factor Rating::1
Methods::25	8
Results::25	8
Discussion/Conclusions::25	9
Application to Risk Assessment Decision-Making::25	8

This a pooled analyses of three agricultural cohorts, and is not a meta-analysis per se. The paper was published in the International Journal of Epidemiology which is a well regarded journal, high impact factor, and good peer-review process in place. The list of authors are highly regarded experts in this field, and the study was funded by the US NIH, and other government funding programs The cohort design of the study (3 separate cohorts: US, France, and Norway) is a strength of the study as well as the identification of 2430 incident cases among a total of 316270 farmers. The cohort study design is better capable of overcoming recall bias, and participation bias that impact measures of associations derived from case control studies. For the two European cohort, 'potential' exposure to glyphosate was determined using a crop exposure matrix (CEM), while for the US study exposure was determined through questionnaires. The estimation of exposure based on potential is not ideal, as this could introduce exposure misclassification, and as the authors point out this could be more of a concern for less commonly used pesticides. The authors do not describe the extent of the missingness of the data (AGRICAN: crop treatment, pesticide treatment, period of production, and period of pesticide treatment task; AHS: pesticides applied). The Heltsche paper, that is cited did look at this issue, and developed a multiple imputation technique, however this was based on a much smaller sample of the AHS cohort. That paper suggests ~37% of AHS participants did not complete a second questionnaire used to derive exposure estimates suggesting that missingness of data may be prevalent. The authors recognize the potential for spurious findings from multiple testing (given the large number of statistical comparisons – 14 chemical groups and 33 active ingredients, and 5 cancer outcomes). The paper did not observe an association between glyphosate and NHL, however noted there was substantial heterogeneity in the measures of association between the three cohorts (I2=57%).

Unfortunately, cohort specific risks estimates for glyphosate were not presented in the paper, not the supplementary materials. Glyphosate was positively associated with an increase risk of Diffuse large B cell lymphoma (HR=1.36, 95% CI:1.00-1.85). Across the three cohorts there was 497 cases of this type of cancer. While the paper does acknowledge that there are etiological differences in the types of NHL, no biologically plausible pathways were put forward to explain why an association would be expected to be stronger for diffuse large B cell lymphoma. The authors draw attention to the fact that some farmers in the reference group may have been exposed to other type of pesticides (non-glyphosate) that could be etiologically relevant. As a result there risk estimates could be understated, and I agree with this assessment. Overall the study found no association with NHL, but an association with diffuse subtype. No biological plausible explanation why a stronger association would be found with this subtype was provided, and overall, there are some potential concerns about multiple testing, and characterization of exposure (potential for exposure in European cohort, and impacts from imputing missing exposure data)

87.5

Expert 4

Factor	Factor Rating::1
Methods::25	9
Results::25	10
Discussion/Conclusions::25	10
Application to Risk Assessment Decision-Making::25	6

First, one must note the Leon et al. (2019) is not a meta-analysis, but a pooled analysis of cohorts. As such, it can be far superior to any meta-analysis because it accounts (or at least can account) for both within- and between-cohort confounding. It does not integrate evidence from other meta-analyses that contain case-control studies, although it is possible to inform pooled analysis of cohorts with results of case-control studies in Bayesian framework that elucidates informative prior from on RR the case-control studies. The analysis of pooled cohorts resulted in the largest currently possible sample size from cohort studies (2430 cases, 1131 of whom were exposed to glyphosate, over 3.5 million person-years). Analytical methods were state of the art and harmonization of exposures through crop-exposure-matrix is likely the best approach to the problem (despite claims to the contrary by Kabat et al., 2021). Missing data was handled using multiple imputations, an entirely appropriate method if is assumptions are met (uncertain). Effect estimates of individual cohorts were pooled used meta-regression, which may not have been as good as modeling random cohort effect and checking cohort-exposure interactions. However, either approach is defensible. Control for multiple comparisons may have been desirable given that exposures were most likely correlated but this is not a fatal flaw when data is interrupted with suitable caution. Hierarchical regression showcased by De Roos et al would have been a superior choice in tackling high dimensionality of the analysis with many exposures that can mutually confound each other. AHS did not dominate the pooled sample (Figure 1), which makes this analysis a useful addition over earlier meta-analyses dominated by AHS. Effect estimate of glyphosate was somewhat heterogenous among cohorts but it is hard to make much of this given how vulnerable such statistic appears to be to analytical choices and the fact that only three cohorts were compared. The effect estimate of pooled cohorts appears to be in line with AHS and lower estimates from meta-analyses of case-control studies; certainly the uncertainty that is captured by the 95%CI is smaller than true uncertainty and this the overlap with meta-RR from case-control studies may well be greater than it appears. The discussion is the typical comprehensive, yet

qualitative treatment of mostly quantitative matters. It is well-articulated but does not help in either causal arguments, or risk assessment. The authors did not undertake any analyses to account for possible latent confounding or misclassification of exposure, which limits their results for risk assessment, as most of the scientific literature in this domain. Given the cohort design, estimation of absolute risks and attributable fractions would have been of interest and certainty feasible. The authors report no evidence of association of having ever used glyphosate with risk of NHL and the effect estimate supporting this is far more precise than that seen from meta-analyses that predated to pooled cohort or excluded it. This precision is misleading, because it does not account for systematic errors, and yet some consistency among cohorts argues against strong biases that would qualitatively alter interpretation in evaluating whether results support existence of association between NHL and glyphosate. The matter of association of glyphosate with subtype of one NHL (diffuse large B-cell lymphoma) is vulnerable to to concerns of latent confounding, plus measurement error in exposure and outcome, which can easily jointly account for HR that is about estimated at 1.4; 95%CI 1-1.9, i.e., in the same range of weak effects as that reported in case-control studies (but not cohorts) of glyphosate and NHL, but with a robust sample size of 434 cases, with 221 exposed cases. Perhaps counter-intuitively, large sample size here would serve to exaggerate any uncontrolled sources of bias, because increase in sample size does nothing to diminish systematic errors, but in fact makes then dominate the observed estimates, as random errors begin to play an ever decreasing role. Consequently, without a through bias analysis, it is premature to interpret this apparent excess of risk.

Comments 3

SCORE
0

Expert 2 *07/16/2021 11:34*

I think the point made by Expert 1 is really important, namely, this is NOT a meta-analysis and systematic review. It should not be referred to as such. It would have been nice to have tried to partition participants in such a way to look at exposure response patterns (even a 3 level category: no exposure, intermediate and high). I think an important concern is that they did not have individual level estimates of exposure per se, they had to rely on a crop exposure metric. Looking at risk based on the potential for exposure is very different that assessing risks based on reported/documented exposure for individuals.

SCORE
2

Expert 6 *07/19/2021 07:20*

As most experts observed, this is not a meta-analysis. Leon used proportional hazard models to calculate pooled HR estimates for a number of cancer sites associated with ever exposure to 34 individual agrochemicals and 13 groups thereof. So, at the end, the weak association with risk of DLBCL, not supported by other studies, seems to pop up just as an outcome of multiple testing. Still, the effort of exploring the association with lymphoma subtypes is worth mentioning. The authors combined three cohort studies: 1) the updated US AHS data set, from which they excluded 3972 commercial applicators, i.e. professional pesticide sprayers. These were

farmers seldom cultivating more than one crop (average 1.1 crops), and 74% had corn crops, so that more than 80% cohort members had been ever exposed to glyphosate. The original studies do not mention crop size, but, based on the selective type of agriculture, it is conceivable that these were very large farms. Although using detailed questionnaire information, the algorithm to calculate intensity of exposure included intensity quartiles as multipliers of lifetime days of use, which might have diluted the condition of more severe exposure. In fact, such strategy relies on the hypothesis of a normal distribution of exposure level, while the right tail of such distribution extends orders of magnitude beyond the upper cut point of the quartile distribution. This might imply that the few highly exposed were diluted in a large group of much less exposed though above the upper quartile cut off (See De Vocht F, Burstyn I, Sanguanchaiyakrit N. J Expo Sci Environ Epidemiol 2015; 25(5):467-473). 2) the French AGRICAN study, included 71,000 farmers cultivating different crops, 36% of whom were considered as exposed to glyphosate based on a local crop-exposure matrix (PESTMAT or PESTEXPO), which resulted from a very large and detailed survey. The study was a mortality follow-up; and 3) the Norwegian CNAP study, by far the largest, including 38% ever exposed to glyphosate, also assessed by a crop exposure matrix, mainly from potato crops, grassland, grain crops and greenhouses. The lack of correlation between the exposure assessed by crop exposure matrices and those assessed by experts in the AHS study was previously discussed (Tomenson JA. Occup Environ Med 2017;74(1):81). Chang & Delzell highlighted other limitations of the AHS study, including missing data. So, while the cohort study design might overcome recall bias and participation bias, they are not exactly free from drawbacks. Leon et al. reported about harmonisation of the exposure data, but I am very much concerned about whether this was possible, particularly because of the heterogeneity between the three cohorts, as highlighted by experts 2 and 5. Although crop-exposure matrices included time as a factor in the assessment, is it not clear when exposure to glyphosate started in France and in Norway. I cannot but agree with expert 5 about the need that future analyses in retrospective studies would include multiple exposure metrics, such as duration, frequency, and intensity of use. At the end, as expert 3 stressed, only the dichotomous exposure variable was analysed, and there was no attempt of investigating exposure response trends.

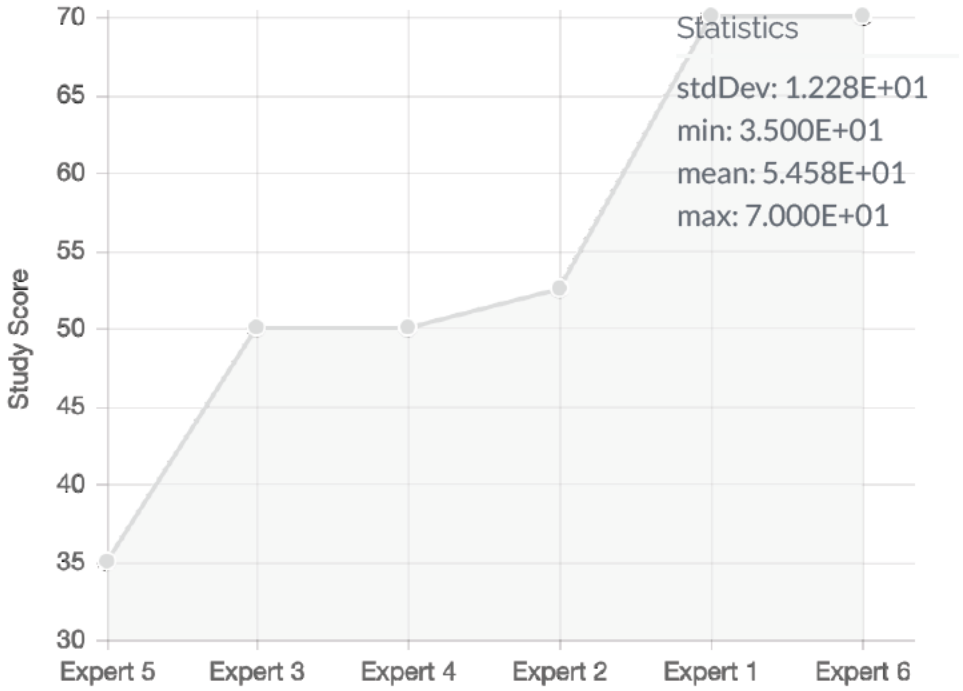
SCORE
0

Expert 1 *07/21/2021 22:09*

I feel that a sub-group analysis of the US AHS data, with the 3972 commercial applicators (professional pesticide sprayers) could have been of value. I agree with expert #6 that those who experienced high levels of exposure to glyphosate may have been diluted out in such a large sample.

Result 1.4 *(ID: 5175)*
Question 1.4 *(ID: 4145)*

On a scale of 1-10 (1=lowest, 10=highest) what is your degree of confidence in the following factors for study D: Zhang et al. (2020). Please provide an explanation for EACH rating in the space below.



Answer Explanations

50.0 — Expert 3

Factor	Factor Rating::1
Methods::25	5
Results::25	5
Discussion/Conclusions::25	7
Application to Risk Assessment Decision-Making::25	3

Methods: The stated purpose was to look at highest exposure estimates. Note that Shinasi suggested that his was "not possible to combine estimates based on multiple categories of exposure in formal meta-anlayses." This was the only study to use a quality scale. However, I didn't see any evidence that the learnings from this assessment were used or incorporated into their assessment(s). Results. The authors present a nice table in which they highlight the RR used in their MA and 3 other MA, showing how influential the selection can be. The authors provided a good summary table, that included weaknesses and notes for each study Discussion. There is extensive discussion on the 2005 and 2018 AHS publications. This is the only publication that summarizes evidence in animal data.

35.0 — Expert 5

Factor	Factor Rating::1
Methods::25	3

Results::25	4
Discussion/Conclusions::25	4
Application to Risk Assessment Decision-Making::25	3

Methods: The use and presentation of the findings based on the fixed effects models versus the random effects model (which are also presented and do not differ substantially from the fixed effects model in the article) is not correctly justified and is not given similar weight in the review. Random effects analysis is a standard method in meta-analyses which takes into account multiple true effects between studies and provides wider confidence intervals. Results: the results should be interpreted with caution based on the limitations of the methodology and the biological hypothesis on which the meta-analysis is based on. Discussion/Conclusions: the authors criticise the data imputation done by Andreotti et al.(2018) as a potential source of weaker RR when included in the meta-analysis. However, the sensitivity analysis performed by the Andreotti authors suggest that the results of the imputed data were similar to the what the results would be if the data were not imputed. Application to Risk Assessment Decision-Making: The authors hypothesis that the highest biologically relevant exposure to glyphosate will lead to increased risk of NHL is not substantiated in the article and therefore raises questions over the certainty of the results. In the article by Andreotti et al. 2018, which was the largest and highest quality study included, the findings showed no trend towards higher ORs with higher exposures.

70.0

Expert 1

Factor	Factor Rating::1
Methods::25	7
Results::25	7
Discussion/Conclusions::25	7
Application to Risk Assessment Decision-Making::25	7

This is a methodologically sound meta-analysis. It uses an 'a priori' hypothesis testing approach. It follows the PRISMA guidelines. It only uses a single database for the literature search (i.e., PubMed, which includes MEDLINE). The search is up to August 2018. There are no language restrictions. Additional sources were searched (i.e., IARC Report, U.S. EPA, WHO JMPR). It included 6 studies (1 cohort, 5 case-control). The total number of participants was almost 65,000. The authors used the Newcastle-Ottawa Scale (NOS) to undertake a quality assessment of the studies included in the meta-analysis. The authors used both a Fixed-effects and Random-effects statistical methodology in conducting their analysis. They assessed publication bias using Egger's regression analysis and Begg's test.

70.0

Expert 6

Factor	Factor Rating::1
Methods::25	7
Results::25	6
Discussion/Conclusions::25	8
Application to Risk Assessment Decision-Making::25	7

This study used an innovative approach by considering the risk estimates associated with the highest exposure level, to overcome the dilution due to inclusion of highly exposed subjects in a much larger number of exposed at very low levels or just occasionally. However, the authors still insisted using the NHL definition, did not consider time since exposure started, did not explore exposure metrics apart from ever exposure and the highest exposure level based on whatever was used in the original papers, and their meta estimates iclude case control and cohort studies. The discussion is articulated and detailed; details are given about differences in exposure information across studies and a comparison to the two previous metanalyses is included

52.5 Expert 2

Factor	Factor Rating::1
Methods::25	4
Results::25	6
Discussion/Conclusions::25	7
Application to Risk Assessment Decision-Making::25	4

This meta-analyses was based on 5 case control studies tat had been used in previous meta-analyses as well as the updated analyses from the AHS study. The derived summary measure of risk between a 'high cumulative exposure' to glyphosate was 1.41 (95% CI=1.13-1.75). When they used the previous version of the AHS the summary risk estimate was 1.45 (95% CI:1.11-1.91). The difference in the result seems counterintuitive because the authors had identified a summary risk estimate of 1.12 from the more updated AHS, and a risk estimate of 0.8 from the earlier AHS study. The fact that the risk estimate increased with a meta-analysis that substituted the earlier AHS risk estimates for the recent one is due to the fact that the weighting for the study decreased from 54.0% to 28.4%. I have substantial concerns about the methods used in this study, and it is my view that inappropriate meta-analysis methods were used. Namely, the authors attempted to derive summary measures of associations using very different measures of exposure across the studies. It my view, they have overstated their conclusions when they indicate there is 'a compelling link between exposure to glyphosate and NHL. The paper also reviewed additional animal and mechanistic studies that they indicated provided support to suggest a carcinogenic potential of GBH. The paper was published in Mutation Research which is a fairly good journal, and several of the authors served as science review board members of the US EPA committee that evaluated glyphosate. In my view, appropriate methods were used, and they had a well defined search strategy and used PRISMA guidelines. That said, I would wager the authors knew precisely what studies would be included beforehand given the few studies in this area, and given that they were familiar with this literature based on their participation on the EPA committee. This paper didn't really characterize associations with NHL subtypes. The authors acknowledged the mixed findings in previous studies between NHL and GBH from previous studies, yet somehow felt inclined to indicate that their meta-analyses provided compelling evidence of an association between GBH and NHL while drawing on these same studies in their meta-analyses. The authors point to meta-analysis being a powerful tool to synthesize findings from across studies to derive a summary risk measure. However, the ability to do so is dependent on the validity of the measures of association from each of the individual studies, and meta-analyses can not overcome poorly designed studies or ones with large inherent bias, and uncertainties that come with each of them. This meta-analysis differs from past ones as it focussed on an apriori hypothesis that targets biologically relevant exposure based on: higher intensity of exposure, long duration, and latency. The

authors argue that measures that incorporate these aspects are better able to identify those with very high exposures, to those of lower exposure. There is some merit to this thinking as the findings of past studies have been inconclusive, so there controversy whether glyphosate increases the risk of NHL in any capacity. Nonetheless, what this also means is that one is summarizing risk estimates from vastly different exposures measures into one summary, and I would argue, uninterpretable, risk estimate. Specifically, it becomes impossible to clearly articulate what change in exposure occurs to produce the summary risk measure of 1.41 (other than it is some mixtures of exposure metrics that incorporate latency, duration, and cumulative measures). So, these values have no real meaning, whereas, the meaning of past relative risk measures (based ever vs never exposures) were more readily interpreted. The studies included in the meta-analysis were De Roose (2003), Orsi (2009), Eriksson (2008), Andreottie (2018), Hardell (2002), and McDuffie. The authors reviewed the quality of the studies using accepted ratings scale (Newcastle) and concluded the AHS was the highest quality study. Now what is interesting about this is that the AHS provides no evidence of an association between glyphosate and NHL – yet still, the authors repeatedly indicate that their meta-analyses provides compelling support for an association between GBL and NHL The authors indicate the process they used to select risk estimate from the study and prioritize using: 1) highest cumulative exposure and longest lags; 2) highest cumulative exposure 3) longest exposure duration and longest lag; 4) longest exposure duration; 5) longest lag or latency, and then 6) ever exposure. While the authors do discuss the issue of exposure measurement error, I am surprised they do not discuss the issues that more distal exposures (historical measures) for several reasons tend to be measured with greater error. Participants may not be able to as accurately recall exposures 20 years ago; data documentation for more distant exposures may not be as complete. So while from an etiological perspective there may be relevance to studying more historical exposures (often because they are higher), the authors are on one hand are critical of exposure characterization used by others, but not taking the same critical view of their own approach. The authors do mention trying to assess publication bias, but as mentioned earlier, this is not possible due to the small numbers of studies – the authors recognize this limitation. The measures of association that the authors drew upon (Table 4) warrant some comment. For the AHS, they used a risk estimate of 1.12 drawn from a 20 year latency in the upper quartile of intensity. It is worth noting that there is no evidence of a dose response pattern across these quartiles (20y latency): specifically the RR's across quartiles are: 1.0, 1.22, 1.15, 0.98 , and 1.12. Perplexingly, given that the meta-analyses weighted this one study very heavily (54%) and this dose not provide support to an exposure-response relationship (as per Bradford Hill criteria) the authors did not draw attention to this. In addition, the AHS study was by and large a null findings. The relative risks across quartiles (0 lagged exposure) for an intensity weighted lifetime analyses were 1.0, 0.83, 0.83, 0.88, 0.87 (Table 2 of Andreotti paper), and for a 5 year lagged exposure the relative risk estimates were 1.0, 0.92, 0.79, 1.03, and 0.87 (Table 3). For the meta-analysis that included the earlier AHS study the authors used the time intensity risk estimate 0.8 (based on a 0 lag). The summary weight for the study (with respect to deriving the summary measure of risk) was reduced from 54% (for the risk estimate of 1.12) down to 28% for the risk estimate of 0.8. The risk estimates provided in Table 4 are vastly different from each other. The exposures include: upper quartile of intensity (relative to lowest) for an intensity weighted lifetime measure, ever vs never exposure, 2 days per year of exposure relative to none, >10 days of lifetime exposure relative to none. I fundamentally disagree to combine these vastly different exposure metrics into a single summary measure, and a student who were to take this approach in my class in meta-analysis would receive a poor grade. The summary measure is uninterpretable in my view – for example, what is the change of exposure that increases risk by 41% - it is some mixture of

lag of exposure, intensity, ever vs never that is based on a series of studies assigned different weights in the derivation of this measure. It is like comparing apples to oranges, and in my view, it becomes impractical to assess heterogeneity in risk estimates when the exposures vary so widely. As described above, I have concerns about the application and the interpretation of the AHS findings. Again, the AHS gives no compelling evidence of an exposure response relationship, most findings are null, and in fact the comparison of the upper quartile of time intensity measures shows a 13% reduction in risk of NHL when compared to the lowest quartiles. Instead the authors model the elevated risk of 12%. The authors themselves note that the AHS is of the highest quality study. They also raise the concern about imputation methods for the AHS and how only 63% of participants completed the 2nd questionnaire. I do not grasp how these authors can say the AHS is the highest quality study, there are potentially severe issues with exposure measurement (which they also published a letter on), while noting that several of the of the risk estimates are null from this study – then they go and conclude that the analyses in their paper (Zhang et al) provide compelling evidence that glyphosate increases the risk of NHL. For the one measure of increased risk that was picked from the 20 year lag, there is no evidence of dose response. At the very least, there is considerable uncertainty here – I see nothing compelling in any of this. Furthermore, particularly when they acknowledge that 'the underling mechanism remains unknown." I am surprised this paper got published given the number of contradictions within it, and being frank, I will use this paper as a teaching example in my course for methods to avoid following. Where were the peer reviewers? In the authors description of biases from the different studies, there is no real mention of participation bias, which I view to be a major concern with some studies. The authors conclude "All meta-analyses conducted to date including our own report the same key findings; glyphosate is associated with an increased risk of NHL". As written, this is a misleading statement because it gives the false impression that there were independent meta-analyses performed. The reality is, all meta-analyses used the same 5 or 6 studies – so one would expect concordance between the published meta-analyses." To me, the more salient issue is the validity of the risk estimates of each of these individual studies and the authors of this meta-analyses identify a series of potential biases, and concerns that would preclude, in my view, compelling evidence.

50.0

Expert 4

Factor	Factor Rating::1
Methods::25	4
Results::25	8
Discussion/Conclusions::25	4
Application to Risk Assessment Decision-Making::25	4

The authors claim that their meta-analysis is different form the two pervious analyses of essentially the same studies because they employ a novel a priori hypothesis. However, the hypothesis boils down to postulating higher risk with higher exposures, which nobody questions, and is still limited to exposure metrics of exposure intensity and duration that cannot be pooled across the six studies under consideration. The key problem with approach adopted by the authors is that "the highest exposure groups" from each study were defined different units and therefore must not be pooled in meta-analysis. The authors make a dubious claim that bias from measurement error and confounding will be smaller when high exposure groups are compared to unexposed, rather then low/medium groups. This is a faulty oversimplification because (a) selective pressures and confounders a more

likely to be similar in groups that do not have extreme differences in exposure (by definition of confounding and selection bias) and (b) differential-misclassification-due-to-dichotomization of exposure (an established phenomena in statistics but little understood by epidemiologists) can be worse when exposure groups are far apart. Using the updated findings from AHS is a major strength of this work compared to other two papers that analyzed essentially the same material. The application of the Newcastle Ottawa Scale is helpful but is not inherently superior to any other analysis of quality of observational studies; the authors do not formally examine the impact of the scale score on meta-RR (by sequential exclusion of studies with lower NOS score). Focus on fixed-effects meta-analysis seems dubious given suspected heterogeneity but then again, such decisions have made no difference to pervious analyses of the same six studies. It is regrettable that I2 is not reported but it is laudable that extensive analysis of publication bias was undertaken. The authors are wrong to conclude that meta-RR of 1.41 (95%CI 1.13-1.75) indicates significant increase in risk where other meta-analyses of the same studies did not (they overly focus on point estimate and ignore total uncertainty, point estimate that is likewise the same as reported by others: 1.3 to 1.5) . The second decimal place in these meta-RR carries no information -- it is statistical noise as is shown by the fact that random choices in meta-analysis on affect the second decimal place in estimated meta-RR's. It is noteworthy that they overall risk that is estimated is the same as in two other studies that examined the same material, despite inclusion of updated AHS data and focus on extreme exposures (extreme effect estimates): the 95%CI of RR is still somewhere between 1 and 2, and the true uncertainty of the analysis (due to systematic bias) is would certainly produce a wider "uncertainty interval". These authors found no evidence of publication bias but this seems to be sensitive to both the choice of summary effect estimates and analytical procedures (as well as subjective judgement), because Donato et al. (2020) reached the opposite conclusion, while Chang & Delzell (2016) and Shinasi & Leon (2014) likewise did not report noticing evidence of publication bias. It is rather hard to make strong conclusions on this matter with only six studies. Sensitivity analyses were extensive and well-motivated, showing that meta-RR was inflated by some factors related to case-control studies; authors argue that Orsi et al. was the main/sole culprit, a useful observation that was not present in other reports. AHS seems to continue to shows results consistent with not effect or effect too weak to measure (and this evidence grows with every update of the AHS cohort and its pooling with other cohorts, in Leon et al (2019)), but case-control analyses pull the meta-estimated away from the null. Tables 5 and 6 implies that there is heterogeneity, as can be suspected when pooling incompatible exposure groups, making it difficult to put much faith into meta-RR given by the authors being appropriate summary of the magnitude of effect estimates (and yet, the effect estimates are the same as from all meta-analyses on these studies: 95% CI between 1 and 2, with true uncertainty certainly greater. The authors provide typical (if superficial) description of threats to validity epidemiological inference, but do not undertake any calculations to account for these, e.g., through probabilistic bias analysis, which is rather trivial to undertake thanks to work of Lash et al., and would certainly we well motivated by the policy implications and large studies with weak associations being pooled (where systematic errors may dominate to manufacture apparent associations where none exist in nature). Overall, it is hard to understand why the authors draw different conclusions in support of causal association when their estimates are essentially the same as that of other reviews of the same materials and their methods far more dubious. The authors over-interpret evidence they analyzed and presented in their final concluding statements: they did not conduct critical analysis of all evidence relevant to classification of glyphosate's role in NHL. The paper does not advance risk analysis and decision making because its is inferior methodologically to what was published before by Schinasi & Leon and Chang & Delzell.

Selection effect estimates for meta-analysis primarily because they are "large" guarantees biased summary effect estimate that is overly precise.

Comments 7

SCORE 2

Expert 3 *07/13/2021 08:51*

I applaud reviewers 2 and 4 for their excellent and comprehensive comments on this MA - and the others. Both highlight the nuance and complexity of summarizing dissimilar datasets.

SCORE 0

Expert 4 *07/15/2021 12:49*

Thank you #3.

SCORE 0

Expert 2 *07/16/2021 11:39*

Yes, thanks expert #3. Having read the other reviews, I still have the most concerns about the methods used in this meta-analysis compared to others. Good point by reviewer #3 about noting the review of animal research - and this helps assess the validity of the summary measure of risk. However, I think the methodology used was flawed to begin with.

SCORE 0

Expert 5 *07/19/2021 01:25*

As already said, very good job by reviewers 2 and 4! My concern is also about the methods used and the combination of dissimilar exposure metrics to reach an over-estimated summary risk estimate. As already pointed out elsewhere, and by expert 3 above, combining datasets with innate differences in population and data collection methodology is challenging.

SCORE 0

Expert 6 *07/19/2021 07:20*

Metanalysis was meant for combining results of randomized trials, not for observational studies. As the glyphosate story shows, these differ non only in the way of assessing exposure, the exposure metrics used, the combination of such exposure metrics in summary indicators (such as the intensity weighted lifetime days of exposure), but also in the definition of the disease outcome. Using the ever exposure might be statistically more appropriate, but it is biologically misleading. Assembling the top level of heterogenous exposure metrics in a summary measure is statistically inappropriate but stresses the point that indicators of high exposure level, whatever they are, give a different picture. In spite of the methodological concerns, while I would not trust

completely its results, I would not simply dismiss the Zhang et al study. They might have overstated their findings (there is nothing compelling out there); still, they called for attention towards the problems in the ever/never categorization of exposure. Their statement that higher intensity of exposure, long duration, and latency are better able to detect effects associated with high exposures, appears something to agree with. Statisticians need well definite variables to rely on: race, education, occupational title, ever exposure, and many other of sociological relevance. Biology does not work this way. What is relevant is the concentration of the xenobiotic at the target organ, following intake and phase I metabolism. Two are the differences between the Zhang et al MA vs the Chang & Delzell meta-analyses: the choice of the measures of association to be used, and the replacing of De Roos et al 2005 analysis of AHS data with the Andreotti et al. analysis. Expert 2 prefers ever/never exposures and is critical against the ther exposure metrics, as they are originated by collection of historical data in case control studies: "Participants may not be able to as accurately recall exposures 20 years ago". But 1) the point is whether there is any differential recall between cases and controls. This might be true for studies on electromagnetic fields, but it does not seem plausible to have occurred in relation to a specific herbicide when data were collected, in times when there was not public concern about its health effects. 2) exposure assessment by experts (industrial hygienists or - in this case – agronomists) prevents recall bias. The Epilymph study (Cocco et al.) relied on local agronomists for identifying the specific chemical agent used in the area, for that crop, that specific phytopathology, at that time (which also explains why the exposed were so few). 3) crop exposure matrices, as job exposure matrices, are useful as a screening tool replacing self-reported exposures. CEMs (and JEMs) have high sensitivity and low specificity, and therefore convey a great deal of exposure misclassification. However, they might allow using different exposure metrics (probability, intensity and frequency of exposure, for instance) which might allow grading misclassification and observing trends. Reliance on the glyphosate results from the AHS study is also questionable. Despite the quality in the planning and collecting data, I do have concerns about how AHS data were assembled in the weighted cumulative lifetime days of exposure, which by the ways I would have simply defined as cumulative exposure. Also, the way of calculating it cumulative exposure in the AHS study considers short time peak exposure levels as similar to long-lasting low-level exposures. As Mitridates was very well aware of, low doses would induce the metabolic enzymes, and reduce or nullify the toxicological properties of poisons. We don't know what exactly Mitridates poisons were, and perhaps pathways to chronic effects, such as cancer, might differ from those leading to acute effects, but we can't forget about toxicological and metabolic implications. Besides, Andreotti et al., excluded commercial applicators. These are those who professionally apply agrochemicals. They might be very

careful in wearing PPE during mixing and spraying, but they do that as their job, and not 2 or 10 days/year. A follow up of the 3792 commercial applicators, 2795 of whom used glyphosate 39 days/year or more, is still to come, and it might be more informative than any existing original study or metanalysis. Finally, my point is that there are still unexplored toxicological implications in the AHS study, that need to be addressed before drawing any conclusive statement on glyphosate.

SCORE
0

Expert 4 *07/19/2021 07:44*

Thank you #6 for informative analysis. But are I do not agree that JEM tend to have "high sensitivity and low specificity". In my experience evaluating these JEM, the reverse is true: near-perfect specificity and low sensitivity. This is by design as for rare exposures this tends to reduce bias from misclassification of exposure. It would be interesting to know if there are evaluations of validity or even reliability of CEM to help illuminate what is going in the glyphosate story (off the top of my head, I am not aware of such assessments). I must confess that I find hope that future analysis of AHS will add clarity overly optimistic. Giving tendency to publish the most extreme results first and epidemiology's decent record of picking up true associations using flawed tools, my sense is that waiting for more definitive analysis before deciding what to do about glyphosate is not defensible. Even though flawed, the available evidence is informative, and remarkable consistent.

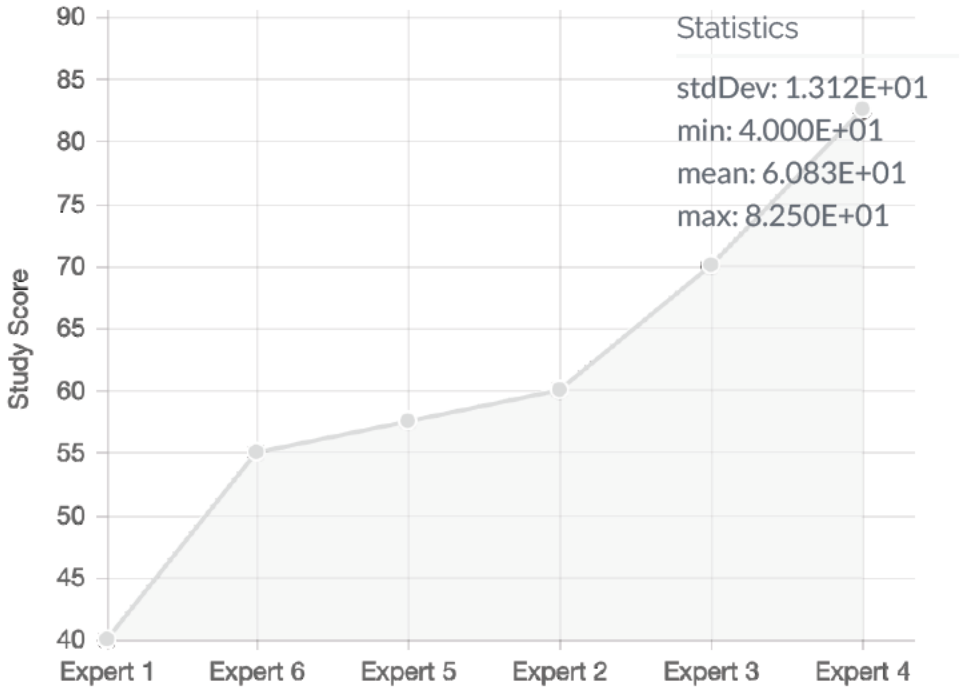
SCORE
0

Expert 1 *07/21/2021 22:25*

I thank experts #2-6 for their informative comments. The animal data is a separate issue but I would expect something to be discussed in a journal such as Mutation Research. Combining the different AHS studies doesn't appear to make difference in terms of the overall point estimate. Again, mixing of different rates (e.g., OR vs. RR) is not a methodologically sound approach. This meta-analysis contains the smallest number of cases (n=173) compared to all the other meta-analyses.

Result 1.5 *(ID: 5176)*
Question 1.5 *(ID: 4146)*

On a scale of 1-10 (1=lowest, 10=highest) what is your degree of confidence in the following factors for study E: Donato et al. (2020). Please provide an explanation for EACH rating in the space below.



Answer Explanations

40.0 — Expert 1

Factor	Factor Rating::1
Methods::25	4
Results::25	4
Discussion/Conclusions::25	4
Application to Risk Assessment Decision-Making::25	4

This is the most recent of the systematic reviews and meta-analyses provided as part of this assessment. It follows the PRISMA guidelines. It uses PubMed, Scopus, and EMBASE database for the literature search. The search is up to May 2019. The languages were restricted to English, Spanish, German, French, and Italian. Two authors independently assessed which studies met the inclusion criteria. It included 7 studies (2 cohort, 5 case-control). The total number of cases was 1,271. There was no quality assessment. The authors used a Random-effects model for analysis. They assessed publication bias using Egger's regression analysis. The conducted sub-group analysis.

57.5 — Expert 5

Factor	Factor Rating::1
Methods::25	6
Results::25	6

Discussion/Conclusions::25	6
Application to Risk Assessment Decision-Making::25	5

Methods: the study gives a good presentation and justification of the statistical analysis conducted. Random-effects models, heterogeneity testing and sensitivity analysis was performed. Results: The authors present fully adjusted models, and showed low heterogeneity between and within studies. However, larger weight in the meta-analysis results was attributed to the AGRICOH consortium, where exposure assessment was differential between the cohorts. Discussion/Conclusions: It is not discussed in the article how the exposure categories were used in the analyses. The authors also suggest that the evidence of publication bias favouring positive studies provide evidence against the association between exposure and NHL. However, most studies included had no positive effect estimates (except De Roos, 2003) and therefore the authors' argument is not valid. If, hypothetically, more studies with negative results were included, it is more likely that the effect estimate would move even more towards the null effect. Application to Risk Assessment Decision-Making: caution is required when applying these results to decision-making as there are issues related to the analysis of the exposure categories and the use of the AGRICOH cohort in which crop-exposure matrices were used which do not provide specific pesticide exposure and no definite conclusion about glyphosate risk on NHL can be discerned.

55.0 Expert 6

Factor	Factor Rating::1
Methods::25	8
Results::25	5
Discussion/Conclusions::25	4
Application to Risk Assessment Decision-Making::25	5

Only ever exposure was considered. NHL and MM were evaluated, with no attempt of exploring subtypes. Mixes case-control and cohort studies. Positive findings from case-control studies are discarded claiming inherent bias, suggesting preconceived negative attitudes, while accepting the introduction of non differential misclassification of exposure through the use of crop exposure matrices in the Norwegian and French cohorts.

70.0 Expert 3

Factor	Factor Rating::1
Methods::25	9
Results::25	7
Discussion/Conclusions::25	9
Application to Risk Assessment Decision-Making::25	3

Methods: Set the standard for having a systematic review protocol and evaluating different components of the MA. However, they lost a point for doing a meta-analysis for highest exposure. Results: reported basic design of each study. Instructive figure (3) on MA by year of publication Discussion: Well written discussing strengths and limitation of the underlying data. Mentioned WHO 2001 classification.

60.0

Expert 2

Factor	Factor Rating::1
Methods::25	7
Results::25	6
Discussion/Conclusions::25	5
Application to Risk Assessment Decision-Making::25	6

This was more recent meta-analysis published in the journal Medicina del Lavoro 2020. This a a lower tier journal with an impact factor of around 1. The authors consisted of researchers at University of Turin, Mount Sinai (NYC), and University of Bologna. One of authors had previously acted as a consultant to glyphosate produced, the others did not declare any conflicts. The background drew attentions to the reviews by the European Food Safety Agency, European Chemical Agency, and the WHO/FAO also of whom concluded it is unlikely that glyphosate is a human carcinogen. The authors followed widely accepted guidelines (PRISMA) when conducting the review. Unlikely previous reviews, they considered papers written in Spanish, German, French or Italian. Two reviewers independently reviewed titles and abstract for relevance and to assess inclusion criteria. The methods used to reduce the number of papers is well described and the methods seem appropriate. The included studies were McDuffie, Hardell, De Roos, Erikssson, Orsi, and Coco. They used the findings by Leon et al which were pooled analyses of three cohorts in Norway, France and the US (AHS). Of note, the risk estimate used for the Coco study was for B cell lymphoma a risk estimate based on only 4 cases and two controls – with a participation rate of approximately 50% on the controls series. B cell lymphomas account for about 90% of NHL. While only a 5% weight was assigned to this study, I have concerns about the validity of this estimate. The majority of the weight used in the meta-analysis came from the pooled analyses of 3 cohorts from the Leon paper (74%). It is unfortunate that study specific estimates for each of these three cohorts was not used, as this would have strengthened efforts to look at heterogeneity, and by using the summary result (which is near the null value) the impact is to give this near null results a far greater weight. This is really important given Leon et al reported that the effects across the 3 cohorts exhibited substantial heterogeneity (I2=57% - page 1530 of their paper). I think the preferable approach would have been to assign a weight to each of the three separate cohorts (similar to how a weight was assigned for each of the case control studies). The failure to do so washes out possible heterogeneity, and moreover, would bias findings towards the null because a weight of 74% was applied to the Leon study in the meta-analysis and the relative risk estimated used was 0.95. The use of this study in this fashion was the main reason why a summary measure so close to the null was obtained 1.03 (95% CI=0.86-1.21). The second most highly weighted study was that by McDuffie et al (~15%). Of note for this study was poor response rates which the original study never formally presented but rather just indicated that a limitation was 'less than optimal response rates'. In my view, high participation rates are important to derive valid measures of association from a case-control study – and another concern is the potential for recall bias from questionnaire based data. So, while this study does not find an association between glyphosate and NHL, I think it is important to acknowledge that there remains uncertainty, and in my view, some of the steps the authors took around the Leon study ultimately served to underestimate some of the risks. I appreciate the authors efforts to explore publication bias, and there is some asymmetry in the funnel plot. That said, the numbers of studies really limit the ability to assess publication bias. Moreover, the authors themselves by providing a single summary measure for 3 cohorts (Leon study)

may have unintentionally contributed to some of this asymmetry. Had they included 3 separate points for these 3 studies, as they included a point for each of the case-control studies may have lost some of the asymmetry. Moreover, there are also other ways to explore publication bias that could have been explored including by doing a cumulative meta-analysis (by study design) – with the thinking that larger studies are less susceptible to publication bias. I do not share the authors view that the studies 'provide evidence that the risk of NHL is not increased in workers exposed to glyphosate'. There are too many other uncertainties to make such a firm conclusion. Moreover, it turns the intended purpose of hypothesis testing on its head. We never set out to prove the null hypothesis, rather we test whether the data provide us enough evidence to reject the null hypothesis. At this point, they do not. This is a subtle but important point.

82.5

Expert 4

Factor	Factor Rating::1
Methods::25	8
Results::25	10
Discussion/Conclusions::25	9
Application to Risk Assessment Decision-Making::25	6

One obvious advantage of the approach of this article is that it did not limit itself to English-language literature, but it does not seem to have yielded any practical advantages in uncovering papers missed by others. The authors extracted information from all relevant publications. taking care to avoid replication. Random-effects model was adopted and is a reasonable default choice, given prior evidence of the possibility of heterogeneity of effect estimates. However, a comparison to fixed-effects model would help reassure the reader that this analytical choice was not consequential. Sensitivity analysis by excluding one study at a time and cumulative meta-regression are among distinct methodological strength, compared to other recent meta-analyses on this topic. The summary meta-RR by ever-exposed status is lower than in other comparable meta-analyses, with little evidence of between-study heterogeneity. This is likely due to dominance of Leon et al. (89% of cases/deaths). This dominance of one study diminishes the value of current meta-analysis and yet is still a fair summary of effect estimate, even if it is not very interesting for understanding source of disagreement among studies. When Leon et al. was excluded, the meta-RR was in line with all similar analyses: 95%CI of RR from 1 to 2. The authors report evidence of publication bias but fail to try to correct it through imputations, as Chang and Delzell did. But the bias they report would lead to reduction of meta-RR, since small negative studies appeared under-reported. Analysis by levels of exposure suffers from lack of harmonization of exposure categories. As such, it adds nothing compared to what was reported in original publications. The discussion is typical of epidemiology papers: comprehensive and yet only qualitative in nature. Good evidence is presented that confounding is important in these studies and their synthesis (this implies the possibility of residual confounding that would further weaken the observed effect estimates). Issues of outcome misclassification are clearly presented and their implications explained; these could have been address in bias analysis. The conclusions appear reasonable. The analysis has limited value to risk assessment because it did not entail quantitative bias and uncertainty analysis, e.g. it did not use cited participation rates and reliability/validity estimates of exposure assessment to evaluate what the resulting bias and cost to certainty are. (Such analyses are rare, and yet are sorely missed if epidemiology is to support risk assessment). The reported effect estimates for glyphosate are precise, but the true uncertainty is certainly greater than

that depicted, as such, any qualitative decision of whether association of glyphosate and NHL is causal has weak support (certainly far less than 95% certainty). One major conclusion suggested by the work is that if we were to use this meta-analysis to increase our certainty about making correct decision about risk of NHL due to glyphosate (that of little evidence of measurable risk), we would be doing so largely based on addition of Leon et al to the body of knowledge.

Comments 4

SCORE
-1

Expert 2 *07/16/2021 11:46*

I have concerns about the study weighting the Leon study so heavily (also noted by expert 5), and the potential links to industry. The risk estimate is much lower than other meta-analyses. Agree with most of comments of Expert 3 in that they followed a good approach for doing the systematic review, although, it would have been helpful if they assessed the quality of studies using OHAT/Newcastle or other tool. I also like the insightful comment about publication bias provided by Expert 5.

SCORE
-1

Expert 6 *07/19/2021 07:16*

The senior author of this meta-analysis, after leaving IARC, has been consulting for corporations, including Monsanto. Another co-author has been acting as a consultant for industries in legal trials, but, as far as I know, never in glyphosate trials. On the other hand, both are senior, expert epidemiologists, the methodology in conducting the study is impeccable, and, differently from Chang & Delzell's metanalysis, this study was not supported by glyphosate manufacturers. Still, the interpretation of their results as providing "evidence that the risk of NHL is not increased in workers exposed to glyphosate" indicates more a kind of "negationist" attitude than a real conflict of interest. Rana et al. (Environ Res. 2020;191:110140), a co-author in the Zhang meta-analysis, criticized Donato et al. study for using the ever/never exposure metric and for not discussing subtypes results. In response, Boffetta et al. (Med Lav. 2021;112(3):194-199) reshaped the previous metanalysis to include the Pahwa pooled analysis of case-control studies (Pahwa M et al. Scand J Work Environ Health 2019;45(6):600–609), discussed and discarded the positive finding on DLBCL (but not other subtypes), and confirmed the same conclusions. Expert 3 noted that Donato et al. mentioned the 2001 WHO classification of lymphoma, but only for raising the point that refraining from using it by evaluating NHL was considered as a limitation, which is correct in my opinion. Despite that statement, however, Donato et al. considered only NHL and MM in their meta-analysis. In updating it, Boffetta et al added DLBCL, but not CLL nor follicular lymphoma. Expert 2 also observed that using the pooled analysis by Leon et al., instead of the individual follow up studies, reduced heterogeneity across results, which exists, and it is one of the reasons for uncertainty in the conclusions and for calling for further

studies on glyphosate, instead of simply denying value to the positive findings. Besides, as previously observed, the AGRICAN and the CNAP studies did use CEMs, and Leon shouldn't have combined these results with those from the AHS study, which used a different exposure assessment approach. The authors also claim evidence of publication bias favouring positive studies, which would provide evidence against the association between exposure and NHL. As pointed out earlier by the other experts, the original studies on glyphosate are too few to assess publication bias, and, the Orsi et al case-control and the AHS study/AGRIGOH pooled analysis are negative for an association. Such choices and comments result in unbalanced conclusions.

SCORE
1

Expert 4

07/19/2021 07:47

I do no find it helpful focusing on who the authors and their sponsors are, as opposed to the work they did. There are valid concerns in all manuscripts we reviewed. Focusing on some of the more flamboyant personalities in the field does not help in objective evaluation of what they did and did not contribute.

SCORE
0

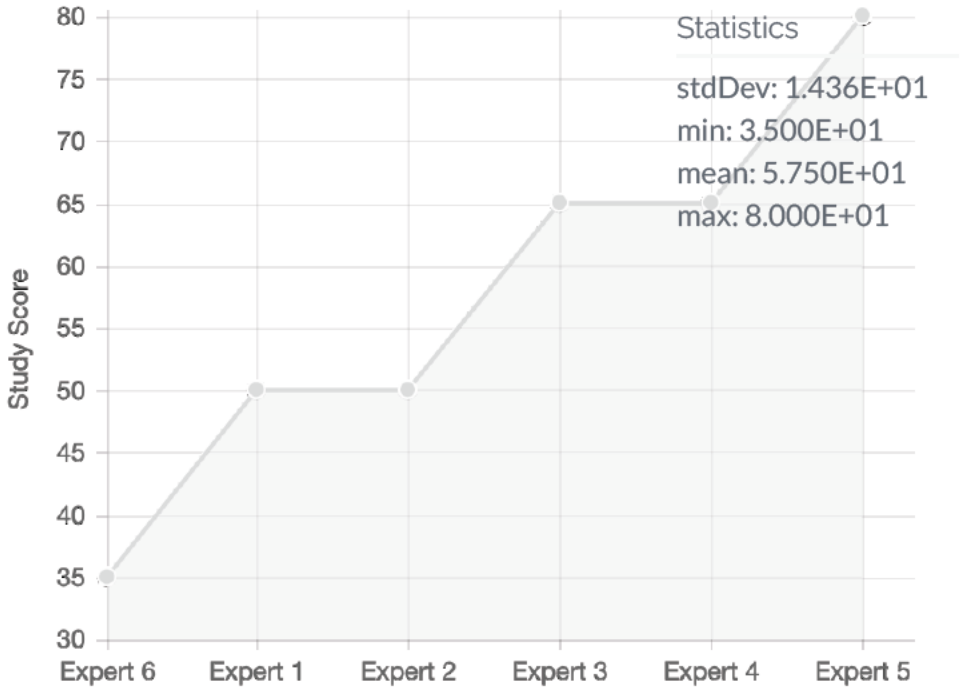
Expert 1

07/21/2021 22:35

I concur with expert #6 regarding the potential conflicts of some authors, not to mention the selection of this 'never heard of journal' as a place to publish the largest sample size meta-analysis on this topic, doesn't that seem strange? Again mixing of different rates and study designs not appropriate.

Result 1.6 *(ID: 5177)*
Question 1.6 *(ID: 4147)*

On a scale of 1-10 (1=lowest, 10=highest) what is your degree of confidence in the following factors for study F: Kabat et al. (2021). Please provide an explanation for EACH rating in the space below.



Answer Explanations

50.0 — Expert 1

	Factor Rating::1
Methods::25	5
Results::25	5
Discussion/Conclusions::25	5
Application to Risk Assessment Decision-Making::25	5

This is not an original meta-analysis of the studies assessing the relationship between glyphosate and non-Hodgkin's lymphoma. This is a re-analysis of the Zhang et el. 2020 meta-analysis for the purpose of conducting sensitivity (sub-group) analysis to further assess the definition of exposure and the latency periods selected. This study states that it re-examined Zhang's study which contained seven studies, however, this re-analysis only included five studies.

35.0 — Expert 6

	Factor Rating::1
Methods::25	4
Results::25	4
Discussion/Conclusions::25	4
Application to Risk Assessment Decision-Making::25	2

Repeats the metanalysis by Zhang et al. Excludes two cohort studies because based con crop-exposure matrices. Mixes the AHS study with the case-control studies. Considers ever exposure as the only exposure metric. Expresses he opinion the case-control studies are biased by default, and hypes the AHS results, not considering its limitations: Andreotti only used days of exposure; exposure data at follow up were missing for 37% cohort members, it was a specialized group of farm workers. AHS applicators had been trained and licensed and they might have been more careful and more frequently users of PPE.

80.0 Expert 5

	Factor Rating::1
Methods::25	8
Results::25	8
Discussion/Conclusions::25	8
Application to Risk Assessment Decision-Making::25	8

Methods: The authors demonstrate clearly the purpose of the study using justifiable criteria for study inclusion and effect estimates. Sensitivity analysis and testing of different scenarios chasing different effect estimates reported by each study were tested. Results: Based on the assumptions of the Kabat et al authors the results are well presented and substantiated. Discussion/Conclusions: Good arguments and substantiated by the results evidence. Application to Risk Assessment Decision-Making: the assumptions used to test the hypothesis, the analysis and the discussion of the results seem appropriate for consideration in a decision-making procedure

65.0 Expert 3

	Factor Rating::1
Methods::25	7
Results::25	7
Discussion/Conclusions::25	9
Application to Risk Assessment Decision-Making::25	3

Methods – they focused more on compare and contrast from Kabat than other MA, such as Chang. I completely agree with their view that ‘the strength of evidence for an association between glyphosate and NHL risk is dependent upon the selection of estimates from various studies” Omitted CNAP and AgriCan because they used crop-exposure matrices. Discussion on evidence of selection bias, exposure misclassification and lack of clarity on duration of latency from exposure to disease. Recommendation that MA not combine case-controls studies and cohort studies.. So true: “meta-analysis cannot improve on the quality of the underlying studies.”

50.0 Expert 2

	Factor Rating::1
Methods::25	4
Results::25	5
Discussion/Conclusions::25	6

Application to Risk Assessment Decision-Making::25	5

The authors had no declared conflicts of interest, and the paper was published in Cancer Cause and Control, which is a fairly good journal (impact factor ~2.3). The senior author (Tarone play a key role in designing the AHS). The authors, in my view fairly, criticized the findings from the Zhang et al meta-analysis because due to different definitions of exposure across studies, evidence of bias in case control studies, uncertainty about the latency of NHL, and selection of risk estimates from AHS and pooled estimate from De Roos's case-control analyses. The paper was undertaken to provide updated risk estimates following Zhang's meta-analysis, and using ever exposure as the key metric for deriving summary risk estimate. These authors (Kabat et al) provided summary risk estimates across different latency windows, and there summar measures demonstrated that risks increased with increasing latency (no lag: RR=1.05 (95% CI=0.87-1.28), 15 year lag RR=1.25, 95% CI: 1.01-1.25; and 20 year lag RR=1.41, 95% CI=1.13-1.76. While seemingly there was no formal search strategy, I have no issue with this given for all intents and purposes the authors sought to update the Zhang et al meta-analysis. This study used the findings from the pooled analyses by Pahwa (of US and Canadian studies). From the perspective of looking at heterogeneity of risk estimates, my preference would have been to look at findings from the individual studies. This also acknowledges that the studies may have different types of bias (from use of different questionnaires, as well as case/control recruitment methods. Apart from that the selection of risk estimates in Table 1 seems well justified, and this meta-analyses of these values produces an RR=1.05, 95% CI=0.87-1.28. I found that the methods used to derive the 15 and 20 year lag exposure-response relationships was not really clearly described in the paper. To me, it appears as though the authors put in, as one of the 6 risk entries for the meta-analysis the risk estimate for the 20 year lag from the AHS study, then repeated this using the risk estimate from the 15 year lag. I think it is important to be clear that there are not 6 risk estimates obtained across 6 studies that provide a measure of association for either of these two lags. I question why the authors would even do this given they expressed concerns about Zhang et al's approach to derive a summary measure across all 6 studies that relied on vastly different exposure metrics. While I can appreciate that some additional sensitivity analyses using different values has some value, I still think the approach to summarize vastly different exposure estimates into a single summary measures is fundamentally flawed. I agree with the authors that the case-control studies of glysophate have important biases, and I also agree with several of the points made by Crump (https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.13440) that these biases may well serve to overestimate possible cancer risks from glyphosate. Recall bias and participation are potentially major sources of bias in these case control studies. Overall, I see this paper adding very little to the literature. Philosophically, I question the incorporation of pooled analyses into a meta-analysis. Meta-analysis is an important tool that can not only generate a summary measure of association (by synthesizing findings across studies) but allows for the assessment of important sources of heterogeneity. This feature is lost when researchers start including pooled analysis into the meta-analytic framework, and this is even worse when there are relatively few studies of a topic (like glyphosate).

65.0

Expert 4

	Factor Rating::1
Methods::25	6
Results::25	8

Discussion/Conclusions::25	7
Application to Risk Assessment Decision-Making::25	5

This is a commentary on Zhang et al. and sensitivity analysis of its results; it is unfortunate that it does not cover analysis of three pooled cohort by Leon et al. As such, it is a bias analysis of somewhat outdated evidence, not synthesis of all available evidence. The authors set out to illustrate how selection of effect estimates influences meta-analysis and they focus on the choice of exposure metrics for which results are summarized, when there are choices other than unambiguous exposed vs unexposed. This is an important and useful step in appraising overall uncertainty as it addresses uncertainty due to analytical choices made by authors, a source of publication bias in situ. The authors blunder when they state that less complex analysis will lead to more stable and precise risk estimate. There is no such guarantee and there is a whole family of "complex" methods that shrink estimates and reduce bias, like the one De Roos used. This strange claim undermines credibility of authors but may be of no practical consequence in this instance. Appropriately, both fixed and random effects meta-regressions were used. Sensitivity analyses involved omitting one study at a time and using alternative summary effect estimate. This is consistent with aims of the investigation. The authors observed no heterogeneity (I2=0), confirming that presence of heterogeneity arises from the analytical choices made by investigators conducting meta-analyses of the NHL-glyphosate association. However, given study's aim, Table 1 seems inappropriate as it falsely represents that the selected effect estimates were the best ones. It would have have been far more useful to conduct sensitivity analysis that captures the full range of possible outcomes, as is done in text (left-hand column p 411, top). The result would have been a range of meta-RR and associated 95%CI. In essence, sensitivity analyses due to to selection of effect estimates is missing from Table 1 and is not captured fully in text. However, the authors do make the point that meta-RR either shows evidence of weak effect (if one uses $p<0.05$ criteria) or no such effect (again using $p<0.05$ criteria), depending on which effect estimates are selected. As such, the analysis still does not tell us any more than that the best estimate of RR is within 95%CI of 1 to 2, without adding further indication of how likely values near 2 are relative to those close to 1.3 (UCL in summary RR in Table 1). I am unsure that this is of any practical importance but such analyses seemed to be within reach of the authors. The discussion is balanced and informative, addressing reasons for heterogeneity between case-control and cohort studies on glyphosate (and pesticides in general) in relation to NHL (the issue of recall bias). Description of alternative choices of summary RR for including in meta-analysis is likewise useful and helps contextualize. The matter of latent confounding and including RR adjusted differently in meta-analysis is likewise illuminated well. The authors tires to defend excluding studies that used crop-exposure-matrix and this is clearly a blunder because it is nonsense to suggest, in the face of decades of empirical and theoretical evidence, that just exposure matrices are "of questionable value" (citing fringe opinion in reference [19] on this matter as evidence is not rigorous). Thus, ironically, the authors' paper suffers form selection of studies in their analysis that may well have the least harmonized exposure estimates. The authors give a sensible overview as to the reasons how publication bias in situ can arise and why; this is important to understanding why meta-analyses of the same 6 reports can have different apparent conclusions (even if numerically equivalent). The authors caution against pooling AHS with case-control studies, but do not provide any suggestions on how to combine evidence quantitatively, e.g. through bias analysis/adjustment prior to pooling of studies afflicted by different information and selection biases (possible if one is to start using methods in textbook of Lash et al). Pooling AHS with case-control studies has been shown to create

heterogeneity in effect estimates, depending on choice of effect estimates, so caution is warranted but caution alone does not offer a practical solution when doing nothing is not an option (such as understanding evidence on carcinogenicity of glyphosate). The authors appear to be apologists for AHS in their concluding statement. This is unhelpful and is an indirect attack on analysis of pooled US, French and Norwegian cohorts. The article makes contribution to risk assessment and decision making by showing that meta-analysis by Zhang et al made questionable choices that likely led to overestimate of meta-RR. The analysis does nothing to help us understand what is the best summary RR, because the authors conducted very limited sensitivity analysis and rejected all studies (which contain the majority of evidence as seen in Leon et al) that used exposure assessment methods they dislike.

Comments 3

SCORE
0

Expert 2 *07/16/2021 11:51*

I don't have other debate comments here, but found the other comments provided by the other experts interesting, and highlights aspects I hadn't considered. The concluding remarks of expert 4 - last 8 lines were insightful, and in my view on target.

SCORE
1

Expert 6 *07/19/2021 10:45*

I need to amend my interpretation of the Andreotti et al. analysis of the AHS data. Please, dismiss my previous comment on this meta-analysis and refer to my second-round comments on the Leon et al. meta-analysis about the calculation of the cumulative exposure indicator in the AHS study. Also, contrary to my impression from the previous hurried reading, commercial applicators were excluded from Andreotti's analysis. These were those who experienced order of magnitude higher exposures to glyphosate. Kabat et al. explored latency, which is a step ahead, although still referring to the dichotomous categorization of exposure. The increase in risk estimates with latency introduces elements of uncertainty about the reliability of results based on the ever-exposed category. There is uncertainty, for sure, which makes statements such as the one by Donato et al about "evidence that the risk of NHL is not increased in workers exposed to glyphosate" definitely unjustified. I insist that diluting exposure, matched with the heterogenous array of NHL as the disease outcome, and raising methodological issues only on the positive findings, results in denying that there might be a problem. Such margin of uncertainty makes the current state of knowledge on glyphosate and lymphoma, based solely on epidemiological studies, not yet applicable to risk assessment and decision making. Glad to see that Kabat et al. shared my view about keeping separate results from case-control studies and from cohort studies and that scoring the quality of papers does not convey superiority to the AHS conclusions. I also like another statement in their discussion about the "…need for sound methodological guidance on

the conduct of meta-analyses of observational studies, which lack the protection of randomization". Kabat et al. seem not that much familiar with the development of occupational epidemiology and retrospective assessment of occupational exposures. Their negative judgment on job-exposure matrices denies the contribution they can provide in going beyond the occupational title to start addressing exposures, although in an explorative way. It is however correct not to combine JEM (CEM in this case) findings with the AHS exposure assessment. I subscribe to the statement of expert 2 about the importance of meta-analysis not only in generating summary measures of association but also for the assessment of heterogeneity across studies. For this reason, the individual studies and not the pooled analysis of their data should be the object of a meta-analysis

SCORE
0

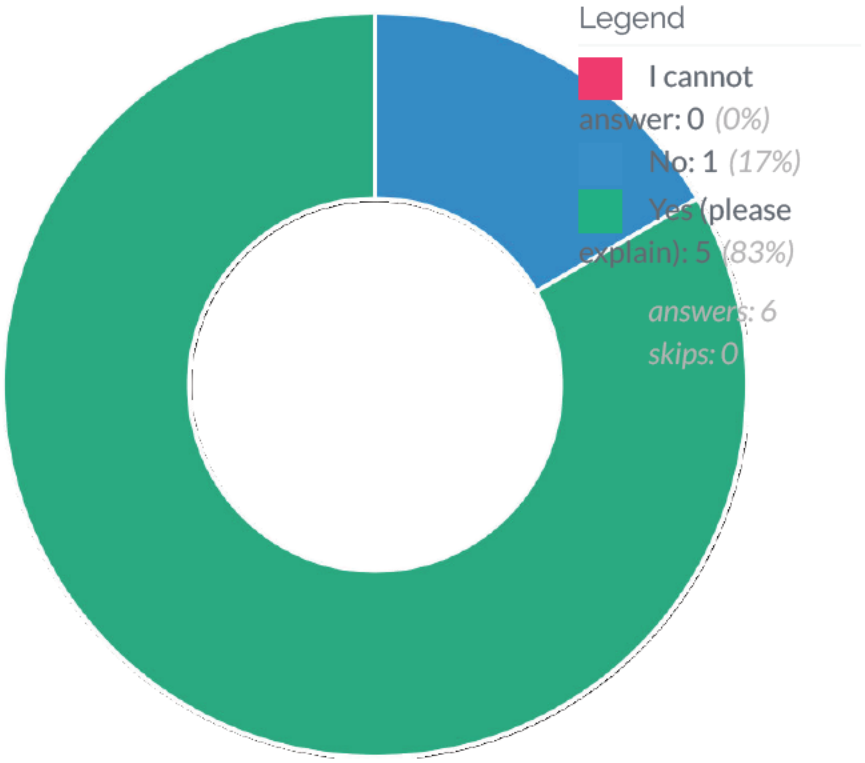
Expert 1

07/21/2021 23:06

I don't have any additional comments regarding this paper. The comments made by the other experts are valid.

Result 2.1 *(ID: 5178)*
Question 2.1 *(ID: 4154)*

Are you aware of any other meta-analyses for glyphosate and non-Hodgkin's lymphoma that should be considered in this review?



Answer Explanations

Yes (please explain) — Expert 1

1. Boffetta P, Ciocan C, Zunarelli C, Pira E. Exposure to glyphosate and risk of non-Hodgkin lymphoma: an updated meta-analysis. Med Lav 2021;112(3):194-199. 2. Barukcic I. Glyphosate and Non-Hodgkin lymphoma: No causal relationship. Journal of Drug Delivery and Therapeutics. 2020;10(1-s):6-29.

Yes (please explain) — Expert 6

The following is an update of Donato et al., 2020 Boffetta P, Ciocan C, Zunarelli C, Pira E.Exposure to glyphosate and risk of non-Hodgkin lymphoma: an updated meta-analysis. Med Lav. 2021 Jun 15;112(3):194-199. doi: 10.23749/mdl.v112i3.11123.

Yes (please explain) — Expert 3

However, there are narrative reviews by MInk et al., 2012 and Acquavella, et al. 2016 I think critical narrative reviews provide important information about the strengths and limitation of the body of evidence, without becoming enamored with a point estimate. I just saw a new review from Boffetta et al. 2021, which is an update to the Denato review.

Yes (please explain) Expert 2

No other reviews that I could find. Of potential note, it appears as though two additional reviews are now underway. From Prospero and it appears Bofetta et al is undertaking another systematic review of this topic: Paolo Boffetta, Francesca Donato. Occupational exposure to glyphosate and risk of non-Hodgkins lymphoma and multiple myeloma – A systematic review and meta-analysis. PROSPERO 2019 CRD42019130300 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019130300. There is an additional review that appears to capture cancer outcomes: Juan Pablo Alzate Granados, Orlando Scopetta, Daniel Rico. Effects of glyphosate on human health: a systematic review of the literature. PROSPERO 2019 CRD42019128975 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019128975

Yes (please explain) Expert 4

De Roos and I proposed a method of assessing if two effect estimates are truly different. We illustrated this in the context of glyphosate and multiple myeloma. The method (not so much the result) can help assess whether all these different meta-RR are all that different: they do not appear to be despite some variation in second decimal place of meta-RR. The reference is https://pubmed.ncbi.nlm.nih.gov/28025514/. StudIies on glyphosate seem to appear all the time and they need to be considered in any future analyses, e.g. https://pubmed.ncbi.nlm.nih.gov/33910586/ and https://pubmed.ncbi.nlm.nih.gov/33920383/, but they are not meta-analyses per se. Work of Crump has relevant observations on bias: https://pubmed.ncbi.nlm.nih.gov/31889327/

Comments 3

SCORE 0 **Expert 2** *07/16/2021 11:56*

Appreciate the citations provided by experts. I think the meta-analyses have really done as much as they can, and really think new individual studies are needed for future assessments. This should be accompanied with a careful assessment of study quality when determined weights to summarize findings across studies.

SCORE 1 **Expert 6** *07/19/2021 10:47*

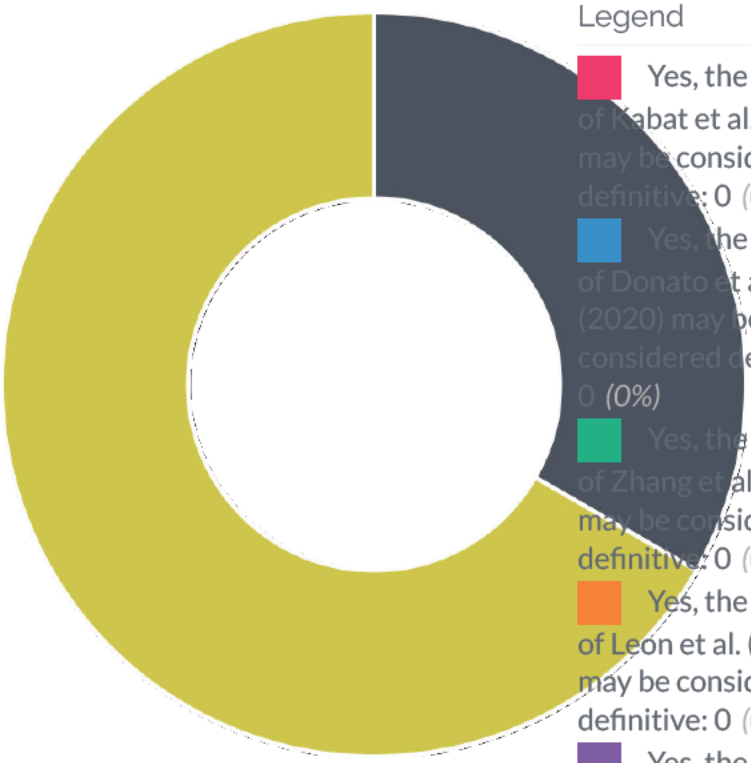
Agreed.

SCORE 0 **Expert 1** *07/21/2021 23:08*

I agree with expert #2, re-analysing the same data in different ways is not progressing this topic. New original research studies are required addressing the limitations of those studies previously published.

Result 2.2 *(ID: 5179)*
Question 2.2 *(ID: 4188)*

Has the definitive meta-analysis for glyphosate and non-Hodgkin lymphoma been performed?



answers: 6
skips: 0

Answer Explanations

Yes, a different analysis may be considered definitive (please explain below) Expert 4

Some hybrid of approach of [Chang and Delzell (2015) or Donato et al. (2020)] and Leon et al. (2019) is warranted, because none of the meta-analyses use all available evidence. Such work would have to include, at the very least, multi-dimensional probabilistic bias analysis following work of Lash et al. In fact, Lash published just such analyses for AHS (https://occup-med.biomedcentral.com/articles/10.1186/1745-6673-2-15), showing that conventional (published) analyses overestimate associations and under-value uncertainty. This approach must be extended to all pertinent studies and their quantitative synthesis. Without conducting calculations that consider total uncertainty, methods used in published meta-analyses, even if applied to all available data are very unlikely lead to a conclusion that perhaps the effect of ever use of glyphosate in agriculture on NHL is as large as RR=2 (97.5 percentile), but is likely barely detectable above 1.0. It seems that such conclusions are not definitive enough to resolve controversy about cancer risk of glyphosate; and perhaps no calculation will resolve this. However, if there is to be any hope to produce a calculation that dramatically alters the debate, it would be the one that adjusts for systematic errors using modern statistical methods, among which probabilistic bias analysis is the most accessible and easily understood. The only other analysis would involve estimation of validity of confounder and exposure assessment to glyphosate and design of prospective cohort studies to evaluate risks, in which there is a plan to adjust results for imperfection of confounder and exposure assessment, ideally via Bayesian approach that integrates all a priori knowledge of the subject matter. Such study would focus on absolute risks and be designed to provide risk estimates for strata of the population most relevant for risk assessment and policy. This is a pie in the sky solution and it is no reason to wait to make policy decisions on glyphosate and NHL: there is plenty of evidence for a decision now.

No, more work is needed (please explain below) Expert 1

The meta-analyses conducted to-date consist of three different types of studies. Firstly, there are the true systematic reviews and meta-analysis of published data. Secondly, there are pooled analysis of combined cohort studies and finally, there are re-analysis (sensitivity/sub-group) analysis of already published systematic reviews and meta-analysis. Some of the studies published so far are heterogeneous and are methodologically flawed, for example, some analyses pool or combine odds ratios and relative risks as if they are the same effect estimate. They are not the same and should not be mixed together. Some of these studies have used or mixed different point estimates from original research studies, for example, some meta-analyses will use overall exposure from some original research studies and combine these with estimates for the high exposure group. This is not methodologically sound and should be avoided. There is missing data! There are original research studies that exist in the literature but have not been included in these systematic reviews and meta-analyses. They should have been identified if the literature search strategy was comprehensive and used multiple databases. These have either been missed during literature searching or if identified the reason for leaving them out has not been justified. In addition, there are other systematic reviews and meta-analyses that have not been included in this review, they should be assessed to determine if they provide superior evidence than those currently available for this review.

No, more work is needed (please explain below) Expert 6

Before conducting new metanalyses, good quality and more powerful original studies are needed. Two aspects need improvement: definition of the disease entity and retrospective exposure assessment. The use of the old NHL definition might continue to result in contradictory findings, and it is confusing

because of inconsistencies in excluding or including subtypes such as CLL. On the other hand, more studies are required using the most updated pathological classifications to explore associations with individual subtypes, with careful consideration of the time frame for exposure and using multiple exposure metrics.

Yes, a different analysis may be considered definitive (please explain below) — Expert 5

Although the meta-analysis of Kabat 2021-compared to the rest- may be considered more appropriate it cannot be considered definite even though it tests several assumptions and hypotheses. Case-control bias (which were the majority of the underlying studies included in all meta-analyses) have inherent limitations, such as recall bias and exposure misclassifications. As is pointed out in the Kabat study, it may be wise to separately analyse case-controls studies from cohort studies, by selecting appropriately the inclusion criteria in the statistical modelling. Analyses taking into account latency period, should report complete (or properly imputed) datasets with appropriate follow-up periods and assessment of co-exposures.

No, more work is needed (please explain below) — Expert 2

At this time, I believe the analysis by Chang and Delzell is the probably the most definitive of the study - however, the number of studies is relatively small, and the question has not been fully resolved. I see limited utility for undertaking new meta-analysis until new results are available. I disagree with the approach to include a measure of association into a meta-analysis that is derived pooled analyses. The preferred approach is to use measures of association from individual studies so as to better understand heterogeneity in the risk estimates across studies. The findings from the AHS , which most authors seem to view as the best study, do not support the hypothesis that glysophate increases the risk of NHL, and the other studies have potentially important sources of bias. As good a tool of meta-analysis is, it can not overcome bias, or lack of validity, from individual-level studies. It is for this reason, I think the Chang and Dezell may well be the best meta-analysis published to date, and that they derived a summary measure of risk among exposed at 1.3 (which was statistically significant) there remains a large degree of uncertainty - but this till not be overcome with additional meta-analysis of the same 5-6 studies. Additional studies would be of value.

No, more work is needed (please explain below) — Expert 3

I disagree with the question. I don't think there is a definitive meta analysis, just as I don't think there is ever a single definitive epid study. The MA by Denato et al. (2020) and Kabat et al. (2021) collectively provide important information about the state of the science for glyphosate.

Comments 4

SCORE 1 — Expert 4 — 07/13/2021 07:45

I tend to agree with #3 that the ask to identify a " definitive meta-analysis" is a trick question: there is not such thing. The implication that meta-analysis is the best way to summarize evidence in this case is also dubious, as another reviewer pointed out. Narrative reviews are quite valuable as they can be more nuanced than brute-force averaging of meta-analysis. If one was seeking better quantitative synthesis of evidence, pooling individual-level data would be advisable, a "collaborative pooled re-

analysis" as Doll called. Pooling cohorts was a step in that direction. There must be a way to also use case-control studies in the same way.

SCORE
0

Expert 2

07/16/2021 11:58

I agree with the comment by expert 4 above. Ultimately I do think however more studies (not meta-analyses, and not just extended follow-up of existing cohorts) are needed to advance knowledge.

SCORE
1

Expert 6

07/19/2021 10:55

Agree with expert 4 and 2, particularly on the need of more brand new studies.

SCORE
0

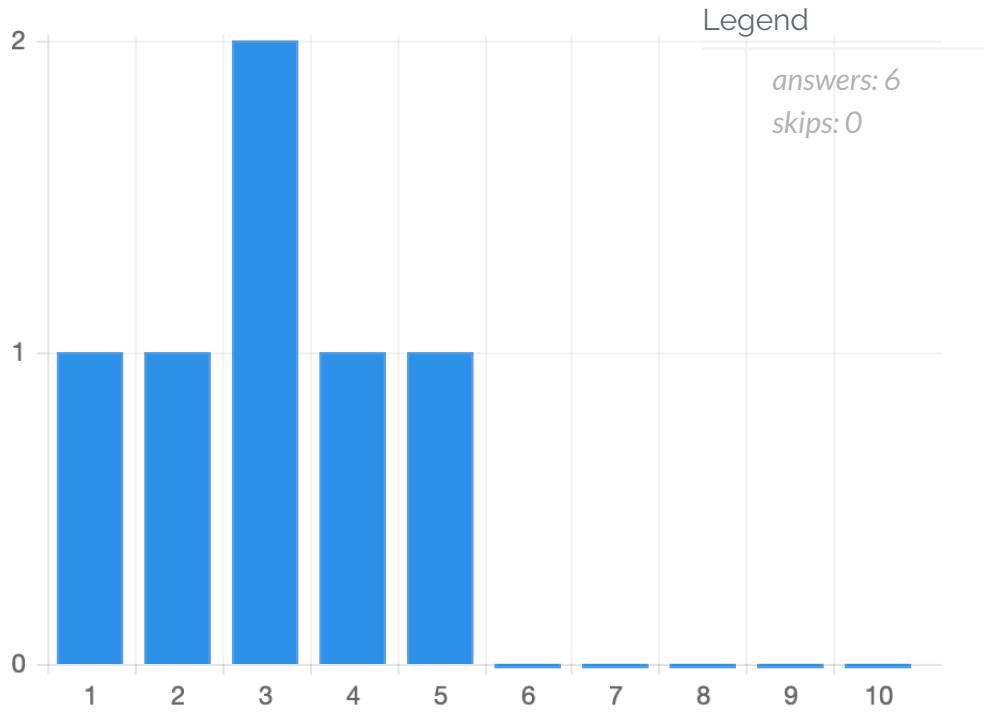
Expert 1

07/21/2021 23:10

I agree with experts #4, #2 and #6 comments, we need new studies, the data remain inconclusive.

Result 2.3 *(ID: 5180)*
Question 2.3 *(ID: 4186)*

Based on your review of the meta-analyses for glyphosate, on a scale of 1-10 what is your overall opinion on a causal relationship between glyphosate exposure and non-Hodgkin's lymphoma (1=no relationship, 10=clear, conclusive relationship; please explain your answer below)?



Answer Explanations

3 Expert 2

Much of my reasoning has already been explained above. To summarize, I have concerns about potential sources of bias (particularly participation and recall bias from the case control studies) in the individual level studies. Moreover, the strongest design (AHS - cohort study) does not provide compelling evidence of an association. There is little support for exposure-response relationships because there are few studies, many did not demonstrate an association to begin with, and given the different exposure measures, it is impossible to harmonize risks across studies. Finally, reviews of the experimental evidence by a number of regulatory agencies in Europe and the US concluded that mechanistically glyphosate is not a carcinogen in humans. There is a possibility of some plausible mechanism where it could cancer (but the same can be said for almost any exposure), but this has not been identified as yet, and any positive summary meta-analysis risk could well be explained by bias. More individual level studies are need to better assess the relevance of glyphosate as a human carcinogen. The existing numbers of studies are simply too small, and many are underpowered (statistically)

5 Expert 1

I think that given the methodological heterogeneity that exists in these specific studies it is difficult to reach a clear conclusion. There needs to be more original studies on the relationship between glyphoate and lymphoma.

1 Expert 6

Overall results for NHL (all subtypes combined) do not support an association with exposure to glyphosate. Positive findings in individual studies might have been driven by the proportion of the cases with specific subtypes.

3 Expert 3

We continue to see low risk estimates for farmers and NHL. After decades of research and trying to target one herbicide after another, we know very little about why this association appears. The AHS was launched to address the problems of case control studies discussed by some of the MA authors. However, the AHS reports their data in seemingly random intervals, which are not helpful for discussions of causality by regulators and risk assessors.

2 Expert 4

It is very hard to exclude a small effect, especially for the subtype that appears to be associated in the pooled cohort. However, there are multiple examples in epidemiology where such weak effects do not withstand either replication or bias analysis. With respect to glyphosate and any NHL, it is safe to say that there is no relationship that modern epidemiology can demonstrate and which is worthy of policy intervention, i.e. a score of 1 to 2 (same as my assessment of 95%CI of meta-RR from most meta-analyses conducted to date).

4 Expert 5

Overall, and as it has been shown by Kabat et al., it seems that the strength of evidence is dependent upon the effect estimate used. Methodological issues, such as study design, exposure and outcome assessment, recall bias and latency period contributed to the uncertainty. Cohort and case-control studies should either be carefully selected and analysed together (which is methodologically very challenging) or analysed separately.

Comments 4

SCORE 0 **Expert 2** *07/16/2021 12:00*

Seems like a pretty symmetric distribution to me! (I am happy to be at the mean/median -lol)

SCORE 0 **Expert 3** *07/16/2021 14:19*

Viva the mean! However, my response looks like an impropriate finger gesture.

SCORE **Expert 6** *07/19/2021 10:56*

0 Oops

SCORE
0

Expert 1 07/21/2021 23:11

I agree with the other experts in their responses.

Result 2.4 *(ID: 5181)*
Question 2.4 *(ID: 4189)*

Some of the meta-analyses also considered other cancer endpoints. Please rate your confidence in those findings (1=low confidence, 10=high confidence)

Endpoint	Confidence score	Total
B-cell lymphoma mRR = 2.0 (1.1-3.6) (Change and Delzell, 2015)	*2* **50.00%** 3 *3* **16.67%** 1 *5* **33.33%** 2	6
Diffuse large B-cell lymphoma mRR = 1.1 (0.5-2.3) (Change and Delzell, 2015)	*1* **16.67%** 1 *2* **33.33%** 2 *3* **16.67%** 1 *5* **33.33%** 2	6
Diffuse large B-cell lymphoma HR = 1.36 (1.00-1.85) (Leon et al. 2019)	*2* **16.67%** 1 *3* **16.67%** 1 *4* **33.33%** 2 *6* **16.67%** 1 *8* **16.67%**	6

	1	
CLL/SLL mRR = 1.3 (0.2-10), 1.9 (0.9-4.0) (Change and Delzell, 2015)	*1* **16.67%** 1 *2* **33.33%** 2 *3* **16.67%** 1 *5* **33.33%** 2	6
CLL/SLL HR = 0.92 (0.69-1.24) (Leon et al. 2019)	*2* **16.67%** 1 *4* **33.33%** 2 *6* **16.67%** 1 *8* **33.33%** 2	6
Follicular lymphoma mRR = 1.7 (0.7-3.9) (Change and Delzell, 2015)	*2* **33.33%** 2 *3* **16.67%** 1 *4* **16.67%** 1 *5* **33.33%** 2	6
	2 **16.67%** 1 *3* **16.67%** 1 *4*	

Follicular lymphoma HR = 0.79 (0.52-1.21) (Leon et al. 2019)	**16.67%** 1 *6* **16.67%** 1 *8* **33.33%** 2	6
Hairy cell leukemia mRR = 2.5 (0.9-7.3) (Change and Delzell, 2015)	*1* **16.67%** 1 *2* **16.67%** 1 *3* **33.33%** 2 *5* **33.33%** 2	6
Hodgkin lymphoma = 1.1 (0.7-1.6) (Change and Delzell, 2015)	*2* **16.67%** 1 *3* **33.33%** 2 *5* **16.67%** 1 *6* **16.67%** 1 *8* **16.67%** 1	6
Multiple myeloma mRR = 1.4-1.5 (0.9-2.1) (Change and Delzell, 2015)	*2* **16.67%** 1 *3* **33.33%** 2 *5* **50.00%** 3	6

Multiple myeloma HR = 0.87 (0.66-1.15) (Leon et al. 2019)	*2* **16.67%** 1 *3* **16.67%** 1 *4* **16.67%** 1 *6* **16.67%** 1 *8* **33.33%** 2	6
Multiple myeloma RR = 1.04 (0.67-1.41) (Donato et al. 2020)	*2* **16.67%** 1 *3* **33.33%** 2 *4* **16.67%** 1 *5* **16.67%** 1 *8* **16.67%** 1	6
Leukemia mRR = 1.0 (0.6-1.5) (Change and Delzell, 2015)	*2* **16.67%** 1 *3* **33.33%** 2 *5* **33.33%** 2 *9* **16.67%** 1	6

Answer Explanations

Expert 2

Endpoint	Confidence score
B-cell lymphoma mRR = 2.0 (1.1-3.6) (Change and Delzell, 2015)	2
Diffuse large B-cell lymphoma mRR = 1.1 (0.5-2.3) (Change and Delzell, 2015)	2
Diffuse large B-cell lymphoma HR = 1.36 (1.00-1.85) (Leon et al. 2019)	6
CLL/SLL mRR = 1.3 (0.2-10), 1.9 (0.9-4.0) (Change and Delzell, 2015)	2
CLL/SLL HR = 0.92 (0.69-1.24) (Leon et al. 2019)	6
Follicular lymphoma mRR = 1.7 (0.7-3.9) (Change and Delzell, 2015)	2
Follicular lymphoma HR = 0.79 (0.52-1.21) (Leon et al. 2019)	6
Hairy cell leukemia mRR = 2.5 (0.9-7.3) (Change and Delzell, 2015)	3
Hodgkin lymphoma = 1.1 (0.7-1.6) (Change and Delzell, 2015)	3
Multiple myeloma mRR = 1.4-1.5 (0.9-2.1) (Change and Delzell, 2015)	5
Multiple myeloma HR = 0.87 (0.66-1.15) (Leon et al. 2019)	6
Multiple myeloma RR = 1.04 (0.67-1.41) (Donato et al. 2020)	3
Leukemia mRR = 1.0 (0.6-1.5) (Change and Delzell, 2015)	3

Please note, it is Chang and Delzell and not Change... Chang and Delzell (2015): B cell lyphoma: this summary measure is based on only 2 studies. I have concerns about the Coco study (poor participation rates). The second study (Erikssson seemingly had high response rates) but I could not find the specific rates in the paper. Recall bias is a potential issue in case-control studies. Given the small number of studies, I have little faith in the validity of this estimate. Chang and Delzell (2015): Diffuse large B cell lymphoma: I also have little confidence in this estimate as it too is based on only 2 studies. Moreover, the confidence intervals of each of these studies are wide 1.22 (0.44 - 3.35) and 1.0 (0.3-2.7) Leon (2019): diffuse large cell lymphoma. I have greater confidence in this estimate given the stronger (cohort) study design that was used. I have some concern given the multiple testing issues (number of active ingredients looked at and number of outcomes) and no really clearly articulated reason why biologically one would expect a stronger association with this subtype. Chang and Delzell (CLL/SLL): As expressed already, this estimate is based on only 2 studies, and the confidence intervals for the individual studies are very broad. Chang and Delzell (Multiple myeloma): I have slightly more confidence in this estimate given it is based on 6 studies, and the measures of associations are somewhat similar across studies. Nonetheless, still potential for some bias in individual studies, thus I am not fully confidence in validity of estimates from individual studies. Chang and Delzell (leukemia): meta-analysis risk estimate based on only 3 study, one of these studies only had one case! Second study only had 15 cases. Confidence intervals were wide and included in them risk estimates that could either double or halve the risk. Chang and Delzell (Hodgkin lymphoma) Little confidence. Only 2 studies used to generate RR estimate. Small number of cases, potential for bias. Donato (Multiple myeloma): Summary measure based on only 3 estimates. Substantial heterogeneity across the studies. Other Leon studies: generally ranked them the same given they are based on cohort data, and provide more information than case-control studies.

Expert 1

Endpoint	Confidence score
B-cell lymphoma mRR = 2.0 (1.1-3.6) (Change and Delzell, 2015)	5
Diffuse large B-cell lymphoma mRR = 1.1 (0.5-2.3) (Change and Delzell, 2015)	5
Diffuse large B-cell lymphoma HR = 1.36 (1.00-1.85) (Leon et al. 2019)	8

CLL/SLL mRR = 1.3 (0.2-10), 1.9 (0.9-4.0) (Change and Delzell, 2015)	5
CLL/SLL HR = 0.92 (0.69-1.24) (Leon et al. 2019)	8
Follicular lymphoma mRR = 1.7 (0.7-3.9) (Change and Delzell, 2015)	5
Follicular lymphoma HR = 0.79 (0.52-1.21) (Leon et al. 2019)	8
Hairy cell leukemia mRR = 2.5 (0.9-7.3) (Change and Delzell, 2015)	5
Hodgkin lymphoma = 1.1 (0.7-1.6) (Change and Delzell, 2015)	5
Multiple myeloma mRR = 1.4-1.5 (0.9-2.1) (Change and Delzell, 2015)	5
Multiple myeloma HR = 0.87 (0.66-1.15) (Leon et al. 2019)	8
Multiple myeloma RR = 1.04 (0.67-1.41) (Donato et al. 2020)	5
Leukemia mRR = 1.0 (0.6-1.5) (Change and Delzell, 2015)	5

Everything here comes back to is the systematic review and meta-analysis conducted in a methodologically sound approach. The Chang and Delzell study has limitations: small sample size, single database for literature search, does not follow the PRISMA Guidelines, does not undertake a quality assessment of the studies, and did not assess publication bias. The Leon study is not a true systematic review and meta-analysis, it is a pooled analysis of three large cohort studies, it has less limitations than the Chang & Delzell study but it is also not directly comparable.

Expert 6

Endpoint	Confidence score
B-cell lymphoma mRR = 2.0 (1.1-3.6) (Change and Delzell, 2015)	5
Diffuse large B-cell lymphoma mRR = 1.1 (0.5-2.3) (Change and Delzell, 2015)	5
Diffuse large B-cell lymphoma HR = 1.36 (1.00-1.85) (Leon et al. 2019)	4
CLL/SLL mRR = 1.3 (0.2-10), 1.9 (0.9-4.0) (Change and Delzell, 2015)	5
CLL/SLL HR = 0.92 (0.69-1.24) (Leon et al. 2019)	4
Follicular lymphoma mRR = 1.7 (0.7-3.9) (Change and Delzell, 2015)	5
Follicular lymphoma HR = 0.79 (0.52-1.21) (Leon et al. 2019)	3
Hairy cell leukemia mRR = 2.5 (0.9-7.3) (Change and Delzell, 2015)	5
Hodgkin lymphoma = 1.1 (0.7-1.6) (Change and Delzell, 2015)	6
Multiple myeloma mRR = 1.4-1.5 (0.9-2.1) (Change and Delzell, 2015)	5
Multiple myeloma HR = 0.87 (0.66-1.15) (Leon et al. 2019)	3
Multiple myeloma RR = 1.04 (0.67-1.41) (Donato et al. 2020)	4
Leukemia mRR = 1.0 (0.6-1.5) (Change and Delzell, 2015)	5

There is still large uncertainty because of the small number of studies that evaluated individual lymphoma subtypes, and the large confidence interval of the risk estimates in each. Results for Hodgkin lymphoma and leukemia were null, but based on a couple of studies each and, as it concerns leukemia, the subtypes included varied across studies (Roos 2005 did not clarify whether chronic lymphatic leukemia (CLL) was included or not; in Brown et al. 1990, CLL was prevalent, and myelodisplasia was also included; in Kaufman et al.2009 only 8/180 cases were CLL). In general, the metanalysis by Change and Delzell was more accurate ed included more details on the original studies and, for MM, also more studies; for these reasons, I scored their estimates a bit higher when more estimates were available for the same malignancy.

Expert 3

Endpoint	Confidence score
B-cell lymphoma mRR = 2.0 (1.1-3.6) (Change and Delzell, 2015)	2
Diffuse large B-cell lymphoma mRR = 1.1 (0.5-2.3) (Change and Delzell, 2015)	2
Diffuse large B-cell lymphoma HR = 1.36 (1.00-1.85) (Leon et al. 2019)	2
CLL/SLL mRR = 1.3 (0.2-10), 1.9 (0.9-4.0) (Change and Delzell, 2015)	2
CLL/SLL HR = 0.92 (0.69-1.24) (Leon et al. 2019)	2
Follicular lymphoma mRR = 1.7 (0.7-3.9) (Change and Delzell, 2015)	2
Follicular lymphoma HR = 0.79 (0.52-1.21) (Leon et al. 2019)	2
Hairy cell leukemia mRR = 2.5 (0.9-7.3) (Change and Delzell, 2015)	2
Hodgkin lymphoma = 1.1 (0.7-1.6) (Change and Delzell, 2015)	2
Multiple myeloma mRR = 1.4-1.5 (0.9-2.1) (Change and Delzell, 2015)	2
Multiple myeloma HR = 0.87 (0.66-1.15) (Leon et al. 2019)	2
Multiple myeloma RR = 1.04 (0.67-1.41) (Donato et al. 2020)	2
Leukemia mRR = 1.0 (0.6-1.5) (Change and Delzell, 2015)	2

None of these analyses provide high confidence given they are based on only a few studies, studies that use crop-exposure matrices or case-control studies with potential for bias. However, given the biological plausibility for specific etiology by subgroup, these indicate opportunities for more focused research.

Expert 4

Endpoint	Confidence score
B-cell lymphoma mRR = 2.0 (1.1-3.6) (Change and Delzell, 2015)	2
Diffuse large B-cell lymphoma mRR = 1.1 (0.5-2.3) (Change and Delzell, 2015)	1
Diffuse large B-cell lymphoma HR = 1.36 (1.00-1.85) (Leon et al. 2019)	3
CLL/SLL mRR = 1.3 (0.2-10), 1.9 (0.9-4.0) (Change and Delzell, 2015)	1
CLL/SLL HR = 0.92 (0.69-1.24) (Leon et al. 2019)	8
Follicular lymphoma mRR = 1.7 (0.7-3.9) (Change and Delzell, 2015)	3
Follicular lymphoma HR = 0.79 (0.52-1.21) (Leon et al. 2019)	8
Hairy cell leukemia mRR = 2.5 (0.9-7.3) (Change and Delzell, 2015)	1
Hodgkin lymphoma = 1.1 (0.7-1.6) (Change and Delzell, 2015)	8
Multiple myeloma mRR = 1.4-1.5 (0.9-2.1) (Change and Delzell, 2015)	3
Multiple myeloma HR = 0.87 (0.66-1.15) (Leon et al. 2019)	8
Multiple myeloma RR = 1.04 (0.67-1.41) (Donato et al. 2020)	8
Leukemia mRR = 1.0 (0.6-1.5) (Change and Delzell, 2015)	9

I use the rule of wider confidence limits be less likely to be replicated as per arguments of Charles Poole (Epidemiology 2001, 12(3):291-294). Higher confidence to pooled cohort; less confidence to case-control studies.

Expert 5

Endpoint	Confidence score
B-cell lymphoma mRR = 2.0 (1.1-3.6) (Change and Delzell, 2015)	3
Diffuse large B-cell lymphoma mRR = 1.1 (0.5-2.3) (Change and Delzell, 2015)	3

Diffuse large B-cell lymphoma HR = 1.36 (1.00-1.85) (Leon et al. 2019)	4
CLL/SLL mRR = 1.3 (0.2-10), 1.9 (0.9-4.0) (Change and Delzell, 2015)	3
CLL/SLL HR = 0.92 (0.69-1.24) (Leon et al. 2019)	4
Follicular lymphoma mRR = 1.7 (0.7-3.9) (Change and Delzell, 2015)	4
Follicular lymphoma HR = 0.79 (0.52-1.21) (Leon et al. 2019)	4
Hairy cell leukemia mRR = 2.5 (0.9-7.3) (Change and Delzell, 2015)	3
Hodgkin lymphoma = 1.1 (0.7-1.6) (Change and Delzell, 2015)	3
Multiple myeloma mRR = 1.4-1.5 (0.9-2.1) (Change and Delzell, 2015)	3
Multiple myeloma HR = 0.87 (0.66-1.15) (Leon et al. 2019)	4
Multiple myeloma RR = 1.04 (0.67-1.41) (Donato et al. 2020)	3
Leukemia mRR = 1.0 (0.6-1.5) (Change and Delzell, 2015)	3

Change and Delzell, 2015: small sample, low power to detect Leon et al. 2019: not clarity in the real effect of the exposure to risk of NHL. Donato et al. 2015: small sample size and large weight is given to the AGRICOH cohort without adjusting properly for heterogeneity within this cohort.

Comments 4

SCORE 0

Expert 2 *07/16/2021 12:01*

I think if there is so much uncertainty around all types of NHL - then likely find just as more (or more) with subtypes - given small numbers.

SCORE 0

Expert 4 *07/19/2021 07:54*

This is probably the case of only massive risks being detectable, like the vinyl-chloride monomer and angiosarcoma of the liver story: if there was an effect calling for intervention, we would have seen it by now. However, we will never be able to exclude the possibility of something going on (no proving of the negative) but we can perhaps agree that there is no _epidemic_ of the subtypes due to glyphosate.

SCORE 0

Expert 6 *07/19/2021 11:01*

Sure. Apart from a few instances, most tumours arise after a long and complex set of interactions between genes and the environment. There is a long list of risk factors for lymphoma subtypes, from viruses to autoimmune diseases, to several occupational exposures. The glyphosate effect (if any) might be just one of these risk factors, and possibly not stronger than any of those.

SCORE 0

Expert 1 *07/21/2021 23:13*

I agree with the comments made, it gets even harder to establish relationships with sub-types of lymphoma given the even smaller sample sizes.

Result 2.5 *(ID: 5182)*
Question 2.5 *(ID: 4190)*

What is your opinion on the primary reason(s) for the differing results between the 6 meta-analyses for glyphosate and non-Hodgkin's lymphoma?

Expert 1

The meta-analyses conducted to-date consist of three different types of studies. Firstly, there are the true systematic reviews and meta-analysis of published data. Secondly, there are pooled analysis of combined cohort studies and finally, there are re-analysis (sensitivity/sub-group) analysis of already published systematic reviews and meta-analysis. Some of the studies published so far are heterogeneous and are methodologically flawed, for example, some analyses pool or combine odds ratios and relative risks as if they are the same effect estimate. They are not the same and should not be mixed together. Some of these studies have used or mixed different point estimates from original research studies, for example, some meta-analyses will use overall exposure from some original research studies and combine these with estimates for the high exposure group. This is not methodologically sound and should be avoided. There is missing data! There are original research studies that exist in the literature but have not been included in these systematic reviews and meta-analyses. They should have been identified if the literature search strategy was comprehensive and used multiple databases. These have either been missed during literature searching or if identified the reason for leaving them out has not been justified. In addition, there are other systematic reviews and meta-analyses that have not been included in this review, they should be assessed to determine if they provide superior evidence than those currently available for this review.

Expert 6

1. generic definition of the disease entity, assembling different nosological entities; 2. dilution of the risk estimate in the ever exposure definition; lack of consideration of important criteria in establishing a causal relationship, such as dose-response analyses. 3. incomplete use of the available exposure information in some studies. 4. under evaluation of well designed crop-exposure matrices. 5. mixing case-control and cohort studies. 6. uncertain consideration of the time frame to define exposure (glyphosate was marketed in 1974 and started being widely used from 1975 first in the US and subsequently in Europe. The suggested overlapping between use of DDT and glyphosate suggested in the CNAP study and reported by Leon et al. 2019, is particularly worrisome, as DDT use for agricultural purposes was ruled out in 1972 and glyphosate came in use not earlier than 3 years later.

Expert 2

First, I think worth noting that there is some concordance in the summary estimates with 3 of the studies yielding mRR estimates between 1.3 and 1.5, while the other series of estimates were around 0.95 to 1.05. These risk estimates differed for many reasons. First, many of the more recent analyses made use of data from extended AHS follow-up (which as discussed above was largely null). In addition, the authors in some cases chose different risk estimates from the same study. In some cases, the authors did not follow their own

procedures and risks seemed to be picked in an adhoc fashion (see Leon et al 2014), in other cases such as Zhang, they leaned more towards picking the strongest exposure, and ultimately, try to force out a summary risk estimate using very dissimilar exposure metric. Some of the latter analyses included a singular measure for pooled analyses and this served to attenuate risk estimates (see Donato and Kabat). While I appreciate the fact that the Chan et al was support in part by industry funding, I felt their paper was the strongest methodologically, and carefully critiqued the findings of the individual studies.

Expert 4

Some authors had different evidence available to them, less complete than that available to authors of work published later. Selection of effect estimates from published work played a role, as well as differences in opinions about value of different exposure assessment methods. It is hard to judge now much role general knowledge about toxicity of glyphosate (if any), hidden biases, and personal values played in the conclusions the authors reached, but it must have played a role. The reason I make this claim is that the results between the 5 meta-analyses are NOT different: they all yield estimate of weak effect for ever vs never exposed that lies within 95%CI of 1 to 2. (Pooled cohort is different: it is not a meta-analysis.) Such effect estimates are known to be vulnerable to many sources of bias, such that definitive qualitative interpretation (is it causal? yes/no) is only possible based on statics alone. Strong effect can be ruled out but the remaining effect may be either absent or too weak to detect using the best methods and resources that epidemiology has mustered; systematic errors remains a plausible explanation for any results that apparently deviate from the null and if these are not objectively appraise, authors are left with additional freedom to infuse interpretations with whatever biases they acquired during lifetime of research and advocacy.

Expert 3

The results are subtly different. Denato's review does a nice job showing the influence of publication year of each study. With the newer, null publications, the MA that included them had lower risk estimates. As Kabat et al. and Chang et al. reviews noted, the selection of which risk estimate to use is highly influential. This speaks to the problem of meta-analysis per se, the point estimate becomes the focal point and the lack of transparency regarding how individual studies were represented in the MA is easily lost.

Expert 5

-Combination of case-control studies and cohort studies - No accurate definition of lags of exposure and exposure misclassification -Recall bias in case-control studies -Definition and use of "proper" latency periods in the only cohort study included. -NHL is a very rare disease and in some reviews (i.e, Zhang et al.) it seems that the number of cases is high.

Comments 6

SCORE
0

Expert 4 *07/13/2021 07:50*

I appreciate concern of rev #1 about pooling of OR and RR, which I glanced over. But is this really a big concern in this specific context of very rare outcome?

Non-portability of OR is a nuisance, but in this situation, it seems that the assumption that OR and RR are not materially different is not too outrageous. As everyone pointed out, there are many other sources of bias at play that nobody seems to be able or willing to address quantitatively, except Lash. And I want to correct my own typo: "Such effect estimates are known to be vulnerable to many sources of bias, such that definitive qualitative interpretation (is it causal? yes/no) is NOT possible based on statics alone."

SCORE
2

Expert 3

07/13/2021 09:11

I agree with reviewer 2, that the results across MA aren't much different. Is too much being made of the differences between 1.3 and 1.5, given the discussions of the reviewers here? Some would say any increase, of any size or precision is a problem, while others focus on sources of bias, chance and confounding. It's interesting that the AHS results - given the cohort design, large sample size, and focus on registered applicators - hasn't dominated the causal discussion (beyond this exercise). Given the millions of tax dollars to fund this 20+year study, I wonder what is the right next step for epid research on this topic.

SCORE
1

Expert 4

07/15/2021 12:52

If AHS is not seen as giving valuable evidence, despite all of its flaws, then no epi study can. I echo that no epi study is every definitive outside of its time and place, i.e. we are not dealing with absolutes by applied messy research that is meant to be useful, not fodder for abstract causal inference debates.

SCORE
0

Expert 2

07/16/2021 12:03

I agree with comment above Expert 4 - AHS is best able to answer the question, and it is messy given rarity of the cancer, and exposure measurement error.

SCORE
0

Expert 6

07/19/2021 11:05

Sorry, I disagree with expert 4. AHS data have not been fully explored as yet. A different combination of the exposure variables and the separate analysis of commercial applicators might provide new insights on glyphosate and lymphoma.

SCORE
0

Expert 1

07/21/2021 23:16

I think everybody has made valuable insights into reasons for differences between these meta-analyses, all valid.

Result 3.1 *(ID: 5183)*
Question 3.1 *(ID: 4185)*

Please provide one question that you would like your fellow panel members to consider (please be sure to phrase in the form of a question)

Expert 1

> Is the pooled analysis of Leon et al. (2019) superior to all the other systematic review and meta-analyses conducted as part of this assessment?

Expert 6

> Previous work from the InterLymph Consortium has provided evidence of the substantial heterogeneity between lymphoma subtypes (Morton L. J Natl Cancer Inst Monogr. 2014;48:130-44). Still, most epidemiologists keep referring to the non-Hodgkin lymphoma disease entity, restricting or extending the list of subtypes included in it, which makes comparisons especially difficult. The same happens with leukemia studies, which most frequently include chronic lymphocytic leukemia, classified as a mature B-cell lymphoma by the REAL classification in 1994. In evaluating etiological factors, would you suggest keeping using the NHL denomination for consistency with the existing literature, or would you rather call for more studies focussing on specific B-cell and T-cell subtypes ?

Expert 2

> To what extent do you think the use of biomonitoring data can be used to support exposure characterization in epidemiological studies in an effort to advance our understanding of whether glyphosate increases the risk of NHL?

Expert 3

> What design of epidemiology or laboratory study would contribute to the question about glyphosate exposure and NHL?

Expert 4

> Do you really think that there is any material difference in summary RR of the five papers that report on case-control studies? Do you think that there is enough evidence to discard case-control studies from consideration until the reason why they systematically differ from cohorts can be explained? Pooled cohort analysis can be improved upon but it seems that it will always dominate epidemiologic evidence. New huge case-control studies are out of the question (in our lifetime). Consequently, is there any value in considering case-control studies at all when their contribution to quantitative synthesis of evidence is dwarfed by a well-conducted pooled cohort analysis?

Result 3.2 *(ID: 5184)*
Question 3.2 *(ID: 4155)*

Do you have any potential conflicts of interest (COI) you would like to declare? If yes, please describe the potential COI and indicate if you believes it impacts your ability to provide your unbiased, scientific opinions.

Expert 1

> No conflicts of interest.

Expert 6

> I have no conflicts of interest.

Expert 2

> In my view, I have no conflicts to declare. It may however be worth noting however that I published papers with several of the authors (Hoppin, Sandler, McLaughlin, Harris, Lubin,) on air pollution epidemiology as well as studies of urban greenness.

Expert 3

> I have worked in the chemical industry for 25 years, with some work related to 24d and NHL. I believe this makes me more aware of the underlying studies and their strengths and limitations but does not impact by ability to provide an unbiased, scientific opinion.

Expert 4

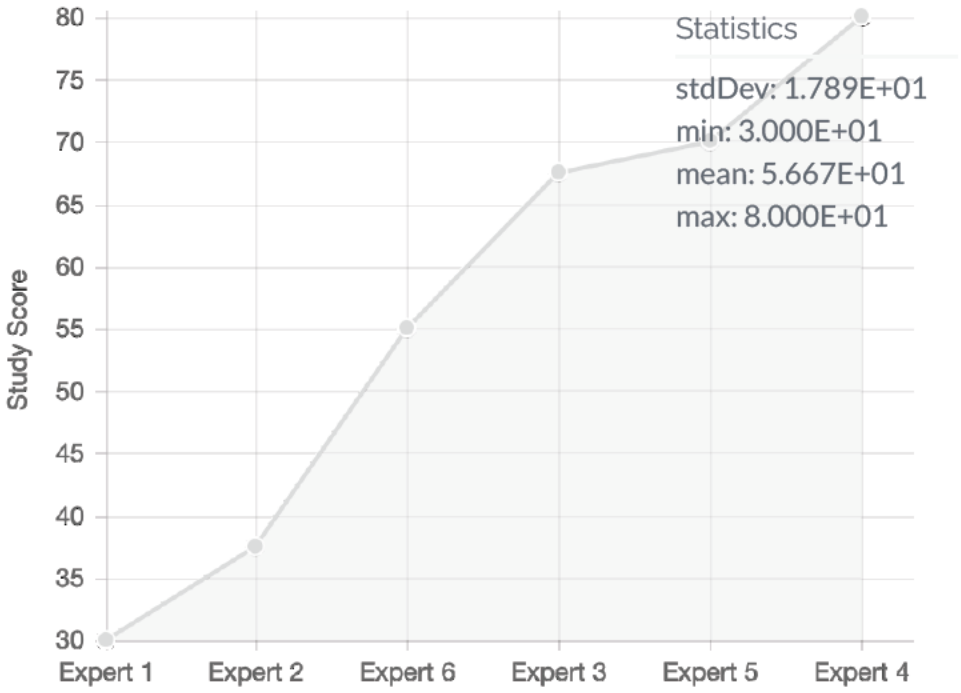
> I declare no conflicts of interest.

Expert 5

> No COI

Result 4.1 *(ID: 5192)*
Question 4.1 *(ID: 4192)*

On a scale of 1-10 (1=lowest, 10=highest) what is your degree of confidence in the following factors for study G: Boffetta et al. (2021; update of Donato et al., 2020)? Please provide an explanation for EACH rating in the space below."



Answer Explanations

30.0 — Expert 1

Factor	Factor Rating::1
Methods::25	3
Results::25	3
Discussion/Conclusions::25	3
Application to Risk Assessment Decision-Making::25	3

This is an updated analysis of the Donato et al. 2020 meta-analysis. It has essentially came up with the same result as the Donato et al. 2020 study. The difference was the change of one study. It is published as a Commentary and not as an Original study. The article contains not Methods section other than to say that they used the same methods as per the Donato et al. 2020 paper. It appears that the main reason for publishing this "updated" review was to address the criticisms by another paper (Rana et al. (reference 2 in the paper) about the Donato paper. I find the re-finding of a null result for a paper published in a never-heard of journal with a senior author who is paid for by Monsanto not so transparent or believable!

70.0 — Expert 5

Factor	Factor Rating::1
Methods::25	7
Results::25	7
Discussion/Conclusions::25	7
Application to Risk Assessment Decision-Making::25	7

The authors conducted an updated meta-analysis in response to issues pointed out by Rana et al (2020). In this updated MA the authors calculated RRs for ever-exposure to glyphosate and the highest category of exposure. They reported no association between glyphosate and NHL in either of the exposure categories, however, not ignoring the potential of publication bias. In addition, the authors did observe an association with DLBCL which cannot be ignored. Interestingly, the authors performed a pooled analysis on the cohorts separately from the case-control studies and found no evidence of an association in contrast to the findings of the case-control studies which showed a moderately increased risk.

55.0 Expert 6

Factor	Factor Rating::1
Methods::25	8
Results::25	5
Discussion/Conclusions::25	4
Application to Risk Assessment Decision-Making::25	5

Basically, Boffetta et al. 2021 is not a new meta-analysis but a replication of that by Donato et al. in response to criticism raised by Rana et al. (Environ Res. 2020;191:110140), a co-author in the Zhang et al.2020 meta-analysis. It was published as a commentary on the same journal that hosted the original paper. The differences from the previous meta-analysis consist 1) in replacing the Mc Duffie et al 2001 and the De Roos et al 2003 case-control studies, with the Pahwa et al. 2019 pooled analysis, which included the Mc Duffie et al paper and the three original studies previously pooled in the De Roos et al. 2003 paper; 2) in discussing results also for DLBL and not only NHL and MM. The authors confirmed their reluctance to accept results from case-control studies and discarded as biased the positive finding on DLBCL, while using an uncritical attitude towards the results of opposite sign from the AHS cohort study. Boffetta et al. used mostly ever exposure as the exposure metric, claiming that if a dose-response existed, it should have been reflected in a weaker increase in the exposure estimate for ever exposure; although based on a very limited number of studies, raised unsubstantiated claims of publication bias in case-control studies; combined the top exposure levels from the case-control studies with those from the AHS cohort study. A positive aspect was their concluding remark about the need of exploring specific lymphoma subtypes (which they keep defining as NHL), but they forget to mention that misclassification of the disease might have contribute to further flattening the association with ever exposure. Finally, they raised the issue of exposure to multiple pesticides in agricultural settings, and that the association with ever exposure to glyphosate was weakened when co-exposure to malathion (an organophosphate insecticide) was considered. However, they did not consider that a combined exposure to herbicides and insecticides, might have not implied exposure to the same level (intensity and/or duration) of two classes of agrochemicals. In fact, it might occur in crops occasionally requiring one herbicide treatment vs multiple, sometimes weekly or more, insecticide treatments. Again, although methodologically sound, the interpretation of results is

affected by the sceptical attitude of the two senior co-authors against positive findings in general, to which their experience as consultants for glyphosate manufacturers and other industries in legal trials might have contributed.

67.5 Expert 3

Factor	Factor Rating::1
Methods::25	8
Results::25	7
Discussion/Conclusions::25	9
Application to Risk Assessment Decision-Making::25	3

Boffetta; This MA is an update to that of Denato et al., and in response to a critique. Since a new publication of pooled data from several case-control studies, the Boffetta et al. MA provides the authors with an opportunity to respond to their critics and add to the growing number of MA on this topic. The North American Pooled Project findings support the prior conclusions of the authors, so their new publication is more of the same. Some additional discussion is made about the differences across study design – case-control vs. cohort. This is the stuff of Epidemiology 101 learnings, but seems to have been lost in the vitriol that has characterized the glyphosate and cancer discussion.

37.5 Expert 2

Factor	Factor Rating::1
Methods::25	4
Results::25	4
Discussion/Conclusions::25	4
Application to Risk Assessment Decision-Making::25	3

The paper by Boffetta et al was another analysis published only a year after the same authors published their other one. I am taken aback that the same authors would publish a 2nd review on the same topic within a year, and include a different list of studies in their meta-regression. I have concerns about the use of pooled studies in the meta-regression. They used a pooled meausure of association from the Pahwa et al article, that included a pooled analyses of 5 case control studies. Many of these studies were included in their previous paper. In addition, they retained the other case-control study (Cocco). Their decision to use findings from pooled studies ignores the differing characteristics of studies being pooled. It is remarkable to me that they make the argument to use a random effect models, and to assess heterogeneity but then decide to include two sets of pooled analyses in their meta-regression (Leon paper, and Pahwa paper). It also weakens their approach to look at publication bias because rather than assessing publication bias based on all published studies, they are throwing many individual studies out the window and substituting these multiple estimates of risk from those derived from the pooled analysis. In my view this is a fundamentally flawed approach, and had I reviewed this paper, I would have told the journal to reject the paper. I also think the authors are too strong in their conclusions described in the abstract. They are dismissive of an association with glyphosate and overall NHL despite deriving an estimate of a 15% increase in risk (that is not statistically significant) and the lack of statistically significance is likely due in part to use a reduce number of studies because they relied on pooled analyses. To me, this is a misleading way to

analyze and present the data. This is also evidenced by their statement “Our conclusion is based on a larger databases than previous reviews and meta-analyses”. So while, yes, it may be true that there are more NHL cases than previous meta-analysis, the precision of the summary estimate from the meta-analysis is driven more by the number of studies, than the cases included – and since they dropped individual studies from the meta-analysis, previous meta-analyses were more powerful.

80.0

Expert 4

Factor	Factor Rating::1
Methods::25	8
Results::25	8
Discussion/Conclusions::25	9
Application to Risk Assessment Decision-Making::25	7

I do not see any issues with this update to earlier meta-analysis. The authors are right to point out that none of the meta-analyses can overcome issues in studies being pooled. No matter how the task is approached, the same estimates are reported and the interpretation only differs depending on subjective views of the authors. My sense is that we learnt what we were going to learn about RR for NHL and glyphosate from available data. We can either learn more about uncertainty by following the path of Lash (quantitative bias analysis or better), or leave it to our intuition to fill the gaps. Neither approach will overcome strongly held biases that different people hold in these debates. The reason that I am comfortable making this predictions is because we have in front of us a collection of papers with essentially the same data and quantitative results, and yet these are taken to support varying opinion, none of which entirely unreasonable, even though I would suggest that Zhang et al is the most fanciful.

Result 4.2 *(ID: 5187)*
Question 4.2 *(ID: 4193)*

PANELIST QUESTION: What design of epidemiology or laboratory study would contribute to the question about glyphosate exposure and NHL?

Expert 1

I don't feel that large prospective studies are the way to find the answer here, they take too long and don't always use the optimal controls. I think that case-control studies offer the quickest, most cost-effective, and methodologically sound approach. Of course these studies also have potential forms of bias. Selection of cases and controls will always be a critical factor in assessing the result of any study. The sample size will have to be adequately Powered. Assessment of potential confounding factors will also be critical. It will be important that various types of data are collected asking about various exposures and timelines. More of these studies from various countries would also provide some insight into the potential differences in risk. As highlighted by one of the experts, an analysis of the cohort studies using the commercial applicators who use glyphosate may also provide some insight into the risks based on different types of exposure.

Expert 5

A cohort study with adequate follow-up periods and response rates would contribute greatly to the investigation of the association between glyphosate exposure and NHL. A successful cohort study, though, depends on good follow-up without "gaps" in meaningful periods of exposure and careful monitoring and registering of changes in exposure classification and the population under study.

Expert 6

First, dismiss the NHL definition. Second, explore cancer incidence among the almost 4,000 commercial applicators who used glyphosate in the AHS cohort using the WHO classification of lymphoma, allowing a period of follow-up long enough for achieving the total person-years required to explore risk of the most frequent B-cell lymphoma subtypes (DLBCL, MM, CLL, and follicular lymphoma). With 25 years of follow-up and about 100,000 person years, this might not achieve the statistical power required to test the hypothesis for MM and FL, but it might be feasible in a reasonable time and it might at least provide useful insights. Second, reanalyze the AHS study after modifying the algorithm for calculating cumulative exposure. In fact, De Roos et al 2005 and Andreotti et al. 2018 used intensity quartiles as multipliers of lifetime days of use, which might have underscored most severe exposures by diluting them in a large group of much less exposed though still above the upper quartile cut off. This would allow to substantiate or drop the methodological concerns on the exposure assessment in the AHS study. Third, conduct a new multicentre case-control study on incident lymphoma cases,using the 2008 WHO classification of lymphoma to define the disease entity, large enough to have a substantial number of cases of diffuse large B-cell lymphoma, multiple myeloma, follicular lymphoma, chronic lymphocytic leukaemia, hairy cell leukaemia, marginal zone lymphoma, and T-cell lymphoma (mostly mycosis fungoides). Give the necessary emphasis to collecting occupational histories. In case of a job in settings where glyphosate might have been used

(such as gardening, and weeding in railway tracks, pavements, and roadsides), support questionnaire information with specific occupational modules gathering details on preparation and use of herbicides. Have local agronomist supporting the assessment of agricultural exposures in each centre, and occupational epidemiologists creating exposure metrics more effectively representative of exposure scenarios in the real world than those used thus far. As far as I know, the experimental animal studies conducted thus far provided sufficient evidence of glyphosate carcinogenicity, and, along with the evidence from mechanistic studies, these are actually the reason why the IARC Working Group classified glyphosate as a probable human carcinogen.

Expert 3

We are at the hypothesis testing phase of this issue. We don't need more of the same. We need exposure studies that will address some of the limitations within self-report (are we asking the right questions?). We need validation studies to better understand the crop exposure matrix, and to include application methods within that tool. Do laboratory data really predict human experience? Since our entire risk assessment structure is built upon this, I hope the toxicologists will do better at reassuring the epidemiology community. No answers from this reviewer…just more questions.

Expert 2

I think either case control of cohort studies could contribute to this question. Case control studies need to be interpreted carefully as they are susceptible to participation bias. They are also susceptible to recall bias. They could conceivably collect more detailed information about exposure than cohort studies. Cohort studies are the best study design, however, one needs to be cautious about exposure measurement error. However, if this exposure measurement error is non differential risk estimates derived from a (well designed) cohort study would provide a lower bound of the risk estimate

Expert 4

I am confused by this question. Any well-designed case-referent or cohort study can help, provided that it has far better exposure and outcome assessment than published work. Recruiting in unstudied, highly exposed populations may help. Validation sub-studies of exposure and outcome would be a MUST, including use of this information in analysis to adjust for these sources of bias. I am not qualified to comment on lab studies.

Result 4.3 *(ID: 5188)*
Question 4.3 *(ID: 4194)*

PANELIST QUESTIONS: (A) Do you really think that there is any material difference in summary RR of the five papers that report on case-control studies? (B) Do you think that there is enough evidence to discard case-control studies from consideration until the reason why they systematically differ from cohorts can be explained? (C) Pooled cohort analysis can be improved upon but it seems that it will always dominate epidemiologic evidence. New huge case-control studies are out of the question (in our lifetime). Consequently, is there any value in considering case-control studies at all when their contribution to quantitative synthesis of evidence is dwarfed by a well-conducted pooled cohort analysis? "

Expert 1

(A) The summary estimates range between (RR: 1.03-1.50), with 3 out of 5 studies not significant. There are differences in the levels of heterogeneity between these studies (range: 0.00% - 53.00%), with 3 of the 5 studies having 0.00%. Generally, there isn't a lot of different in the summary estimates and they are very impressive in terms of cancer risk. (B) No I don't think there is enough evidence to discard case-control studies from the equation. I generally find that case-control studies provide higher estimates than cohort studies as occurs in this situation. The differences are in part related to the summary estimates used (OR vs. RR) and the adjusted analysis usually used in the cohort studies, not so much in case-control studies. (C) Slicing the analysis of cohort studies a different way isn't going to make a substantial difference to the pooled estimates. This is particularly true if the cohort selection is biased or not appropriate in order to compare. I think case-control studies always add value to assessing the likelihood of developing cancer when exposed to an environmental factor.

Expert 5

A. I don't think the RR reported in the case-control studies are in essence much different. B.In some of the case-control studies included also here, issues related to study design, section bias or recall bias have not been adequately proven. As long as they are not definitely proven to be related to the risk estimates (in any direction) I don't think that there is enough evidence to discard the information provided by the case-control studies. C. If a "well-conducted" pooled cohort exists then the case-control study would always be of less (but not negligible) value. No evidence should be disregarded unless methodological pitfalls are proven as the aetiology of the outcome observed. In addition, even in a well designed cohort, issues such as data harmonisation due to heterogeneity leads to compromises in sample size and even investigation of confounding. Therefore, everything has to be taken into account when deciding over 'superiority' of study designs.

Expert 6

(A) If results on NHL are considered, the meta-analyses of case-control studies on glyphosate are fairly consistent in showing an excess risk, at least once duration and latency are appropriately considered. The use of top exposure estimates in the Zhang et al. meta-estimates, although criticized for combining different exposure metrics, suggests that using the ever exposure metric might flatten differences and mask a true association. Once leaving apart the NHL definition, positive findings come up for one or the other subtypes. If one or two such subtypes were truly associated, differences in the numbers of each subtype within the NHL cases included in each original study might account for inconsistent findings. The Orsi et al. case-control study, for instance, did not use the WHO classification. Their results were null for NHL, however, they did not explore DLBCL (107 cases), nor follicular lymphoma (50 cases); combined CLL (usually combined with small lymphocytic lymphoma) and hairy cell leukaemia, both mature, indolent forms of B-cell lymphoma, but different in respect to risk factors (Morton L et al. J Natl Cancer Inst Monogr. 2014;48:130-44) ; and did observe a non significant excess risk for MM, which the authors did not comment upon. (B) No, positive findings cannot simply be discarded because it is difficult to explain them. I do have concerns on the results of the AHS study, as I have concerns on those of the case-control studies. A new analysis of the AHS data, and a new very large, multicentre case-control study are mandatory requirements to achieve a final consensus about the epidemiological evidence of glyphosate carcinogenicity. (C) I disagree with the statement that a new very large, multicentre case-control study cannot be completed in the next 10 years. The European Epilymph study took 8 years from getting funded through the first publications. I also argue against the statement that the well conducted cohort studies would dwarf the opposite findings from case-control studies, until the effort in assembling the AHS exposure data to create valuable exposure metrics would match that spent in planning and conducting that study.

Expert 3

A) My short answer “no”. The glyphosate debate, among others, highlights the question around statistical relevance vs. biological relevance. This question pushes the philosophical. This is particularly important as we debate the merits of larger studies – to achieve statistical significance – but with inherently weaker exposure assessment. Is the trade off informative at all? (B) Do you think that there is enough evidence to discard case-control studies from consideration until the reason why they systematically differ from cohorts can be explained? Tempting but not advisable. For now, I think it is important to compare the two approaches…AND to compare what we know about exposure and genotoxicity from the laboratory studies. We want to stay “in our lane” so to speak and evaluate the evidence from our own discipline. There is a reason the EPA and risk-based organizations haven’t concluded that glyphosate is a carcinogen. The other problem is that these case-control studies are really old. When 2,4-D was on the chopping block in the 1990’s there was a lot of work done to better understand the systematic errors – such as that the proxies tended to over report 2,4-D vs. self-respondents. It might be possible to dig into the data for the NCI and Canada studies with specific questions to this effect. Clearly the files are still round since they are being used for the North America Pooled Project. Maybe the question is can we quantify bias so that they can be used? (C) Pooled cohort analysis can be improved upon but it seems that it will always dominate epidemiologic evidence. New huge case-control studies are out of the question (in our lifetime). Consequently, is there any value in considering case-control studies at all when their contribution to quantitative synthesis of

evidence is dwarfed by a well-conducted pooled cohort analysis? " I am a broken record on this topic. Big is not always better. Case-control studies CAN be informative. We want to look at NHL subtypes, so we are going to need these specific cases. Further, how can we identify individuals with internal exposure, and not just passive use. With that being said, maybe we have been asking the wrong question in the past 2 decades. Instead of focusing on a specific pesticide (to target or exonerate), we have an opportunity to work out why NHL seems to be associated with agricultural groups.

Expert 2

I think the numbers of studies are very small. Dismissing the case control studies, in my view implies one needs also to dismiss a meta-analysis approach. This means relying strictly on cohort studies of which there are very few.

Expert 4

A: No, all estimates are about the same. B. No. C. Yes but not through simple averaging of effect estimates. Case-control studies will always be part of evidence that cannot be dismissed unless they are shown to be fatally flawed.

Result 4.4 *(ID: 5189)*
Question 4.4 *(ID: 4195)*

PANELIST QUESTION: To what extent do you think the use of biomonitoring data can be used to support exposure characterization in epidemiological studies in an effort to advance our understanding of whether glyphosate increases the risk of NHL?

Expert 1

I think anything that may add value in assessing the exposure to glyphosate in the risk of NHL would be of value. I don't know if it will add significantly in terms of exposure assessment. For this question, for me it isn't so much about amount of exposure, it is really exposed vs. never exposed.

Expert 5

It is always the question whether there is a difference between exposure to the substance and the dose of this substance that is absorbed. In fact, the dose may be more important in studying causal relationships rather than exposure. A correlation of exposure and dose would potentially help to understand whether glyphosate is linked to NHL and if so at what level.

Expert 6

Glyphosate and its metabolites are strongly polar. Unchanged glyphosate would be a more suitable biomarker (Faniband MH et al. Int J Hyg Environ Health 2021;231:113657). However, its half-life following oral ingestion would be 6-9 hours in the first rapid phase and up to 33 hours in the slow phase. Another study of dermal and inhalation exposure showed that background levels were restored 24 hours after exposure (Pierce JS et al. Inhal Toxicol 2020; 32(8): 354-67), which makes it useful to access current exposure, but useless as a biomarker of past exposure. Relying on biomonitoring data would have sense if part of the initial phase of a follow up study to characterize glyphosate exposure, and repeated along the time, to corroborate the exposure assessment in relation to different working procedures. However, inter-individual variation is expected due to gene polymorphisms in phase I and phase II metabolic genes, personal attitudes towards workplace exposures, personal hygiene, and environmental circumstances (temperature, humidity, strength and direction of wind) during spraying. A more careful exposure assessment and use of the exposure information in the analysis would serve better the scope.

Expert 3

I think biomonitoring can add significantly to this issue, and others of pesticide exposure. As a short-lived chemical, biomonitoring isn’t the right choice to determine exposure retrospectively. However, information learned from biomonitoring can be extremely valuable. Most studies that form the basis of the glyphosate discussion around glyphosate and NHL are based upon questionnaire data, with “dose levels” based upon number of applications and years of use. These studies highlight our constraints as researchers to capture relevant exposure and heterogeneity across participants. There is an opportunity

> going forward to validate assumptions with biomonitoring. The crop exposure matrix assumes exposure based upon growing practices. Using (i.e. applying) a pesticide is equated with homogenous exposure. Small and focused studies incorporating multiple exposure methods (such as environmental samples, crops, proximity, biological monitoring, silicone wrist bands), particularly if timed with an application (capturing behaviors and weather conditions) would begin to tell us about the risk (and direction) of bias in existing exposure approaches.

Expert 2

> It could be used a tool to validate exposure measures in cohort studies. I think it is impractical to conduct a large cohort with regular biomonitoring of all. But, biomonitoring could potentially be used to validate job exposure matrices, or assess changes in exposure over time in similarily held jobs.

Expert 4

> It can be used but short half-lives make it near-impossible in retrospective studies. However, questionnaires can be validated by biomarkers versus current exposures. NCI and NIOSH did this work (in relation to AHS) but I do not think it was ever published.

Result 4.5 *(ID: 5190)*
Question 4.5 *(ID: 4196)*

PANELIST QUESTION: Previous work from the InterLymph Consortium has provided evidence of the substantial heterogeneity between lymphoma subtypes (Morton L. J Natl Cancer Inst Monogr. 2014;48:130-44). Still, most epidemiologists keep referring to the non-Hodgkin lymphoma disease entity, restricting or extending the list of subtypes included in it, which makes comparisons especially difficult. The same happens with leukemia studies, which most frequently include chronic lymphocytic leukemia, classified as a mature B-cell lymphoma by the REAL classification in 1994. In evaluating etiological factors, would you suggest keeping using the NHL denomination for consistency with the existing literature, or would you rather call for more studies focussing on specific B-cell and T-cell subtypes?

Expert 1

> This is really an issue related to adequate sample size and statistical Power. Obviously, larger numbers of NHL cases is great in defining risk in terms of exposure to glyphosate. However, conducting sub-group analysis on much smaller groups of lymphoma sub-types may not be methodologically sound or appropriate. For me, the focus should be on NHL. I think it is always worthwhile in collecting information on NHL sub-types for analysis, but it will always be limited.

Expert 5

> I think that there is not enough research yet to fully understand the differences among the subtypes of NHL. If we want to know, for example, why B-cell subtype is more common, therefore assessing etiological factors, then more studies are needed on this topic. This could help also with the personalised design of treatment.

Expert 6

> As stated above, I would definitely call for new studies investigating the whole spectre of lymphomas by individual subtype. When aiming to compare results with previous reports, subtypes might be subsequently combined according the NHL definition. However, special care is necessary to the confusing combination of subtypes in the NHL category in the previous literature, sometimes including CLL, and oftentimes excluding it, sometimes excluding T-cell lymphomas (B-NHL) and most og the times including them.

Expert 3

> Lumping or splitting? There are several rationales to “lump” outcomes like NHL. A) I’ve had

to do it because of lack of the data on subtypes, particularly when using cases diagnosed in the past. B) It is similarly tempting because it gives the research more statistical power. The number of exposed cases becomes very small as one evaluates subtypes. C) Lastly, evaluating all the subtypes then opens the study results to issues of multiple comparisons. In general, we see also the "lumpers" combine outcomes or exposures in systematic reviews. It's an exercise that is flooding the literature and allows authors to promote a hypothesis under the umbrella of being systematic. However, is it scientific? Not really. The rationales to split are better detailed by Morton et al. As readers and critics, we must be willing to give the authors a break for having imprecise data.

Expert 2

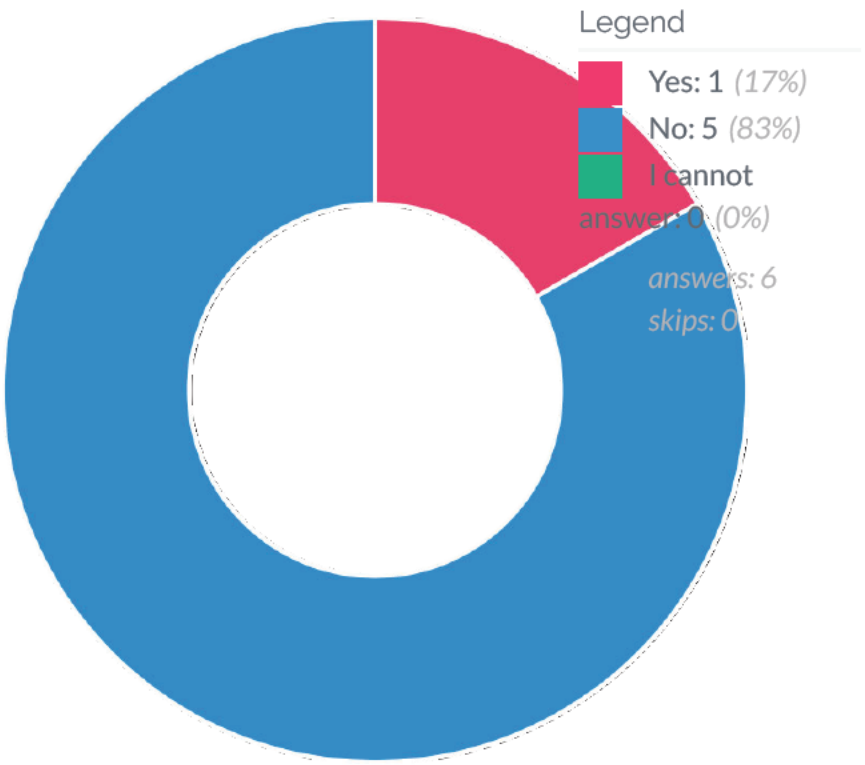
Ideally, would like this decision to be driven by knowledge of mechanistic pathways (recognizing this is a challenge). I do see value on studies focussing on specific B -cell and T-cell subtypes, but worry this constrains several cohorts due to inadequate sample size which may make several risk estimates unstable.

Expert 4

Until we know more about mechanism of these outcomes, pooling them cannot be excluded. Reporting on subtypes is always a good idea but outcome misclassification becomes a bigger issues as you look at ever more exotic and rare diagnoses. There is not simple fix, but validation studies are a must if we are to seriously consider sub-types. This is not a new issue but comes up for every outcome, e.g. asthma, lung cancer, stroke, autism (all are heterogeneous entities).

Result 4.6 *(ID: 5191)*
Question 4.6 *(ID: 4197)*

PANELIST QUESTION: Is the pooled analysis of Leon et al. (2019) superior to all the other systematic review and meta-analyses conducted as part of this assessment (please explain your answer below)?



Answer Explanations

No — Expert 1

This analysis has its limitations, the pooling of different cohort studies is much more the mixing of apples and oranges than the meta-analysis of the case-control studies. Of course it all looks good when you have a combined large sample size and the ability to adjust for confounding factors, but there was no quality assessment undertaken, which could have been done as it would be for individual cohort studies. There was substantial heterogeneity reported. I believe that this analysis could be improved upon.

Yes — Expert 5

Through cohort-specific exclusions the resulting study population included in the combined study of meta-analysis was homogeneous. The harmonisation of the exposure data allowed for the increase of the sample size and therefore the ability to detect associations if present. However, the heterogeneity was, relatively, high and perhaps due to some innate differences between the three cohorts. An advantage of the study is the large number of exposed cases and the inclusion of prospective cohorts which minimised a differential recall bias. However, differences were observed in the exposure information used in the meta-analysis. Only one of the three cohorts used actual exposure information based on self-report. In addition, the three cohorts differed in other variables, such as age

and gender while each of the cohorts adjusted for different factors, depending on the availability of the information in each cohort. Therefore, it is uncertain whether the correct adjustments were made to account for the observed differences among the cohorts. Although the were innate differences among the cohorts, a relatively good harmonisation of the data was described. The authors mention that a future in-depth analysis will improve the specificity of the exposure information by adding parameters such as duration, frequency, and intensity of use.

No

Expert 6

The value of the Leon et al. pooled analysis is diminished by the drawbacks of the cohort studies included in it: the AHS cohort was supported by a very good exposure assessment, but a criticisable use of that information, the French and the Norwegian cohort studies used crop exposure matrices (CEMs). CEMs are sensitive but not much specific when relying on dichotomous categories. The authors reported about harmonisation of the exposure data, but I have concerns about whether this was possible. Besides, only the dichotomous exposure variable was analysed, and there was no attempt of investigating exposure response trends.

No

Expert 3

Like all projects, the pooled analysis of Leon et al. has strengths and limitation. First, since it is based on 3 cohorts, it has the advantage of large sample size and reduction of the selection and recall biases that characterize the case-control studies (that form the basis of the other MA). The CNAP and AGRICAN cohorts bring other countries into the mix, increasing the generalizability. As pooled data the results are not limited by the publication transparency of each study. However, since the two European cohorts use a different approach to exposure assessment, I think it is important for the authors to present the results of the 3 studies individually, since this hasn't been part of individual publications. The nice thing about a MA is that the approach allows one to see if one study has a greater impact/weight than another. With pooled data, this is not transparent. In this case, the association of glyphosate and DLBCL in the AGRICOH analyses seems to be restricted to the CNAP cohort and not elevated in the French (AGRICAN) or US (AHS) cohorts. Is this a random observation or are the exposures higher in the CNAP cohort? The lack of an exposure-response within the AHS suggests the former (a random observation). A strength and limitation that these studies did not focus only on glyphosate. With multiple pesticides comes multiple comparisons. The findings of this pooled analysis should be discussed in this context. The paper evaluated 14 chemical groups and 33 active ingredients with 5 outcomes. This publication was not specific to glyphosate in its focus of analyses or discussion. In this context, random associations would be expected.

No

Expert 2

I think it is one of the stronger studies, but also reminded that it does rely on a crop exposure measure - so no direct measure of individual level exposure. The characterization of exposure is better in some of the case-control studies, and I don't think that should be entirely dismissed (even acknowledging there may be potential biases in recall and participation).

No

Expert 4

I already explained that no single one of these analyses is an absolute clear winner.