ELSEVIER

LETTER TO THE EDITOR

Regarding "Meta-analysis and Causal Inference: A Case Study of Benzene and Non-Hodgkin Lymphoma": An Incomplete Analysis

Dear Editors:

This Journal recently published an article by Douglas Weed (1) dedicated to criticizing our meta-analysis, which identified evidence of a causal association between non-Hodgkin lymphoma (NHL) and benzene and between NHL and petroleum refinery work (2). The whole of Weed's article is devoted to criticizing our paper, although under the pretext of using our work as a "case study." Although we appreciate Weed's interest in our paper, his major points are either incomplete or incorrect.

Weed suggests that meta-analysis can only play three roles in evaluating causal inference: increased precision, evaluations of heterogeneity, and publication bias. But meta-analysis involves so much more. For example, our meta-analysis involved an extensive literature search, evaluation of hundreds of studies, and quantitative or qualitative evaluations of precision, magnitude of effect, heterogeneity, biologic plausibility, healthy worker effect, confounding, multiple sensitivity analyses, publication bias, review of comparable work…and so on. Many other meta-analyses have quantitatively assessed dose-response (3), chance (e.g., *p* values), latency (4), effect modification (5), biologic plausibility (6), temporality (7), outcome misclassification (8), exposure misclassification (9), diagnostic bias (10), recall bias (11), selection bias (12, 13), confounding (14), funding source (15), and many of the other elements of causal inference. We don't claim that every meta-analysis does all things. However, the suggestion that the process of meta-analysis is limited to the three things described by Weed is incorrect and even inconsistent with some of his previous statements (16).

Weed states that because of a "lack of consistency (by which he means statistically significant heterogeneity), weak associations, no evidence of dose-response, no effort to provide an assessment of biological plausibility, and no new epidemiological evidence" we should not have concluded that the elevated relative risks we identified provide "further evidence that benzene exposure causes NHL." However, his review of each of these factors is wrong or incomplete. For example, it is implied that because our meta-analyses of NHL and refinery work involved statistically significant tests for heterogeneity, the data are inconsistent and "combining them was not warranted". However, numerous authors have warned against over-interpreting these tests (14, 17). In Weed's own words from a previous article, "statistical heterogeneity…need not preclude a conclusion of consistency" (16). In our meta-analysis, as in every meta-analysis of observational data, the studies are a mix of study designs, populations, time periods, lengths of follow-up, exposure and outcome assessment methods, statistics, researchers, funding sources, and so on. Given this broad mix, one should expect statistical heterogeneity and maybe even be surprised when it isn't there. If we arbitrarily decide that only studies meeting imperfect statistical criteria can be summarized, many true effects will be missed. A truly thoughtful and complete interpretation of consistency should acknowledge the fact that different epidemiologic studies really are different. As an aside, although our refinery work–NHL results involve statistical heterogeneity, none of our benzene-NHL results did. As such, by Weed's standards, all of our benzene-NHL results meet his key causal criteria of consistency.

In his Table 2, Weed highlights the point that clear dose-response relationships are not seen in the individual studies included in our meta-analysis. But, the largest study with the highest average exposures, Hayes et al. (18), did report statistically significant dose-response trends for average exposure, exposure duration, and cumulative exposure (all p trends < 0.05). Weed says two of the studies have dose-response trends are "U-shaped" and one is in the opposite direction of expectation. However, this is not true when one takes into account the very wide confidence intervals (CIs) in almost all strata (e.g., a 95% CI of 0.0–10.7). Regardless, a statistically significant trend in dose-response is not an absolute requirement for causation. *Modern Epidemiology* (cited several times in Weed's article) provides several examples that "imply that the existence of a monotonic association is neither necessary nor sufficient for a causal relation" (19). Weed's table provides another classic example: every study had a very small number of cases. In fact, 75% of the exposure strata had four or fewer cases. Using the National Cancer Institute's Power Program, it can be seen that all of these studies had very poor statistical power to identify dose-response trends (20).

Weed asserts that the relative risks (RRs) we reported are low and therefore could be due to confounding. First, we don't agree that an RR of 2.26 ($p = 0.002$), which we estimated for high exposures, is low. Second, although low RRs can be due to confounding, they can also represent true effects. Third, a thoughtful evaluation of confounding would provide some idea of what the confounder might be and provide at least some evidence that it is strongly enough related to benzene and to NHL to cause the effects

© 2011 Elsevier Inc. All rights reserved.
360 Park Avenue South, New York, NY 10010

1047-2797/$ - see front matter
doi:10.1016/j.annepidem.2010.06.004

AEP Vol. 21, No. 1
January 2011: 67–69

identified. If Weed has evidence that the confounder is toluene, styrene, or another solvent, then refineries may want to reduce their workers' exposures to these agents. If it's an unknown mystery confounder, then the likelihood that an unknown mystery confounder could cause an RR greater than 2.0 should be discussed. Although Weed did not do any of this, we did, and found it unlikely.

Weed implies that we provide only "a single unreferenced sentence" on biologic plausibility. However, most of the fifth paragraph of the discussion section in our paper and much of our companion article (cited several times by Weed) discusses biological plausibility (21). Weed also states that we provide no new evidence. However, our meta-analysis was the first to quantitatively assess high exposure groups, the first to include analyses of both benzene exposure and refinery work, and the first to quantitatively evaluate the healthy worker effect. Since none of this was published before, all of it is new evidence.

Two other points: first, Weed suggests that the healthy worker effect had little impact on our results, but these adjustments increased some excess RRs by 70% to 100%. Second, in his Table 1, Weed implies that our conclusion is wrong because it is inconsistent with the conclusions of some other reviews. Is this now a popularity contest, based simply on a tally of votes? Did all of these reviews quantitatively assess high exposures, the healthy worker effect, and all of the other aspects of causal inference we assessed? No, they did not. In fact, three provide fewer than four sentences on the subject of benzene and NHL (22–24), and four don't even include the word "benzene" in the review at all (25–28).

In conclusion, all of Weed's major points regarding our meta-analysis involve incomplete or inaccurate evaluations of the issues. Although causal inference for exposures resulting in high relative risks does not require meta-analysis, the meta-analysis process provides a vital tool as we seek to make inference about exposures which create smaller increases in risks. We have been teaching this for many years in our course at the University of California, Berkeley, School of Public Health, titled "Causal inference and meta-analysis."

Finally, we note that Weed's acknowledgement section states that financial support came from an innocent-sounding organization called the CONservation of Clean Air and Water in Europe (CONCAWE). Although not stated in Weed's article, CONCAWE was established by refinery companies and "full membership is currently open to companies that own crude oil refining capacity" (29). A letter to the editor in this Journal in 2006 aptly made the point that, "Too often, data reanalyses and reinterpretations are attempts to manufacture uncertainty, a strategy used by polluters and manufacturers of dangerous products to avoid or delay regulation and civil liability" (30). Although we believe that we have addressed all points of substance in Weed's criticisms, we also believe that readers need to be aware of funding sources when they see literature published, which suggests that there is no causal evidence, just an association.

Craig Steinmaus, MD*
Allan H. Smith, PhD
Martyn T. Smith, PhD
School of Public Health
University of California
Berkeley, CA 94720

*Address correspondence to: Craig Steinmaus
50 University Hall, University of California
Berkeley, CA, 94720-7360. Tel: (510) 504-5395
E-mail: craigs@berkeley.edu

REFERENCES

1. Weed DL. Meta-analysis and causal inference: a case study of benzene and non-Hodgkin lymphoma. Ann Epidemiol. 2010;20:347–355.
2. Steinmaus C, Smith AH, Jones RM, Smith MT. Meta-analysis of benzene exposure and non-Hodgkin lymphoma: biases could mask an important association. Occup Environ Med. 2008;65:371–378.
3. Berlin JA, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. Epidemiology. 1993;4:218–228.
4. Goodman M, Morgan RW, Ray R, Malloy CD, Zhao K. Cancer in asbestos-exposed occupational cohorts: a meta-analysis. Cancer Causes Control. 1999;10:453–465.
5. Miller MD, Marty MA, Broadwin R, Johnson KC, Salmon AG, Winder B, et al. The association between exposure to environmental tobacco smoke and breast cancer: a review by the California Environmental Protection Agency. Prev Med. 2007;44:93–106.
6. Weed DL. Meta-analysis under the microscope. J Natl Cancer Inst. 1997; 89:904–905.
7. Chak E, Rutherford GW, Steinmaus C. The role of breast-feeding in the prevention of Helicobacter pylori infection: a systematic review. Clin Infect Dis. 2009;48:430–437.
8. Juni P, Altman D, Egger M. Assessing the quality of randomised controlled trials. In: Egger M, Davey Smith G, Altman D, eds., Systemic reviews in health care: Meta-analysis in context. London: BMJ Publishing Group; 2001:100–101.
9. Zhuo H, Smith AH, Steinmaus C. Selenium and lung cancer: a quantitative analysis of heterogeneity in the current epidemiological literature. Cancer Epidemiol Biomarkers Prev. 2004;13:771778.
10. Zhang L, Steinmaus C, Eastmond DA, Xin XK, Smith MT. Formaldehyde exposure and leukemia: a new meta-analysis and potential mechanisms. Mutat Res. 2009;681:150–168.
11. Steinmaus CM, Nunez S, Smith AH. Diet and bladder cancer: a meta-analysis of six dietary variables. Am J Epidemiol. 2000;151:693–702.
12. Gillum LA, Mamidipudi SK, Johnston SC. Ischemic stroke risk with oral contraceptives: A meta-analysis. JAMA. 2000;284:72–78.
13. Guha N, Merletti F, Steenland NK, Altieri A, Cogliano V, Straif K. Lung cancer risk in painters: a meta-analysis. Environ Health Perspect. 2010; 118:303–312.
14. Greenland S. Meta-analysis. In: Rothman K, Greenland S, eds. Modern epidemiology. 2nd ed. Philadelphia: Lippincott Raven; 1998:643–673.

15. Barnes DE, Bero LA. Why review articles on the health effects of passive smoking reach different conclusions. JAMA. 1998;279:1566–1570.
16. Weed DL. Interpreting epidemiological evidence: how meta-analysis and causal inference methods are related. Int J Epidemiol. 2000;29:387–390.
17. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327:557–560.
18. Hayes RB, Yin SN, Dosemeci M, Li GL, Wacholder S, Travis LB, et al. Benzene and the dose-related incidence of hematologic neoplasms in China. Chinese Academy of Preventive Medicine–National Cancer Institute Benzene Study Group. J Natl Cancer Inst. 1997;89:1065–1071.
19. Rothman K, Greenland S. Causation and causal inference. In: Rothman K, Greenland S, eds. Modern epidemiology. 2nd ed. Philadelphia: Lippincott Raven; 1998:7–28.
20. Garcia-Closas M, Lubin JH. Power and sample size calculations in case-control studies of gene-environment interactions: comments on different approaches. Am J Epidemiol. 1999;149:689–692.
21. Smith MT, Jones RM, Smith AH. Benzene exposure and risk of non-Hodgkin lymphoma. Cancer Epidemiol Biomarkers Prev. 2007;16:385–391.
22. Ekstrom-Smedby K. Epidemiology and etiology of non-Hodgkin lymphoma–a review. Acta Oncol. 2006;45:258–271.
23. Grulich AE, Vajdic CM. The epidemiology of non–Hodgkin lymphoma. Pathology. 2005;37:409–419.
24. Muller AM, Ihorst G, Mertelsmann R, Engelhardt M. Epidemiology of non-Hodgkin's lymphoma (NHL): trends, geographic distribution, and etiology. Ann Hematol. 2005;84:1–12.
25. Baris D, Zahm SH. Epidemiology of lymphomas. Curr Opin Oncol. 2000; 12:383–394.
26. Fisher SG, Fisher RI. The epidemiology of non-Hodgkin's lymphoma. Oncogene. 2004;23:6524–6534.
27. Rodriguez-Abreu D, Bordoni A, Zucca E. Epidemiology of hematological malignancies. Ann Oncol. 2007;18(Suppl. 1):i3–i8.
28. Swerdlow AJ. Epidemiology of Hodgkin's disease and non-Hodgkin's lymphoma. Eur J Nucl Med Mol Imaging. 2003;30(Suppl. 1):S3–12.
29. CONCAWE. CONCAWE Membership, 2003. Available at: http://www.concawe.be/Content/Default.asp?PageID=9. Accessed May 5, 2010.
30. Michaels D. Regarding "Phenylpropanolamine and hemorrhagic stroke in the hemorrhagic stroke project": mercenary epidemiology–data reanalysis and reinterpretation for sponsors with financial interest in the outcome. Ann Epidemiol. 2006;16:583–585, author reply 586.