REVIEWERS' BIAS AGAINST THE NULL HYPOTHESIS: THE REPRODUCTIVE HAZARDS OF BINGE DRINKING

Gideon Koren, Andrea Fernandes

The Ivey Chair in Molecular Toxicology, Department of Medicine, University of Western Ontario, London, Canada; Motherisk Program, The Hospital for Sick Children, Toronto, Canada

ABSTRACT

We examined whether scientific reviewers exhibit bias in scoring a simulated "positive" study (i.e. showing adverse fetal effects) as compared to a simulated "negative" study on the fetal effects of binge drinking. The reviewers of the "negative" study tended to reject it more commonly, to give it lower scores, and there was significantly more variability from the median in their scores. Scientific journals should make an effort to eliminate this source of bias against negative results.

Bias against the null hypothesis is defined as preferential publication of studies showing statistically significant results ("positive") as compared to those not finding differences between or among study groups ("negative"). In the context of maternal-fetal toxicology, this means that studies showing a drug or a chemical to cause fetal risk may be more likely to be published, presented in meetings and cited by the media, than "negative" studies.[1] Such bias causes serious threat to the interpretation of risk-safety of drugs and chemicals in pregnancy. If positive studies are more likely to be published, then safe drugs may be presented as unsafe, leading women not to be treated appropriately, or even consider termination of pregnancy when exposed.[2]

While potential bias by medical journals against negative results has been suggested repeatedly[3], providing proof of the source of such bias has not been possible as it necessitates access to confidential editorial source data. In an attempt to explore the role of the editorial process, we simulated "real time" editorial submission of two identical papers on binge drinking in pregnancy, one with adverse fetal outcomes and the other with no adverse outcomes. In reality, both "positive" and "negative" studies on this topic have been published.[4] The objective of the study was to compare reviewers' ratings and acceptance of these identical "negative" and "positive" studies.

METHODS

After approval by the Research Ethics Board of the University of Western Ontario, two anonymous versions of the same original scientific paper were sent to independent reviewers by the journal Fetal Alcohol Research (www.cjcp.ca). Invited reviewers were experts in the area of fetal alcohol syndrome, maternal fetal pharmacology and developmental pharmacology. Using the usual editorial process they were asked to rank the importance of the topic from zero to 10, to recommend rejection, resubmission with revisions or acceptance "as is." The reviewers were not aware of the simulation nature of this experiment or its objectives, and believed it was a genuine original paper.

The two versions of the paper compared a group of non addicted pregnant women, exposed to binge drinking in the first trimester of pregnancy, and a comparison group of non exposed women. The neurobehavioral tests conducted on the children included cognitive, language and behavioral tasks. The two versions differed in the results, showing "positive" or "negative" effects on the offspring (Table 1). The Discussion of each paper differed to reflect the different results. We compared the reviewers' evaluation of the two papers using the Mann Whitney U test.

RESULTS

Sixty-two reviewers were approached and randomly assigned to the "positive" and "negative" studies. 22 reviewers returned their evaluations (n=11 for each paper). Overall, all 11 reviewers of the "positive" paper recommended publication with or without revision. In contrast, 2 reviewers of the "negative" paper recommended rejection. The median reviewers' score of the "positive" paper (7.5) tended to be higher than that of the "negative" paper (6.5) (p=0.15). There was significantly more variability of scores from the median in the "negative" paper as compared to the "positive" one (median 2.5 vs. 0.5, p=0.023).

DISCUSSION

The bias against "negative" results leads to the "file drawer" phenomenon, whereby "negative" studies are less likely to be published[3] or reported by the media.[5] In the area of maternal-fetal toxicology another risk is added to this "natural" bias against negative results: studies of drugs of abuse and alcohol use by mothers, showing no adverse effects – are less likely to be accepted or presented[6] suggesting that either reviewers, editors, or both are unwilling to support publication of "good news" about "evil drugs".

To examine the role of the scientific review, we simulated the routine process of scientific review practiced in medical journals, employing experts in the fields of maternal-fetal toxicology and the fetal effects of alcohol. The same paper was offered once with strong "positive" results, outlining major adverse neurobehavioral effects, and once not finding such effects.

The "positive" studies received high marks and these assessments were remarkably homogenous. The "negative" study, on the other hand, received both high as well as very low marks, with significantly more variability among the evaluations, and with two rejections.

These results support the hypothesis that reports of favorable fetal outcome after binge drinking cause some reviewers to oppose publication on the basis of "socially irresponsible science." Any effort should be made to mitigate such bias, thus to ensure that timely publication of "negative" results can effectively inform decision making by health professionals in highly sensitive areas.

TABLE 1 Characteristics of the Positive and Negative Papers

	"Positive"	"Negative"
Women	binge, but non alcoholic women	binge, but non alcoholic women
Sample Size	40 binge exposed, 40 healthy	40 binge exposed, 40 healthy
Outcome Measures	IQ, language, behaviors	IQ, language, behaviors
Results	highly adverse statistical and clinical effects on cognition, language and behavior	No differences between binge exposed and controls (values set to be almost identical)
Discussion	"This study confirms the fetal risk of binge drinking"	"This study fails to show fetal risk of binge drinking in non-alcoholic women"

Corresponding Author: **gkoren@sickkids.ca**

REFERENCES

1. Koren G. Bias against the null hypothesis in maternal-fetal pharmacology and toxicology. Clin Pharmacol Ther 1997; 62: 1-5.
2. Koren G, Pastuszak A, Ito S. Drugs in Pregnancy. N Engl J Med 1998; 338: 1128-37.
3. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet 1991; 337: 867-72.
4. Hendersen J, Kesmodel U, Gray R. Systematic review of the fetal effects of prenatal binge-drinking. J Epidemiol Community Health 2007;61:1069-73
5. Koren G, Klein N. Bias against negative studies in newspaper reports of medical research. JAMA 1991; 266: 1824-6.
6. Koren G, Graham K, Shear H, Einarson T. Bias against the null hypothesis: the reproductive hazards of cocaine. Lancet 1989; 2: 1440-2.