Published by Oxford University Press on behalf of the International Epidemiological Association

International Journal of Epidemiology 2007;**36**:944–945
doi:10.1093/ije/dym174

Commentary: On 'Quality in epidemiological research: should we be submitting papers before we have the results and submitting more hypothesis generating research?'

Sander Greenland

Accepted 7 August 2007

Lawlor's excellent proposals to mitigate publication bias[1] complements other ideas, such as (truly) blind review of completed studies by blanking-out study results. Nonetheless, despite my general agreement with Lawlor, I think certain conventional presumptions in the editorial are in error and have blinded discussants to important sources of publication bias.

Most authors write as if publication bias is chiefly about tendency to not publish null results. Studies linking products with hazards provide examples to the contrary, as in the Vioxx debacle (one that happened to come to prominence).[2] Typically, the sponsor or conductor is the affected industry, with the power to either suspend study activity or else to press for further data collection or analysis before publication. A strategy that extends a study or analysis only when a particular type of result (e.g. a positive finding) emerges at first is biased against reporting that type of result. Of course, some researchers follow similarly biased strategies when their initial findings contravene strongly held beliefs, whether those beliefs are in the null or an alternative. In any event, it is time to recognize that publication bias has many forms, some away from the null, others toward the null.

For similar reasons, it is time to stop using the term 'hypothesis generating' for studies that screen multiple associations, because the term conceals the true nature and potential bias of such studies. Even in the most speculative fishing expedition, the fact that an association was examined means a hypothesis about it was generated and approved for examination at the outset. These studies do not create hypotheses any more than screening programs create people. What they do is screen a selected set of hypotheses,[3] usually according to dubious statistical criteria (such as multiple-testing procedures that aggravate publication bias by reducing alpha levels at the expense of power).

To understand why my complaint concerns more than terminology, consider that the manner in which the hypotheses are selected for examination opens an avenue for publication bias, in the broad sense of bias in what we see in the published literature. As with patient selection for medical screening, there can be considerable (and faulty) judgement in what hypotheses get selected for examination and how they are examined, beginning with data-collection decisions and continuing through the analysis.[4] Certain questions may be avoided for fear of the answer, but most often are ignored due to certainty about the answer. We all do the latter; for example, we fail to ask a person's favourite movies in the firm belief that such information would provide no useful health-risk information beyond demographics, diet and exercise.

Essential as pre-data hypothesis selection may be, these decisions are based on political, academic and personal judgements about pursuit of certain questions, and thus subject to the biases prevalent in their context. The fact that hypotheses are generated and screened before a study is done needs to be acknowledged, instead of concealed by confused terminology. Consider the Los Angeles Heart Study.[5] Begun in the 1940s, it neglected to obtain baseline smoking information. In fairness to the authors, this was at a time when, in the Anglo-American world, a smoking effect on lung cancer was still regarded with scepticism[6]; an effect on heart disease would probably have been lumped by some in the same category as flying saucers. Nonetheless, the narrow range of data collected helped ensure that this study would play only a minor role in the landmark discoveries of cardiovascular risk factors during subsequent decades, and reduced the pool of early evidence regarding smoking and heart disease.

As for biased publication of collected data, submission and review innovations cannot substitute for mandatory study registration at the outset of data collection, for otherwise the very existence of data could be missed. Ethical screening bodies such as Institutional Review Boards and Research Ethics Committees may hold the key to preventive measures. Via ethical guidelines, such boards exert extensive control over study approval. Those guidelines can be used to argue that studies collecting data on human subjects should be required to register their approved proposal or protocol in a searchable public database.

Mandatory registration is not enough to prevent publication bias, however; study results must be available to the public. Rationales offered for not publishing results have included protection of proprietary information or patient confidentiality (the latter has been invoked when the study has few subjects or few cases). If the data holder is unresponsive to inquiries, other

Departments of Epidemiology and Statistics, University of California, Los Angeles, USA. E-mail: lesdomes@ucla.edu

parties will be left without necessary information if they cannot afford costly discovery litigation. Although mandating publication of results might address this problem, such mandates could worsen bias if they only required publication upon hazard detection, and would involve many practical problems (especially if the study was of little interest to journals). An alternative would require final study reports and dissertations generated from a study be deposited in a searchable public database. Deposited reports could then be supplemented by investigator annotations, along with references to published articles as those appear.

References

[1] Lawlor DA. Quality in epidemiological research: should we be submitting papers before we have the results and submitting more hypothesis generating research? *Int J Epidemiol* 2007;**36**(5).

[2] Curfman GD, Morissey S, Drazen JM. Expression of concern reaffirmed. *N Eng J Med* 2006;**354:**1193.

[3] Rothman KJ, Greenland S. *Modern Epidemiology*. 2nd edn, Philadelphia: Lippincott, 1998; p. 78.

[4] Phillips CV. Publication bias in situ. *BMC Med Res Methodol* 2004;**4:**20. Published online 2004 August 5. doi: 10.1186/1471-2288-4-20.

[5] Chapman JM, Goerke LS, Dixon W, Loveland DB, Phillips E. Measuring the risk of coronary heart disease in adult population groups: IV. The clinical status of a population group in Los Angeles under observation for two to three years. *Am J Public Health Nations Health* 1957;**47**(4 Pt 2):33–42.

[6] Wynder EL. Tobacco and health: a review of the history and suggestions for public health policy. *Public Health Reports* 1988;**103:**8–18.