

The Registration of Observational Studies— When Metaphors Go Bad

The Editors

Should epidemiologists be required to pre-register their studies and hypotheses? Such a proposal was made at a recent workshop. Lest you think this is just an odd idea that will go away, 2 major clinical journals have already endorsed it.^{1,2}

The workshop report³ itself has not been made public before now. We have been given permission by the workshop conveners (the European Center for Ecotoxicology and Toxicology of Chemicals) to post the report on our Web site (<http://links.lww.com/EDE/A415>). With this report as a starting point, we have invited five epidemiologists (including one who participated in the workshop) to respond. Their commentaries follow this editorial.^{4–8}

The workshop report states that journal editors play “a critical role in promoting the process by ‘selling’ this to potential authors or readers.” As editors, we won’t be selling this proposal—but we do have some thoughts about it.

There are conceivable benefits from the registration of protocols and hypotheses of observational studies. For one, the information that a hypothesis has previously been explored (even if the results were never published) might help other investigators decide whether the hypothesis is worth further study. Another possible benefit is that researchers working on publicly-accessible data could be informed that others were doing similar analyses.⁹ Researchers whose results might damage the business interests of their employers would be less able to hide those results (an argument that led to the registration of clinical trials).

That said, we do not agree that registration would provide the benefits claimed by the workshop. Those purported benefits include a reduction in “publication bias,” improved research “transparency” and fulfillment of “ethical obligations.” We address these points in turn.

Reduction in “Publication Bias”

There is a common perception that negative studies fall into publication limbo (or are never written up), whereas positive studies (including false-positives) get a free pass to publication. As supporting evidence, the workshop report refers repeatedly to “observational studies [that] are not published, either because they are not submitted for publication or because they were rejected by journals.” Such studies exist—but are they evidence of a problem? If there is any validity to the self-critical judgment of researchers and the process of peer reviews of their peers, then we should expect some portion of study results to remain unpublished. Some findings are fatally flawed, while others are of little importance. The mere existence of unpublished analyses is not an argument that unpublished analyses are as valid as those published.

Furthermore, the timing of a research hypothesis (whether before or after data collection) is irrelevant to its validity. In their commentaries, Charlie Poole⁷ and Tim Lash⁵ both discuss the intellectual history of this mistaken perception.

A more empirically-based concern is that hypotheses stated in advance may be less prone to false-positive findings. Put another way, findings supported by previous evidence

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Copyright © 2010 by Lippincott Williams & Wilkins
ISSN: 1044-3983/10/2105-0607
DOI: 10.1097/EDE.0b013e3181eafbcf

are more likely to be proved valid than findings having no previous evidence. If this is true, we shouldn't be surprised. Prior evidence should not determine whether results are published.

Improved "Transparency"

"Transparency" is an unassailable scientific virtue. But what does transparency mean in this context? Does transparency require the prior registration of the "universe" of observational data from all sources?^{3,p2} Unlike clinical trials, which are rigidly designed to address narrow questions, observational studies exploit a wide range of methods and data sources to address evolving public health questions. It would be a daunting task to identify every relevant population registry and health-care-records system—and every possible potentially important hypothesis—let alone register them.

Furthermore, the potential "transparency" of such registered information could easily be clouded by the complexity of assembling the information. The workshop report downplays this problem, saying that the bureaucracy for universal registration would be "very minor."^{3,p17} This is a disingenuous statement, given the scope of what registration might entail. For example, the report suggests that researchers be obliged to register "any amendments" to their protocols or hypotheses over time.^{3,p1} The report also suggests that the STROBE guidelines¹⁰ for reporting epidemiologic results (these guidelines themselves comprising 31 pages of dense print) could form the basis for the kind of information investigators should register.^{3,p9} What are the chances that the registration of observational studies would be a "very minor" encumbrance? Judging from the cancerous growth of bureaucracies to protect human subjects in observational studies, we are not optimistic.

The best place for transparency may not be at the onset of a study, but at its culmination. Data become "results" when they have been sifted through careful analytic minds and subjected to the rigors of peer review. As we have argued before,¹¹ transparency (in the form of data availability) makes good sense once results have been published.

"Ethical Obligation"

Ethics is another broad virtue that no one can dispute, but that can be applied in sloppy ways. The use of "ethics" in the present context acquires its perceived force (once again) from an analogy with clinical trials. Is this analogy apt? In trials, people who are appropriately informed put themselves at potential harm for the sake of contributing to the future welfare of others. Such altruistic risk on the part of participants compels researchers to plan their studies carefully and publish their findings fully. In contrast, the cost to human subjects in epidemiologic studies ranges from near-zero (having one's medical records compiled into a research database) to mild inconvenience (responding to repeated questionnaires in longitudinal studies). The risk added by such participation

rarely (if ever) comes close to the jeopardy taken on by volunteers in a clinical trial. To equate this jeopardy to the risk in observational studies is to denigrate the contribution of clinical-trial volunteers.

Epidemiologists have real and serious obligations with regard to publication of findings, ethical behavior and transparency. But the various metaphors that equate those obligations with those of clinical trials reveal a lack of understanding of the deep differences between experimental and observational research.

Journal Policies

Coming closer to home, the workshop report suggests that journal editors should at some point begin to decline manuscripts if the hypothesis has not been pre-registered. Such pre-registration, so the argument goes, provides an indicator of "quality" of study results. We could not disagree more. The test of a predefined hypothesis is but one path to scientific discovery. Another is through exploration of large and carefully-collected data sets. For evidence of this, we need look no further than the field of human genetics. As our genetics colleagues embrace the agnostic world of hypothesis-free exploration, they accept false-positive results as a logical part of the scientific process. The problem for epidemiologic research is more subtle: the number of hypotheses that authors may have explored is rarely recorded and impossible to infer. This is a legitimate worry in the interpretation of epidemiologic results. For geneticists, the solution is replication. Replication is not so simple for epidemiologists—every observational study is unique. For this reason, epidemiologists have an even greater obligation to judge the extent to which new results are consistent with existing evidence of all types.

We do not wish to be apologists for the shortcomings of epidemiology. To the contrary, epidemiologists must constantly expose and debate the limitations of our craft. But we don't see new bureaucratic structures as the solution. At its best, epidemiology is a process of synthesizing evidence within data sets and across populations, with due respect for biologic plausibility and a skeptic's eye for alternative explanations. Papers that contribute to this process are what we value at *EPIDEMIOLOGY*, and these are the papers we will strive to publish. On such matters, registration has no bearing.

REFERENCES

1. Loder E, Groves T, MacAuley D. Registration of observational studies: The next step towards research transparency. *BMJ*. 2010;340:375–376.
2. The editors. Should protocols for observational studies be registered? *Lancet*. 2010, 375:348.
3. Workshop: Enhancement of the Scientific Process and Transparency of Observational Epidemiology Studies, 24–25 September 2009, London. Workshop Report No. 18, Brussels, November 2009, European Centre for Ecotoxicology and Toxicology of Chemicals. Available at: <http://links.lww.com/EDE/A415>.

4. Samet J. To register or not to register. *Epidemiology*. 2010;21:610–611.
5. Lash TL. Preregistration of study proposals is unlikely to improve the yield from our sciences, but other strategies might. *Epidemiology*. 2010;21:612–613.
6. Takkouche B, Norman G. Meta-analysis protocol registration: Sed quis custodiet ipsos custodes? [*But who will guard the guardians?*]. *Epidemiology*. 2010;21:614–615.
7. Poole C. A vision of accessible epidemiology. *Epidemiology*. 2010; 21:616–618.
8. Vandembroucke JP. Pre-registration of epidemiologic studies: An ill-founded mix of ideas. *Epidemiology*. 2010;21:619–620.
9. The Editors. On the death of a manuscript. *Epidemiology*. 2002;13:495–496.
10. Vandembroucke JP, von Elm E, Altman DG, et al; for the STROBE Initiative. Strengthening the reporting of observational studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology*. 2007;18:805–835.
11. Hernán M, Wilcox AJ. Epidemiology, data sharing, and the challenge of scientific replication. *Epidemiology*. 2009;20:167–168.