Letters to the Editor

Journals Should Publish All ''Null'' Results and Should Sparingly Publish ''Positive'' Results

To the Editor: The editorial by Rebbeck et al. (1) is timely and important. Here, I share some thoughts on this debate. The X team (real but anonymous here) meets successfully most proposed criteria. X has published nine articles on mostly brand new (but also some replicated) gene-disease associations in journals with an average impact factor exceeding 10. All publications show statistically significant (''positive'') results; include fascinating arguments about biological plausibility and pathways; emphasize population structure (most ''positives'' pertain to ''racial'' subgroups); insist on heterogeneity and interactions; and have decent power for main effects, although it is interactions/subgroups that claim importance. I wish I knew what is post hoc versus a priori here. X team accurately represents, quite well, the research that currently dominates high-impact journals.

What is the credibility of this literature? I think very low. At a minimum, there are heavy reporting biases. With small effects and trillions of possible, mostly false, hypotheses, the chance that any team comes across nine ''positive'' and no ''null'' findings approaches 0% (2). This literature may be less credible than the small genetic association studies in Chinese language journals wherein 50% of reports show statistically significant results and the genetic effects exceed anything seen in the average English language literature (3).

Epidemiology has been criticized for nonreplication because it is a robust science where measurements are amenable to quantitative refutation. ''Basic science'' reasoning is often as cloudy, beautiful, and impossible to refute as poetry. Few promises made in basic science journals are fulfilled even 20 to 25 years later (4). Biological evidence is extremely interesting but we should try to measure better how (well) it works. Functional data are usually derived after epidemiologic hints are provided (5). We do not know what the concordance of epidemiology and biology really means (6). Why promote spurious concordance seeking?

I also feel uncomfortable about chasing heterogeneity, multigene models, and interactions as primary targets. I firmly believe in heterogeneity (7) but, in my limited experience, when one accumulates more and better evidence, heterogeneity often goes away and interactions disappear, leaving simple main effects.

The large majority of molecular epidemiology analytic results should be ''null'' (2, 8, 9), representing truly null findings. A credible scientific journal should publish all studies with ''null'' results provided they acknowledge their limitations. Conversely, such a journal should be cautious about publishing ''positive'' results, most of which are false. Independent replication is important and should be done by different teams, preferably by competitors. ''Null'' results should be published promptly in print in short versions, with more extensive details in web-based files. ''Positive'' results should be published equally promptly, but only on the web, pending independent replication; once refuted, the original article and the refutation could be printed as a single nice null report; the rare validated findings should appear in print with full details.

Even for simple main effects, we need large-scale evidence and a healthy environment where selective reporting is minimized. The Network of Networks, a HuGENet initiative, was recently launched as an effort to attain this goal (10). It is important to know who is doing what in one field, join forces across teams working in the same field, and use standard methods for phenotypes, genotypes, exposures, and analyses. If we firmly document main effects, we can then also pursue second- and third-order analyses.

John P.A. Ioannidis
Clinical and Molecular Epidemiology Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece and Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts University School of Medicine, Boston, Massachusetts

References

1. Rebbeck TR, Martinez ME, Sellers TA, Shields PG, Wild CP, Potter JD. Genetic variation and cancer: improving the environment for publication of association studies. Cancer Epidemiol Biomarkers Prev 2004;13:1985–6.
2. Ioannidis JP. Why most published research findings are false. PLoS Med 2005;2:e124.
3. Pan Z, Trikalinos TA, Kavvoura FK, Lau J, Ioannidis JPA. Local literature bias in genetic epidemiology: an empirical evaluation of the Chinese literature. PLoS Med. In press.
4. Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications. Am J Med 2003;114:477–84.
5. Pharoah PD, Dunning AM, Ponder BA, Easton DF. The reliable identification of disease-gene associations. Cancer Epidemiol Biomarkers Prev 2005; 14:1362.
6. Rebbeck TR, Spitz M, Wu X. Assessing the function of genetic variants in candidate gene association studies. Nat Rev Genet 2004;5:589–97.
7. Lau J, Ioannidis JP, Schmid CH. Summing up evidence: one answer is not always enough. Lancet 1998;351:123–7.
8. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42.
9. Sterne JA, Davey Smith G. Sifting the evidence—what's wrong with significance tests? BMJ 2001;322:226–31.
10. Ioannidis JP, Bernstein J, Boffetta P, et al. A network of investigator networks in human genome epidemiology. Am J Epidemiol 2005;162:302–4.

doi:10.1158/1055-9965.EPI-05-0921

Journals Should Publish All "Null" Results and Should Sparingly Publish "Positive" Results

John P.A. Ioannidis

Cancer Epidemiol Biomarkers Prev 2006;15:186.