Modeling and Research on Research

John P.A. Ioannidis1,2,3,4*

**Featured Article:** Ioannidis JP. Why most published research findings are false. PLoS Med 2005;2:e124.5

The article featured here emerged as an effort to create an overarching modeling framework for proposed significant research findings and their validation, in line with accumulating empirical evidence on the replication rates of different types of research designs and settings. The main challenges behind the essay had occupied me and several other scientists over many years before its first draft was written in 2004.

Some initial impetus was offered by the advent of systematic approaches to evidence, such as systematic reviews and metaanalyses. Since the early 1990s, there was growing excitement that these approaches would allow piecing together the fragments of published literature and obtain summary results with higher accuracy and less uncertainty. In theory, this could revolutionize the translation of best science into best practice (1). For example, the Cochrane Collaboration adopted this vision (2) and was populated by volunteers with the noblest intentions for doing good—and no harm. Evidence-based medicine aimed to become the new basic science of health (3).

However, as more and more researchers started taking careful looks at the scientific literature in large scale, serious deficiencies increasingly became apparent. Most people who ventured to look intensely at the big picture quickly felt nauseated from what they saw. Doug Altman (4) had already claimed in 1994 that poor medical research was a scandal, and Drummond Rennie (5) had already started the peer review congresses in 1989, revealing more of the major problems eroding credibility in published biomedical articles. For some of us, studying the processes and biases that caused this trouble became more important than just struggling to claim yet another fundable discovery that might be false, useless, or both, or trying to get a summary result that would further promote biased evidence. If one could find a way to “fix” even just 1% of the scientific literature by identifying and correcting some biases, it would amount to an equivalent of hundreds of thousands of papers being affected and corrected. It is unlikely that any siloed discovery in any discipline where scientists try to “focus, focus, focus” (the classic way of getting grants, promotion, and major awards instead of looking at the big picture) could have such a major impact.

However, the task has been ambitious and difficult. Even bias tests are biased (6). Initially, the tools of metaresearch were sparse. A whole new scientific discipline had to be defined and go through the growing pains of how to perform empirical research on research. But now we have plenty of material to work with. Whereas epistemology until the late 20th century had used mostly theoretical, philosophical arguments and concepts and had tested them mostly through case studies, now there are millions of scientific investigations and trillions of analyses. Some bias concepts were developed in diverse scientific domains long ago, mostly in the social and psychological sciences, but their adoption and use have been erratic. It is increasingly apparent that problems such as publication bias, selective analysis and outcome reporting, and data dredging affect disciplines as remote as clinical medicine, omics, animal studies, economics, social sciences, psychology, and neurosciences. Best practices may also be possible to share or transplant across disciplines.

Very little of this emerging picture of convergence was apparent 10 years ago. Most investigators had difficulty understanding even what a simple metaanalysis was, let alone what it meant to study empirically large domains of evidence. Prestigious senior experts would use the word “metaanalysis” mostly to express their total contempt for anything quantitative. The rule of thumb for what was perceived as “good science” in their world was lots of expert opinion, biological plausibility speculation, limited data, and no replication. Since then, we have accumulated substantial empirical evidence to argue for larger and better-designed studies, rigorous replication mechanisms, international consortia, fewer conflicts, transparency, and data sharing. Some of these ideas are more widely applied than others. Still, occasionally they have been misconstrued, as when experts now use the term “evidence-based medicine” whenever they want to give prestige to their pure

---

1 Stanford Prevention Research Center, Department of Medicine, and 2 Department of Health Research and Policy, Stanford University School of Medicine; 3 Department of Statistics, Stanford University School of Humanities and Sciences; 4 Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA.

* Address correspondence to this author at: Stanford Prevention Research Center, Medical School Office Bldg., Rm. X306, 1265 Welch Rd., Stanford, CA 94305. E-mail jioannid@stanford.edu.

Received June 9, 2014; accepted June 11, 2014. Previously published online at DOI: 10.1373/clinchem.2013.218453

This article has been cited >1200 times since publication.
expert opinion, when randomized trials and meta-analyses are used as marketing tools by the industry, and when many predatory journals masquerade as open-access noble endeavors. Studying science with scientific methods remains highly exciting and challenging.

**Author Contributions:** All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

**Authors’ Disclosures or Potential Conflicts of Interest:** No authors declared any potential conflicts of interest.

**References**