The Thin Line Between Hope and Hype in Biomarker Research

Patrick M. M. Bossuyt, PhD

BIOMARKERS HAVE BECOME A POPULAR TOPIC IN medicine, and investigations of putative molecular indicators of a specific biological state have started to occupy a considerable part of health research. In the past decades, advances in molecular biology coupled with progress in genomics, proteomics, and metabolomics have fueled hope for the development of new medical tests. Biomarkers should enable clinicians to make an earlier or more definitive diagnosis, identify persons at risk of developing disease, develop more precise estimates about prognosis, and fine-tune treatment selection, thereby approaching a form of stratified, or even personalized, medicine.

With few exceptions, most of these promises have yet to be fulfilled. Only a small number of biomarkers are being used in routine clinical practice.[1] No new major cancer biomarkers have been approved for clinical use for at least 25 years.[2] Most clinical decisions still rely on more conventional forms of medical testing, such as existing laboratory measurements and imaging studies.

There are several reasons for the relatively slow progress. For example, molecular biomarkers for many conditions have yet to be identified. Other issues involve problems with characterization and control of the preanalytical variability[2] and suboptimal design of studies used for marker discovery and validation. Many biomarker studies have major methodological shortcomings, in particular in the selection of appropriate study groups; for instance, some studies include only extreme cases and contrast them with healthy controls. Despite these concerns, hope has been high, and hype has never been far away.

In this issue of *JAMA*, Ioannidis and Panagiotou[3] demonstrate that frequently cited biomarker studies reported effect sizes that were often higher than effect sizes reported in subsequent larger studies of the same biomarker and were more extreme than summary estimates reported in a meta-analysis of that biomarker. For 29 of 35 studies included in their analysis, the subsequently published meta-analysis reported a less optimistic effect size estimate than the highly cited study.

See also p 2200.

For instance, in a 1994 study on cancer risk in 33 families with evidence of linkage to *BRCA1* carriers, the authors compared cancer cases other than breast or ovarian cancer with national incidence rates and reported a 4.11 relative excess risk for colon cancer among *BRCA1* carriers.[4] A study published 11 years later, in which data were summarized from more than 30 epidemiologic studies on cancer incidence in *BRCA1* mutation carriers, found that all of the studies on colon cancer that had appeared after the 1994 study had reported smaller, and often nonsignificant, relative risks.[5] One of these, published in 2004, reported a nonsignificant odds ratio of 1.24.[6] However, this study has received only 26 citations so far, compared with 1051 for the initial 1994 article.[4]

Likewise, in a 1991 article, the authors reported high peak serum levels of homocysteine in 16 of 38 patients with cerebrovascular disease, in 7 of 25 with peripheral vascular disease, and in 18 of 60 with coronary vascular disease, but in 0 of 27 normal adults, and reported a statistically significant odds ratio of 23.9 for coronary vascular disease in patients with hyperhomocysteinemia.[7] A meta-analysis of hyperhomocysteinemia, published 9 years later and including 33 studies and more than 16 000 patients,[8] reported a summary odds ratio for cardiovascular disease of 1.58. The initial report has received 1451 citations, whereas to date, the meta-analysis has had 37 citations.

It is difficult to estimate how often a study that publishes a more extreme effect receives more attention than larger studies of the same marker, or than meta-analyses, which provide a summary estimate based on all available evidence, after critical appraisal. The review by Ioannidis and Panagiotou[3] is not based on an "inception cohort," ie, a group of studies of a biomarker defined from the first evaluation. The authors first selected highly cited studies and then tried to find a matching meta-analysis published after the highly cited study. They used an arbitrary threshold of 400 citations, regardless of the date of publication of the index study, and were able to match less than half of the highly cited studies to a meta-analysis.

Author Affiliations: Department of Clinical Epidemiology, Biostatistics, and Bioinformatics, University of Amsterdam, the Netherlands.
Corresponding Author: Patrick M. M. Bossuyt, PhD, Department of Clinical Epidemiology, Biostatistics, and Bioinformatics, Academic Medical Center, Room J1b-214; PO Box 22700; 1100 DE Amsterdam; the Netherlands (p.m.bossuyt@amc.uva.nl).

Another factor involves the dynamics of study initiation, a condition for inclusion in such an analysis. New studies to evaluate biomarkers do not occur at random and are not initiated without knowledge of previous studies. Investigators design a study and seek funding because they feel they have something useful, promising, or otherwise worthwhile to evaluate. A large number of issues drive the biomarker research agenda, and while some of these issues can be quite mundane, they do not occur at random. For instance, the likelihood of an additional evaluation of an individual biomarker from a list of all potential biomarkers is not distributed evenly across biomarkers, nor should it be. This may help explain why a study with a smaller effect size more often follows a study with a large effect size than the other way around. In the latter case, the second study may never be started. Analogously, the likelihood that a meta-analysis will be conducted for a specific biomarker is not homogeneous for all biomarkers.

Even more complicated processes affect the citation of previous studies. Studies that are published early and appear to report novel and promising findings have a citation advantage and also have more time to accrue citations. However, the exact reasons for citing previous articles, and the hurdles before switching citations to more complete or more precise studies, deserve further study. The science of the scientific reception of biomarker evaluation studies is still in its infancy.

The report by Ioannidis and Panagiotou[3] is a convincing case study demonstrating that more extreme, often early associations receive considerable attention and continue to do so, despite the availability of subsequent studies or meta-analyses with more precise estimates. The authors do not explain the citation advantage of the highly cited studies but refer to several mechanisms that may be responsible for the inflated effects. For instance, many studies were small, so chance plays an important role, and several used a case-control design, which is known to generate inflated results.[9] Most of these deficiencies can be remedied by using better study designs.[10] In addition, the completeness and transparency of reporting may be improved through the use of various standardized checklists, enabling editors, reviewers, and readers to more easily assess the studies and detect study weaknesses.[11]

It would be premature to doubt all scientific efforts at marker discovery and unwise to discount all future biomarker evaluation studies. However, the analysis presented by Ioannidis and Panagiotou should convince clinicians and researchers to be careful to match personal hope with professional skepticism, to apply critical appraisal of study design and close scrutiny of findings where indicated, and to be aware of the findings of well-conducted systematic reviews and meta-analyses when evaluating the evidence on biomarkers.

Conflict of Interest Disclosures: The author has completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

REFERENCES

1. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol*. 2006;24(8):971-983.
2. Diamandis EP. Cancer biomarkers: can we turn recent failures into success? *J Natl Cancer Inst*. 2010;102(19):1462-1467.
3. Ioannidis JPA, Panagiotou OA. Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses. *JAMA*. 2011;305(21):2200-2210.
4. Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE; Breast Cancer Linkage Consortium. Risks of cancer in *BRCA1*-mutation carriers. *Lancet*. 1994;343(8899):692-695.
5. Friedenson B. *BRCA1* and *BRCA2* pathways and the risk of cancers other than breast or ovarian. *MedGenMed*. 2005;7(2):60.
6. Niell BL, Rennert G, Bonner JD, Almog R, Tomsho LP, Gruber SB. *BRCA1* and *BRCA2* founder mutations and the risk of colorectal cancer. *J Natl Cancer Inst*. 2004;96(1):15-21.
7. Clarke R, Daly L, Robinson K, et al. Hyperhomocysteinemia: an independent risk factor for vascular disease. *N Engl J Med*. 1991;324(17):1149-1155.
8. Cleophas TJ, Hornstra N, van Hoogstraten B, van der Meulen J. Homocysteine, a risk factor for coronary artery disease or not? a meta-analysis. *Am J Cardiol*. 2000;86(9):1005-1009.
9. Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM. Case-control and two-gate designs in diagnostic accuracy studies. *Clin Chem*. 2005;51(8):1335-1341.
10. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *J Natl Cancer Inst*. 2008;100(20):1432-1438.
11. Bossuyt PM, Reitsma JB, Bruns DE, et al; Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326(7379):41-44.