Editorial

Statistical Methods for Test and Biomarker Evaluation Studies: A *Clinical Chemistry* Series

James C. Boyd,[1*] Nader Rifai,[2,3] and Thomas Annesley[4]

Studies related to the development and evaluation of biomarkers constitute a major portion of the published work in *Clinical Chemistry*. In 2011, for example, approximately 70% of original articles considered these topics. Owing to the importance of biomarker evaluation, knowledge of statistical approaches for assessing biomarker data has become increasingly necessary, and the surprisingly high citation rates for many articles on such statistical approaches [e.g., *(1–3)*] would appear to support this view. To comprehend the rapidly growing literature regarding the utility of new biomarkers, readers must develop familiarity with a growing list of statistical concepts, not only of how biomarker studies should be conducted and evaluated but also of the statistical methods used to evaluate the data from such studies. The latter can range from simple measures of diagnostic accuracy to complex multivariate models for assessing the additional value of a new marker when used in conjunction with existing markers.

To facilitate a better understanding of these methods, the editors invited a group of respected epidemiologists and biostatisticians to prepare a series of review articles that would carefully examine the statistical methods currently used for biomarker evaluation and that would provide a foundation for interpreting data from studies in the current literature that evaluate diagnostic tests. This issue of *Clinical Chemistry* contains the first of 4 articles written by these experts.

The authors of this series approached this task systematically. They conducted multiple meetings—both face-to-face and by telephone—to outline the concepts they thought important to include. Their outlines were reviewed and approved by the editorial group. The subsequent process of writing and revising the 4 articles has taken much labor and many months. Each of these articles has been reviewed by external experts on the subject matter and revised in response to the comments of these experts.

The first article in the series, by Linnet and co-workers, reviews statistical methods used for the evaluation of single biomarkers *(4)*. These methods provide a groundwork for developing an initial understanding of the performance characteristics and diagnostic value of individual tests; however, performance characteristics for diagnostic tests must be understood not only for the individual test but also in the context of what additional diagnostic information a new biomarker might provide beyond what is available with current biomarkers. Moons and colleagues cover this topic in the second article in the series *(5)*. The third article, by Reitsma and colleagues, summarizes the methods used and the elements necessary for the preparation and evaluation of systematic reviews of biomarkers *(6)*. The fourth and final article, by Bossuyt and colleagues, examines the importance of moving beyond simple measures of sensitivity, specificity, and diagnostic accuracy to an assessment of how results of biomarker testing contribute to and/or change the course of patient care *(7)*. Such measures of the clinical utility of a biomarker ultimately determine the true importance of that biomarker.

We and the authors of this series acknowledge that it is impossible to cover all aspects of this broad topic within the confines of 4 articles. It is their hope and ours that these articles will provide readers with a strong starting point for understanding the most important concepts in this rapidly evolving area.

[1] Department of Pathology, University of Virginia Health System, Charlottesville, VA; Departments of [2] Laboratory Medicine and [3] Pathology, Boston Children's Hospital and Harvard Medical School, Boston, MA; [4] University of Michigan Health System, Ann Arbor, MI.
* Address correspondence to this author at: Department of Pathology, University of Virginia Health System, Box 800168, Charlottesville, VA 22908. E-mail jboyd@virginia.edu.

Editorial

References

1. Bland JM, Altman DG. Statistical method for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
2. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993;39:561–77.
3. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. Clin Chem 2008;54:17–23.
4. Linnet K, Bossuyt PM, Moons KG, Reitsma JB. Quantifying the accuracy of a diagnostic test or marker. Clin Chem 2012;58:1292–1301.
5. Moons KG, de Groot JA, Linnet K, Reitsma JB, Bossuyt PM. Quantifying the added value of a diagnostic test or marker. Clin Chem. Forthcoming 2012;58.
6. Reitsma JB, Moons KG, Bossuyt PM, Linnet K. Systematic reviews of studies quantifying the accuracy of diagnostic tests and markers. Clin Chem. Forthcoming 2012;58.
7. Bossuyt PM, Reitsma JB, Linnet K, Moons KG. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. Clin Chem. Forthcoming 2012;58.