

Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK)

Lisa M. McShane, Douglas G. Altman, Willi Sauerbrei, Sheila E. Taube, Massimo Gion, Gary M. Clark for the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics

Despite years of research and hundreds of reports on tumor markers in oncology, the number of markers that have emerged as clinically useful is pitifully small. Often, initially reported studies of a marker show great promise, but subsequent studies on the same or related markers yield inconsistent conclusions or stand in direct contradiction to the promising results. It is imperative that we attempt to understand the reasons that multiple studies of the same marker lead to differing conclusions. A variety of methodologic problems have been cited to explain these discrepancies. Unfortunately, many tumor marker studies have not been reported in a rigorous fashion, and published articles often lack sufficient information to allow adequate assessment of the quality of the study or the generalizability of study results. The development of guidelines for the reporting of tumor marker studies was a major recommendation of the National Cancer Institute–European Organisation for Research and Treatment of Cancer (NCI-EORTC) First International Meeting on Cancer Diagnostics in 2000. As for the successful CONSORT initiative for randomized trials and for the STARD statement for diagnostic studies, we suggest guidelines to provide relevant information about the study design, preplanned hypotheses, patient and specimen characteristics, assay methods, and statistical analysis methods. In addition, the guidelines suggest helpful presentations of data and important elements to include in discussions. The goal of these guidelines is to encourage transparent and complete reporting so that the relevant information will be available to others to help them to judge the usefulness of the data and understand the context in which the conclusions apply. [J Natl Cancer Inst 2005;97:1180–4]

Despite years of research and hundreds of reports on tumor markers in oncology, the number of markers that have emerged as clinically useful is pitifully small (1–3). Often, initially reported studies of a marker show great promise, but subsequent studies on the same or related markers yield inconsistent conclusions or stand in direct contradiction to the promising results. It is imperative that we attempt to understand the reasons that multiple studies of the same marker lead to differing conclusions. A variety of problems have been cited to explain these discrepancies, such as general methodologic differences, poor study design, assays that are not standardized or lack reproducibility, and inappropriate or misleading statistical analyses that are often based on sample sizes too small to draw meaningful conclusions (4–11). For example, in retrospective studies, patient populations are often biased toward patients with available tumor specimens.

Specimen availability may be related to tumor size and patient outcome (12), and the quantity, quality, and preservation method of the specimen may affect feasibility of conducting certain assays. There can also be biases or large variability inherent in the assay results, depending on the particular assay methods used (13–17). Statistical problems are commonplace. These problems include underpowered studies or overly optimistic reporting of effect sizes and significance levels due to multiple testing, subset analyses, and cutpoint optimization (18).

Unfortunately, many tumor marker studies have not been reported in a rigorous fashion, and published articles often lack sufficient information to allow adequate assessment of the quality of the study or the generalizability of study results. Such reporting deficiencies are increasingly being highlighted by systematic reviews of the published literature on particular markers or cancers (19–25).

The development of guidelines for the reporting of tumor marker studies was a major recommendation of the National Cancer Institute–European Organisation for Research and Treatment of Cancer (NCI-EORTC) First International Meeting on Cancer Diagnostics (From Discovery to Clinical Practice: Diagnostic Innovation, Implementation, and Evaluation) that was convened in Nyborg, Denmark, in July 2000. The purpose of the meeting was to discuss issues, accomplishments, and barriers in the field of cancer diagnostics. Poor study design and analysis, assay variability, and inadequate reporting of studies were identified as some of the major barriers to progress in this field. One of the working groups formed at the Nyborg meeting was charged with addressing statistical issues of poor design and analysis and with reporting of tumor marker prognostic studies. The guidelines that we present in this commentary are the product of that committee. The Program for the Assessment of Clinical Cancer Tests (PACCT) Strategy Group of the U.S. NCI has also strongly endorsed this effort (<http://www.cancerdiagnosis.nci.nih.gov/assessment>).

Affiliations of authors: Biometric Research Branch (LMM) and Cancer Diagnosis Program (SET), National Cancer Institute, Bethesda, MD; Medical Statistics Group, Cancer Research UK, Center for Statistics in Medicine, Wolfson College, Oxford, UK (DGA); Institut fuer Medizinische Biometrie und Medizinische Informatik, Universitaetsklinikum Freiburg, Germany (WS); Centro Regionale Indicatori Biochimici di Tumore, Ospedale Civile, Venezia, Italy (MG); OSI Pharmaceuticals, Inc., Boulder, CO (GMC).

Correspondence to: Lisa M. McShane, PhD, National Cancer Institute, Biometric Research Branch, DCTD, Rm. 8126, Executive Plaza North, MSC 7434, 6130 Executive Blvd., Bethesda, MD 20892-7434 (e-mail: lm5h@nih.gov).

See “Notes” following “References.”

DOI: 10.1093/jnci/dji237

© The Author 2005. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oupjournals.org.

The guidelines that we present in this commentary build on earlier suggestions (21,26–29) and on educational publications (30–33). They recommend elements and formats for presentation with the objectives of facilitating evaluation of the appropriateness and quality of study design, methods, analyses, and improving the ability to compare results across studies. As for the successful CONSORT initiative for randomized clinical trials (34), and the STARD statement for studies of diagnostic test accuracy (35), these guidelines suggest relevant information that should be provided about the study design, preplanned hypotheses, patient and specimen characteristics, assay methods, and statistical analysis methods. In addition, the guidelines suggest helpful presentations of data and important elements to include in discussions. To be published elsewhere, in an explanatory document, are specific justifications for the need for each of the elements of the recommendations.

We have developed these reporting guidelines primarily for studies evaluating a single tumor marker of interest, often including adjustment for standard clinical prognostic variables. They are largely relevant for studies exploring more than one marker, but they are not intended to specifically address statistical considerations in development of prognostic models from very large numbers of candidate markers. The reason we chose to emphasize prognostic marker studies is that they represent a large proportion of the tumor marker literature and tend to be particularly fraught with problems because they are often conducted on retrospective collections of specimens, and analyses may contain substantial exploratory components. For this commentary, we define prognostic markers to be markers that have an association with some clinical outcome, typically a time-to-event outcome such as overall survival or recurrence-free survival. (Some individuals adhere to a stricter definition of prognostic marker as applying only to the natural history of patients who received no treatment following local therapy.) Prognostic markers may be considered in the clinical management of a patient. For example, they may be used as decision aids in determining whether a patient should receive adjuvant chemotherapy or how aggressive that therapy should be. Predictive markers are generally used to make more specific choices between treatment options. Predictive markers are used as indicators of the likely benefit to a specific patient of a specific treatment. For example, a predictive marker might indicate that a patient expressing the marker will benefit more from a new treatment than from standard treatment, whereas a patient not expressing the marker will derive little or no benefit from the new treatment. Predictive marker studies usually occur later in the marker development process, and there are far fewer published examples. Knowledge of specific treatments received and of how those treatment decisions were made become even more critical. In our judgment, the issues in reporting predictive marker studies are complex and different enough from those of prognostic marker studies that we are not willing to claim that these guidelines give predictive marker studies adequate coverage, although we believe that most of the guidance is relevant to such studies also.

The goal of these guidelines is to encourage transparent and complete reporting so that the relevant information will be available to others to help them to judge the usefulness of the data and understand the context in which the conclusions apply. These guidelines are not intended to dictate specific designs or analysis strategies. In general, there is more than one acceptable approach to the design or analysis of a particular study, although these

guidelines should help to eliminate some clearly unacceptable options, as have been discussed in other papers (7,26,33,36). For example, unacceptable options include reporting statistical significance of a marker's prognostic effect without acknowledging that the significance testing was preceded by extensive manipulations involving derivation of data-dependent cutpoints or variable selection procedures. High-quality reporting of a study cannot transform a poorly designed or -analyzed study into a good one, but it can help to identify the poor studies, and we believe it is an important first step in improving the overall quality of tumor marker prognostic studies.

MATERIALS AND METHODS

Initial ideas for key elements to be addressed in the guidelines were assembled from literature citing empirical evidence of inadequate reporting or problematic analysis methods (9,18,36,37) that are based on published reviews of tumor marker studies. Ideas were also generated by reviewing similar reporting guidelines that have been produced for other types of medical research studies (CONSORT, QUOROM, MOOSE, and STARD) (34,35,38,39). Three individuals from the working group (L.M., D.A., and G.C.) wrote a first draft to serve as a starting point for discussion by the full group. Comments on drafts were made by the full group on a conference call and through multiple e-mail exchanges. A very preliminary draft was presented to the PACCT Strategy Group in January 2001. In response to comments, the guidelines were shortened, reformatted, and recirculated to the full committee. They were posted to the PACCT website (<http://www.cancerdiagnosis.nci.nih.gov/assessment/progress/clinical.html>) for public comment and circulated to attendees of the NCI-EORTC Second International Meeting on Cancer Diagnostics (Conference on the Development of New Diagnostic Tools for Cancer) that was held in Washington, DC, in June 2002. In February 2003, three committee members (D.A., L.M., and W.S.) met for 2 days to make further revisions. The version produced in that February meeting was sent to the full committee for final comment. The version presented here incorporates those final comments and was approved by the full committee.

RESULTS

Table 1 shows the recommendations for reporting studies on tumor markers. Specific items are grouped under headings Introduction, Materials and Methods, Results, and Discussion, reflecting the relevant sections of a published scientific article. Further details about the recommendations and explanatory material will be provided elsewhere.

As noted in item 12, a diagram may be helpful to indicate numbers of individuals included at different stages of a study. As a minimum, such a diagram could show the number of patients originally in the sample, the number remaining after exclusions, and the numbers incorporated into univariate and multivariable analyses.

DISCUSSION

The reporting guidelines in this commentary are the result of a collaborative effort among statisticians, clinicians, and laboratory scientists who are committed to improving and accelerating

Table 1. Reporting recommendations for tumor marker prognostic studies (REMARK)

INTRODUCTION

1. State the marker examined, the study objectives, and any prespecified hypotheses.

MATERIALS AND METHODS

Patients

2. Describe the characteristics (e.g., disease stage or comorbidities) of the study patients, including their source and inclusion and exclusion criteria.
3. Describe treatments received and how chosen (e.g., randomized or rule-based).

Specimen characteristics

4. Describe type of biological material used (including control samples) and methods of preservation and storage.

Assay methods

5. Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study endpoint.

Study design

6. State the method of case selection, including whether prospective or retrospective and whether stratification or matching (e.g., by stage of disease or age) was used. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.
7. Precisely define all clinical endpoints examined.
8. List all candidate variables initially examined or considered for inclusion in models.
9. Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.

Statistical analysis methods

10. Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.
11. Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.

RESULTS

Data

12. Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the numbers of patients and the number of events.
13. Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumor marker, including numbers of missing values.

Analysis and presentation

14. Show the relation of the marker to standard prognostic variables.
15. Present univariate analyses showing the relation between the marker and outcome, with the estimated effect (e.g., hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analyzed. For the effect of a tumor marker on a time-to-event outcome, a Kaplan–Meier plot is recommended.
16. For key multivariable analyses, report estimated effects (e.g., hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.
17. Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their statistical significance.
18. If done, report results of further investigations, such as checking assumptions, sensitivity analyses, and internal validation.

DISCUSSION

19. Interpret the results in the context of the prespecified hypotheses and other relevant studies; include a discussion of limitations of the study.
 20. Discuss implications for future research and clinical value.
-

the process by which tumor markers that provide useful information for management of cancer patients are adopted into clinical practice. In addition to the authors of this commentary, we gratefully acknowledge the contributions of many individuals with whom we have had informal discussions regarding these guidelines and who have been supportive of this effort. All of us participating in the development of these guidelines are actively involved in the design, conduct, and analysis of studies involving tumor markers. We serve as editors and reviewers for many scientific journals that publish tumor marker studies. We serve on program committees for international meetings, as decision-makers for funding agencies, and as participants in national and international committees charged with evaluating and prioritizing tumor markers for further study or making recommendations for clinical use. We also are actively involved in our own research involving tumor markers. As editors, reviewers, and program and advisory committee members, we have struggled with having to make decisions when insufficient information is provided about study design or analysis methods. As individual investigators, we

have experienced the frustration of trying to interpret often confusing literature to guide our own research programs.

There are consequences of poor study reporting for the research community as a whole. Poorly designed or inappropriately analyzed studies can attract undeserved attention when they produce very dramatic but unfortunately incorrect results. In contrast, some carefully designed and analyzed studies have been overlooked because they produced less dramatic but perhaps more accurate and realistic results. The poor quality of reporting of prognostic marker studies may have contributed to the relative scarcity of markers whose prognostic influence is well supported. Thorough reporting is required no matter what methods of design and analysis are used. Thorough reporting does not solve problems of poor design or analysis that are being reported; rather, it just fairly describes what problems may exist and need to be considered in interpretation. It is our hope that these guidelines will be embraced and used by journal editors, reviewers, funding agencies, decision-making bodies, and individual investigators.

These guidelines have been labeled as applying to clinical prognostic studies. Not all of the elements apply to studies conducted in earlier phases of marker development (40), for example, early marker studies seeking to find an association between a new marker and other clinical variables or existing prognostic factors. However, our recommendation is that investigators conducting early marker studies strive to adhere to as many of the reporting guidelines as applicable in their situation, and the guidelines might also suggest issues that will be important for them to consider in planning follow-up studies on their investigational markers. Studies of markers that can be used to predict the success of particular therapies, such as molecular targeted therapies, need additional considerations. It is our opinion that predictive marker studies should generally be conducted within randomized trials and should require a sufficient (usually larger) effective sample size and that assays should be in a more advanced state of development. The CONSORT statement for randomized clinical trials can serve as a starting point for reporting guidelines for predictive marker studies, but more issues relating to the marker assays must be addressed. It is our feeling that more stringent and specific guidelines need to be developed for reporting studies of predictive markers. Such studies will be considered in somewhat more detail in the planned explanatory paper to be published elsewhere.

It may not be possible to report every detail for every study. For example, it is often difficult to provide detailed patient inclusion/exclusion criteria or treatment information in retrospective prognostic marker studies using archived tumor specimens. The impact of such missing information must be judged in the specific context of the study and its stated conclusions. For example, a “pure” prognostic study should be conducted in a group of patients who have not received any systemic adjuvant therapy, but treatment information is often missing or unreliable in retrospective studies. In these cases, it is important to recognize that apparent “prognostic” effects may be influenced by potential treatment by marker interactions. The key point is that there must be a clear statement of what is and is not known. In addition, it was beyond the scope of these guidelines to recommend specific details that should be reported for each of the major classes of marker assays, for example, immunohistochemistry, in situ hybridization methods, or DNA-based assays. There is an ongoing effort to define such assay-specific checklists by another working group evolving from the NCI-EORTC International Meetings on Cancer Diagnostics.

Some of the reviewers suggested that the guidelines should promote full public access to data, possibly even individual-level data. We have chosen not to include this issue in the current scope of the guidelines even though we view movement in this direction as generally positive. One concern is that if a study was poorly designed or inadequately reported, making its data publicly available may simply propagate bad science. Good study design and data quality have to come first. We do recognize the potential benefits of promoting full public access to good-quality data. It would allow verification of published analysis methods and results and would facilitate alternative analyses and meta-analyses. Attainment of these goals would be helped substantially if guidelines 10 and 11 were strictly applied so that statistical analysis methods were described in sufficient detail to allow an individual independent of the original research team to reproduce the results of the study if supplied with the raw data. For extensive analyses, it is possible that some of this information would have to be provided as supplementary material available outside of the main published report, for example on the journal’s or author’s Web site.

Although some might view adherence to these guidelines as yet another burden in trying to publish or obtain funding, we would argue that use of these guidelines is more likely to reduce burdens on the research community. Making clear what is considered relevant and important to report in journal articles or funding proposals will likely reduce review time, reduce requests for revisions, and help to ensure a fair review process. Furthermore, we consider it as a prerequisite for a thoughtful presentation and interpretation of the results of a specific study and a key aid for a summary assessment of the effect of a marker in a review paper. Most importantly, what greater reduction in burden could there be than to eliminate some of the false leads generated by poorly designed, analyzed, or reported studies that send researchers down unproductive paths, wasting years of time and money?

The ultimate usefulness of these guidelines will rely on how widely they are adopted. We are heartened by the enthusiastic responses that we received from the several journals who have agreed to simultaneously publish this paper. There is a clear recognition in the community that the time has come (if not long overdue) to improve the quality of tumor marker study reporting and conduct. We hope that many journals will adopt these guidelines as part of their editorial requirements. To the extent that does not happen immediately, we have to rely on authors of journal articles and reviewers of those articles to initiate the movement toward adherence to these guidelines.

We expect that just as tumor marker research will evolve, these guidelines will have to evolve to address new study paradigms and new assay technologies. It is our hope that publication of these guidelines will generate vigorous discussion leading to continually improved versions and, ultimately, improved quality of tumor marker studies.

The guidelines presented in this paper are available at <http://www.cancerdiagnosis.nci.nih.gov/assessment/progress/clinical.html>, as will be other recommendations from the group in due course. As noted, a detailed explanatory paper is to be published elsewhere, following the model of similar articles relating to the CONSORT and STARD statements (41–42).

REFERENCES

- (1) Hayes DF, Bast RC, Desch CE, Fritsche H Jr, Kemeny NE, Jessup JM, et al. Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. *J Natl Cancer Inst* 1996;88:1456–66.
- (2) Bast RC Jr, Ravdin P, Hayes DF, Bates S, Fritsche H Jr, Jessup JM, et al. for the American Society of Clinical Oncology Tumor Markers Expert Panel. 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. *J Clin Oncol* 2001;19:1865–78.
- (3) Schilsky RL and Taube SE. Introduction: Tumor markers as clinical cancer tests—are we there yet? *Semin Oncol* 2002;29:211–2.
- (4) McGuire WL. Breast cancer prognostic factors: evaluation guidelines. *J Natl Cancer Inst* 1991;83:154–5.
- (5) Fielding LP, Fenoglio-Preiser CM, and Freedman LS. The future of prognostic factors in outcome prediction for patients with cancer. *Cancer* 1992;70:2367–77.
- (6) Burke HB, Henson DE. Criteria for prognostic factors and for an enhanced prognostic system. *Cancer* 1993;72:3131–5.
- (7) Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. *Ann Intern Med* 1993;118:201–10.
- (8) Gasparini G, Pozza F, Harris AL. Evaluating the potential usefulness of new prognostic and predictive indicators in node-negative breast cancer patients. *J Natl Cancer Inst* 1993;85:1206–19.
- (9) Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. *Br J Cancer* 1994;69:979–85.

- (10) Gasparini G. Prognostic variables in node-negative and node-positive breast cancer. *Breast Cancer Res Treat* 1998;52:321–31.
- (11) Hall PA, Going JJ. Predicting the future: a critical appraisal of cancer prognosis studies. *Histopathology* 1999;35:489–94.
- (12) Hoppin JA, Tolbert PE, Taylor JA, Schroeder JC, Holly EA. Potential for selection bias with tumor tissue retrieval in molecular epidemiology studies. *Ann Epidemiol* 2002;12:1–6.
- (13) Thor AD, Liu S, Moore DH II, Edgerton SM. Comparison of mitotic index, in vitro bromodeoxyuridine labeling, and MIB-1 assays to quantitate proliferation in breast cancer. *J Clin Oncol* 1999;17:470–7.
- (14) Gancberg D, Lespagnard L, Rouas G, Paesmans M, Piccart M, DiLeo A, et al. Sensitivity of HER-2/neu antibodies in archival tissue samples of invasive breast carcinomas. Correlation with oncogene amplification in 160 cases. *Am J Clin Pathol* 2000;113:675–82.
- (15) McShane LM, Aamodt R, Cordon-Cardo C, Cote R, Faraggi D, Fradet Y, et al., and the National Cancer Institute Bladder Tumor Marker Network. Reproducibility of p53 immunohistochemistry in bladder tumors. *Clin Cancer Res* 2000;6:1854–64.
- (16) Paik S, Bryant J, Tan-Chiu E, Romond E, Hiller W, Park K, et al. Real-world performance of HER2 testing—National Surgical Adjuvant Breast and Bowel Project Experience. *J Natl Cancer Inst* 2002;94:852–4.
- (17) Roche PC, Suman VJ, Jenkins RB, Davidson NE, Martino S, Kaufman PA, et al. Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. *J Natl Cancer Inst* 2002;94:855–7.
- (18) Altman DG, De Stavola BL, Love SB, Stepniowska KA. Review of survival analyses published in cancer journals. *Br J Cancer* 1995;72:511–8.
- (19) Brundage MD, Davies D, Mackillop WJ. Prognostic factors in non-small cell lung cancer: a decade of progress. *Chest* 2002;122:1037–57.
- (20) Mirza AN, Mirza NQ, Vlastos G, Singletary SE. Prognostic factors in node-negative breast cancer: a review of studies with sample size more than 200 and follow-up more than 5 years. *Ann Surg* 2002;235:10–26.
- (21) Riley RD, Abrams KR, Sutton AJ, Lambert PC, Jones DR, Heney D, et al. Reporting of prognostic markers: current problems and development of guidelines for evidence-based practice in the future. *Br J Cancer* 2003;88:1191–8.
- (22) Riley RD, Burchill SA, Abrams KR, Heney D, Sutton AJ, Jones DR, et al. A systematic review of molecular and biological markers in tumours of the Ewing's sarcoma family. *Eur J Cancer* 2003;39:19–30.
- (23) Burton A, Altman DG. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. *Br J Cancer* 2004;91:4–8.
- (24) Papat S, Matakidou A, Houlston RS. Thymidylate synthase expression and prognosis in colorectal cancer: a systematic review and meta-analysis. *J Clin Oncol* 2004;22:529–36.
- (25) Riley RD, Heney D, Jones DR, Sutton AJ, Lambert PC, Abrams KR, et al. A systematic review of molecular and biological tumor markers in neuroblastoma. *Clin Cancer Res* 2004;10:4–12.
- (26) Altman DG, Lyman GH. Methodological challenges in the evaluation of prognostic factors in breast cancer. *Breast Cancer Res Treat* 1998;52:289–303.
- (27) Gion M, Boracchi P, Biganzoli E, Daidone MG. A guide for reviewing submitted manuscripts (and indications for the design of translational research studies on biomarkers). *Int J Biol Markers* 1999;14:123–33.
- (28) Altman DG. Systematic reviews of evaluations of prognostic variables. In: Egger M, Davey Smith G, Altman DG, editors. *Systematic reviews in health care. Meta-analysis in context*. 2nd ed. London (UK): BMJ Books; 2001. p. 228–47.
- (29) Altman DG. Systematic reviews of evaluations of prognostic variables. *BMJ* 2001;323:224–8.
- (30) McShane LM, Simon R. Statistical methods for the analysis of prognostic factor studies. In: Gospodarowicz MK, Henson DE, Hutter RV, O'Sullivan B, Sobin LH, Wittekind Ch, editors. *Prognostic factors in cancer*. 2nd ed. New York (NY): Wiley-Liss; 2001. p. 37–48.
- (31) Simon R. Evaluating prognostic factor studies. In: Gospodarowicz MK, Henson DE, Hutter RV, O'Sullivan B, Sobin LH, Wittekind Ch, editors. *Prognostic factors in cancer*. 2nd ed. New York (NY): Wiley-Liss; 2001. p. 49–56.
- (32) Biganzoli E, Boracchi P, Marubini E. Biostatistics and tumor marker studies in breast cancer: Design, analysis and interpretation issues. *Int J Biol Markers* 2003;18:40–8.
- (33) Schumacher M, Hollander N, Schwarzer G, Sauerbrei W. Prognostic factor studies. In: Crowley J, editor. *Handbook of statistics in clinical oncology*. New York (NY): CRC Press; 2005. p. 307–51.
- (34) Moher D, Schulz KF, Altman D for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285:1987–91.
- (35) Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Standards for Reporting of Diagnostic Accuracy*. *Clin Chem* 2003;49:1–6.
- (36) Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using “optimal” cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst* 1994;86:829–35.
- (37) Hilsenbeck SG, Clark GM, McGuire WL. Why do so many prognostic factors fail to pan out? *Breast Cancer Res Treat* 1992;22:197–206.
- (38) Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup D for the QUOROM Group. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet* 1999;354:1896–900.
- (39) Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group*. *JAMA* 2000;283:2008–12.
- (40) Hammond ME, Taube SE. Issues and barriers to development of clinically useful tumor markers: a development pathway proposal. *Semin Oncol* 2002;29:213–21.
- (41) Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. for the CONSORT Group. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663–94.
- (42) Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. *Standards for Reporting of Diagnostic Accuracy. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration*. *Clin Chem* 2003;49:7–18.

NOTES

We are grateful to the U.S. National Cancer Institute and the European Organization for Research and Treatment of Cancer for their support of the NCI-EORTC International Meetings on Cancer Diagnostics, from which the idea for these guidelines originated. We thank the U.K. National Translational Cancer Research Network for financial support provided to D. G. Altman.

Members of the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics are: **Douglas G. Altman**, DSc (Co-chair), Medical Statistics Group, Cancer Research UK, Centre for Statistics in Medicine, Wolfson College, Oxford OX2 6UD, UK; **Lisa M. McShane**, PhD (Co-chair), Biometric Research Branch, US National Cancer Institute, Bethesda, MD 20892; **Gary M. Clark**, PhD, OSI Pharmaceuticals, Inc., Boulder, CO 80301; **Jose Costa**, MD, Yale Cancer Center, New Haven, CT 06510-3202; **Angelo Di Leo**, MD, PhD, Department of Oncology, Hospital of Prato, 59100 Prato, Italy; **Massimo Gion**, MD, Centro Regionale Indicatori Biochimici di Tumore, Ospedale Civile, 30122 Venezia, Italy; **Robert J. Mayer**, MD, Dana-Farber Cancer Institute, Boston, MA 02115; **Willi Sauerbrei**, PhD, Institut fuer Medizinische Biometrie und Medizinische Informatik, Universitaetsklinikum Freiburg, 79104 Freiburg, Germany; and **Sheila E. Taube**, PhD, Cancer Diagnosis Program, US National Cancer Institute, Bethesda, MD 20892.

Manuscript received December 28, 2004; revised June 14, 2005; accepted July 21, 2005.