JNCI J Natl Cancer Inst (2018) 110(8): djy088

doi: 10.1093/jnci/djy088
First published online June 4, 2018
Commentary

COMMENTARY

Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): An Abridged Explanation and Elaboration

Willi Sauerbrei, Sheila E. Taube, Lisa M. McShane, Margaret M. Cavenagh, Douglas G. Altman

See the Notes section for the full list of authors' affiliations.
Correspondence to: Prof. Dr. Willli Sauerbrei, Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Stefan-Meier-Str. 26, 79104 Freiburg, Germany (e-mail: wfs@imbi.uni-freiburg.de).

Abstract

The Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) were developed to address widespread deficiencies in the reporting of such studies. The REMARK checklist consists of 20 items to report for published tumor marker prognostic studies. A detailed paper was published explaining the rationale behind checklist items, providing positive examples and giving empirical evidence of the quality of reporting. REMARK provides a comprehensive overview to educate on good reporting and provide a valuable reference for the many issues to consider when designing, conducting, and analyzing tumor marker studies and prognostic studies in medicine in general. Despite support for REMARK from major cancer journals, prognostic factor research studies remain poorly reported. To encourage dissemination and uptake of REMARK, we have produced this considerably abridged version of the detailed explanatory manuscript, which may also serve as a brief guide to key issues for investigators planning tumor marker prognostic studies. To summarize the current situation, more recent papers investigating the quality of reporting and related reporting guidelines are cited, but otherwise the literature is not updated. Another important impetus for this paper is that it serves as a basis for literal translations into other languages. Translations will help to bring key information to a larger audience world-wide. Many more details can be found in the original paper.

The Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) (1) were developed to address widespread deficiencies in the reporting of such studies. The REMARK checklist consists of 20 items to report in published tumor marker prognostic studies. A detailed explanation and elaboration paper was published explaining the rationale behind checklist items, providing positive examples, and giving empirical evidence of the quality of reporting (2,3). REMARK provides a comprehensive overview to educate on good reporting and provide a valuable reference for the many issues to consider when designing, conducting, and analyzing tumor marker studies and, with minimal adjustment, in prognostic studies in medicine in general.

The purpose of the REMARK checklist is to encourage investigators to properly report prognostic marker research (Figure 1). Careful reporting of what was done and what results were obtained facilitates the assessment of study quality and aids understanding of the relevance of the study conclusions. Despite support for REMARK from major cancer journals, prognostic factor research studies remain poorly reported.

Convincing evidence for this unfortunate situation is given in recent papers that reported investigations of the reporting quality of prognostic factor studies (4,5). The authors of the latter paper show that many key items are still very poorly reported and conclude that improvement seems to require more pressure on authors, reviewers, and editors.

The original focus of the REMARK recommendations was on studies of prognostic tumor markers that reported measurements of the biological molecules found in tissues, blood, and other body fluids. However, REMARK generally applies to any studies involving prognostic factors, not only in cancer.

Prognostic marker studies typically evolve through a series of steps beginning with exploratory discovery studies and proceeding through a series of studies addressing increasingly demanding hypotheses to elucidate a marker's prognostic value.

Received: November 21, 2017; **Revised:** March 16, 2018; **Accepted:** April 17, 2018

INTRODUCTION

1 State the marker examined, the study objectives, and any pre-specified hypotheses.

MATERIALS AND METHODS

Patients

2 Describe the characteristics (e.g., disease stage or co-morbidities) 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 univariable 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 pre-specified hypotheses and other relevant studies; include a discussion of limitations of the study.

20 Discuss implications for future research and clinical value.

Figure 1. The REMARK checklist (1–3).

The REMARK recommendations attempt to recognize these stages of development.

To encourage dissemination of the REMARK checklist, we have produced this considerably abridged version of the detailed explanation and elaboration manuscript. As well as a guide for investigators planning tumor marker prognostic studies, it is intended as a basis for literal translations into other languages. Many more details can be found in the original paper—the examples of good reporting may be especially helpful.

Checklist Items

Each checklist item should be addressed somewhere in a report, even if it can only be addressed by an acknowledgment that the information is unknown. We do not prescribe a precise location or order of presentation as this may be dependent upon journal policies and is best left to the discretion of the authors. We recognize that authors may address several items in a single section of text or in a table. Authors may find it convenient to report some of the requested items in a supplementary material section, rather than in the body of the manuscript, to allow sufficient space for adequate detail to be provided.

Authors may find it helpful to use the REMARK checklist reporting template, which can be downloaded from http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumour-marker-prognostic-studies-remark/.

Item 1: Introduction

The markers to be examined, the study objectives, and any prespecified hypotheses should be provided early in the study report. The description of the marker should include the biological aspects of the marker and the time in a patient's clinical course when it is to be assessed.

Objectives are goals one hopes to accomplish by conducting the study. Typical objectives for tumor marker prognostic studies include evaluation of the association between marker value and clinical outcome or determination of whether a tumor marker contributes additional information about likely clinical outcome beyond the information provided by standard clinical or pathologic factors.

Hypotheses should be formulated in terms of measures amenable to statistical evaluation. They represent tentative assumptions that can be supported or refuted by the results of the study. Prespecified hypotheses are those that are developed and stated before the study is initiated. Analyses performed to address new hypotheses suggested by inspection of the data are exploratory and should be reported as such, as it affects their interpretation (6).

Materials and Methods (Items 2–11)

Patients

Item 2: Patient Characteristics

A description of the specific population from which data were collected is needed to place the study in a clinical context. The source of the patients should be specified, for example, from a clinical trial population, a health care system, a clinical practice, or all hospitals in a certain geographic area. Furthermore, patient eligibility and exclusion criteria, usually based on clinical or pathologic characteristics, should be clearly stated.

Item 3: Patient Treatments

It is critical to report which treatments the patients received and at what time relative to specimen collection. This is because different treatments might alter the disease course in different ways, and biological characteristics of a tumor may be altered by therapies to which it was exposed before the specimen collection (Item 4). Conversely, the impact of a treatment might depend on the biological characteristics of the tumor, the essence of predictive marker research (Supplementary Box 1, available online; all five boxes can be found in the Supplementary Material, available online).

Item 4: Specimen Characteristics

Authors should report what types of specimens were used for the marker assays: tumor tissue; tumor cells or tumor DNA isolated from blood, bone marrow, urine, or sputum; serum or plasma. As much information about the source of the specimen as possible should be included, for example, whether a tumor sample was obtained at surgery or from a biopsy procedure such as core needle biopsy or fine needle aspirate. For patients with advanced disease, it should be clearly stated whether tumor samples assayed came from the primary tumor site or from a current metastatic lesion and whether the patient had received any prior cancer-directed therapies (Item 3).

Information about specimen processing and handling might only be ascertainable indirectly through knowledge of standard operating procedures of the pathology departments involved. The "Biospecimen Reporting for Improved Study Quality" (BRISQ) guidelines provide comprehensive recommendations for what information should be reported regarding specimen characteristics and methods of specimen processing and handling when publishing research involving the use of biospecimens (7).

Criteria for acceptability of biospecimens for use in marker studies, such as percent tumor cellularity, RNA integrity number, percent viable cells, or hemolysis assessment, should be established before initiating the study. These criteria should be reported, along with the percentage of specimens that met the criteria and therefore were included in the study. The numbers of specimens examined at each stage in the study should be recorded in the suggested flowchart and, particularly, in a study profile (Item 12).

Item 5: Assay Methods

Assay methods should be reported with a level of detail that would enable another laboratory to reproduce the measurement technique. The term "assay" is used broadly to mean any measurement process applied to a biological specimen that yields information about that specimen. It has been demonstrated for many markers that different measurement techniques can produce systematically different results (8–10).

It is important to report the minimum amount of specimen that was required to perform the assay (eg, a 5 micron section, 5 micrograms of DNA). Any additional specimen preprocessing steps required (eg, microdissection, polymerase chain reaction amplification) should also be stated. Any strategies used to address measurement error should be reported.

It is important to report whether marker assessments were made blinded to clinical outcome, as there may be a risk of introducing bias when a patient's clinical outcome is known by the individual making the marker assessment.

Study Design (Items 6–9)

Item 6a: Patient Selection

An understanding of how patients were selected is critical to identifying potential biases and to determine an appropriate statistical analysis approach. Reliance on a label of "prospective" or "retrospective" is inadequate because these terms are ill defined (11). It should be stated whether patients were recruited prospectively as part of a planned marker study or were identified retrospectively through search of an existing database. Authors should describe exactly how and when clinical, pathologic, and follow-up data were collected for the identified patients.

In situations where more complex case selection strategies are used, those approaches must be carefully described. Given the small size of most prognostic studies (Item 9), it is sometimes desirable to perform stratified sampling to ensure that important subgroups (eg, different stages of disease or different age groups) are represented.

Item 6b: Time Period

Knowing when a study took place and over what period participants were recruited places a study in historical context regarding typical clinical care. In most studies where the outcome is the time to an event, follow-up of all participants is ended on a specific date. This date should be given, and it is also useful to report the median duration of follow-up.

Item 7: Clinical End Points

Survival analysis is based on the elapsed time from a relevant time origin, often the date of diagnosis, surgery, or random assignment, to a clinical end point. That time origin should always be specified.

Most prognostic studies in cancer examine a few end points, mainly death, recurrence of disease, or both, but these end points are often not clearly defined (Supplementary Box 2, available online). Analyses of time to death may be based on either death from any cause or only cancer-related deaths. The end point should be defined precisely, and not referred to just as "survival" or "overall survival." If there was a specific rationale for choosing the primary clinical end point, it should be stated.

Item 8: Candidate Variables

It is important for readers to know which marker measurements or other clinical or pathological variables were initially considered for inclusion in statistical models, including variables not ultimately used. The reasons for lack of inclusion of variables should be addressed, for example, variables with large amounts of missing data (Supplementary Box 3, available online). Authors should fully define all variables and explain how they were measured.

All of the variables considered for standard survival analyses should be measured at or before the study time origin (eg, the date of diagnosis) (12,13). Variables measured after the time origin, such as experiencing an adverse event, should more properly be considered outcomes, not predictors (14).

Item 9: Rationale for Sample Size

Sample size has generally received little attention in prognostic studies, perhaps because these studies are often performed using preexisting specimen collections or data sets. The most important factor influencing power and sample size requirement for a study with a time-to-event outcome is the number of observed events (effective sample size), not the number of patients. For a binary outcome, the effective sample size is the smaller of the two frequencies "event" or "nonevent."

Choice of an end point that includes recurrence as an event in addition to death will result in more observed events and higher power than death alone, an important reason why disease-free survival is often preferred as an end point (15). Authors should explain the considerations that led to the sample size, whether based on a formal statistical calculation or determined by practical considerations, such as the availability of tumor samples or cost.

Sample size requirements will differ depending on the goal of the study and stage of development of the marker. For markers early in the development process, investigators may be most interested in detecting large effects unadjusted for other variables. As a prognostic marker advances in the development process, it will typically be studied in the context of regression models containing other clinically relevant variables, as discussed in Item 10d. These situations will require larger sample sizes to account for the diminished size of marker effects adjusted for other (potentially correlated) variables.

Statistical Analysis Methods (Items 10–11)

Item 10: All Statistical Methods

All of the statistical methods used should be reported. A sound general principle is "describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results" (16). It is additionally valuable if the reader can understand the reasons for the approaches taken.

For prognostic marker studies, there are many possible analysis strategies, and choices are made at each step. If many different analyses are performed, and only those with the best results are reported, inferences can be very misleading. It is essential to give a comprehensive overview of the range of analyses that have been undertaken in the study (see also discussion of the REMARK profile in Item 12). Details can be given in the supplementary material.

Analysis of a marker's prognostic value is usually more complex than the analysis of a randomized trial, for which statistical principles and methods are well established and primary analysis plans are generally prespecified. Nonetheless, when there are prespecified analyses in marker studies, it is helpful to identify them in order to distinguish them from exploratory analyses.

Reporting of key features of an analysis is important to allow readers to understand the reasons why the specific approach was chosen and to interpret the results. In the following sections, we consider specific aspects of analyses under eight

headings. Not all aspects will be relevant for some studies. More extensive discussions of statistical analysis methods for binary outcome and for survival data can be found elsewhere (17–21).

a. Preliminary Data Preparation

Some assessment of the data quality usually takes place before the main statistical analyses, which address the primary objectives, and some data values may be modified or removed if they are deemed unreliable. Selective deletion or modification of data values with the intent to make results look more striking must be avoided, as this would introduce bias. Any data manipulations and premodeling decisions should be reported (22–24).

b. Association of Marker Values With Other Variables

Early steps in an analysis may include examination of the relationship of the marker to other variables being considered in the study. These variables might include established clinical, pathologic, and demographic covariates (Items 13 and 14). Analytic or graphical methods used to conduct these examinations should be explained.

c. Methods to Evaluate a Marker's Univariable Association With Clinical Outcome

The first evaluation of the marker's value will usually be conducted without adjustment for additional variables, that is, a univariable analysis. The analysis strategy used should be described. Any variable codings or groupings, or transformations of continuous values applied to the marker variable or any other variables, should be stated to allow for proper interpretation of the estimated associations (Item 11; Supplementary Box 4, available online).

d. Multivariable Analyses

Univariable analyses are useful but, except in early studies, are generally insufficient because of the possible relationship of the marker with other variables. Thus the prognostic value of the marker after adjustment for established prognostic factors, as estimated from a multivariable model (Item 17), will be of major interest. To facilitate comparison of unadjusted and adjusted measures of association, it is helpful to report results from univariable analyses that used the same general approach as used for multivariable analysis. Multivariable methods can also be used to build prognostic models involving combinations of several candidate markers or even many hundreds of markers (eg, gene expression microarray data).

e. Missing Data

Almost all prognostic studies have missing marker or covariate data for some patients. Authors should report the number of missing values for each variable of interest. Including only cases with complete data may greatly reduce the sample size and potentially lead to biased results if the likelihood of being missing is related to the true values (Supplementary Box 3, available online) (25,26). Modern statistical methods exist to allow estimation (imputation) of missing observations.

f. Variable Selection

The main multivariable model may sometimes be prespecified, which helps to avoid biases caused by data-dependent model selection. More often, however, many candidate variables are available, and some type of variable selection procedure is sensible in order to derive simpler models that are easier to interpret and may be more generally useful (27). It is particularly important to state if the variables included in a model were determined using variable selection procedures (Item 16).

Sometimes several multivariable models containing different subsets of variables are considered. The rationale for these choices and details of model selection strategies used should be described. The REMARK profile can provide a concise summary of all analyses performed (Item 12).

g. Checking Model Assumptions

Any statistical model, univariable or multivariable, makes certain assumptions about the distributions of variables or the functional relationships between variables. Assumptions need careful checking to ensure they are not seriously violated, and methods used for this purpose should be reported. Alternative models evaluated for sensitivity analyses should also be described (Item 18).

h. Model Validation

The strongest evidence for the validity of results is confirmation of the findings on data not involved in the original analysis (28,29). The ideal approach is to confirm findings from the main (final) model on completely independent data, preferably collected by different investigators. If successful, this approach would indicate that the results are transportable to other settings. This would be a type of "external validation." A completely independent data set (a "similar" study) often will not be available, but "internal" validation procedures such as cross-validation, bootstrapping, or other data resampling methods (27,30) are useful to give insights into critical issues such as bias of regression parameter estimates, overoptimism of prognostic model discriminatory ability, or stability of the model derived (see also Item 18).

Item 11: Handling of the Marker

A central question is how to analyze continuous variables, including how to incorporate them in a multivariable model. Often this applies to the marker and to several standard variables, such as age and tumor size.

Two main approaches are to keep the variables as continuous (but not necessarily assume a linear relation with the outcome) or to group the data into categories. Although categorization is ubiquitous in cancer studies, there are some major concerns about that approach (Supplementary Box 4, available online).

The authors should report how each continuous variable was incorporated into the analyses. For categorized variables, cut-points used should be specified, along with an explanation of how they were chosen. For continuous variables, authors should clarify whether the data were kept on the original scale or how they were transformed.

Results (Items 12–18)

Data (Items 12–13)

Item 12: Flow of Patients and Multiplicity of Analyses

The interpretation of prognostic studies depends on having a good understanding of the patients included, the methods used and the analyses conducted, and the amount of data available

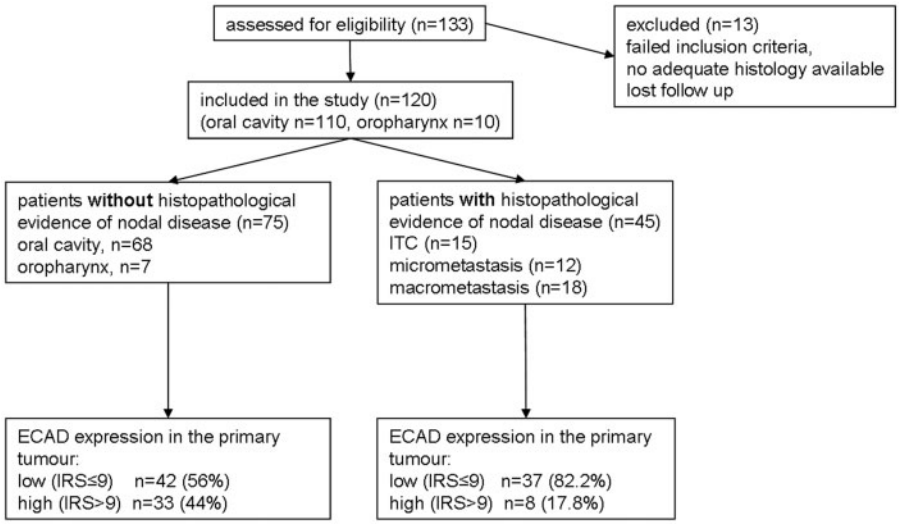

Figure 2. Example of a participant flow diagram (31). ECAD = E-Cadherin; IRS = Intensity Reactivity score; ITC = isolated tumour cells.

a) Patients, treatment and variables

Study and marker	Remarks
Marker (If non-binary: how was marker analyzed? continuous or categorical. If categorical, how were cutpoints determined?)	M = ploidy (diploid, aneuploid)
Further variables (variables collected, variables available for analysis, baseline variables, patient and tumor variables)	v1 = age, v2 = histologic type, v3 = grade, v4 = residual tumor, v5 = stage, v6 = ascites[a], v7 = estrogen[a], v8 = progesterone[a], v9 = CA-125[a]

Patients	n	Remarks
Assessed for eligibility	257	*Disease*: Advanced ovarian cancer, stage III and IV *Patient source*: Surgery 1982 to 1990, University Hospital Freiburg *Sample source*: Archived specimens available
Excluded	73	General exclusion criteria[b], non-standard therapy[b], coefficient of variation >7%[b]
Included	184	Previously untreated. *Treatment*: all had platinum based chemotherapy after surgery
With outcome events	139	Overall survival: death from any cause

b) Statistical analyses of survival outcomes

Analysis	Patients	Events	Variables considered	Results/remarks
A1: Univariable	184	139	M, v1 to v5	Table 2, Figure 1
A2: Multivariable	174	133	M, v1, v3 to v5	Table 3 [v2 omitted because many missing data; Backward selection, see text]
A3: Effect for ploidy adjusted for v4	184	139	M, v4	Figure 2 [Based on result of A2]
A4: Interaction: ploidy and stage	175	133	M, v1, v2, v4, v5	See text
A5: Ploidy in stage subgroups				
v5 = III	128	88	M	Figure 3
v5 = IV	56	51	M	Figure 4

[a]Not considered for survival outcome as these factors are not considered as 'standard' factors and/or number of missing values was relatively large;
[b]values not given in the paper.
doi:10.1371/journal.pmed.1001216.t002

Figure 3. Example of the REMARK profile using data from a study of ploidy in patients with advanced ovarian cancer (2,3).

at each stage. Typically, several analyses are conducted. To avoid selective reporting, which gives rise to biased results and biased interpretation, it is important to report all analyses, preferably by giving key information via a structured display. We suggest two complementary displays that authors can use to summarize key aspects of a prognostic study.

First, a flow diagram provides an easy-to-follow view of the major changes in the study sample as the study proceeds. An example is shown in Figure 2. Second, the two-part REMARK profile summarizes key aspects, especially the derivation of the sample, and details of the analyses performed (32). We provide an example in Figure 3.

The upper part gives details about how the marker of interest was handled in the analysis and which further variables were available. In addition, key information is provided in this part about the patient population, inclusion and exclusion

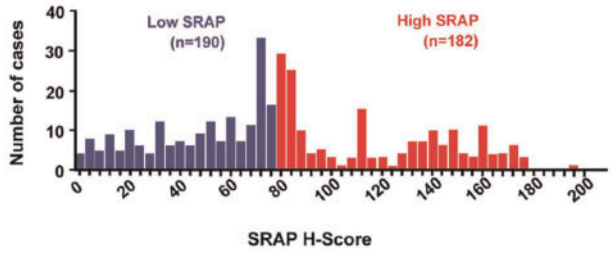

Figure 4. Frequency distribution of steroid receptor RNA activator protein H-scores in 372 breast tumors, showing median of 76.67 used to delineate low and high subgroups (33). SRAP = steroid receptor RNA activator protein.

criteria, number of eligible patients, and numbers of events for each outcome in the full data set. The number and reasons for patients excluded are given.

The lower part shows details of each analysis performed. In addition to the number of patients and events in a study, it is important to know the amount of data available for each analysis. Missing values (Supplementary Box 3, available online) are common, and the complete case analysis is the most widely used method despite its potential drawbacks (Item 10e). The number of patients and events will often vary across analyses according to the outcome measure, the choice of adjusting variables, and whether the analysis was restricted to a subgroup. These numbers are key determinants of the statistical reliability of any analysis. A standard format for reporting all analyses performed would be extremely helpful to reduce selective reporting bias (Supplementary Box 5, available online) and is strongly recommended.

Item 13: Distribution of Demographic Characteristics

Inclusion and exclusion criteria (Item 2) describe the target patient population. The group of patients included in the study is a sample from that population. Distributions of basic demographic variables and standard prognostic variables should be reported to characterize the group of patients that was studied. The number of patients with missing values should be reported for each variable. These demographic and standard prognostic variables are often the variables considered for inclusion in multivariable analyses (Item 8). A thorough description of the distribution of the marker of interest should also be provided, preferably graphically (Figure 4).

Analysis and Presentation (Items 14–18)

Item 14: Relation Between Marker and Standard Variables

The association of the tumor marker with standard prognostic variables should be described. A new marker is most useful if it provides clinically important information beyond that given by existing prognostic variables or indices or it offers an advantage because it is easier or less expensive to measure or quantify. Often a new marker has at least a modest association with some other standard prognostic markers (Item 10d). Graphical displays can be particularly helpful in conveying the nature of associations.

Item 15: Univariable Analyses

A marker's simple association with outcome should be shown first, without adjustment for other clinical or pathologic characteristics, to get some feeling about its prognostic strength. Precision of the estimates should be indicated, for example, by providing confidence intervals. *P* values may also be presented. Similar analyses are useful for showing the relation to outcome of all other variables being assessed.

Univariable measures of association of the marker with outcome and differences between Kaplan-Meier curves might be heavily influenced by other prognostic variables that are correlated with the marker. However, those analyses are still useful to report as they provide a baseline against which to compare measures of association adjusted for other variables (Item 16).

Item 16: Multivariable Analyses

As a tumor's biological characteristics are not controllable experimentally, a study examining the prognostic value of a tumor marker is subject to the challenges inherent in analysis of observational studies such as adjustment for the effect of potential confounding factors. Some of these factors are standard variables that are generally accepted as being related to prognosis, whereas others might be candidate variables that happen to be available. Any of these variables might be considered for inclusion in multivariable models (Items 12 and 17).

Often the multivariable analysis involves model building that begins with what we will designate as the "full model" and after several data-dependent modeling steps may result in identification of a "final model." The full model is a model containing all the available candidate variables (Item 8). Usually the full model contains too many variables to be readily interpretable and reporting results of the full model is often nonessential. Multivariable methods can be used to derive a (sparser) prognostic model (Item 10d) (34). The final model, which is a more parsimonious model, should be presented. The "standardized model" (for explanation, see Item 17) is another important multivariable model that should be examined. The precision of estimated effects should be provided (eg, confidence intervals), at least for the final model for all variables in the model.

Item 17: Adjustment for Standard Variables

In many clinical situations, some standard variables have previously been demonstrated to have prognostic value and are generally measured. Typical standard variables include stage and its constituent elements, such as tumor size and nodal status, and sometimes patient demographic variables such as age or sex. It is important to evaluate whether the new marker maintains some association with clinical outcome after accounting for these standard prognostic variables. Evaluation of a marker's effect adjusted for standard variables is generally accomplished by examining what we will call the "standardized model." Results of fitting this standardized model should be explicitly reported, as they facilitate the comparison of estimated effects of the marker across studies.

Item 18: Further Investigations

Results of many prognostic studies rely on the validity of statistical models, and inherent in any model are certain assumptions (eg, proportional hazards, linear effects of covariates,

missing data mechanisms). Prognostic analysis results will have greater credibility if arguments can be made that the modeling assumptions are justifiable or that the results are not unduly sensitive to certain assumptions. At least by providing a brief summary, the report should mention the results obtained from any additional check of assumptions or investigation of robustness of results (Item 10g; Supplementary Box 4, available online). In some situations, modeling assumptions cannot be empirically verified. Sensitivity analyses can illustrate whether alternative assumptions still lead to similar conclusions.

Discussion (Items 19 and 20)

Item 19: Results in Context of Prespecified Hypotheses

The discussion is the appropriate section to interpret the data and suggest further research that might be needed. It should begin by briefly restating the purpose of the study and recalling any prespecified hypotheses. A simple summary of the major findings should follow.

A critical evaluation of the reported results should include acknowledgment of biases or inconsistencies in the data, limitations of the assay methods, limitations of the design or data analysis methods, and assumptions made. Any unexpected findings should be identified. A thorough and open discussion will maximize the value of the study to the broader community regardless of the study's results.

Item 20: Implications

The rationale for studying any marker is to gain relevant information about the biology of the disease, to find new tools to aid in clinical decision-making, or to develop new treatments. Observation of a statistically significant association between a marker and an outcome may be encouraging, but in the long term, the difference in outcome should have clinically important implications for patient care. Note that even if a prognostic marker does not provide added value to existing prognostic information, it may nevertheless be useful if it can be assessed more easily, more reproducibly, or at lower cost than markers currently used to provide clinically meaningful information.

Final Comments

Physicians seek information about tumor markers to inform therapeutic decisions for individual patients. Availability of a marker that can distinguish subsets of patients with varying prognosis may also influence the design of clinical trials. In order for information about the utility of tumor markers to be appropriately evaluated, the methods used to study the markers and the results generated must be fully reported. The REMARK recommendations were designed to help authors ensure that reports of their tumor marker studies contain the information that readers need. Good reporting reveals the strengths and weaknesses of a study and facilitates sound interpretation and application of study results. The REMARK recommendations may also aid in planning new studies and may be helpful for peer reviewers and editors in their evaluation of manuscripts.

Although we have primarily focused on studies of single prognostic markers, most of the recommendations apply equally to other types of prognostic studies, including studies of multiple markers and studies of markers to predict response to treatment. A guideline for reporting studies that develop or validate a multivariable prediction model (TRIPOD statement) was published recently (35,36).

The REMARK recommendations were initially targeted at tumor marker prognostic studies. However, the recommendations are equally relevant to specialties other than cancer and have begun to be used more widely (37–41). We hope that this document will encourage further use of REMARK across many specialties.

REMARK is not intended to dictate standards for assessing the quality of research, and it should not be used as such. However, it can be a useful tool to help assemble the information needed in order to assess the quality and relevance of research.

In the Supplementary Materials (available online), we have added a list with check boxes, proposing three possible responses (see page/paragraph, not applicable, not available) (Supplementary Figure 1, available online). We propose that journals use this checklist for the first submission.

Several cancer journals ask authors to follow the REMARK recommendations in their instructions to authors; we encourage more journals to follow this example. Recent research has shown that more pressure on authors, reviewers, and editors seems to be needed (4,5,42).

To improve reporting, it is necessary that authors consult REMARK before submitting their paper. Journals may consider using our table with check boxes (Supplementary Appendix, available online), proposing three possible responses (see page/paragraph, not applicable, not available), for the first submission. A useful further enhancement would be to implement pop-up windows containing the short text associated with each item. This may remind researchers of the relevance and expectations of information to be reported in the manuscript.

Up-to-date information on REMARK and numerous other reporting guidelines can be found on the website of the EQUATOR Network (www.equator-network.org/library/).

Funding

DGA was supported by a grant from Cancer Research UK (C5529) in the development of REMARK.

Notes

Affiliations of authors: Institute for Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany (WS); Cancer Diagnosis Program (SET [retired], MMC) and Biometric Research Program (LMM), Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD; Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK (DGA).

The funder had no role in the writing of this Commentary or the decision to submit it for publication.

The authors declare that they have no competing interests.

Four of the authors (WS, SET, LMM, DGA) jointly wrote and agreed to the original text over several years, with MMC's assistance, including several face-to-face meetings and numerous telephone conference calls. All the authors jointly wrote and agreed to the abridged text over several more years, including many telephone conferences.

We are grateful to the US National Cancer Institute and the European Organization for Research and Treatment of Cancer for their support of the first National Cancer Institute–European Organisation for Research and Treatment of Cancer International Meeting on Cancer Diagnostics, from which the

idea for these recommendations originated. We thank the US National Cancer Institute for continued support of the work on REMARK publications. The original paper (1) was published in 2005 simultaneously in the *Journal of the National Cancer Institute*, the *British Journal of Cancer*, the *European Journal of Cancer*, the *Journal of Clinical Oncology*, and *Nature Clinical Practice Oncology*. In 2006, it was also published in *Breast Cancer Research and Treatment* and *Experimental Oncology*. In 2012, we published the related "Explanation and Elaboration" paper (2,3), which is the basis for this abridged version.

We thank Tim Haeussler and Lisa Martin for administrative assistance.

References

1. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst*. 2005;97(16): 1180–1184.
2. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): Explanation and elaboration. *PLoS Med*. 2012;9(5):e1001216.
3. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting recommendations for tumor marker prognostic studies (REMARK): Explanation and elaboration. *BMC Med*. 2012;10(1):51.
4. Jankova L, Dent OF, Molloy MP, et al. Reporting in studies of protein biomarkers of prognosis in colorectal cancer in relation to the REMARK guidelines. *Proteomics Clin Appl*. 2015;9(11–12):1078–1086.
5. Sekula P, Mallett S, Altman DG, Sauerbrei W. Did the reporting of prognostic studies of tumour markers improve since the introduction of REMARK guideline? A comparison of reporting in published articles. *PLoS One*. 2017;12(6): e0178531.
6. Vandenbroucke JP, von Elm E , Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *PLoS Med*. 2007;4(10):e297.
7. Moore HM, Kelly AB, Jewell SD, et al. Biospecimen reporting for improved study quality (BRISQ). *Cancer Cytopathol*. 2011;119(2):92–101.
8. McShane LM, Aamodt R, Cordon-Cardo C, et al. Reproducibility of p53 immunohistochemistry in bladder tumors. National Cancer Institute, Bladder Tumor Marker Network. *Clin Cancer Res*. 2000;6(5):1854–1864.
9. Romero H, Schneider J. Different detection rates of HER-2/NEU overexpression in ovarian carcinoma using two different commercially available detection kits. *Eur J Cancer*. 1995;31A(6):1020–1021.
10. Press MF, Hung G, Godolphin W, Slamon DJ. Sensitivity of HER-2/neu antibodies in archival tissue samples: Potential source of error in immunohistochemical studies of oncogene expression. *Cancer Res*. 1994;54(10):2771.
11. Vandenbroucke JP. Prospective or retrospective: What's in a name? *BMJ*. 1991; 302(6771):249–250.
12. Beyersmann J, Wolkewitz M, Schumacher M. The impact of time-dependent bias in proportional hazards modelling. *Statist. Med*. 2008;27:6439–6454.
13. van Walraven C, Davis D, Forster AJ, Wells GA. Time-dependent bias was common in survival analyses published in leading clinical journals. *J Clin Epidemiol*. 2004;57(7):672–682.
14. Rochon J. Issues in adjusting for covariates arising postrandomization in clinical trials. *Drug Inf J*. 1999;33:1219–1228.
15. Feinstein AR. *Multivariable Analysis: An Introduction*. New Haven, CT: Yale University Press; 1996.
16. International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. http://icmje.org/icmje-recommendations.pdf. Updated December 2016. Accessed September 2017.
17. Royston P, Sauerbrei W. *Multivariable Model-Building: A Pragmatic Approach to Regression Analysis Based on Fractional Polynomials for Continuous Variables*. Chichester, UK: Wiley; 2008.
18. Hosmer DW, Lemeshow S, May S. *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. 2nd ed. Hoboken, NJ: Wiley-Interscience; 2008.
19. Schumacher M, Holländer N, Schwarzer G, Binder H, Sauerbrei W. prognostic factor studies. In: J HA Crowley, ed. *Handbook of Statistics in Clinical Oncology*. 3rd ed. Boca Raton, FL: Chapman & Hall/CRC; 2012:416–469.
20. Vittinghoff E, Glidden DV, McCulloch CE, Shiboski SC. *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*. 2nd ed. New York: Springer; 2012.
21. Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. 2nd ed. Cham: Springer International Publishing Switzerland; 2015.
22. Bailar JC. How to distort the scientific record without actually lying: Truth, and the arts of science. *Eur J Oncol*. 2006;11:217–224.
23. Chatfield C. Confessions of a pragmatic statistician. *J R Stat Soc D Sta*. 2002; 51(Part 1):1–20.
24. Altman DG. Preparing to analyse data. In: DG Altman, ed. *Practical Statistics for Medical Research*. Boca Raton, FL: Chapman & Hall/CRC; 1991:122–151.
25. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ*. 2009;338: b2393.
26. Burton A, Altman DG. Missing covariate data within cancer prognostic studies: A review of current reporting and proposed guidelines. *Br J Cancer*. 2004; 91(1):4–8.
27. Sauerbrei W. The use of resampling methods to simplify regression models in medical statistics. *Appl Stat*. 1999;48:313–329.
28. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: Validating a prognostic model. *BMJ*. 2009;338:b605.
29. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med*. 2000;19(4):453–473.
30. Davison AC, Hinkley DV. *Bootstrap Methods and Their Application: Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge University Press; 1997.
31. Huber GF, Züllig L, Soltermann A, et al. Down regulation of E-Cadherin (ECAD) - a predictor for occult metastatic disease in sentinel node biopsy of early squamous cell carcinomas of the oral cavity and oropharynx. *BMC Cancer*. 2011;11:217:1–8.
32. Mallett S, Timmer A, Sauerbrei W, Altman DG. Reporting of prognostic studies of tumour markers: A review of published articles in relation to REMARK guidelines. *Br J Cancer*. 2010;102(1):173–180.
33. Yan Y, Skliris GP, Penner C, et al. Steroid receptor RNA activator protein (SRAP): A potential new prognostic marker for estrogen receptor-positive/ node-negative/younger breast cancer patients. *Breast Cancer Res*. 2009;11(5): R67.
34. Sauerbrei W, Royston P, Bojar H., Schmoor C., Schumacher M. Modelling the effects of standard prognostic factors in node-positive breast cancer. *Br J Cancer*. 1999;79(11–12):1752–1760.
35. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1–W73.
36. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med*. 2015;162(1):55–63.
37. Mupparapu M, Kim IH. Calcified carotid artery atheroma and stroke: A systematic review. *J Am Dent Assoc*. 2007;138(4):483–492.
38. Whiteley W, Chong WL, Sengupta A, Sandercock P. Blood markers for the prognosis of ischemic stroke: A systematic review. *Stroke*. 2009;40(5): 380–389.
39. Hemingway H, Philipson P, Chen R, et al. Evaluating the quality of research into a single prognostic biomarker: A systematic review and meta-analysis of 83 studies of C-reactive protein in stable coronary artery disease. *PLoS Med*. 2010;7(6):e1000286.
40. Sigounas DE, Tatsioni A, Christodoulou DK, Tsianos EV, Ioannidis JPA. New prognostic markers for outcome of acute pancreatitis: Overview of reporting in 184 studies. *Pancreas*. 2011;40(4):522–532.
41. Sutaria S, Philipson P, Fitzpatrick NK, et al. Translational phases of evidence in a prognostic biomarker: A systematic review and meta-analysis of natriuretic peptides and the prognosis of stable coronary disease. *Heart*. 2012; 98(8):615–622.
42. Caron JE, March JK, Cohen MB, Schmidt RL. A survey of the prevalence and impact of reporting guideline endorsement in pathology journals. *Am J Clin Pathol*. 2017;148(4):314–322.
43. Hudis CA, Barlow WE, Constantino JP, et al. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: The STEEP System. *J Clin Oncol*. 2007;25(15):2127–2132.
44. Donders ART, van der Heijden GJMG, Stijnen T, Moons KGM. Review: A gentle introduction to imputation of missing values. *J Clin Epidemiol*. 2006;59(10): 1087–1091.
45. Greenland S. Tests for interaction in epidemiologic studies: A review and a study of power. *Stat Med*. 1983;2(2):243–251.
46. Royston P, Sauerbrei W. A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. *Stat Med*. 2004;23:2509–2525.
47. 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(11):829–835.
48. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: A bad idea. *Stat Med*. 2006;25(1):127–141.
49. Riley RD, Sauerbrei W, Altman DG. Prognostic markers in cancer: The evolution of evidence from single studies to meta-analysis, and beyond. *Br J Cancer*. 2009;100(8):1219–1229.
50. Ioannidis JPA. Why most discovered true associations are inflated. *Epidemiology*. 2008;19(5):640–648.
51. Andre F, McShane LM, Michiels S, et al. Biomarker studies: A call for a comprehensive biomarker study registry. *Nat Rev Clin Oncol*. 2011;8(3): 171–176.