

Challenges and Standards in Reporting Diagnostic and Prognostic Biomarker Studies

Francisco Azuaje, Ph.D.¹, Yvan Devaux, Ph.D.¹, and Daniel Wagner, Ph.D., M.D.^{1,2}

Abstract

The discovery of novel biomarkers is a crucial goal in translational biomedical research. A complete and accurate reporting of biomarker studies, including quantitative prediction models, is fundamental to improve research quality and facilitate their potential incorporation into clinical practice. This paper reviews key problems, guidelines, and challenges in reporting biomarker studies, with an emphasis on diagnostic and prognostic applications in cardiovascular research. Recent advances and recommendations for aiding in peer review, research quality assessment, and the reproducibility of findings, such as diagnostic biosignatures, are discussed. An examination of research recently published in the area of cardiovascular biomarkers was implemented. Such a survey, which was based on a sample of papers deposited in PubMed Central, suggests that there is a need to improve the documentation of biomarker studies in terms of information completeness and clarity, as well as the application of more rigorous quantitative evaluation techniques. There is also room for improving practices in reporting data analysis and research limitations. This survey also suggests that, in comparison with other research areas, the cardiovascular biomarker research domain may not be taking advantage of existing standards for reporting diagnostic accuracy. The review concludes with a discussion of the challenges and recommendations.

Keywords: cardiovascular disease, computers, diagnosis, translational research, “omic” biomarkers

Introduction

The development of novel biomarkers for diagnostic and prognostic applications is currently at the front of a renewed bidirectional interaction between basic and clinical research. The combination of different “omic” approaches to measure molecular activity at different levels of complexity is accelerating the pace at which new advances relevant to translational research are being reported. Therefore, an accurate documentation and reporting of biomarker studies is fundamental to improve research quality and facilitate their potential incorporation into clinical practice.

According to the Biomarkers Definitions Working Group of the US National Institutes of Health, a biomarker is “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.”¹ Three main types of biomarkers may be defined: type 0 biomarkers, which correlate with the emergence or development of a disease; type 1 biomarkers, which reflect the action of a therapeutic intervention; and type 2 biomarkers, which may be used as surrogate clinical endpoints. Diagnostic and prognostic biomarkers fall into the type 0 category. The former refers to biomarkers that predict the occurrence of a disease in subjects suspected of having the disease. The latter aims to predict the outcome (e.g., complications and clinical response) of a patient exhibiting disease symptoms.

A complete and accurate reporting of biomarker studies goes beyond the listing of biosignatures (e.g., genes), the description of standard experimental protocols, or the presentation of statistical significance probability values. It requires a more rigorous and detailed specification of different qualitative and quantitative aspects. These aspects are required for assisting in quality assessment, peer review, and external validation of results. Good reporting practices should provide enough information to allow the reproducibility of experimental findings and the reconstruction of quantitative prediction models. Incomplete or inaccurate reporting of findings, quantitative models, or predictive evaluation may potentially lead to misinterpretation of hypotheses, exaggerated claims, and inadequate interpretation

of quantitative analysis results. Thus, a more rigorous reporting of diagnostic and prognostic models will prevent premature efforts in translational research, and contribute to the overall goal of improving research quality and human health.

The following sections will include an overview of “omic” biomarker research, a discussion on the challenges and limitations in the reporting of biomarker studies, community-based recommendations, and an examination of the cardiovascular research literature.

“Omic” biomarkers

“Omic” biomarker studies may be categorized on the basis of the complexity of the data investigated. Under such a scheme, four major types of studies can be defined according to: (a) the number of biomarkers that are (or can be potentially) used as inputs to a prediction model and, (b) the characteristics of these data. Single marker studies aim to report the predictive capacity of a single biomarker, such as a protein or risk factor, which could be used to distinguish between risk or disease groups. Multiple marker studies emphasize the description of a list of different biomarkers with potential diagnostic or prognostic value. This may involve the identification of several genes or proteins that may be applied as potential predictors of disease or clinical outcome. A typical approach is to identify a list of biomarkers by performing quantitative analysis based on assumptions of biomarker independence, followed by additional procedures to find potentially meaningful associations, for example, data clustering. Integrated prediction models combine different biomarkers, which may be derived from different experiments or bioinformatic/statistical techniques. A distinguishing methodological feature is that biomarkers represent the inputs to an integrated mathematical model (e.g., variable regression and sample classification) that takes into account both between-marker and marker-outcome relationships. This category may also be divided according to the characteristics and diversity of data sources. Thus, the biomarkers may originate from the

¹Laboratory of Cardiovascular Research, Centre de Recherche Public-Santé, Luxembourg; ²Division of Cardiology, Centre Hospitalier de Luxembourg, Luxembourg.

Correspondence: F Azuaje (francisco.azuaje@crp-sante.lu)

DOI: 10.1111/j.1752-8062.2008.00075.x

Type	Characteristic and application	Examples
SM	The predictive capacity of a single biomarker, e.g., protein or risk factor, to distinguish between risk or disease groups is analyzed.	Protein marker for heart failure after myocardial infarction ² Polymorphism in a gene associated with metabolic syndrome and mortality ³
MM	List of different biomarkers, e.g., genes or proteins, is analyzed independently as potential predictors of disease or clinical outcome, followed by additional analyses, e.g., data clustering.	Plasma levels of proteins associated with coronary artery calcification ⁴ Gene expression patterns associated with arterial hypertension ⁵
IPM-Hom	The combination of different biomarkers derived from the same resource type, e.g., gene expression only, is implemented using a computational prediction model.	Combination of multiple SNPs polymorphisms to detect esophageal cancer ⁶ Integration of multiple gene expression variables to estimate coronary collateralization ⁷
IPM-Het	Biomarkers derived from diverse types of data, e.g., gene expression and pathway information, are integrated into a computational prediction model.	Integration of information extracted from medical records and gene expression measurements to identify biomarkers of aging ⁸ Integration of information extracted from gene expression profiles, gene annotation, and functional networks to identify dilated cardiomyopathy ⁹

SM = single marker; MM = multiple markers; IPM = integrated prediction models; Hom = homogeneous information sources; Het = heterogeneous information sources.

Table 1. Main types of “omic” diagnostic or prognostic biomarker studies according to model complexity.

same “omic” resource type, for example, gene expression only, or represent diverse types of data, for example, gene expression and pathway information. *Table 1* summarizes this categorization together with corresponding examples of recent studies and applications.

Diagnostic and prognostic biomarkers may also be categorized in terms of the type of “omic” information that they represent. The major information domains are: genomics, transcriptomics, proteomics, and metabolomics. A significant expansion of genetic variability studies over the past 3 years has led to a stream of diverse sets of gene–disease associations. This includes single nucleotide polymorphisms (SNPs) from candidate gene and genome-wide association studies.³ Transcriptomic-based biomarkers mainly comprised the identification of diagnostic or prognostic signatures using gene expression data for patient classification, risk estimation, and the discovery of treatment response sub-classes.¹⁰ Diagnostic and prognostic biomarkers originating from the plasma and serum proteome, as well as metabolite profiling, have also been investigated for risk assessment and subject classification.^{11,12} *Table 2* summarizes typical applications and

examples in the cardiovascular research area according to these information types. Reviews by Sawyers¹³ and Gerszten and Wang¹⁴ discuss different methodologies, as well as applications, in cancer and cardiovascular research, respectively.

Challenges in reporting and documenting biomarker studies

The proposal of reporting guidelines in different biomedical research areas has been mainly motivated by the need to improve the quality of published papers. Previous investigations have highlighted that some methodological errors are still commonly present, together with exaggerated or incorrect interpretations of key factors, such as the meaning of statistical results.^{15–17} It has been shown that, in many cases, the reporting of diagnostic models using, for example, gene signatures lacks important information to allow their external validation.¹⁸ Although the specification of standards has facilitated the incremental improvement of the reporting of biomarker studies over the past 5 years,¹⁹ there is still a need to promote better levels of documentation accuracy and completeness in diagnostic and prognostic research.

Examples of some of these challenges include the need to increase the acknowledgment of methodological limitations. This may require a discussion of the meaning and potential influence and implications of errors, as well as potential problems and bias of validation procedures. In the area of biomarker discovery, this is fundamental to understand results in context and achieve higher levels of credibility. But the need to improve the quality of reporting of limitations is also apparent in different areas of life and physical sciences. For instance, a recent survey involving articles from journals that received the highest impact factors in 2005 shows that less than 20% of the articles referred to potential limitations of their studies.¹⁷ Some journals have started to encourage authors to report limitations as part of their guidelines or even as specific formatting instructions.

Recent research has also pointed out that diagnostic and prognostic models, such as those based on microarray gene expression, may be lacking key information to allow external researchers to re-apply the model or reproduce predictions.¹⁸ It has been demonstrated that even an exact specification of quantitative

Omic information	Typical applications	Examples
Genomics	SNPs associated with disease or clinical outcome	SNPs associated with metabolic syndrome and mortality ⁶
Transcriptomics	Gene expression biosignatures for patient classification	Gene expression signature in peripheral blood to detect thoracic aortic aneurysm ¹⁰
Proteomics	Biosignatures, e.g., from plasma proteome, for risk assessment and disease diagnosis	Different biomarkers for ischemic heart disease ¹¹
Metabolomics	Metabolite profiling of serum or plasma for risk stratification	Biomarkers for myocardial ischemia ¹²

Table 2. Main types of “omic” information used for implementing diagnostic or prognostic biomarker studies in cardiovascular research.

Aspect or factor	Typical problem or limitation	Further reading
Study limitations	Limitations or caveats are not adequately acknowledged	Ref. 17
Data pre-processing	Sufficient information on data pre-processing (e.g., diagnostic biosignatures) is not provided to allow external validation or model reconstruction. Normalization or data transformation is not reported	Refs. 16 and 17
Hypothesis testing	Incomplete reporting of statistical significance test, e.g., only <i>p</i> values are reported. Incorrect interpretations of “significance” and test results	Ref. 15
Comparative analysis	Proposed diagnostic or prognostic models are not compared with reference or traditional models	Ref. 21
Diagnostic or prognostic accuracy evaluation	Biased and incorrect estimation of diagnostic accuracy. Lack of data sampling or cross-validation procedures	Ref. 21

Table 3. Examples of common shortcomings observed in the reporting and documentation of diagnostic or prognostic biomarker studies.

prediction models may not be sufficient to define a biomarker signature unambiguously. These problems may be in part the product of incomplete information about data pre-processing (i.e., normalization and data transformation), prediction model implementation, and cross-validation procedures applied. This is also supported by several studies showing how prediction outcomes may be affected by the selection of normalization methods.^{16,18}

Other aspects that deserve greater attention are the reporting of hypothesis testing results and their interpretation, software packages applied, and prediction evaluation including comparisons with alternative models. For instance, the practice of referring to “significant” and “not significant” results without providing statistic values and exact probability values may hinder an unbiased and correct interpretation of results.^{15,16} There is also the need to report prediction accuracy results in the context of data sampling and cross-validation.²⁰ Efforts should be maintained to ensure that diagnostic or prognostic studies report predictive performance results obtained when a model is built (trained) and tested on independent samples using standard cross-validation procedures, such as k-fold and leave-one-out cross-validation. *Table 3* outlines common shortcomings and aspects that deserve further attention when reporting diagnostic or prognostic biomarker studies.

Reporting guidelines and tools

Reporting guidelines are community-driven initiatives that offer specific advice on how to report methodologies and findings. In different areas, they foster a clear and accurate presentation of information. Different guidelines have been proposed that are relevant to biomarker development, and that cover different (sometimes overlapping) aspects of

scientific research and documentation. Their applicability ranges from clinical trials, the reporting of the accuracy of diagnostic studies, to meta-analyses.

The community of clinical trials pioneered efforts to improve the accurate reporting of health research through the Consolidated Standards of Reporting Trials (CONSORT), which has been widely endorsed by journals and international publishing organizations.²² The CONSORT illustrates how evolving, collective initiatives can improve the reporting of scientific research and its credibility. Moreover, these and ongoing efforts in other areas can assist authors, reviewers, and editors in increasing the readers’ confidence in the quality and validity of the studies published.²³

The Standards for Reporting of Diagnostic Accuracy (STARD) are of particular relevance to biomarker discovery research.²¹ The two main products of this initiative are a checklist for reporting diagnostic studies and recommendations for helping describe the study design and implementation using flow diagrams. The STARD checklist is organized as a set of sections (e.g., methods, statistical methods, and results) and topics with specific recommendations on

“what” should be reported. Since the publication of the STARD statement in January 2003, more than 200 scientific journals have endorsed it.²⁴

Other guidelines and tools for supporting the reporting of diagnostic studies are: the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool,²⁵ the Quality of Reporting of Meta-analyses (QUORUM) guidelines,²⁶ and the Meta-analysis Of Observational Studies in Epidemiology (MOOSE).²⁷ *Table 4* summarizes the applicability and content of these reporting guidelines. For additional information on these and related projects, the reader is referred to the EQUATOR Network

Guideline	Applicability	Examples of reporting factors covered
STARD	Reports of diagnostic accuracy studies	Guidelines for documenting participants, statistical analysis, results, and discussions
QUADAS	Reports of diagnostic studies based on systematic reviews	Conceived as an evaluation tool in which items and scores for the assessment of the quality of reporting of reference standards, sample selection, and bias in interpretations
QUORUM	Reports of meta-analyses of randomized controlled trials	Recommendations on reporting searching methodology, synthesis of quantitative analysis, analysis flow, and summary of findings
MOOSE	Reports of meta-analyses in observational studies	Checklist of key reporting factors, e.g., search strategies, methodology reporting, provision of appropriate graphical information, publication bias, and future work

Table 4. Key guidelines and tools for aiding in the reporting of diagnostic or prognostic studies.

Reporting aspect	Meaning
Report of study modality	Does the paper explicitly report whether the study was either prospective or retrospective? (Yes/No)
Report of predictive accuracy	Were predictive quality measures discussed (i.e., diagnostic or prognostic capability estimation)? Any of the following: sensitivity, specificity, receiver operating curves, etc.
Comparative assessment	Is the proposed model compared with outcomes obtained using traditional or alternative models? (Yes/No)
Specification of pre-processing	Does the paper report pre-processing procedures applied (i.e., normalization, filtering, or transformation) of data prior to analyses? (Yes/No) Specification of pre-processing?
Documentation of prediction model	Was a diagnostic or prognostic model specified? (Yes/No)
Report of missing data	Does the paper report missing values? (Yes/No), Method used for dealing with missing values, e.g., value imputation?
Report of software	Was software used for statistical analysis, model implementation, or evaluation reported? (Yes/No)
Complete report of statistical significance	Are <i>p</i> values reported together with statistical values? (Yes/No)
Discussion of limitations	Are limitations (including caveats) of the study discussed? Section dedicated to limitations? (Yes/No)
Data availability	Are the datasets available to external researchers? (Yes/No)
Model cross-validation	Were the diagnostic/prognostic models tested on data not used for training (building) the models?
Reference to reporting guidelines	Does the study refer to any reporting guidelines, i.e., STARD, CONSORT, QUOROM, or QUADAS? Does it fully or partially follow their recommendations? (Yes/No)
Discussion of clinical applicability	Does the paper discuss potential clinical applicability of the biomarkers investigated or specific relevance to translational research? (Yes/No)

Table 5. Reporting aspects assessed in a survey of biomarker publications in the cardiovascular research literature.

website.²³ This network aims to improve reporting quality and the reliability of the health research literature by promoting good practice and offering resources and training for the application of existing reporting guidelines.

An examination of the cardiovascular research literature

An examination of papers on cardiovascular biomarkers with diagnostic or prognostic relevance was conducted to estimate the current state of reporting practices in this area. We focused on papers deposited in PubMed Central (PMC) and published from 1st January to 31st December, 2007. The PMC was queried using the following terms: “cardiovascular AND biomarker AND (diagnosis OR prognosis).” This was also constrained to papers written in English and findings obtained from human samples. This initial search produced a list of 73 papers. A further filtering was applied to concentrate on research papers, that is, removal of reviews and editorials. Additional manual and automatic searches were implemented to remove other studies

whose main application was not cardiovascular research, for example, main focus on cancer biomarkers. The final list of papers examined consisted of 53 papers.

The next step involved the definition of a set of common reporting aspects to be assessed in the (full-text) papers retrieved. On the basis of key factors included in the existing guidelines (Table 4) and the critical problems pointed out above, we concentrated on the aspects listed in Table 5. Their definitions and the questions asked when examining the papers are also included in the table.

None of the examined papers provided evidence that their findings or prediction models were cross-validated or externally evaluated, including those studies whose main goal was to implement new diagnostic or prognostic models. Similarly, none of the papers explicitly referred to its compliance with any reporting standard. Only three papers explicitly discussed the potential clinical application of their findings. The same number of studies reported the presence or handling of missing data. Only about 17% of the papers specified the pre-processing procedures applied (i.e., normalization, filtering, or transformation) prior to data analyses. The same proportion of papers reported sufficient information to interpret the statistical significance of the findings, beyond the simple presentation of *p* (probability) values. Around 83% of the papers did not report information about the corresponding statistical test values, dispersion, or uncertainty estimation. It was also observed that the reporting of *p* estimates based on restricted or approximate ranges (e.g., $p < 0.05$), rather than on the actual values observed, is still a common practice.

Only about a quarter of the papers reported information on how to access the data analysed. Microarray gene expression and SNPs data deposited in centralized repositories (e.g., NCBI websites) represented the most common data dissemination mechanism. About 38% of the papers explicitly compared their findings with alternative models, biomarkers, or previous studies. Only 18 papers specified or documented diagnostic or prognostic computational models (i.e., regression or classification) based on the biomarkers investigated, or with sufficient information so as to allow re-implementation.

A bigger proportion of studies (45%) explicitly discussed potential limitations of their findings or caveats to assess their relevance. Twenty-five papers (47%) explicitly stated whether their studies were conducted either prospectively or retrospectively. Almost half of the papers reported quantitative estimates of the diagnostic or prognostic accuracy of the biomarkers or prediction models investigated. More than half of the papers specified the software tools applied to implement statistical analysis and prediction models or evaluate findings. Figure 1 graphically summarizes the results of this assessment in terms of the proportions of papers complying with each of the reporting aspects shown in Table 5.

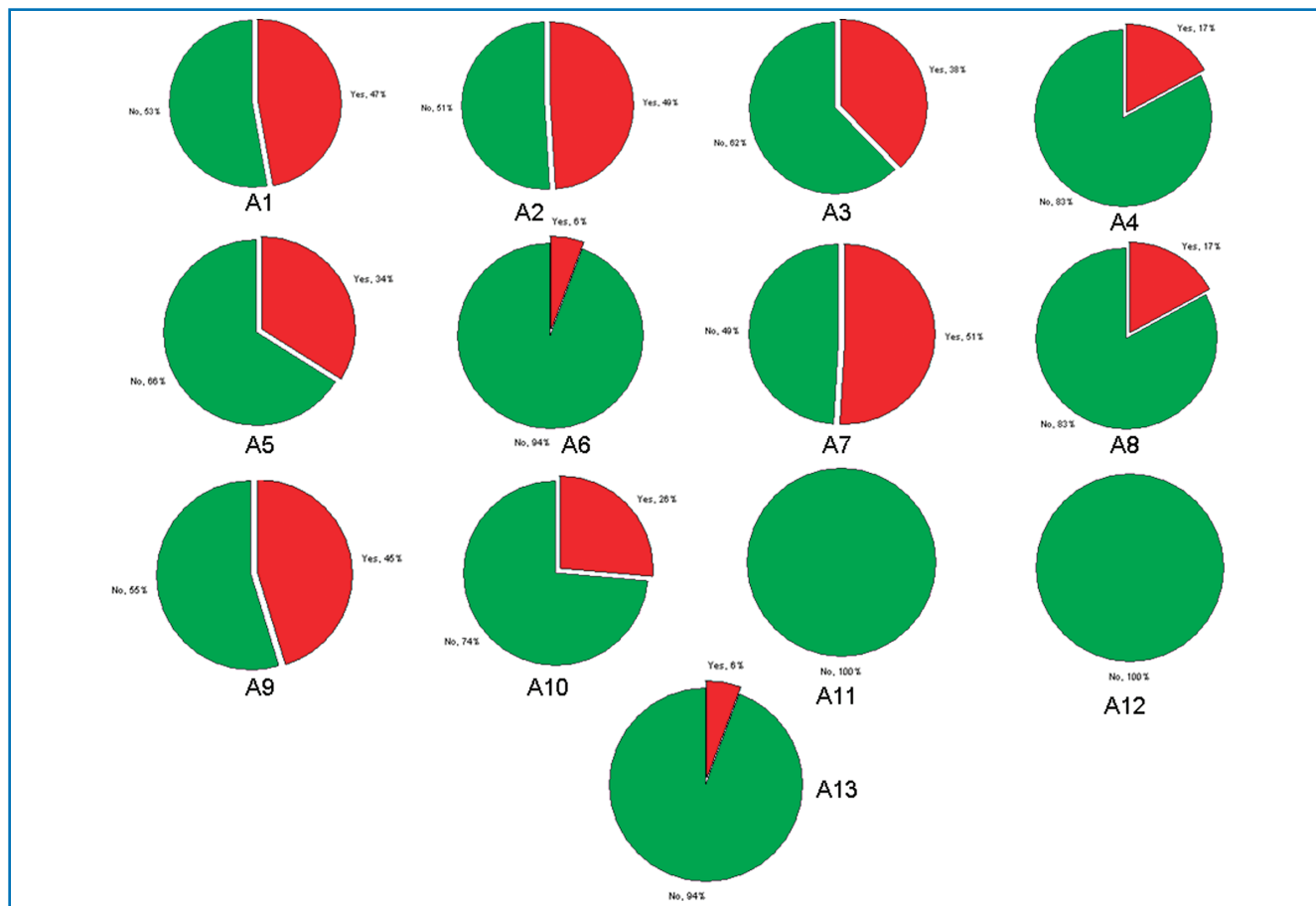


Figure 1. Summary of survey results in terms of the proportions of papers complying with each of the reporting aspects shown in Table 5. Yes = evidence found to suggest compliance with reporting aspect; A1 = report of study modality; A2 = report of predictive accuracy; A3 = comparative assessment; A4 = specification of pre-processing; A5 = documentation of prediction model documentation; A6 = report of missing data; A7 = report of software; A8 = complete report of statistical significance; A9 = discussion of limitations; A10 = data availability; A11 = model cross-validation; A12 = reference to reporting guidelines; A13 = discussion of clinical applicability.

Conclusions

This review concentrated on the problems, challenges, and tools relevant to the reporting of biomarker studies. Despite significant advances accelerated by new experimental and computational technologies over the past 10 years, the concern is the relative lack of successful external replication and validation of published diagnostic and prognostic models.¹⁸ Moreover, inadequate reporting of quantitative findings may increase the likelihood of publishing spurious predictions. Therefore, the reporting of complete and accurate information for re-implementing and interpreting findings from biomarker studies is a central challenge in translational research.

This investigation also offered a view of important reporting patterns observed in studies with diagnostic and prognostic applications. Although it was not intended to be an exhaustive analysis, the results of this survey highlight common practices in reporting and documenting biomarker studies in general, some of which have also been observed previously. To the best of our knowledge, this is the first analysis of reporting practices focused on contributions from the cardiovascular biomarkers area. In general, and in comparison with previous findings,^{16,17} this survey indicates that the completeness and accuracy of scientific reporting is gradually improving. For instance, researchers are becoming more aware of the need of clearly documenting the different research design and data analysis phases, including

pre-processing and software tools used. The open access to data and information is quickly becoming the norm in many research areas and publishing domains. These changes have been driven in part by journals, editors, and public funding organizations. Efforts led by the research community have been crucial for improving awareness of fundamental factors in scientific reporting through the definition of guidelines and recommendations.

However, there is plenty of room for improvement. There is little evidence in this survey to suggest that, at least in the area of cardiovascular biomarker research, important aspects of scientific reporting are being sufficiently (or clearly) covered. There is a need to continue improving the presentation of statistical information in general. In particular, the reporting of data pre-processing, statistical significance results, and prediction model evaluation deserves more attention. Furthermore, it is necessary to motivate a more rigorous documentation of diagnostic and prognostic biosignatures to allow its independent re-implementing and evaluation. When reporting automated prediction models, this may require going beyond the description of standardized laboratory experimental protocols or the listing of the genes defining the biosignature.

To achieve a more balanced reflection of the survey results, the reader may consider the following caveats. First, the sample of papers examined was, apart from being limited in size, focused

on papers available through the PMC. Moreover, this repository may not include papers from some of the leading journals in cardiovascular and translational research. Second, in some cases, missing evidence of reporting aspects (e.g., discussion of clinical applicability and limitations) might be explained by a lack of clarity of expression, rather than by their complete absence. Third, although all the examined papers report findings of relevance to cardiovascular biomarker discovery, it is possible that, in some cases, their main objective may not have been to report findings directly intended for diagnostic or prognostic applications. Despite these potential limitations, the survey stresses the need for a clearer and more accurate specification of methods, results, and context in biomarker studies.

Reporting guidelines should be seen as tools for aiding in the quality assessment of papers before their submission and during peer review. There is empirical evidence that the introduction of some of these guidelines and subsequent endorsement by journals have improved the quality of the reporting of biomedical research.^{19,28} A more active participation of computing and statistical researchers in peer review and editorial tasks may also contribute to the improvement of the reporting of biomarkers research. On the basis of a randomized trial, researchers have recently demonstrated that the inclusion of reviewers with a solid statistical background can significantly enhance the quality of research manuscripts with diagnostic and prognostic applications.²⁹

We hope that this review will motivate the reader to discuss existing or alternative mechanisms for achieving a more accurate and transparent reporting of translational research findings in general, and of biomarker studies in particular.

References

1. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001; 69: 89–95.
2. Wagner DR, Delagardelle C, Ernens I, Rouy D, Vaillant M, Beissel J. Matrix metalloproteinase-9 is a marker of heart failure after acute myocardial infarction. *J Card Fail.* 2006; 12: 66–72.
3. Lindholm E, Melander O, Almgren P, Berglund G, Agardh CD, Groop L, Orho-Melander M. Polymorphism in the MHC2TA gene is associated with features of the metabolic syndrome and cardiovascular mortality. *PLoS ONE.* 2006; 1: e64.
4. Tang T, Pankow JS, Carr JJ, Tracy RP, Bielinski SJ, North KE, Hopkins PN, Kraja AT, Arnett DK. Association of sICAM-1 and MCP-1 with coronary artery calcification in families enriched for coronary heart disease or hypertension: the NHLBI Family Heart Study. *BMC Cardiovasc Disord.* 2007; 7: 30.
5. Timofeeva AV, Goryunova LE, Khaspekov GL, Kovalevskii DA, Scamrov AV, Bulkina OS, Karpov YA, Talitskii KA, Buza VV, Britareva VV, Beabealashvili RSh. Altered gene expression pattern in peripheral blood leukocytes from patients with arterial hypertension. *Ann N Y Acad Sci.* 2006; 1091: 319–335.
6. Xie Q, Ratnasinghe LD, Hong H, Perkins R, Tang ZZ, Hu N, Taylor PR, Tong W. Decision forest analysis of 61 single nucleotide polymorphisms in a case-control study of esophageal cancer. *BMC Bioinformatics.* 2005; 6(Suppl 2): S4.
7. Chittenden TW, Sherman JA, Xiong F, Hall AE, Lanahan AA, Taylor JM, Duan H, Pearlman JD, Moore JH, Schwartz SM, Simons M. Transcriptional profiling in coronary artery disease: indications for novel markers of coronary collateralization. *Circulation.* 2006; 114: 1811–1820.
8. Chen DP, Weber SC, Constantinou PS, Ferris TA, Lowe HJ, Butte AJ. Novel integration of hospital electronic medical records and gene expression measurements to identify genetic markers of maturation. *Pac Symp Biocomput.* 2008; 243–254.
9. Camargo A, Azuaje F. Identification of dilated cardiomyopathy signature genes through gene expression and network data integration. *Genomics.* 2008; 92: 404–413.
10. Wang Y, Barbacioru CC, Shiffman D, Balasubramanian S, Iakoubova O, Tranquilli M, Albornoz G, Blake J, Mehmet NN, Ngadimo D, Poulter K, Chan F, Samaha RR, Eleftheriades JA. Gene expression signature in peripheral blood detects thoracic aortic aneurysm. *PLoS ONE.* 2007; 2: e1050.
11. Arab S, Gramolini AO, Ping P, Kislinger T, Stanley B, van Eyk J, Ouzounian M, MacLennan DH, Emili A, Liu PP. Cardiovascular proteomics: tools to develop novel biomarkers and potential applications. *J Am Coll Cardiol.* 2006; 48: 1733–1734.
12. Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, Wiegand R, Berriz GF, Roth FP, Gerszten RE. Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation.* 2005; 112: 3868–3875.
13. Sawyers CL. The cancer biomarker problem. *Nature.* 2008; 452: 548–552.
14. Gerszten RE, Wang TJ. The search for new cardiovascular biomarkers. *Nature.* 2008; 451: 949–952.
15. Sterne JA, Davey-Smith G. Sifting the evidence—what's wrong with significance tests? *BMJ.* 2001; 322: 226–231.
16. Jafari P, Azuaje F. An assessment of recently published gene expression data analyses: reporting experimental design and statistical factors. *BMC Med Inform Decis Mak.* 2006; 6: 27.
17. Ioannidis JP. Limitations are not properly acknowledged in the scientific literature. *J Clin Epidemiol.* 2007; 60: 324–329.
18. Kostka D, Spang R. Microarray based diagnosis profits from better documentation of gene expression signatures. *PLoS Comput Biol.* 2008; 4(2): e22.
19. Smidt N, Rutjes AW, van der Windt DA, Ostelo RW, Bossuyt PM, Reitsma JB, Bouter LM, de Vet HC. The quality of diagnostic accuracy studies since the STARD statement: has it improved? *Neurology.* 2006; 67(5): 792–797.
20. Wood IA, Visscher PM, Mengersen KL. Classification based upon gene expression data: bias and precision of error rates. *Bioinformatics.* 2007; 23: 1363–1370.
21. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ.* 2003; 326: 41–44.
22. The CONSORT statement. Available at: <http://www.consort-statement.org>. Accessed on 8 December 2008.
23. The Equator Network. Available at: <http://www.equator-network.org>. Accessed on 8 December 2008.
24. STARD statement. Available at: <http://www.stard-statement.org>. Accessed on 8 December 2008.
25. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003; 3: 25.
26. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet.* 1999; 354: 1896–1900.
27. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 2000; 283: 2008–2012.
28. Smidt N, Rutjes AW, van der Windt DA, Ostelo RW, Bossuyt PM, Reitsma JB, Bouter LM, de Vet HC. Reproducibility of the STARD checklist: an instrument to assess the quality of reporting of diagnostic accuracy studies. *BMC Med Res Methodol.* 2006; 6: 12.
29. Cobo E, Selva-O'Callaghan A, Ribera JM, Cardellach F, Dominguez R, Vilardell M. Statistical reviewers improve reporting in biomedical articles: a randomized trial. *PLoS ONE.* 2007; 2(3): e332.