British Journal of Cancer www.nature.com/bjc

PERSPECTIVE

Molecular Diagnostics

REMARK guidelines for tumour biomarker study reporting: a remarkable history

Daniel F. Hayes[1]✉, Willi Sauerbrei[2] and Lisa M. McShane[3]

In 2005, several experts in tumor biomarker research publishe the REporting recommendations for Tumor MARKer prognostic studies (REMARK) criteria. Coupled with the subsequent Biospecimen Reporting for Improved Study Quality (BRISQ) criteria, these initiatives provide a framework for transparently reporting of the methods of study conduct and analyses.

British Journal of Cancer (2023) 128:443–445; https://doi.org/10.1038/s41416-022-02046-4

As the field of medical oncology evolved over the last 70 years, tumour biomarker tests (TBT) have become indispensable to determine prognosis, estimate prediction of benefit from therapy, or monitor patients over time [1]. For example, in breast cancer, estrogen receptor (ER) content in tissue is highly predictive of whether anti-estrogen therapy is or is not likely to work (ER positive), and its use is now standard of care for all patients with this disease [2]. However, prior to the early 1990s, there were few if any guidelines regarding how to judge the reliability of TBT or the research studies purportedly demonstrating benefit from their use. During that decade, a few commentaries began to highlight the need for better quality evidence and to suggest particular areas for improvement, but most published reports of tumour biomarkers were difficult to evaluate or lacked sufficient information to draw any meaningful conclusions [3–5].

In general, diagnostic tests, including TBT, should be used to guide medical decisions only if they are reliable and are shown to improve outcomes [6, 7]. These decisions require both analytical validity and clinical utility. Analytical validity includes addressing both pre-analytical processing and archiving issues as well as specifics regarding the actual assay performance and its interpretation [8]. Clinical Utility requires demonstration of high levels of evidence documenting that patients for whom results of the TBT are used to guide clinical decisions will have superior outcomes compared to those for whom the test was not considered [7]. Such high levels of evidence can be generated either within prospective randomized clinical trials in which the utility of the TBT is the primary objective [9, 10] or from performing rigorous prospective-retrospective studies using archived specimens collected from previously performed prospective trials in a clinical setting relevant to how the TBT might be used [11].

As in all scientific endeavors, evidence cannot be properly evaluated to determine whether criteria are met to establish clinical utility unless there is transparency about how the studies were designed, conducted, and analyzed. Since most TBT studies are retrospective, observational studies, they are prone to a variety of biases due to confounding factors and extensive data analyses, which may lead to generation of spurious results. Development of guidelines for the reporting of TBT studies was initiated by a group of experts at the recommendation of attendees at the First International Meeting on Cancer Diagnostics (From Discovery to Clinical Practice: Diagnostic Innovation, Implementation, and Evaluation), convened in Nyborg, Denmark in July 2000 by the US National Cancer Institute and the European Organisation for Research and Treatment of Cancer (NCI-EORTC). These experts subsequently issued the REporting recommendations for Tumor MARKer prognostic studies (REMARK), published simultaneously by this and other journals in 2005–2006 [12–19].

To enhance understanding and increase adherence to the reporting guidelines, the REMARK group subsequently published a comprehensive explanation and elaboration (E&E) companion paper [20, 21] followed by an abridged version [22]. The latter publication reduced the more highly technical content of the original REMARK report and was more amenable to translation into other languages (Chinese version available at https://mp.weixin.qq.com/s/pNUJ9o9e_6vtcXI9CMKzFg). Moreover, the REMARK guidelines contain a profile template for use with journal submissions to promote the reporting of all 20 REMARK items (https://www.equator-network.org/wp-content/uploads/2016/10/REMARK-checklist-for-EQUATOR-website-002.docx). Furthermore, to promote structured reporting of analyses, the REMARK profile is proposed in item 12 [20, 21]. This profile consists of two parts: (A) patients and treatment variables and (B) statistical analysis of survival outcomes. The statements of the reporting elements have stood the test of time and remained unchanged throughout the years.

Later, the Biospecimen Reporting for Improved Study Quality (BRISQ) criteria were developed to provide more detailed

[1]University of Michigan Rogel Cancer Center, Ann Arbor, MI, USA. [2]Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Stefan-Meier-Str.26, 79104 Freiburg, Germany. [3]National Cancer Institute, Bethesda, MD, USA. ✉email: hayesdf@umich.edu

Received: 27 September 2022 Revised: 17 October 2022 Accepted: 21 October 2022
Published online: 7 December 2022

guidance on what information should be reported about biospecimen collection, processing, and archiving [8, 23, 24]. BRISQ was designed to assess whether biomarker assays applied to specimens are likely to produce reliable and interpretable results.

Taken together, these structured reporting guidelines were intended to have an impact on development of TBT for clinical use. Recently, the impact of simplified but structured reporting was evaluated by Sauerbrei et al, who reviewed 15 papers published in 5 journals [25]. They reported that structured profiles can, indeed, enable readers to quickly and accurately understand the aims of the paper, the patient population, and all statistical analyses performed, including their weaknesses.

Despite these efforts, actual adherence to REMARK and the quality of reporting still leave much room for improvement. For example, in a systematic review of prognostic factor (including TBT) studies published in high-impact journals, Kempf et al. demonstrated evidence of incomplete and selective reporting [26]. Sekula et al. observed that after publication of REMARK, details of reporting were slightly better in articles citing REMARK compared to those that did not [27]. However, when comparing pre- to post-REMARK, they found that, regardless of whether REMARK was cited, the change in reporting quality was barely perceptible, and many key items were still very poorly reported. These authors concluded that the potential overall improvement was possibly diluted by lack of attention to REMARK and they urged a concerted effort from authors, editors, reviewers and methodologists to improve the situation. Interestingly, a study by Botos suggested that (1) authors' claims of adherence to reporting guidelines in general (not specifically REMARK) tend to be overstated compared to review and editorial staff judgment; and (2) judgement by reviewers and editorial staff that reporting in an article is high quality tends to associate more closely with favorable editorial decisions compared to authors' judgments of their reporting [28].

Although REMARK was initially aimed at tumour marker prognostic studies, efforts to expand upon some of the REMARK reporting elements and recognition of their applicability to other types of TBT studies within oncology and even other diseases soon followed. Moreover, the concept of structured reporting of analyses can be easily transferred to many other types of biomedical and methodological research studies [25]. Going beyond the focus on prognosis, Collins *et al* developed the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) statement, noting that the two reporting guidelines most closely related to prediction models are REMARK and the Genetic Risk Prediction Studies (GRIPS) statement [https://www.equator-network.org/reporting-guidelines/strengthening-the-reporting-of-genetic-risk-prediction-studies-the-grips-statement/] [29]. The latter is aimed at studies of genetic risk prediction involving large numbers of genetic variants. TRIPOD explicitly covers all types of predictors in all medical domains. Currently there is an effort underway to update and expand the REMARK reporting guidelines to include any type of factors used for diagnostic or prognostic purposes, both within and outside of oncology (https://www.equator-network.org/library/reporting-guidelines-under-development/reporting-guidelines-under-development-for-other-study-designs/#REMARK).

Complete, transparent and unbiased reporting of research is critical to enable evaluation of appropriateness and quality of study design, methods, and analysis, to promote replication and facilitate comparison across studies, and to understand the context in which study conclusions apply. All research expends resources and effort, and thorough reporting is essential to realize its full potential. Just as the research takes effort, so does good reporting. Greater attention by researchers, journals, and research funding bodies to reporting guidelines like REMARK, and the myriad other reporting guidelines of relevance to health research catalogued by the EQUATOR Network (https://www.equator-network.org), is needed to maximize the value of health research.

Although reporting guidelines are not intended to instruct researchers in how to design, conduct, or analyze their studies, reviewing the REMARK elements can serve as a useful reminder of issues that are important to consider in these aspects. Attention to reporting guidelines when planning a study may therefore have the potential to improve study quality through indirect means. Ultimately, the goal of any biomedical study is to improve patient care, but application of any technology to guide clinical practice cannot occur without reliable information regarding analytical validity and clinical utility, which can only be evaluated with complete and proper knowledge provided by transparent reporting of the methods used to identify patients, collect and handle specimens, and analyze results.

DISCLAIMERS

The views presented here are those of the authors and should not be viewed as official opinions or positions of the National Cancer Institute, National Institutes of Health, or U.S. Department of Health and Human Services; nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

REFERENCES

1. Hayes DF, Allen J, Compton C, Gustavsen G, Leonard DG, McCormack R, et al. Breaking a vicious cycle. Sci Transl Med. 2013;5:196cm196.
2. Allison KH, Hammond MEH, Dowsett M, McKernin SE, Carey LA, Fitzgibbons PL, et al. Estrogen and Progesterone Receptor Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Guideline Update. Arch Pathol Lab Med. 2020;144:545–63.
3. McGuire WL. Breast cancer prognostic factors: Evaluation guidelines. J Natl Cancer Inst. 1991;83:154–5.
4. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. Br J Cancer. 1994;69:979–85.
5. Altman DG, Lyman GH. Methodological challenges in the evaluation of prognostic factors in breast cancer. Breast Cancer Res Treat. 1998;52:289–303.
6. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genet Med. 2009;11:3–14.
7. Hayes DF. Defining Clinical Utility of Tumor Biomarker Tests: A Clinician's Viewpoint. J Clin Oncol. 2021;39:238–48.
8. Moore HM, Kelly AB, Jewell SD, McShane LM, Clark DP, Greenspan R, et al. Biospecimen reporting for improved study quality (BRISQ). J Proteome Res. 2011;10:3429–38.
9. Sargent DJ, Conley BA, Allegra C, Collette L. Clinical trial designs for predictive marker validation in cancer treatment trials. J Clin Oncol. 2005;23:2020–7.
10. Freidlin B, McShane LM, Korn EL. Randomized clinical trials with biomarkers: design issues. J Natl Cancer Inst. 2010;102:152–60.
11. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J Natl Cancer Inst. 2009;101:1446–52.
12. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumour MARKer prognostic studies (REMARK). Br J Cancer. 2005;93:387–91.
13. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst. 2005;97:1180–4.
14. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumor marker prognostic studies (REMARK). J Clin Oncol. 2005;23:9067–72.
15. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, et al. REporting recommendations for tumor MARKer prognostic studies (REMARK). Nat Clin Pr Oncol. 2005;2:416–22.
16. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, et al. REporting recommendations for tumour MARKer prognostic studies (REMARK). Eur J Cancer. 2005;41:1690–6.
17. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, et al. REporting recommendations for tumor MARKer prognostic studies (REMARK). Nat Clin Pr Urol. 2005;2:416–22.

18. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumor MARKer prognostic studies (REMARK). Breast Cancer Res Treat. 2006;100:229–35.
19. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, et al. Reporting recommendations for tumor marker prognostic studies (remark). Exp Oncol. 2006;28:99–105.
20. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. PLoS Med. 2012;9:e1001216.
21. Altman, DG, McShane, L, Sauerbrei, W & Taube, SE Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. BMC Med. 2012;10.
22. Sauerbrei W, Taube SE, McShane LM, Cavenagh MM, Altman DG. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): An Abridged Explanation and Elaboration. J Natl Cancer Inst. 2018;110:803–11.
23. Moore HM, Kelly AB, Jewell SD, McShane LM, Clark DP, Greenspan R, et al. Biospecimen reporting for improved study quality (BRISQ). Cancer Cytopathol. 2011;119:92–101.
24. Moore HM, Kelly AB, Jewell SD, McShane LM, Clark DP, Greenspan R, et al. Biospecimen reporting for improved study quality (BRISQ). Biopreservation Biobanking. 2011;9:57–70.
25. Sauerbrei W, Haeussler T, Balmford J, Huebner M. Structured reporting to improve transparency of analyses in prognostic marker studies. BMC Med. 2022;20:184.
26. Kempf E, de Beyer JA, Cook J, Holmes J, Mohammed S, Nguyen TL, et al. Overinterpretation and misreporting of prognostic factor studies in oncology: a systematic review. Br J Cancer. 2018;119:1288–96.
27. 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:e0178531.
28. Botos J. Reported use of reporting guidelines among JNCI: Journal of the National Cancer Institute authors, editorial outcomes, and reviewer ratings related to adherence to guidelines and clarity of presentation. Res Integr Peer Rev. 2018;3:7.
29. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. Br J Surg. 2015;102:148–58.

ACKNOWLEDGEMENTS

The authors wish to recognize the vision and leadership of the late Professor Douglas G. Altman whose passionate commitment to the quality of health research and its reporting were a driving force behind the REMARK guidelines. Additionally, we would like to acknowledge other members of the REMARK writing committee who made important contributions to the development of the guidelines and explanatory materials.

AUTHOR CONTRIBUTIONS

All three authors contributed equally to preparation of this commentary.

COMPETING INTERESTS

DFH has no conflicts specifically related to this study. DFH reports support unrelated to this study but provided to his institution in the last 24 months during conduct and analysis of this study from Astra Zeneca, Menarini Silicon Biosystems, Merrimack Pharmaceuticals, and Pfizer. DFH reports personal income related to consulting or advisory board activities from BioVeca, Cellworks, Cepheid, EPIC Sciences, EXACT Sciences, Freenome, Guardant, L-Nutra, Macrogenics, Oncocyte, Predictus BioSciences, Tempus, Turnstone Biologics, and Xilis. The University of Michigan holds a patent for which DFH is the named investigator and which is licensed to Menarini Silicon Biosystems from whom UM and DFH receive annual royalties. DFH reports personally held stock options from InBiomotion. LMM declares no conflicts. WS declares no conflicts.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Daniel F. Hayes.