Evidence-Based Health Informatics
E. Ammenwerth and M. Rigby (Eds.)

doi:10.3233/978-1-61499-635-4-304

Publishing Health IT Evaluation Studies

Elske AMMENWERTH[a,1] and Nicolet F. de KEIZER[b]
[a] *UMIT – University for Health Sciences, Medical Informatics, Hall in Tirol, Austria*
[b] *Academic Medical Center, Amsterdam, Netherlands*

Abstract. Progress in science is based on evidence from well-designed studies. However, publication quality of health IT evaluation studies is often low, making exploitation of published evidence within systematic reviews and meta-analysis a challenging task. Consequently, reporting guidelines have been published and recommended to be used. After a short overview of publication guidelines relevant for health IT evaluation studies (such as CONSORT and PRISMA), the STARE-HI guidelines for publishing health IT evaluation studies are presented. Health IT evaluation publications should take into account published guidelines, to improve the quality of published evidence. Publication guidelines, in line with addressing publication bias and low study quality, help strengthening the evidence available in the public domain to enable effective evidence-based health informatics.

Keywords. Medical informatics, publishing, evaluation studies, guideline .

1. Introduction

Progress in science is based on evidence from well-designed studies, normally in individual peer-reviewed publications, and also sometimes in repositories of studies. This evidence is often collected and aggregated in the form of systematic literature reviews and meta-analyses. A systematic literature review typically involves a detailed and comprehensive plan and search strategy derived *a priori*. The goal is to add strength and reduce selection bias by identifying, appraising, and synthesizing all relevant studies on a particular topic. A meta-analysis, in addition, comprises statistical method to synthesize the data from several studies into a single quantitative estimate or summary effect size [1].

However, while preparing systematic reviews and meta-analyses on health IT evaluation studies, reviewers have been confronted with three major challenges leading to possible bias and low quality of published evidence: Publication bias, summarizing the problem that studies with unfavourable outcome may not be published due to stakeholder pressure or related political reasons [2];[2] low quality of the conducted evaluation study;[3] and poor reporting quality of the published evaluation study, where often important information needed to understand, interpret, reproduce or generalize

[1] Corresponding author: Prof. Dr. Elske Ammenwerth, Institute for Biomedical Informatics, Eduard Wallnöfer Zentrum 1, 6060 Hall in Tirol, Austria, elske.ammenwerth@umit.at.

[2] See also: A. Georgiou, Finding, appraising and interpreting the evidence, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

[3] See also: P. Nykänen et al., Quality of health IT evaluations, in: ibid.

the findings of a study is missing in a study paper. This contribution will address this last challenge: low reporting quality.

2. Reporting quality of health IT evaluation studies

The interpretation of studies of health IT is highly context-specific [3]. Thus, especially information on the study context is important to allow the reader to judge the generalizability or relevance to their setting of the published study. Missing context information endangers the evidence-base of health informatics [4]. As Shekelle writes: "The generalizability of evidence will remain low unless more systematic, comprehensive, and relevant descriptions and measurements are made regarding how the technology is utilized, the individuals using it, and the environment it is used in". [3]

Nevertheless, health IT evaluation studies often show insufficient reporting quality. For example, while reviewing 23 randomized health IT trials, Jamal et al. found an insufficient description of the health IT intervention, of allocation or randomization procedures, or of data collection procedures [5]. Likewise, while reviewing 257 health IT studies, Chaudhry et al. found insufficient description of the health IT intervention, the implementation process and the organizational context while reviewing health IT evaluation studies [6]. Eisenstein et al. analysed 134 economic health IT evaluations and found that many studies did not report on key information such as invested financial and personal resources or cost elements. Talmon et al. analyzed the reporting quality of 47 health IT trials and found that title and abstract often missed important information such as the type of evaluated health IT [7]. Shekelle et al. analysed 258 health IT evaluation studies and found that only very few studies reported sufficient information on the organizational and technical context, including health IT usage and users [3].

In a study specifically analysing publication quality, de Keizer et al. reviewed the quality of 120 randomly chosen health IT evaluation studies [8]. They found varying degrees of reporting quality. Often, the evaluated health IT intervention (including functionality, usage, and workflow), the involved study population, and methods or instruments for data collection or data analysis were not described in sufficient detail. Also, no improvement in reporting quality was visible between 1980 and 2005.

Consequently, several reviewers expressed the strong need to improve reporting quality and to develop reporting standards for publication of health IT evaluation studies [5,6,8]. In medical science, guidelines to improve publication quality such as CONSORT [9] or PRISMA [10] have existed for many years. While these guidelines may be helpful on a general basis, they do not cover specific aspects of health IT evaluation studies. Therefore, in 2009, STARE-HI was proposed as a specific guideline for health IT evaluation papers [11]. In 2011, in addition, CONSORT-eHealth [12] for specific types of health IT evaluation studies was published.

In this contribution, we will first present and discuss the applicability of the guidelines from medical sciences such as CONSORT. We will then present the motivation and details of the STARE-HI guideline in more detail.

3. General publication guidelines

This problem of insufficient publication quality is well-known in the medical sciences and several publication guidelines have been developed in the last few years for several clinical study types.

Due to the rising number of publication guidelines in the medical sciences, the EQUATOR network was launched in 2006 [13]. On its website, EQUATOR collects available guidelines and makes them easily accessible. As of June 2015, the website already contained 276 publication guidelines. Many of these guidelines are also of relevance for health IT evaluation publications.

The publication guidelines included in the EQUATOR network have different adoption rates in the scientific community. Some of them are very well known and frequently used. Some of them have even been adopted by major medical journals; submitting authors have to indicate which guideline applies to their submission, and how they follow this guideline. Some of these broadly adopted guidelines include CONSORT for reporting of randomized controlled trials [14], STARD for reporting of diagnostic studies [15], STROBE for reporting of observational studies [16], and PRISMA for systematic reviews [10]. We will give a summary of these guidelines in this section, and discuss their applicability for health IT evaluation studies.

3.1. CONSORT

CONSORT (Consolidated Standards of Reporting Trials) addresses the problems arising from inadequate reporting of randomized controlled trials. The CONSORT Statement is a minimum set of recommendations for reporting randomized trials [9]. The CONSORT 2010 checklist includes 25 items that have to be included in a report of a randomized trial, including information on objectives, design, participants, outcomes, blinding, patient flow, harms, and limitations. A detailed explanation and elaboration paper is available [17]. Several adaptations of the CONSORT statements for specific situations have been published, e.g. for reporting of cluster randomized trials [18] or for reporting of patient-reported outcomes [19]. CONSORT has been endorsed by more than 600 biomedical journals [20]. More details on CONSORT are available at http://www.consort-statement.org.

Health IT evaluation studies which use a randomized controlled trial design are recommended to use CONSORT when reporting their results. In addition, a precise description of the health IT system under evaluation and the context in which the intervention is implemented should be provided.

3.2. STARD

The objective of STARD (Standards for Reporting of Diagnostic Accuracy) is to improve the completeness and transparency of reporting of studies of diagnostic accuracy and to allow assessing internal and external validity [15]. The STARD checklist comprises 25 items, including participant recruitment, data collection, study population, estimates of diagnostic accuracy, adverse events, and discussion of clinical applicability. An explanation and elaboration paper is available [21]. STARD has been endorsed by more than 200 biomedical journals [22]. More details on STARD are available at http://www.stard-statement.org.

Health IT evaluation studies which focus on diagnostic accuracy, e.g. accuracy of teledermatology systems or clinical decision support systems, are recommended to use STARD when reporting their results.

3.3. STROBE

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement supports the dissemination of observational studies [16]. Observational studies comprise, for example, cohort studies, case-control studies, and cross-sectional studies. The STROBE checklist comprises 22 items, including objectives, study design, setting, participants, variables, statistical methods, outcome data, and key results. An explanation and elaboration paper is available [23]. STROBE has been endorsed by around 200 biomedical journals [24]. More details on STROBE are available at http://www.strobe-statement.org.

Many health IT evaluation studies have an observational nature, monitoring for example the effect of an health IT system in a before-after study or in a time series study. These studies are recommended to use STROBE when reporting their results.

3.4. PRISMA

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement is a minimum set of items for reporting of systematic reviews and meta-analyses [10]. The PRISMA checklist comprises 27 items, including objectives, eligibility criteria, search, study selection, data collection, risk of bias, synthesis of results, summary of evidence, and limitations. An explanation and elaboration paper is available [25]. PRISMA has been endorsed by more around 200 biomedical journals [26]. More details on PRISMA are available at http://www.prisma-statement.org.

All systematic reviews and meta-analyses on health IT topics are recommended to use PRISMA when reporting their results.

4. Publication guidelines for health IT evaluation studies

We will now look at two publication guidelines specifically developed for health IT evaluation studies.

4.1. STARE-HI

The Statement on Reporting of Evaluation Studies in Health Informatics (STARE-HI) addresses writing and assessing evaluation reports in Health Informatics. Its goal is to improve the quality of published evaluation studies in Health Informatics, and thus to improve the evidence-base of Health Informatics [11]. The STARE-HI checklist comprises 30 items, including objective of the study, organizational setting, system details and system in use, study design, study flow, outcome measures, unexpected observations, and meaning and generalizability of results. An explanation and elaboration paper is available [27] as well as a shortened version for conference paper [28].

STARE-HI has been endorsed by major health informatics journals as well as by the International Medical Informatics Association (IMIA) and the European Federation for Medical Informatics (EFMI) [24]. Furthermore, STARE-HI has been included in the EQUATOR network [13]. More details on STARE-HI are available at http://iig.umit.at/efmi/starehi.htm. Table 1 presents the content of STARE-HI in more detail.

Table 1. The STARE-HI principles: Items recommended to be included in health informatics evaluation reports [11].

1 Title
2 Abstract
3 Keywords
4 Introduction
4.1 Scientific background
4.2 Rationale for the study
4.3 Objectives of study
5 Study context
5.1 Organizational setting
5.2 System details and system in use
6 Methods
6.1 Study design
6.2 Theoretical background
6.3 Participants
6.4 Study flow
6.5 Outcome measures or evaluation criteria
6.6 Methods for data acquisition and measurement
6.7 Methods for data analysis
7 Results
7.1 Demographic and other study coverage data
7.2 Unexpected events during the study
7.3 Study findings and outcome data
7.4 Unexpected observations
8 Discussion
8.1 Answers to study questions
8.2 Strengths and weaknesses of the study
8.3 Results in relation to other studies
8.4 Meaning and generalisability of the study
8.5 Unanswered and new questions
9 Conclusion
10 Authors' contribution
11 Competing interests
12 Acknowledgement
13 References
14 Appendices

4.2. CONSORT-eHealth

CONSORT-eHealth aims at improving and standardizing evaluation reports of web-based and mobile health interventions. The authors argue that "RCTs of web-based interventions pose very specific issues and challenges, in particular related to reporting sufficient details of the intervention" [12], and therefore they developed CONSORT-eHealth based on CONSORT. CONSORT-eHealth comprises 53 additional sub-items explaining or enhancing the original CONSORT items, such as type of system, bug fixes and down items, computer literacy of participants, names of sponsors, revisions and updating, level of human involvement, intensity of use, and safety and security

procedures. CONSORT-eHealth has been endorsed by the Journal of Medical Internet Research [12].

5. Discussion and Conclusion

Publication of health IT evaluation studies needs to comprise sufficient information to be understandable and generalizable. Low publication quality as found in many reviews impedes evidence-based health informatics, devalues the work that has been done, reduces the value to the reader eager to learn and apply the knowledge, and compromises the potential value of subsequent wider systematic reviews and meta-analyses.[4]

Medical science has a long history of publication guidelines, the major ones being endorsed by many larger biomedical journals. However, there are specific issues of health IT evaluation studies that are often insufficiently reported, such as details of the health IT intervention; IT related characteristics of the system users; and the technical and organizational setting. STARE-HI and other specific guidelines attempt to address this by offering specific guidance for publishing health IT evaluation studies. They do not replace, but complement, other established guidelines. Health IT evaluation publication should apply these other relevant guidelines where appropriate – for example, CONSORT for randomized trials. The adoption of STARE-HI by larger health informatics journals as well as by international health informatics organizations stresses the importance of ensuring the quality of publication, and thus in turn the evidence base, for health informatics and its applications.

Publication guidelines are one of the three means of improving publication of studies and thus strengthening evidence available in the public domain to enable effective evidence based health informatics (EBHI). Publication guidelines are the easiest of the three to implement, as they are focussed on the reporting of studies which have been undertaken. Improving the quality of studies also requires use of other guidelines which address planning and conduct of studies, such as GEP-HI, the Guidelines for Good Evaluation Practice in Health Informatics.[5] Publishing all studies clearly and effectively, and not just positive ones, faces a number of challenges [29], and is a problem also being faced in the clinical and pharmaceutical domains, but is a moral duty and should be addressed by all who believe in effective health IT support to health care.

Recommended further readings

1. J. Talmon, A. Ammenwerth, J. Brender, N. de Keizer, P. Nykänen, M. Rigby, STARE-HI - Statement on Reporting of Evaluation Studies in Health Informatics, *Int J Med Inform* **78**(1) (2009), 1-9.

[4] See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

[5] See also: P. Nykänen et al., Quality of health IT evaluations, in: ibid.

Food for thought

1. Do you routinely assess the quality of the publication of studies which you use either in evidence for decision-making, or as a basis for designing further studies?
2. Imagine different situations or use contexts where you are reading or writing a paper. Is it possible that in these different situations, the importance of items to be covered in a publication may differ?
3. How could you access whether the quality of publications increased after a journal endorsed a certain guideline? Look for evidence on this question!

References

[1] L.S. Uman, Systematic reviews and meta-analyses, *J Can Acad Child Adolesc Psychiatry* **20**(1) (2011), 57-9.
[2] C. Friedman, J. Wyatt, Publication bias in Medical Informatics, *J Am Med Inform Assoc* **8**(2) (2001), 189-191.
[3] P. Shekelle, S. Morton, M. Keeler, Costs and Benefits of Health Information Technology, *Evid Rep Technol Assess (Full Rep)* **132** (2006), 1-71.
[4] E. Ammenwerth, P. Schnell-Inderst, U. Siebert, Vision and challenges of Evidence-Based Health Informatics: a case study of a CPOE meta-analysis, *Int J Med Inform* **79**(4) (2010), e83-8.
[5] A. Jamal, K. McKenzie, M. Clark, The impact of health information technology on the quality of medical and health care: a systematic review, *Him J* **38**(3) (2009), 26-37.
[6] B. Chaudhry, J. Wang, S. Wu, M. Maglione, W. Mojica, E. Roth, S.C. Morton, P.G. Shekelle, Systematic review: impact of health information technology on quality, efficiency, and costs of medical care, *Ann Intern Med* **144**(10) (2006), 742-52.
[7] J. Talmon, E. Ammenwerth, T. Geven, The quality of reporting of health informatics evaluation studies: a pilot study, *Stud Health Technol Inform* **129**(Pt 1) (2007), 193-7.
[8] N.F. de Keizer, E. Ammenwerth, The quality of evidence in health informatics: how did the quality of healthcare IT evaluation publications develop from 1982 to 2005? *Int J Med Inform* **77**(1) (2008), 41-9.
[9] G. Antes, The new CONSORT statement, *BMJ* **340** (2010), c1432.
[10] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, P. Group, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, *Int J Surg* **8**(5) (2010), 336-41.
[11] J. Talmon, A. Ammenwerth, J. Brender, N. de Keizer, P. Nykänen, M. Rigby, STARE-HI - Statement on Reporting of Evaluation Studies in Health Informatics, *Int J Med Inform* **78**(1) (2009), 1-9.
[12] G. Eysenbach, CONSORT-eHealth: improving and standardizing evaluation reports of Web-based and mobile health interventions, *J Med Internet Res* **13**(4) (2011), e126.
[13] EQUATOR, *The Equator Network: Enhancing the Quality and Transparency of Health Research*, http://www.equator-network.org, last access 11 February 2016.
[14] K.F. Schulz, D.G. Altman, D. Moher, C. Group, CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials, *BMJ* **340** (2010), c332.
[15] P.M. Bossuyt, J. Reitsma, D. Bruns, C. Gatsonis, P. Glasziou, L. Irwig, J. Lijmer, D. Moher, D. Rennie, H. de Vet, Towards Complete and Accurate Reporting of Studies of Diagnostic Accuracy: The STARD Initiative, *Ann Int Med* **138**(1) (2003), 40-44.
[16] E. von Elm, D.G. Altman, M. Egger, S.J. Pocock, P.C. Gotzsche, J.P. Vandenbroucke, S. Initiative, The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies, *Lancet* **370**(9596) (2007), 1453-7.
[17] D. Moher, S. Hopewell, K.F. Schulz, V. Montori, P.C. Gotzsche, P.J. Devereaux, D. Elbourne, M. Egger, D.G. Altman, CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials, *BMJ* **340** (2010), c869.
[18] M.K. Campbell, G. Piaggio, D.R. Elbourne, D.G. Altman, Consort 2010 statement: extension to cluster randomised trials, *BMJ* **345** (2012), e5661.
[19] M. Calvert, J. Blazeby, D.G. Altman, D.A. Revicki, D. Moher, M.D. Brundage, Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension, *Jama* **309**(8) (2013), 814-22.
[20] Consort, *Impact of CONSORT,* http://www.consort-statement.org/about-consort/impact-of-consort, last access 11 February 2016.

[21] P.M. Bossuyt, J.B. Reitsma, D.E. Bruns, C.A. Gatsonis, P.P. Glasziou, L.M. Irwig, D. Moher, D. Rennie, H.C. de Vet, J.G. Lijmer, The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. The Standards for Reporting of Diagnostic Accuracy Group, *Croat Med J* **44**(5) (2003), 639-50.

[22] STARD, *News about STARD*, http://www.stard-statement.org, last access 11 February 2016.

[23] J.P. Vandenbroucke, E. von Elm, D.G. Altman, P.C. Gotzsche, C.D. Mulrow, S.J. Pocock, C. Poole, J.J. Schlesselman, M. Egger, S, Initiative, Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration, *PLoS Med* **4**(10) (2007), e297.

[24] STROBE, *STROBE Supporting journals & organizations*, http://www.strobe-statement.org/index.php?id=strobe-endorsement, last access 11 February 2016.

[25] A. Liberati, D.G. Altman, J. Tetzlaff, C. Mulrow, P.C. Gotzsche, J.P. Ioannidis, M. Clarke, P.J. Devereaux, J. Kleijnen, D. Moher, The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration, *BMJ* **339** (2009), b2700.

[26] PRISMA, *PRISMA Endorsers*, http://www.prisma-statement.org/Endorsement/EndorsePRISMA.aspx, last access 11 February 2016.

[27] J. Brender, J. Talmon, N. de Keizer, P. Nykanen, M. Rigby, E. Ammenwerth, STARE-HI - Statement on Reporting of Evaluation Studies in Health Informatics: explanation and elaboration, *Appl Clin Inform* **4**(3) (2013), 331-58.

[28] N.F. de Keizer, J. Talmon, E. Ammenwerth, J. Brender, M. Rigby, P. Nykanen, Systematic Prioritization of the STARE-HI Reporting Items. An Application to Short Conference Papers on Health Informatics Evaluation, *Methods Inf Med* **51**(2) (2012), 104-11.

[29] M. Rigby, *Evaluation: 16 Powerful Reasons Why Not to Do It - And 6 Over-Riding Imperatives*, in: V. Patel, R. Rogers, R. Haux, editors. Proceedings of the 10th World Congress on Medical Informatics (Medinfo 2001), IOS Press, Amsterdam, 2001. pp. 1198-202.