MEDINFO 2010
C. Safran et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-588-4-1206

Mini Stare-HI: Guidelines for reporting health informatics evaluations in conference papers

Nicolette F. de Keizer[a], Jan Talmon[b], Elske Ammenwerth[c], Jytte Brender[d], Pirkko Nykanen[e], Michael Rigby[f]

[a] *Department of Medical Informatics, Academic Medical Center, Amsterdam, The Netherlands*
[b] *School for Prevention and Primary Care – CAPHRI, Maastricht University, Maastricht, The Netherlands*
[c] *UMIT – University for Health Sciences, Medical Informatics and Technology, Institute for Health Information Systems, Hall in Tyrol, Austria*
[d] *Department of Health Science and Technology, Aalborg University, and Virtual Centre for Health Informatics, Aalborg, Denmark;*
[e] *Department of computer sciences, University of Tampere, Tampere, Finland*
[f] *School of Public Policy and Professional Practice, Keele University, Keele, United Kingdom*

Abstract

Background: To improve the quality of reports of health informatics evaluations we recently devised and published a guideline named STARE-HI, now formally endorsed by IMIA. Objective: To develop a prioritization framework of ranked items (a mini-STARE-HI) to assist authors when reporting health informatics evaluation studies in a restricted space conference paper. Method: We invited 111 editors of health informatics conference proceedings and reviewers and authors of health informatics evaluation studies to score 106 reporting items on a scale ranging from "0 - not necessary" through to "10 - essential" by a web-based survey. Results: The response rate for the survey was 63% (70 out of 111). The most important items (score >9) were "Interpret the data and give an answer to the study question", "Whether it is a laboratory, simulation or field study" and "Description of the outcome measure/evaluation criteria". Four items had a mean score <6. Conclusion: It has been possible to produce a ranking of reporting items from STARE-HI according to their prioritized relevance for inclusion in space-limited conference papers. We believe that this prioritization framework can improve quality and utility of conference papers on health informatics evaluation studies.

Keywords:

Standards, Reporting, Evaluation studies, Conference papers, Evidence

Introduction

Modern healthcare tends increasingly to depend on health informatics applications such as electronic patient records, order entry and image processing systems. Within the last two decades, the viewpoint that interventions in healthcare should be evidence-based has become the accepted norm. In this context it is more than ever imperative to also assess the impact of health informatics investments, as concluded in [1,2]. Until now this has seldom been the case, and in those cases where health informatics applications have been assessed reports on these evaluations are often of limited value because essential information is not properly communicated – possibly because reporting standards were missing. High quality reporting is essential as it serves the target audience better and assists in the build-up of a robust evidence base. To improve reports of health informatics evaluations we recently devised and published a Statement on Reporting of Evaluation Studies in Health Informatics - STARE-HI [3]. This is now endorsed by IMIA, and listed by the EQUATOR initiative [4].

About half of all health informatics evaluation studies are published in health informatics journals and conference proceedings as shown by [5]. Too many studies report their results in a conference paper without a full journal report, as noted by [6]. Therefore, conference papers provide valuable information for systematic reviews about studies that are not otherwise published, the exclusion of which from the review would seriously weaken the evidence base as well as possibly introduce publication bias. For these reasons, conference papers should contain sufficient information about the evaluation study to serve as an accurate record of its conduct and findings, providing optimal information about the study within the space constraints of the conference paper format so that they are of sufficient quality and utility to contribute to the evidence base. However, they are inevitably constrained by prescribed limits on length, providing authors with a severe challenge in meeting the STARE-HI guidelines. This paper describes the development of a prioritization framework, called Mini STARE-HI, to assist authors in considering which items to include, and which to exclude on space grounds, when reporting health informatics evaluation studies in a conference paper.

Materials and Methods

The scope and purpose of the full STARE-HI is to provide guidelines for reporting evaluations in Health Informatics, independent of the evaluation method used. Therefore, these guidelines have a general character, with a main focus on the description of the context in which the study took place including a description of the system being evaluated, the description of the methodology, the systematic reporting of results, and the structuring of the discussion [3]. STARE-HI formed the basis for the development of this Mini STARE-HI, which had two steps.

First, all members of the core editorial team of STARE-HI independently identified in STARE-HI all items in the various reporting (sub) areas that they believed authors should consider when reporting an health informatics evaluation study. For example, the sub area "system details" contains items such as "type of system", "aim of the system", "profession and number of users", etc. We included in the next step all items identified by one or more editorial team members.

Next, a total of 111 participants, consisting of authors of health informatics evaluation studies, reviewers with an interest in health informatics evaluation and editors of Medical Informatics conference proceedings, were invited by email to participate in a web-based survey and rate the importance of each of the checklist items. Respondents were asked to score the items with the common limitation of a 5-6 pages conference paper such as MIE, Medinfo and AMIA in mind. The scoring system was a scale ranging from "0 - not necessary" to "10 - essential" to include the item in a report of a health informatics evaluation study in a conference paper. Furthermore, respondents were invited to provide any items per reporting (sub) area they felt important but were missing in the checklist.

Results

One-hundred-and-six items were identified in the STARE-HI guideline in step one and included in the survey. After the original invitation and two reminders the response rate for the survey was 63% (70/111).

Table 1 shows the items per reporting area ordered on their mean score. The most important items (score >9) were "Interpret the data and give an answer to the study question" (in the discussion section), "Whether it is a laboratory, simulation or field study" (in the methods/study design section) and "Description of the outcome measure/evaluation criteria" (in the methods/study design section). In contrast, the items "Name of the health care organization" (in the methods/study context section), "Authors' contribution", "Formal permission and ethical concerns" (in the introduction) and "study limitations" (in the abstract) were considered the first candidates to ignore in case of lack of reporting space (score <6). The overall mean score per reporting area was highest for the discussion (8.3).

In total 19 different comments were given for adding new or "how to" items, rephrasing items, or for rearrangement of order. The latter group counts for more than half (11/19) of the comments.

Table 1- Survey results on the prioritization score of health informatics evaluation items derived from the STARE-HI guideline when applied to a conference paper.

Reporting area (mean score)	Reporting items (generated from STARE-HI, [3], re-ordered per reporting area based on the mean score)	Mean score (SD)
Content of Title (7.1)	Study question	7.5 (2.8)
	Type of information system	7.5 (2.9)
	Study design	6.6 (2.9)
	The term "evaluation" or "assessment"	6.6 (3.1)
Abstract (8.0)	Describe the major results	9.0 (1.1)
	Include the objective	8.9 (1.4)
	Include a conclusion	8.7 (1.7)
	Describe/define the outcome measures	8.5 (1.6)
	Describe the methods - study design	8.5 (1.6)
	Describe the setting	7.7 (2.1)
	Describe the participants	7.6 (1.9)
	Be structured	7.6 (2.4)
	Describe study limitations	5.2 (2.8)
Keywords (7.1)	Refer to the type of system being evaluated	7.8 (2.3)
	Include "evaluation" or "assessment"	7.3 (2.8)
	Refer to the outcome measure	7.1 (2.4)
	Based on MeSH terms	6.9 (2.4)
	Refer to the study design	6.9 (2.5)
	Refer to the setting	6.5 (2.6)

Table 1 (continued)

Introduction (6.9)	Study questions and hypotheses	8.7 (1.8)
	Motivation for the study	8.1 (2.1)
	What is already known about the type of system	7.4 (2.0)
	A description of the system (e.g. function)	7.0 (2.4)
	Position of this study in a larger study/project	6.5 (2.4)
	Which stakeholders viewpoint(s) is/are used	6.3 (2.4)
	Potential influence of the study	6.1 (2.3)
	Formal permissions and ethical concern (e.g. ethical board)	5.0 (3.0)
Methods – study context (7.1)	Kind of facility (e.g. outpatient clinic, hospital)	8.2 (1.4)
	Aim of the system	8.0 (1.9)
	Type of system	7.8 (1.9)
	Type of information managed	7.6 (1.7)
	Clinical or other tasks of the system	7.4 (1.9)
	How long the system is used	7.4 (1.9)
	How wide spread the system is used	7.4 (2.0)
	Description of how the system works	7.3 (2.2)
	Which facilities/department(s)	7.1 (2.1)
	Professions and number of users	7.0 (2.3)
	Reference to a full technical description of the system	6.7 (2.4)
	Geographical location of the health organization	6.4 (2.6)
	Name of the health organization	4.3 (2.9)
Methods-study design (7.9)	Whether it is a laboratory, simulation or field study	9.1 (1.1)
	Description of the outcome measure/evaluation criteria	9.1 (1.5)
	Study type (e.g. case study, (quasi) experimental etc)	9.0 (1.5)
	Methods to select participants	8.5 (1.6)
	Allocation strategy in controlled trials	8.3 (1.8)
	Definition of the key concepts e.g. Medical error, Adverse Drug Event	8.2 (2.1)
	Focus of the researchers in case of qualitative concepts	8.0 (1.8)
	Entry criteria	8.0 (1.9)
	Description of the study flow	7.7 (2.0)
	Sample size calculation in controlled trials	7.6 (2.3)
	Start and end dates of the study	7.6 (2.2)
	Date of intervention(s)	7.3 (2.5)
	Theory on which the study is based (e.g. the user acceptance model that guided a quantitative survey)	6.9 (2.5)
	Biases following from the chosen study design	6.8 (2.6)
	Motivation for the study design	6.3 (3.0)
Methods-data collection (7.8)	Methods used per outcome measure	8.8 (1.4)
	Retrospective or prospective data collection	8.6 (1.6)
	Validity of the measurement (e.g. use of a validated questionnaire)	8.3 (1.5)
	Blinding of observer and participants	8.1 (1.9)
	Number and type of interviews	8.0 (1.8)
	Type and duration of observations	8.0 (1.8)
	Details about new measurement tools	7.7 (2.2)
	Location and setting where data is collected	7.7 (2.3)
	Full disclosure of new measurement tools in appendix	6.8 (2.8)
	Professional background of the interviewer	6.2 (2.7)
Methods-data analyses (7.7)	Analysis methods for qualitative data	8.7 (1.3)
	Statistical techniques for quantitative data	8.7 (1.5)
	Kind of triangulation used	7.4 (2.0)
	Awareness of any analysis bias	7.2 (2.6)
	Analysis software used	6.5 (2.7)

Table 1 (continued)

Results (8.1)	Basic numbers of the study (e.g. no. of observations, response rate etc)	8.9 (1.5)
	Quantitative data in tables and figures	8.6 (1.2)
	Most important results in the text	8.6 (1.6)
	Sufficient data for all outcome measures	8.6 (1.6)
	Baseline demographic data / characteristics of participants	8.3 (1.7)
	Special notice to any unexpected striking result	8.3 (1.6)
	Any unintended side effect (positive or negative) of the system	8.1 (1.8)
	Absolute numbers and not just relative numbers	7.9 (2.1)
	Characteristics and qualities of the participants in qualitative studies	7.8 (2.0)
	Influence of unexpected events on the study findings	7.8 (2.0)
	Number and type of drop outs	7.7 (2.0)
	Quotes to illustrate any major qualitative points	7.6 (1.9)
Discussion (8.3)	Interpret the data and an answer to the study question	9.5 (0.9)
	Strong and weak points of the study	9.0 (1.2)
	Meaning/implications of the study	8.8 (1.4)
	New insight from this study	8.7 (1.3)
	Generalizability/ applicability of the study results	8.5 (1.7)
	Discuss any biases	8.4 (1.7)
	What is novel compared to other studies	8.3 (1.6)
	Reasons for disagreement with other studies	8.0 (1.8)
	Critically discuss the methods used	8.0 (2.1)
	Agreement of findings with other studies	7.8 (1.9)
	Comparability of the setting of other studies	7.3 (2.2)
	New future research questions	7.3 (2.3)
Conclusion (7.7)	Impact of the findings	8.4 (2.1)
	Summarize the findings	8.3 (2.6)
	Relation of the findings to the big picture	7.7 (2.1)
	Recommendations of the authors	7.4 (2.4)
	Future research to be done	6.8 (2.6)
References should be included according to the conference guidelines		8.9 (1.5)
Acknowledge any financial or other support		8.0 (2.2)
Financial or other interests which may influence the design or interpretation of the results		7.6 (2.3)
An appendix can be used to describe any supporting material		6.6 (2.7)
Authors' contributions		5.1 (3.1)

Discussion

In this study we used opinions from key stakeholders to develop a ranked list of reporting items for health informatics evaluation studies. Rather than presenting a rigid list of items to report in a conference paper, the prioritization framework of "Mini STARE-HI" assists authors in meeting the principles of the STARE-HI guideline within the constraints of a conference paper, and related to their study topic.

We were somewhat surprised by the low mean score (5.0) for "Formal permission and ethical concern" in the introduction section of a paper. In medical research formal approval of a study by an ethics committee or Internal Review Board is mandatory. Based on the premise that evaluation of health informatics interventions is ethically imperative (as stated in [1]), we should be careful that all participants in such studies are properly protected. This is clearly an issue that requires a wider discussion in the health informatics community, in particular since about an equal number of our respondents found this item either unnecessary (6 scored 0) or essential (8 scored 10).

We asked the respondents to score the items with the common limitation of a 5-6 pages conference paper such as MIE, Medinfo and AMIA in mind. Conference papers and abstracts for medical conferences are often even more restrictive in space (250 to 300 words). As the basic principle of Mini-STARE-HI is to prioritize items to report instead of urging what should be or should not be reported we believe these guidelines are also applicable to health informatics evaluation studies presented as the more restrictive short medical conference abstracts. Similar methods as applied in our study were used to develop comparable guidelines for RCT abstracts based on CONSORT [7]

and for observational study conference abstracts [8]. Our response rate of 63% was comparable to [7].

A weakness of our study is that we only sent out the questionnaire once. We did not give feedback to the participants and did not ask for a potential revision of their position as is commonly done in Delphi studies.

For our ranking purpose, we computed the mean of the scores of the respondents. In principle our measurement scale is of ordinal type and medians and percentiles are the most appropriate way to represent the characteristics of the underlying distribution. Ranking on medians, however is more problematic, since the median can only take the integer values assigned to the response categories. Since we have 11 response categories, the response categories approach an interval scale. Hence we considered taking means as a good alternative for ranking purposes.

We plan to measure the quality of health informatics evaluation conference papers published in the past. Authors, reviewers and editors of health informatics evaluation papers are encouraged to use the results of this study to improve the quality of conference papers. In the future, we will monitor the effect of (Mini) STARE-HI on the publications as some studies (i.e. [9]) show an increase of publication quality after the publication of similar reporting guidelines.

Conclusion

It has been possible to produce a ranking of reporting items from STARE-HI according to their prioritized relevance for inclusion in space-limited conference papers. Only a few items were considered to be (nearly) essential for inclusion in a conference paper, some of the items from STARE-HI that add credibility to full paper publications were considered of less relevance to be included in a conference paper on a health informatics evaluation study and can be left out. Which of the other items to select to be included in a report is the responsibility of the authors, but the ranking that resulted from our study will help them to make an informed decision.

We believe that by guiding authors in prioritizing what information is important to report within the given constraints of a conference paper, quality and utility of such publications can be improved.

Acknowledgments

We thank all respondents to our survey for their opinions on the reporting items.

References

[1] Ammenwerth E, Brender J, Nykanen P, Prokosch HU, Rigby M, Talmon J. Visions and strategies to improve evaluation of health information systems. Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, Int J Med Inform 73 (June (6)) (2004) 479–491.

[2] Catwell L, Sheikh A (2009) Evaluating eHealth Interventions: The Need for Continuous Systemic Evaluation. PLoS Med 6(8): e1000126. doi:10.1371/journal.pmed.1000126

[3] Talmon J, Ammenwerth E, Brender J, de Keizer NF, Nykänen P, Rigby M. STARE-HI -- Statement on reporting of evaluation studies in Health Informatics. Int J Med Inform 78 (2009) 1-9.

[4] www.equator-network.org/resource-centre/library-of-health-research-reporting/reporting-guidelines/other-reporting-guidelines/ (accessed 3 September 2009)

[5] Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. Cochrane Database of Systematic Reviews 2007, Issue 2. Art. No.: MR000005. DOI: 10.1002/14651858.MR000005.pub3

[6] Ammenwerth E, de Keizer NF. An inventory of evaluation studies of information technology in health care trends in evaluation research 1982–2002. Meth Inf Med 44 (1) (2005) 44–56.

[7] Altman D. Better reporting of randomised controlled trials: the CONSORT statement, BMJ 313 (1996) 570–571.

[8] Kho ME, Eva KW, Cook DJ, Brouwers MC. The Completeness of Reporting (CORE) index identifies important deficiencies in observational study conference abstracts. J Clin Epid 61 (2008) 1241-1249.

[9] Moher D, Jones A, Lepage L. CONSORT Group (Consolidated Standards for Reporting of Trials). Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA 285 (15, Apr 18) (2001) 2006-7.

Address for correspondence

Nicolette de Keizer, PhD
Academic Medical Center
Department of Medical Informatics, J1b-115
P.O.Box 22700
1100 DE Amsterdam
The Netherlands
n.f.keizer@amc.uva.nl