

STARE-HI -Statement on Reporting of Evaluation Studies in Health Informatics

J. Talmon¹, E. Ammenwerth², J. Brender³, N. de Keizer⁴, P. Nykänen⁵, M. Rigby⁶

¹ Center for Research, Innovation, Support and Policy-CRISP, Maastricht University, Maastricht, The Netherlands

² UMIT-University for Health Sciences, Medical Informatics and Technology, Hall in Tyrol, Austria

³ Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

⁴ Department of Medical Informatics, Academic Medical Center, Amsterdam, The Netherlands

⁵ Department of Computer Sciences, University of Tampere, Tampere, Finland

⁶ Centre for Health Planning and Management, Keele University, Keele, United Kingdom

Summary

Objective: Development of guidelines for publication of evaluation studies of Health Informatics applications.

Methods: An initial list of issues to be addressed in reports on evaluation studies was drafted based on experiences as editors and reviewers and as authors of systematic reviews, taking into account guidelines for reporting of medical research. This list has been discussed in several rounds by an increasing number of experts in Health Informatics evaluation during conferences and by using e-mail.

Results: A set of STARE-HI principles to be addressed in papers describing evaluations of Health Informatics interventions is presented. These principles include formulation of title and abstract, of introduction (e.g. scientific background, study objectives), study context (e.g. organizational setting, system details), methods (e.g. study design, outcome measures), results (e.g. study findings, unexpected observations) and discussion and conclusion.

Conclusion: A comprehensive list of principles relevant for properly describing Health Informatics evaluations has been developed. When manuscripts submitted to Health Informatics journals and general medical journals adhere to these aspects, readers will be better positioned to place the studies in a proper context and judge their validity and generalisability. STARE-HI may also be used for study planning and hence positively influence the quality of evaluation studies in Health Informatics. We believe that better publication of (both quantitative and qualitative) evaluation studies is an important step toward the vision of evidence-based Health Informatics.

Limitations: This study is based on experiences from editors, reviewers, authors of systematic reviews and readers of the scientific literature. The applicability of the principles has not been evaluated in real practice. Only when authors start to use these principles for reporting, shortcomings in the principles will emerge.

Keywords

Medical Informatics, guidelines, publishing standards, research design

Yearb Med Inform 2009; 23-31

Introduction

Evidence-based¹ application of technologies, methods, and interventions is recognised as good practice and ethically important in all aspects of health care [1]; the technologies, methods, and interventions have to be proven to be safe, effective, and the most appropriate compared with other methods and alternative solutions.

Given the essential role of information technology (IT) systems on the delivery of modern health care, and the dependence of health professionals and organizations on them, it is imperative that they are thoroughly assessed through robust evaluations as with any other form of health process or technology. This principle is advocated and elaborated in the Declaration of Innsbruck [2].

In the past decades it has been demonstrated that IT systems can not only be beneficial, but also can have unintended, potentially detrimental effects as documented by [3, 4], just like other interventions in (social) environments as extensively described in [5]. It is imperative to monitor IT system implementations and their effects during the whole

life cycle of the system. Unintended and intended consequences as well as the socio-technical circumstances under which these occur are important to report since they provide insight that can inform system developers and implementations of similar systems elsewhere. Although there might be pressure to suppress bad news, there are several overriding imperatives for reporting unintended consequences [6]. Also assessment of the development and implementation process itself provides indications on what is good practice and hence contributes to successful Health Informatics applications [7].

When it comes to decisions how best (if at all) to use IT systems for a particular task in the health care delivery process, objective appraisal of opportunities and options requires access to available evidence. Part of this evidence can be found in the scientific literature. Besides the fact that the volume and coverage of evaluation studies in the literature is small given the importance and impact of Health Informatics on modern health care practice and delivery [8], the evidence that is available is to some extent difficult to appreciate due to poor reporting practice [9].

There is growing published evidence of the impact of Health Informatics² on

¹ We use the term evidence in a broad sense. It is not necessarily restricted to some outcome measure like morbidity, mortality or compliance, but includes also user experiences as collected with questionnaires or qualitative studies and results from the analysis of implementation processes, for example. Consequently, evidence-based health informatics is based on both quantitative and qualitative research results.

² The term Health Informatics is used rather than Medical Informatics, since the latter has the annotation that it concerns the processing of information related to the medical condition of patients, but there are applications that cover a broader scope, including the health of a broader part of the population.

health care [10], and, increasingly, reviews appear that summarize the available evidence in the form of a narrative review, a systematic review, or a formal meta-analysis (e.g. [11]). However, the wide variety in the kinds of IT systems and their application domains as well as the various kinds of outcome measures have limited the generalisability of findings, and thus have hampered meta-analysis, as reported for example by [11]. Other studies too have shown that publications on the evaluation of IT interventions in health care have several shortcomings that severely hamper the proper appraisal of these publications, see the review in [12]), p. 243-323.

During a specially convened expert workshop on Health Informatics evaluation - HISEVAL - sponsored by the European Science Foundation and held in Innsbruck [2], the concern was raised that without proper guidelines for the design, planning, execution, and reporting of evaluation studies in Health Informatics, it would be difficult to build up a proper evidence base that can be used to make informed decisions regarding IT interventions in health care.

In other domains of medicine, these problems have also been identified. Work done in the early 1990s lead to the publication of the CONSORT statement in 1996 [13]. The CONSORT statement provides guidelines for the publication of randomized controlled clinical trials (RCTs). This statement has been adopted by many medical journals. Later this statement was revised [14] and has been extended to cover specific kinds of RCT designs such as cluster RCTs [15] or RCTs assessing nonpharmacologic treatments [16] and has been shortened for reporting RCTs in journal and conference abstracts [17]. Several other guidelines have been developed following the approach of the CONSORT statement; for example the QUOROM statement for reporting of meta-analyses [18]), STARD for reporting of diagnostic studies [19] and STROBE for observational studies [20]. An overview of the various guidelines has been published in

[21]. Such an overview is also available at the EQUATOR-network website [22].

CONSORT has proven its value over time. Studies have demonstrated that there is more quality in the reporting of controlled studies after the introduction of the CONSORT statement [23, 24]. Guidelines for good reporting of studies are likely to have an influence on the quality of the studies themselves as well, because of the requirement for a clear demonstration of sound scientific methodology.

Health Informatics is a significant area of health systems investment, and potentially affects every professional and patient. It is therefore evident that Health Informatics should adopt similar robust guidelines as to build a more solid evidence base. Health Informatics applications potentially have effects on health care organizations, health care delivery and outcomes, therefore a Health Informatics application may not directly affect the medical condition of the patient – as drugs do – but it will generally have an indirect effect by assisting the care givers in their decisions and their patient management. The study designs that are covered by CONSORT and other reporting guidelines are not always the most appropriate in Health Informatics evaluation research [2, 25, 26]. Other study designs, both qualitative and quantitative ones, are frequently used.

These observations have led us to the development of guidelines for reporting of evaluation studies in Health Informatics, which build upon work of others, yet also take into account some specific issues that are often central to Health Informatics evaluation studies. They are intended to be applicable for a spectrum of quantitative and qualitative study designs as found in Health Informatics research.

Objectives

The objective of STARE-HI (STatement on the Reporting of Evaluation studies in Health Informatics) is to provide

guidelines for writing evaluation reports in Health Informatics which can be reliably interpreted by subsequent readers; and by doing this to improve the quality of published evaluation studies in Health Informatics; and thus to improve the evidence base of Health Informatics.

These objectives are achieved by presenting guidelines for reporting, which are formatted as a checklist with enough detail to guide authors without making reporting too complex.

This paper describes the procedure followed to come to these guidelines. It gives in tabular form the issues that should be included in a report of an evaluation of an IT intervention in health care. Next, the various items are discussed in more detail. A description of the rationale of each item and examples will be published in a subsequent separate paper.

Method

An initial set of items was drafted by the editorial team (represented by the authors of this paper) based on discussions at the HISEVAL workshop and on their experience with assessing the quality of papers for either a review or a meta-analysis or as part of the editorial process as reviewers and editors. The CONSORT statement [13, 27], criteria for reviewers of biomedical informatics manuscripts [28], the QUOROM statement [18], the STARD statement [19] and other more general recommendations on publication quality such as [29] were used as further reference material.

For each item, a brief description was made. Preliminary versions of the list of items and their descriptions were presented to and discussed with colleagues world-wide at workshops at Medinfo2004, MIE2005, MIE2006, AMIA 2006, and AMIA2007.

The list of items and the descriptions was refined by the editorial team based on the comments and suggestions received. The resulting draft was then published

on the website of the EFMI- working group on evaluation (EFMI-WG EVAL), and all members of that working group were informed of the call for comments at the end of 2006. These comments were assessed by the editorial team and incorporated in the draft document. In spring 2007, a second draft of STARE-HI was put on the website of EFMI-WG EVAL, and potential authors were invited to use and validate STARE-HI. This version of STARE-HI was also used in a pilot study to assess the quality of reporting of evaluation studies in a selected group of Medical Informatics journals in one particular year [30]. Both the usage of the draft of STARE-HI and this assessment study provided feedback on various aspects of STARE-HI.

All received comments have been taken into account leading to a final draft of STARE-HI. This final draft was once again made available through the EFMI-WG EVAL website and distributed among members of the AMIA Working Group on Evaluation for a final round of comments, which have been included in the guidelines as presented here.

STARE-HI Recommendations

The scope of STARE-HI is to provide guidelines for the reporting of 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, the description of the methodology, general recommendations for the reporting of results, and the structuring of the discussion. In cases where a study design has been selected for which already a more specific guideline exists (e.g. a randomized trial), STARE-HI should be used in a way which is complementary to the more specific guideline.

We consider the Health Informatics application as the object of evaluation.

This assessment can cover the technical artefact (or system) as such including its functionality, but also the effect the artefact has on the surrounding organizational environment, the working procedures and social and psychological issues, as well as its effect on patients' health and outcomes. STARE-HI is not primarily developed for assessing information resources that are made available through IT, for example patient guides on the Internet, or for studies on general attitudes of health care providers and patients toward the use of IT in health care.

In Table 1, we list the preferred structure of a publication on an evaluation study. In the next section, we then describe the items in more detail. For the rationale and for examples of good practice, we refer to a subsequent paper that is being developed.

The reader should be aware that STARE-HI is intended as a guideline, not a rigid structural standard. We do recognise that the order of items and the detail to which the various items may be described in a report or paper may depend on the audience, the type of paper, the available space, and the type of study. The target reader must also be aware of the more detailed instructions of the respective scientific journal that may affect the order of the sections and impose restrictions on their length. In addition, not all issues presented may be of relevance for a report for a particular study.

Description of Items

1. Title

The title should give a clear indication of the type of system evaluated, the study question and the study design. The use of the term "evaluation" (or "assessment" or "study") preceded by a specification of the type of study in the title helps to detect evaluation studies (e.g. "Evaluation of the effect of a CPOE system on medication errors: a retrospective record analysis").

Table 1 The STARE-HI principles: items recommended to be included in Health Informatics evaluation reports

Item #	Item
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 content
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

2. Abstract

The abstract should preferably be structured and must clearly describe the objective, setting, participants, measures, study design, major results, and conclusions. In case there are major limi-

tations in the study, these should be mentioned as well [31]. Structured abstracts are already mandatory for the International Journal of Medical Informatics, Methods of Information in Medicine, JAMIA for their research papers, and most medical journals. A given scientific journal may have detailed guidelines for organization of the abstract.

3. Keywords

To secure proper indexing in, for example, Medline it is essential to provide good keywords. Among the keywords should be “evaluation” and keywords describing the type of system evaluated (e.g. EHR, LIS, telemedicine), the setting (e.g. primary care, secondary care), the outcome measures, the study design (e.g. RCT, before-and-after, field study). It is advised to use MESH terms as provided by [32] whenever possible since that will enhance retrievability of the paper in searches.

4. Introduction

The introduction should provide the reader with the background for the rest of the paper. Hence, it should address the following issues:

4.1 Scientific Background

The scientific background is a description of what is already known about the (type of) system that is object of study. We use the term system in a broad sense. It does not necessarily refer to a complete system, but can be restricted to a certain functionality of a larger system, or the usage of a more general application in a specific domain or for a specific purpose, etc. The term covers both hardware and software systems, functionalities and algorithms, and the organizational and social environment where relevant.

4.2 Rationale for the Study

Describes briefly the motivation for the study: what are the specific reasons to perform the study (scientific interest, justification for expenditure, insight into problems, addressing open research questions)? Is the study part of a larger research, development or implementation project? From which stakeholder viewpoint (if any) is the study performed? If possible, it should also be mentioned what influence the findings of the study may have.

4.3 Objectives of the Study

The specific study questions and hypotheses must be described as concisely as possible. It should then be stated where appropriate whether any formal permission was obtained for example from institutional review boards (IRBs), ethics committees, staff committees, and the like.^{3,4}

5. Study Context

Information on study context is important for the later assessment of generalisability⁵ of results. Clearly, giving all details on the context is not feasible—the authors of an evaluation paper have to decide to which extent information is needed to secure the validity and generalisability of the paper.

³ Some journals may require that this information is provided at the end of the manuscript.

⁴ Requirements with respect to IRB approval may vary from country-to-country. Whenever participants run a certain risk by participating in a study, IRB approval should be sought. Currently, it is often considered appropriate to seek IRB approval for studies that ask health care professionals for their opinions about IT applications in a specific situation (e.g. how they perceive the CPOE system that was currently installed in their organization).

⁵ We use the term generalisability in a broad sense. It not only covers the notion that is common for quantitative studies, but also the assessment of the findings by a reader on the applicability of the results in his or her own environment. A qualitative study on the perceived usefulness of a particular system in primary care may not be generalizable to/ applicable in a tertiary setting.

5.1 Organizational Setting

This should describe the health organization(s) where the system is being evaluated, including its geographical location and preferably its name. It should indicate what kind of health care facility it is (primary, secondary, tertiary care, home care, etc). In case the whole organization is not involved, it should be made clear which departments have been involved in the evaluation (e.g. a 12-beds intensive care unit of a 320-beds urban referral hospital.).

5.2 System Details and System in Use

The system description must permit the reader to understand how the system works (or is intended to work). The authors may refer to a technical description provided elsewhere, and this may reduce the technical description, but sufficient detail is still needed for the evaluation study report to be self-contained. Systems details should comprise the aim of the IT system (e.g. laboratory system, administrative system, nursing documentation system, CPOE system), type of system (home-grown, open source or commercial system), the type of information that is managed by the system (e.g. drug orders, nursing care plans), and the clinical or other tasks supported by the system (e.g. ordering processes, nursing documentation processes).

The description should also include information on (1) how wide-spread the system is used in the facility in which the system is evaluated, for how long and for what purpose and (2) number and professions of the users of the system in that facility.

Any additional information to detail relevant aspects of the context in which the study was conducted should be mentioned (e.g. customization of software, user training, additional attention to the study, system only implemented shortly before the evaluation).

6. Methods

This section of a paper describes in sufficient detail the study design and the

methods used in the study. STARE-HI contains items which have been dealt with in more detail in guidelines for medical studies; where appropriate we will refer to these.

6.1 Study Design

This describes the overall study design and the motivation for choosing it. The description of study design comprises the type of study, for instance observational study (case study), quasi-experimental study (e.g. before-after, with or without control; interrupted time-series with or without on-off-design), or experimental study (RCT). For an RCT, the authors should follow the CONSORT guidelines for all parts of Section 6. In case of non-RCT, the taxonomy presented in (33) can be used to describe the type of study. It should also be specified whether it is a laboratory study, simulation, or actual field study. If applicable, state reasons for choice of level of observation/participation (patients, professionals, clinics, hospitals, etc.).

Throughout the description of the study design, the authors must show their awareness of specific and potential study design biases (for this purpose see the review of biases in [12], p. 257-289 and in [34], p. 209-215) by making a statement on their potential impact and how that is handled.

6.2 Theoretical Background of the Study

Where appropriate, state the theories – with sufficient references – on which the study is based, which guided the selection of the measurement instruments used and which form the basis for interpretation of the results (e.g. the user acceptance model that guided a quantitative survey or the organizational theories that guided a qualitative study).

6.3 Participants

Describe the methods of selection of participating users, patients, units, hospitals, etc., including - if applicable - inclusion and exclusion criteria for

each type of participant in a study. In case of a controlled trial, it should be specified how participants were allocated to intervention and control groups (randomization or other approaches - refer to CONSORT for more details [14, 16]). State the basis of sample size calculations if applicable (power calculations).

6.4 Study Flow

Give sufficient details on date of beginning and end of the overall study and any study periods; give clear description and date of intervention (in experimental studies). In case of a study in which several methods have been used, specify when each method was used for which group. A flow diagram should be used to summarize the experimental study designs (like the RCT flowchart as required in CONSORT [13]). For observational studies, use a diagram showing study activities over time. In each case, indicate time line and mark any important dates such as beginning of study, intervention, end of study, where appropriate compared to development milestones of the system (place/phase in life cycle).

6.5 Outcome Measures

Clearly state outcome measures or other evaluation variables of interest that were used in the study. Define to a sufficient detail the key concepts in the study such as medication error or user satisfaction. In open qualitative studies, no pre-defined outcome measure can be defined; however, when certain aspects are more in the focus of the researcher than others, these can be stated here (see also section 6.2).

6.6 Methods for Data Acquisition and Measurement

This section should provide sufficient detail such that others are able to duplicate the study or to use some of the methods for other studies.

All relevant aspects of applied methods (e.g. questionnaire, interview, ob-

servations, log file analysis, chart review) should be described. Examples of aspects are location and setting of data collection, number and type of interviews, type and duration of observations, whether data collection was retrospective or prospective, professional background of the interviewers, blinding of observer and/or participants and/or analysts, etc. It should also be identified which outcome measures are covered by each of the selected methods.

For every measurement or observation, information on their validation has to be given, with references to earlier work where necessary (e.g. was there a pre-test with assessment of inter-rater reliability? Was the questionnaire previously validated?). Newly designed measurement tools should be described in more detail, full disclosure of such tools should be given in the appendices or as supplementary material.

6.7 Methods for Data Analysis

This section describes the methods used for data analysis. The selection of those methods depends on data acquisition methods and study questions. When several methods are used, combine the description of data acquisition and data analysis for each method.

For quantitative data, state the statistical techniques that have been used for analysis. For the analysis of qualitative data, indicate the analysis methods in detail. For all data analysis methods, indicate any software product used.

Triangulation may be used to combine data from various sources [35, 36]. When triangulation is used, it should be specified what kind of triangulation was applied (methods, measures, data, investigator or theory triangulation), and how the data were combined.

Throughout the description of the data acquisition and data analysis, the authors must show their awareness of specific and potential data analysis biases (for more information, see [12], p. 290-313 and [34] p. 209-215).

7. Results

The result section presents mainly the data obtained from applying the methods as described in Section 6. Depending on the type of study, the organization and naming of sections may be done in a different way. The interpretation and discussion of the results should be left to the discussion section.

Sections 7.2 and 7.4 are of special importance for but not limited to qualitative studies, as one of their objectives is to obtain new insight in, for example social and organizational aspects of Health Informatics applications.

7.1 Demographic and other Study Coverage Data

Give basic numbers on the size of the study, for example number of users observed/interviewed, documents or medical records analysed, distribution and return rate of questionnaires, number of observation days, pages of transcripts analysed, etc.

When the study measures are related to persons, baseline demographic data and/or (clinical) characteristics of study participants (users, patients, and units) should be given, such as age groups, professions, usage patterns, patients' diagnostic scores, etc. In particular in qualitative studies the characteristics and qualities of the participants may be of more importance than sheer numbers. Information on number and type of drop-outs should be added as well with identification of reasons. Where appropriate, baseline data must be given for relevant groups separately (e.g. for control and intervention groups in trials, or for different professions or age groups when that is relevant).

7.2 Unexpected Events During the Study

Any event that may have influenced study design and/or results has to be described (e.g. deviations from timeline, system updates during the study, staff changes, educational interventions, system failure, high drop-out rate in one group and changes in management or organizational strategy during the

study period). If possible, these events should be related to the timeline of the study. The authors should indicate to what extent these unexpected events might influence (bias) the study findings.

7.3 Study Findings and Outcome Data

This is the major section, presenting the results of the study. For each study question, outcome variable and evaluation criterion, sufficient data should be presented.

Qualitative data may be presented as text. Quotations from participants should be used to illustrate major points. These quotes should be anonymous, but have an indication of the type of person being quoted (e.g. Nurse 3).

Quantitative data can be presented in tables and figures.⁶ Typically, each table, figure, etc., should be referenced in the text. The most important or relevant results should be emphasised here, and special notion should be given to unexpected or striking results such as differences between groups. Absolute numbers should always be provided; not only relative numbers.

7.4 Unexpected Observations

Any unintended (positive or negative) side-effects of the system that were not in the focus of the study but that seem remarkable should be reported here. This could be, for example, the observation of bottlenecks in the clinical workflow after system implementation, severe organizational problems that seem related to the new system, or persistent unsolicited responses on a specific effect of a system during a qualitative study that focussed on other aspects. Here the authors may report about the difference between the intended use as described in section 5.2 and the observed actual use.

⁶ Detailed advice on how to present data in tables and graphs is out of scope of STARE-HI. Guidance on this topic can be found in several publications, for example in E.R. Tufte, *The Visual Display of Quantitative Information*, Graphics Press.

The difference with section 7.2 is that unexpected events may influence the findings (and may have caused changes of the study protocol), while unexpected observations relate to issues that arose during the study that could lead to additional insights, further recommendations, potential explanations for the findings or future research topics.

8. Discussion

The discussion should be a critical interpretation and assessment of the study results and the study itself in view of the study questions. We suggest authors should make the discussion structured with the following clear subheadings [37].

8.1 Answer to Study Questions

Interpret the data and answer your study question(s). Whereas in Section 7 the results are presented in detail, in this part of the report the answers to the study questions are in focus. Make explicit reference to the specific study questions either by restating them or by other cross-reference mechanisms.

8.2 Strengths and Weaknesses of the Study

This section contains a critical discussion of the methods used. Describe the strong and weak points of the study, for example concerning the study design, comparability of intervention and control group with respect to baseline data, study execution, confounders, internal and external validity of findings, completeness of acquired data, drop-out of participants, representativeness of the participants, low or high response rates, etc. Refer here also to the information presented in Section 7.2. Discuss any biases that could be present and that would influence the findings of the study or the interpretation of the data.

8.3 Results in Relation to Other Studies

Make clear what exactly is novel about your results. Describe to what extent the results are in agreement with findings of others

and in this light provide information about the comparability with the study setting. When there is disagreement with findings of others, discuss possible reasons.

8.4 Meaning and Generalisability of the Study

Describe the meaning of the study findings, both for the various stakeholders in the study, for other institutions and for Health Informatics in general. In this context, discuss the generalisability/applicability of the study for other organizations. Refer here also to the information given in Section 5.1 and 5.2, and in 8.2.

8.5 Unanswered and New Questions

Discuss whether the study has shed new light on an issue and/or has raised new questions. Describe what research should/could be performed in the future to further improve our knowledge about the system and its effects. Refer here also to the information presented in Section 7.4.

9. Conclusion

The conclusion summarizes the main findings, discusses the impact of the findings and how they relate back to the big picture described in the Section 1, gives recommendations by the authors and provides a short outlook for future research.

10. Authors' Contribution

An increasing number of journals require making the contributions of the authors to a paper explicit or at least to make clear that each author qualifies for authorship[29]. It is recommendable to make that information part of the paper and not only in the covering letter while submitting the manuscript.

11. Competing Interests

A statement of the interests, financial or otherwise, the authors may have with respect to subject of study and which

may – but not necessarily have – influenced the design of the study and/or the interpretation of the results [29]. This information should also make clear the relation of the study team towards the system being evaluated (developers, users, operators, internal quality assurance, external team, etc.).

12. Acknowledgements

Acknowledge any financial or other support you got when conducting the study or writing the paper [29].

13. References

References should adhere to the journal guidelines. When no strict guidelines are provided by the journal, the style recommended by the NLM is preferred [38]. The authors should include only references that are really needed for their arguments. It is improper to provide a long list of references only to demonstrate that many papers have dealt with a specific issue. Only the key papers should be included. Preferably such lists should be restricted to references the contents of which are discussed in the paper.

14. Appendices

Any supporting material should be included in appendices. This includes necessary, detailed descriptions of specific measurement methods/tools (e.g. a questionnaire), specific data analysis techniques and detailed study results. Currently, more and more journals allow such additional material being stored on the publisher's website. This also opens up the ability to include (parts of) audio-visual and multi-dimensional material collected in the study.

Discussion

STARE-HI was developed to provide guidelines for writing and interpreting

evaluation reports in Health Informatics, by doing this to improve the quality of published IT evaluation studies in Health Informatics; and thus to improve the evidence base of Health Informatics. It encourages transparency in reporting of IT evaluation studies.

STARE-HI was developed in an iterative process involving volunteer experts from various Health Informatics domains. No formal procedures (e.g. Delphi technique) or voting procedure was used to develop this version, but all colleagues interested could submit comments that were assessed by the editorial group. Several presentations of drafts of STARE-HI at various workshops and conferences were used to make sure no important aspect was missed.

STARE-HI borrows from earlier work in other domains such as CONSORT for reporting of RCTs; however, to our knowledge, STARE-HI is the first approach to develop specific guidelines for the reporting of Health Informatics evaluation studies.

Although the guidelines are written for the publication of evaluation studies in the scientific literature, they can also be used for other reporting purposes. Although the structure might require some changes, for example the Abstract could become an executive summary and the conclusions could become a set of recommendations, the issues are largely valid for any kind of reporting of studies.

Conclusion

We present STARE-HI as a guideline to report IT evaluation studies in health care with detailed recommendations for each aspect that is particularly relevant for an evaluation study.

Whether STARE-HI is feasible for the broad range of (quantitative and qualitative) Health Informatics evaluation papers can only be shown when it is used by authors and editors. We invite anybody to report their experience

that may be incorporated in subsequent updates of STARE-HI. Subsequent more rigid evaluation studies on this question are planned.

We perceive higher quality publications of evaluation studies to be an important step towards the vision of scientifically valid evidence-based Health Informatics. Such an evidence base will help

- to demonstrate the value of Health Informatics applications,
- to assist decision makers in finding evidence on the most appropriate systems and approaches,
- to guide the field towards developments that have maximal beneficial impact on health care delivery, and
- to demonstrate that Health Informatics has a solid scientific basis and hence is a discipline that can contribute to further improvement of the quality of care.

Authors' Contributions

The idea for STARE-HI was raised during the HISEVAL workshop in Innsbruck [2]. J.T. took the initiative to develop STARE-HI, he is the guarantor of the study. J.T. and E.A. drafted a first list of issues and drafted the first version of the manuscript. J.B., N.d.K., P.N., and M.R. all contributed by critically assessing the items and their descriptions in several iterations. They have made suggestions for expansion and provided various parts of the text. J.T. and E.A. integrated the various contributions and edited the final version of the manuscript. All authors have approved this final version.

Competing Interests

J.T. is editor of the International Journal of Medical Informatics. He was chair of the working group on Technology Assessment and Quality Development of the International Medical Informatics

Association (IMIA) with J.B. as co-chair. Currently N.d.K. chairs this working group.

E.A. is chair of the Working Group on Assessment of Health Information Systems of the European Federation of Medical Informatics (EFMI). J.B. and P.N. are co-chairs of this working group.

Acknowledgment

This work has been validated and enriched by comments from many during its preparation and development. In particular we acknowledge the contributions of: Jos Aarts, Emily Campbell, Petra Knaup, Christof Machan, Zahra Niazkhani, Christian Nøhr, Habib Pirnejad, Joshua Richardson, Rainer Röhrig, Dean F. Sittig, Murat Sincan, Christa Wessel, Johanna Westbrook and Jeremy Wyatt.

Endorsements

EFMI has formally endorsed STARE-HI in their board meeting of 30 May 2007 in Brijuni, Croatia. In November 2007, the working group on Evaluation of the American Medical Informatics Association (AMIA) endorsed STARE-HI during the AMIA 2007 Annual Meeting. The General Assembly of IMIA has adopted the STARE-HI paper as a formal IMIA document on 25 May 2008.

References

1. Sackett D, Rosenberg W, Gray J, Haynes RB, Richardson S. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312(7023):71-2.
2. 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* 2004 Jun 30;73(6):479-91.
3. Ammenwerth E, Shaw NT. Bad health informatics can kill - is evaluation the answer? *Methods Inf Med* 2005;44(1):1-3.

4. Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *J Am Med Inform Assoc* 2004 Mar-Apr;11(2):104-12.
5. Tenner E. Why things bite back: technology and the revenge of unintended consequences. New York, NY: Vintage Books; 1997.
6. Rigby M. Evaluation: 16 Powerful Reasons Why Not to Do It - And 6 Over-Riding Imperatives. In: Patel V, Rogers R, Haux R, editors. *MEDINFO 2001*; 2001: IOS Press; 2001. p. 1198-202.
7. Talmon J. Evaluation and implementation: A call for action. *Methods Inf Med*. 2006;45(Suppl 1):S11-S5.
8. Rigby M. Evaluation – the Cinderella Science of ICT in Health. *Methods Inf Med* 2006;45(Suppl 1):S114-S20.
9. Whitten PS, Mair FS, Haycox A, May CR, Williams TL, Hellmich S. Systematic review of cost effectiveness studies of telemedicine interventions. *BMJ* 2002 Jun 15;324(7351):1434-7.
10. Ammenwerth E, de Keizer N. An inventory of evaluation studies of information technology in health care trends in evaluation research 1982-2002. *Methods Inf Med*. 2005;44(1):44-56.
11. Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med* 2006 May 16;144(10):742-52.
12. Brender J. *Handbook of Evaluation Methods for Health Informatics*. New York: Academic Press; 2006.
13. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996 Aug 28;276(8):637-9.
14. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001 Apr 17;134(8):663-94.
15. Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004 Mar 20;328(7441):702-8.
16. Boutron I, Moher D, Altman D, Schulz KF, Ravaud P. Methods and Processes of the CONSORT group: Example of an extension for trials assessing nonpharmacologic treatments. *Ann Intern Med* 2008 19 February 2008;148(4):W-60-W-6.
17. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman D, et al. Consort for reporting randomized trials in journal and conference abstracts. *Lancet* 2008 January 26, 2008;371:281-3.
18. 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 Nov 27;354(9193):1896-900.
19. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Standards for Reporting of Diagnostic Accuracy*. *Clin Chem*. 2003 Jan;49(1):1-6.
20. von Elm EA, Altman D, Egger M, Pocock SJ,

- Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement. *Lancet* 2007.
21. Falagas ME, Pitsouni EI. Guidelines and consensus statements regarding the conduction and reporting of clinical research studies. *Arch Intern Med* 2007 May 14;167(9):877-8.
 22. The Equator-Network Website. 2008 [cited 26 May 2009]; Available from: <http://www.equator-network.org/>
 23. Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001 Apr 18;285(15):1992-5.
 24. Egger M, Juni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001 Apr 18;285(15):1996-9.
 25. Heathfield HA, Buchan IE. Current evaluations of information technology in health care are often inadequate. *BMJ* 1996 Oct 19;313(7063):1008.
 26. Kaplan B. Evaluating informatics applications—some alternative approaches: theory, social interactionism, and call for methodological pluralism. *Int J Med Inform* 2001 Nov;64(1):39-56.
 27. Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ* 1996 Sep 7;313(7057):570-1.
 28. Ammenwerth E, Wolff AC, Knaup P, Ulmer H, Skonetzki S, van Bommel JH, et al. Developing and evaluating criteria to help reviewers of biomedical informatics manuscripts. *J Am Med Inform Assoc* 2003 Sep-Oct;10(5):512-4.
 29. International Committee of Medical Journal Editors. Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication. 2007 [cited 26 May 2009]; Available from: www.ICMJE.org
 30. Talmon J, Ammenwerth E, Geven T. The quality of reporting of health informatics evaluation studies: A pilot study. *MEDINFO* 2007; 2007; Brisbane, Australia; 2007.
 31. Haynes RB, Mulrow CD, Huth EJ, Altman DG, Gardner MJ. More informative abstracts revisited. *Ann Intern Med*. 1990 Jul 1;113(1):69-76.
 32. National Library of Medicine. Medical Subject Headings. 2008 [cited 26 May 2009]; Available from: <http://www.nlm.nih.gov/mesh/meshhome.html>
 33. Harris AD, McGregor JC, Perencevich EN, Furuno JP, Zhu J, Peterson DE, et al. The use and interpretation of quasi-experimental studies in medical informatics. *J Am Med Inform Assoc* 2006 Jan-Feb;13(1):16-23.
 34. Friedman C, Wyatt J. *Evaluation Methods in Biomedical Informatics*. 2 ed. New York, NY: Springer-Publishing; 2005.
 35. Greene J, McClintock C. Triangulation in evaluation: Design and analysis issues. *Evaluation review*. 1985;9(5):523-45.
 36. Kaplan B, Duchon D. Combining qualitative and quantitative approaches in information systems research: a case study. *MIS Quarterly* 1988;12(4):571-86.
 37. Docherty M, Smith R. The case for structuring the discussion of scientific papers. *BMJ* 1999 May 8;318(7193):1224-5.
 38. Patrias K. *Citing medicine: the NLM style guide for authors, editors and publishers* [Internet]. 2007 [cited 26 May 2009]; 2nd ed.: Available from: <http://www.nlm.nih.gov/citingmedicine>

Correspondence to:

Jan L. Talmon
 Center for Research, Innovation, Support and Policy
 Maastricht University
 PO Box 616
 6200MD Maastricht
 The Netherlands
 Tel: +31 43 388 2243
 E-mail: Talmon@ML.unimaas.nl or jan.talmon@mumc.nl

This article was published in the *International Journal of Medical Informatics* 2009;78(1)1-9. Talmon J, Ammenwerth E, Brender J, de Keizer N, Nykänen P, Rigby M. STARE-HI - Statement on reporting of evaluation studies in Health Informatics. Copyright Elsevier 2009