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

Jan Talmon 1*, Elske Ammenwerth 2, Jytte Brender 3, Nicolette de Keizer 4, Pirkko Nykänen 5, Michael Rigby 6

1 Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands
2 UMIT – University for Health Sciences, Medical Informatics and Technology, Hall in Tyrol, Austria
3 Aalborg University, Aalborg, Denmark
4 Dept Medical Informatics, Academic Medical Center, Amsterdam, The Netherlands
5 Tampere, Finland
6 Keele University, Keele, United Kingdom

* Corresponding Author
Care and Public Health Research Institute (CAPHRI)
Maastricht University
PO Box 616
6200 MD Maastricht
The Netherlands
+31-43-3882243
talmon@mi.unimaas.nl

THIS VERSION IS THE DRAFT FOR PUBLICATION.
CONTRIBUTERS ARE LISTED ON UMIT’s WEBSITE
THIS IS WORK IN PROGRESS

Any comments welcome to elske.ammenwerth@umit.at or talmon@MI.unimaas.nl.

The most recent version of STARE-HI can be found at http://iig.umit.at/efmi.
Information on further work on STARE-HI is provided through the mailing list of the EFMI-WG on Assessment of Hospital Information Systems (http://iig.umit.at/efmi), free membership to the mailing list through http://listman.umit.at/mailman/listinfo/eval.
Abstract

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 of journals in Health Informatics and as authors of systematic reviews of Health Informatics studies, 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 and has been put up for comments on the web.

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 of an IT evaluation paper.

Conclusion: A comprehensive list of principles relevant for properly describing Health Informatics evaluations has been developed. When manuscripts submitted to Health Informatics journals as well as 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. It will also be possible to judge better whether papers will fit in the scope of meta-analyses of Health Informatics interventions. 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 towards the vision of evidence-based Health Informatics.

Limitations: This study is based on experiences from editors, 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
Introduction

In all aspects of health care it is recognised as good practice and ethically important that application of technologies, methods, and interventions are evidence-based.\(^{(1)}\) In other words, that they are proven to be safe and effective, against 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 organisations on them, it is imperative that they should be thoroughly assessed through robust evaluations as with any other form of health process or technology, as advocated in the Innsbruck Declaration \(^{(2)}\).

In the past decades it has been demonstrated that IT systems are not only beneficial, but also can have unintended, potentially detrimental effects as documented by \(^{(3, 4)}\), just like other interventions in social environments \(^{(5)}\). It is imperative to monitor closely IT system implementations and their effects during the whole life-cycle of the system. Unintended as well as intended consequences are important to report since they provide insight which 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 studies in the literature is small given the importance and impact of Health Informatics on modern healthcare practice and delivery \(^{(8)}\), the evidence that is available is partly difficult to appreciate due to poor reporting practice\(^{(9)}\).

There is growing published evidence of the impact of Health Informatics\(^{1}\) on health care \(^{(10)}\), and increasingly reviews appear that summarize available evidence\(^{2}\) 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 \(^{(11)}\). Also other studies have

\(^{1}\) We use the term Health Informatics rather than Medical Informatics. The latter has the annotation that it deals with the medical condition of patients, but there are applications that cover a broader scope, including the health of a broader part of the population.

\(^{2}\) 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 e.g. user experiences as collected with questionnaires or qualitative studies and results from the analysis of implementation processes. That means that evidence-based health informatics is based on both quantitative and qualitative research results.
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.

The authors and contributors of this paper act as editors and/or reviewers for several Medical Informatics journals. Some of the authors conducted systematic reviews on evaluation of Health Informatics applications. We have noted that many manuscripts and papers dealing with the evaluation of IT interventions in health care are of questionable quality. Often, important information such as the functionality of the IT system evaluated, the clinical setting or the evaluation methods were not sufficiently described. Discussions were often incomplete with regard to validity and generalisability of findings.

In other domains of medicine these problems have also been identified. Work done in the early nineties lead to the publication of the CONSORT statement in 1996 (13). The CONSORT statement provides guidelines for the publication of randomized controlled clinical trials (RCT). This statement has been adopted by many medical journals. Later, in 2001, this statement was revised (14) and has been extended to cover other study designs such as cluster RCTs (15). Several other guidelines have been developed along the lines of the CONSORT statement; for example the QUOROM statement for reporting of meta-analyses (16) and STARD for reporting of diagnostic studies (17). An overview of the various guidelines has recently been published in (18).

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 (19, 20). Guidelines for good reporting of studies are also likely to have an influence on the quality of the studies themselves as well, as a result of the requirement for 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 so as to build up a more solid evidence base. However, Health Informatics applications potentially have effects on health care organisations, 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 (this exception is diminishing with the advent of distributed and ubiquitous technologies (21, 22)). This means that the study designs (mostly RCTs) that are covered by CONSORT and comparable approaches are not always the most
STARE-HI v0.91

appropriate in Health Informatics evaluation research (2, 23, 24), where other study
designs, including qualitative ones, are frequently used.

These observations have led us to the decision to develop guidelines for reporting of
evaluation studies in Health Informatics, which build upon work of others, but that take
into account some specific issues that are typical for Health Informatics evaluation
studies and that cover the broader spectrum of (quantitative and qualitative) study designs
as found in Health Informatics research.

Objectives

The objectives of STARE-HI (STAtement on the Reporting of Evaluation studies in
Health Informatics) are to provide guidelines for writing and assessing evaluation reports
in Health Informatics; 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 which are formatted as a checklist
for publication 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. Then, the various items are discussed in more detail.

A description of the rationale of each item as well as extended examples will be
published in a subsequent separate paper.

Method

An iterative refinement approach has been followed for the development of STARE-HI.
A first set of items was drafted by a core editorial group (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. The CONSORT statement (13, 25), criteria for reviewers of biomedical
informatics manuscripts (26), the QUOROM statement (16), the STARD statement (17)
and other more general recommendations on publication quality such as (27) were used
as further reference material.

The approach and the first list of items were then further developed; preliminary versions
were presented to and discussed with colleagues world-wide at various events, namely
during workshops at Medinfo2004, MIE2005, MIE2006, and AMIA 2006,

The resulting draft was then openly published on the website of the EFMI-WG EVAL,
and all members were informed of the call for comments at the end of 2006. In spring
2007, STARE-HI was put on the EFMI WG Website, and potential authors were invited
to use and validate STARE-HI. The most recent 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 (28). This study provided feedback on various aspects of STARE-HI.

All received comments have been taken into account, leading to the version presented in this paper.

**STARE-HI recommendations**

The scope of STARE-HI is to provide guidelines for the reporting of evaluations in Health Informatics, independent of whether it is a qualitative, quantitative or mix-method evaluation.

We consider as the object of evaluation the Health Informatics application that is assessed. This assessment can cover the technical artefact as such with 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 outcome. STARE-HI is not developed for assessing information resources that are made available through IT, for example patient guides on the internet, nor for studies on general attitudes of health care providers and patients towards the use of IT in health care.

In Table 1 we list the preferred structure of a publication on an evaluation study. In the following section we then describe the items in more detail and provide examples.

The reader should be aware that STARE-HI is intended as a model guideline, not a rigid structural standard. We do recognize that the order of items as well as the detail to which the various items are described 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 instruction 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.
Table 1: THE STARE-HI PRINCIPLES: Items recommended to be included in any Health Informatics evaluation
Description of items

1. Title
The title should give a clear indication of the type of evaluated system and the study question as well as 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”).

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 limitations in the study, these should be mentioned as well (29). Structured abstracts are already mandatory for the International Journal of Medical Informatics, JAMIA, Methods of Information in Medicine, 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 (30) 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
This is a description of what is already known about the (type of) system that is object of study, what are the open research questions and why is there a need to answer them. 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 application 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, as well as 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)? 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.\(^3\)

5. **Study context**
Information on study context is important for the later assessment of generalisability 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.

5.1 **Organisational setting**
This should describe the health organisation(s) where the system is being evaluated, including its name and its geographical location. It should give an indication of 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 level of detail of 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 related information as well as the methods used in the study. STARE-HI contains items which

---

\(^3\) 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).
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 (e.g. any time, budget or organizational constraint). 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). In the case of RCTs the authors should follow the CONSORT guidelines for all parts of section 6. In case of non-RCT, the taxonomy presented in (31) 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 allocation unit (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. 243-323) by making an account of 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 Technology Acceptance Model (32) that guided a quantitative survey, or 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 all kinds of participants (e.g. hospitals, departments, users, patients). 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 (13)). Also state the basis of sample size calculation 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
may be used, 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. For each outcome measure or aspect of the study, details of methods used to obtain the outcome data (e.g. questionnaire, interview, observation, log file analysis, chart review) should be provided (e.g. 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 observers and/or participants and/or analysts etc.).

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.

**6.7 Methods for data analysis**

This section describes the methods used for data analysis. The selection of those methods clearly 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 which statistical techniques 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 (33, 34). 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. 243-323).

**7. Results**

The result section presents mainly the data obtained from applying the methods as described in chapter 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 get new insight for example social and organisational aspects of Health Informatics applications.

**7.1 Demographic data**

Baseline demographic data and clinical characteristics of study participants (users, patients, and units) and of the study should be given first, such as age groups, professions, usage patterns, patients’ diagnostic scores etc. 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 group separately (e.g. for control and intervention groups in trials, or for different professions or age groups when that is relevant). Give basic numbers on the size of the study, for example number of users observed or interviewed, documents or medical records analysed, distribution and return rate of questionnaires, number of observation days, pages of transcripts analysed etc.

7.2 Unexpected events during the study
Any event that may have influenced study design and/or results has to be described in detail (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 organisational strategy during the study period). The authors should indicate to what extent these unexpected events may 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. Quantitative data can be presented in tables and figures. 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. Make sure to always provide absolute numbers and not only relative numbers. Qualitative data should 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).

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 comprises all observations on unintended effects. This could be, for example, the observation of bottlenecks in the clinical workflow after system implementation, severe organizational problems which 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. 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 deal with issues that arose during the study that could lead to additional insights, further recommendations or future research topics.

8. Discussion
The discussion should be a critical interpretation and assessment of the study results as well as the study itself in view of the study questions. We suggest authors should make the discussion structured with the following clear subheadings (35).

8.1 Answer to study questions
Interpret the data and answer your study question(s). While in section 7 the results are presented in detail, here, in the discussion, the answers to the study questions are in the
centre. Make sure that it is clear which study questions and hypotheses are addressed either by restating them or by other cross-reference mechanisms.

8.2 Strengths and weaknesses of the study
This section comprises a critical discussion of the methods used. Describe the strong and weak points of the study, for example with regard to study design, comparability of intervention and control group with regard to baseline data, study execution, confounders, internal and external validity of findings, completeness of acquired data, drop-out of participants 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; for this purpose see the description of biases in (12).

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 as well as for other institutions and for Health Informatics in general. In this context, discuss the generalisability 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 introduction section, gives recommendations by the author and provides a short outlook for future research.

10. Authors’ contribution
Medline requires journals to make explicit the contributions of the authors to a paper or at least to make sure that each author qualifies for authorship. 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 interest, financial or otherwise, the authors may have with respect to the outcome of the study. It should also be made clear what interest the authors of the paper, designers of the study, those who execute the study, and those who do the interpretation of the results have with respect to the outcome of the study. This
information should also include the description of the relation of the study team towards the system being evaluated (developers, users, operators, internal quality assurance, external team).

12. Acknowledgements
Acknowledge any financial or other support you got when conducting the study or writing the paper.

13. References
References should adhere to the journal guidelines. When no strict guidelines are provided by the journal, the Vancouver style is preferred. 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 material collected in the study.

Discussion
STARE-HI was developed to provide guidelines for writing and assessing IT 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. In encourages to 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 voting procedure was used to come to this version 1, but all colleagues interested could submit comments that were dealt with by the editorial group. Several presentations of drafts of STARE-HI at various workshops and conferences were used to make sure no important aspects were missed.

STARE-HI borrows from earlier work in other fields such as CONSORT; 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

---

5 For items 10, 11 and 12 the reader is also referred to the website of the International Committee of Medical Journal Editors for their recommendations: http://www.icmje.org. We expect that Medical Informatics Journals will require adherence to these requirements as well.
executive summary and the conclusions could become a set of recommendations, the issues are largely valid for other forms 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 now 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 that will help
a) to demonstrate the value of Health Informatics applications,
b) to assist decision makers in finding evidence on the most appropriate systems and approaches,
c) to guide the field towards developments that have maximal beneficial impact on health care delivery, and
d) to demonstrate that Health Informatics has a solid scientific basis and hence is a serious discipline that can contribute to further improvement of the quality of care.

**Authors’ contributions**

The idea for STARE-HI was raised during the ESF workshop in Innsbruck (36). JT took the initiative to develop STARE-HI, he is the guarantor of the study. JT and EA drafted a first list of issues. EA and JT drafted the first version of the manuscript. JB, NdK, PN, and MR all contributed by critically assessing the items and their descriptions. They have made suggestions for expansion and provided various parts of the text. JT and EA integrated the various contributions and wrote the final version of the manuscript. All authors have approved this final version.

**Competing Interests**

JT is editor of the International Journal of Medical Informatics. He is also chair of the IMIA working group on Technology Assessment and Quality Improvement.
EA is chair of the EFMI working group on Evaluation. JB, NdK and PN are co-chairs of these working groups.
Acknowledgment

We are indebted to all persons that provided helpful input to this paper. The European Federation of Medical Informatics has formally endorsed STARE-HI in their board meeting of May 30th, 2007 in Brijuni, Croatia.

References