



OPEN ACCESS

The SQUIRE Guidelines: an evaluation from the field, 5 years post release

Louise Davies,^{1,2,3} Paul Batalden,³ Frank Davidoff,³ David Stevens,³ Greg Ogrinc^{4,5}

For numbered affiliations see end of article.

Correspondence to

Dr Louise Davies, VA Outcomes Group –111B, Department of Veterans Affairs Medical Center, 215 North Main Street, White River Junction, VT 05009, USA; Louise.Davies@dartmouth.edu

Received 25 February 2015

Revised 16 May 2015

Accepted 23 May 2015

Published Online First

18 June 2015

ABSTRACT

Background The Standards for Quality Improvement Reporting Excellence (SQUIRE) Guidelines were published in 2008 to increase the completeness, precision and accuracy of published reports of systematic efforts to improve the quality, value and safety of healthcare. Since that time, the field has expanded. We asked people from the field to evaluate the Guidelines, a novel approach to a first step in revision.

Methods Evaluative design using focus groups and semi-structured interviews with 29 end users and an advisory group of 18 thinkers in the field. Sampling of end users was purposive to achieve variation in work setting, geographic location, area of expertise, manuscript writing experience, healthcare improvement and research experience.

Results Study participants reported that SQUIRE was useful in planning a healthcare improvement project, but not as helpful during writing because of redundancies, uncertainty about what was important to include and lack of clarity in items. The concept "planning the study of the intervention" (item 10) was hard for many participants to understand. Participants varied in their interpretation of the meaning of item 10b "the concept of the mechanism by which changes were expected to occur". Participants disagreed about whether iterations of an intervention should be reported. Level of experience in writing, knowledge of the science of improvement and the evolving meaning of some terms in the field are hypothesised as the reasons for these findings.

Conclusions The original SQUIRE Guidelines help with planning healthcare improvement work, but are perceived as complicated and unclear during writing. Key goals of the revision will be to clarify items where conflict was identified and outline the key components necessary for complete reporting of improvement work.

INTRODUCTION

In 2008, the Standards for Quality Improvement Reporting Excellence (SQUIRE) Guidelines were published, designed to support the scholarly publication of healthcare improvement work. The Guidelines resulted from a nearly 3-year development period. The work included a statement of purpose, followed by a working draft, a modified Delphi process to identify key content items and consensus meetings to reach a final version.^{1 2} The goal of SQUIRE was to increase the breadth and frequency of published reports; improve the utility of the reports by enhancing transparency, comprehensiveness and rigour; and encourage reflection on the epistemology and emerging science of improvement.³

In 2012, we began a revision of the SQUIRE Guidelines, anticipating the creation of a new version, to be published in 2015. Why was this necessary so soon after their initial release? SQUIRE is guiding a field that is growing and changing—the science of healthcare improvement is dynamic: its methods and boundaries are still under development.^{4–7} The field is still determining to what degree the work should be considered research,⁸ and the definition of the science has not reached consensus.^{9 10}

Because of this environment, we decided the SQUIRE Guidelines revision process should explicitly include efforts to understand the issues in the field as experienced by those working in improvement. While guidance for the developers of reporting guidelines suggests consumers might be involved as members of the development team,¹¹ asking consumers to evaluate existing guidelines, to our knowledge, has not been done before.



CrossMark

To cite: Davies L, Batalden P, Davidoff F, et al. *BMJ Qual Saf* 2015;**24**:769–775.

We report here the findings from an evaluation performed to obtain people's views on SQUIRE as a tool for writing scholarly healthcare improvement manuscripts. Our research question was, "What are people's experiences with and impressions of the SQUIRE Guidelines?" These findings will support the development of SQUIRE 2.0—the updated set of guidelines that will be released in late 2015.

METHODS

The evaluative design had three objectives: determine how people interpret and apply the items of SQUIRE; outline barriers to writing about healthcare improvement work; and identify key emerging issues for those doing scholarly healthcare improvement writing. We sought input from both end users of the guidelines and an advisory group of thinkers in the field of healthcare improvement.

End user evaluations of SQUIRE were obtained through four 1½ h focus groups (one in person and three via interactive two-way video mechanisms) and nine 1 h semi-structured interviews (all by telephone) performed between October 2012 and June 2013. A topic guide was used to conduct the first part of the interview or focus group, and then the participants were offered the opportunity to comment on the Guidelines item by item. Interviews and focus groups were digitally recorded and professionally transcribed. All interviews and focus groups were completed by a researcher with extensive experience in these methods (LD). The researcher was knowledgeable about healthcare improvement but uninvolved in the development or dissemination of the SQUIRE Guidelines prior to beginning the research, and she introduced herself as such at the start of each interview and focus group. For all but two of the interactions, this was the only person present with the participants; the other authors (all of whom had been part of the development of the Guidelines) were not present. The two focus groups conducted in Sweden were conducted in the presence of two people who had been involved in the Guideline development to manage technical issues (PB) and allow for immediate iterations if necessary of the interview guide (GO). Prior to completing the interviews, the main interviewer (LD) had no personal relationship with any of the interviewees.

Input from thinkers in the field was obtained from an advisory group of 18 people. Thinkers were defined as people who had contributed to healthcare improvement by developing the field through action or writing, disseminating key ideas and/or teaching extensively about healthcare improvement. Participants in this group were identified based on prior involvement in the SQUIRE Guideline development or publications related to the methods and science of healthcare improvement. The advisory group contributed in three ways in the style of participatory action research—in which participants both help direct and are the subject

of the research.¹² First, they provided an approach for us to triangulate emerging findings from end users—triangulating is the process of checking veracity of emerging findings by seeing if they are present in other data sources.¹³ Second, they contributed data about existing and emerging issues around SQUIRE item definitions and interpretation; third, they helped with data interpretation and the process of translating the interpretation into changes in the SQUIRE Guidelines.

The sampling strategy for composing the end users group was purposive,¹⁴ employing a 'maximal variation' sampling technique.¹⁴ Maximal variation sampling means participants are chosen across a variety of characteristics, rather than just one or two. We sought variety in (1) work settings—for example, both front-line care providers and improvement consultants; (2) healthcare specialties—that is, medical specialties as well as different professions; (3) training—for example, social sciences, administration and healthcare; (4) experience in writing and in healthcare improvement work; and (5) geographic location. Sampling for the advisory group focused on achieving balance by gender, profession and training. Nineteen individuals were invited to achieve a group of 18.

Invitations to end users were given by email, describing the request for an interview or participation in a focus group as a research project. Candidates for interviews and focus groups were drawn from lists of alumni and faculty of improvement fellowships, programmes and organisations, and lists of healthcare improvement conference attendees. To locate authors who were outside the range of people already on the lists above, we completed a Google Scholar search of people citing SQUIRE in their improvement publication. Sampling was considered complete when thematic saturation was reached—no new ideas or concepts were emerging.¹³

Interviews and focus groups proceeded in small waves, with an interim analysis between each wave. Triangulation of the data from end users was completed by asking later interviewees and members of the advisory group to confirm or disconfirm key statements from prior interviewees. We analysed the data using a grounded theory approach¹³ through close reading of transcripts and written comments, followed by initial coding of data into major categories. Finally, through a process of data reduction, important findings were collapsed into major themes. Interim analyses of emerging findings were presented monthly to study co-investigators ([redacted]) and twice to the advisory group described above. The goal of the presentations was to triangulate emerging findings, identify further areas for investigation in subsequent data collection from end users, interpret the data and reach consensus when there were conflicts in coding or analysis.

The COREQ Guidelines were used to guide the reporting of this work.¹⁵

RESULTS

Twenty-nine end users of SQUIRE participated in focus groups (n=20) and semi-structured interviews (n=9). Participants were located in the USA, Canada, Sweden, the UK and Norway. Forty-two people were approached to achieve this sample size. Work settings represented included government, private business, academic and community. Experience with healthcare improvement ranged from self-taught authors, to current students in fellowship programmes, to faculty supporting healthcare improvement work—themselves with varied levels of experience in healthcare improvement. Nurses, physicians, physical therapists, administrators and doctorally prepared social scientists participated. Medical specialties represented were internal medicine, obstetrics and gynaecology, rheumatology, neonatology, paediatrics, infectious disease, psychiatry, anaesthesia and critical care. Of the 29 end users, 11 were not native English speakers. Major findings about how SQUIRE has been experienced by users fell into three major themes, described below.

Arranging SQUIRE items into the traditional framework of a scientific manuscript is challenging for end users

The SQUIRE Guidelines were uniformly praised as a document that was very useful for planning a healthcare improvement project, but harder to use in the task of writing about the work. Said participant 12, a PhD social scientist and healthcare improvement consultant: “We use SQUIRE a lot for planning—we complete the sections up through the methods at the time we design the study...[but] SQUIRE creates sort of long reports if followed exactly.” Participant 21, a graduate of a 2-year fellowship in healthcare improvement, said, “The Guidelines tell me everything [to think about], but they don’t tell me what is important to include [in my manuscript]... there is no hierarchy...”

Guideline items 5, 6, 9–11 and 13–19 elicited substantial reactions that were informative to the revision process (table 1). Common concerns included requests for seemingly similar information in both the methods and results section, a sense that some items were in the wrong place, and a lack of clarity about what information was being requested in certain items. In item 10, the concept of reporting the ‘study of the intervention’ was simply not comprehensible to several participants. These participants suggested tools or visuals were needed to distinguish between ‘the work of the intervention’ and ‘the study of the intervention’ components to a project. Last, some felt the Guidelines seemed to presume a linearity to improvement work that might not be present.

Among end users, usability perceptions of the SQUIRE

Guidelines varies with experience in scholarly medical writing

Participants with less experience in scholarly medical writing found the SQUIRE Guidelines harder to

understand and apply than those with more such experience. For example, participant 8, a PhD social scientist new to improvement work reporting, gave an initial impression of the Guidelines as follows: “Everything [in the SQUIRE checklist] is in such small pieces. You do not get the whole picture of what you are supposed to be doing.” She further explained that it felt hard to understand the whole of the task of writing about a healthcare improvement project because the checklist approach did not help her understand which parts of the work should be reported where or in what order. This observation was echoed by many others in the sample with a similar background in non-medically oriented scholarly writing.

Among those with more experience in scholarly medical writing but new to scholarly writing about healthcare improvement work, the SQUIRE Guidelines made more sense, and they identified similarities to the experience of using other publication guidelines. Participant 9, a nurse researcher with experience in both mixed methods research and clinical trials, noted that she found the Guidelines generally useful and clear, but had to draw on her mixed methods background to develop ways to teach the concept of reporting ‘the work’ as well as the ‘study of the work’ of healthcare improvement: “The doctors [I was advising]—they only knew statistics and quantitative work—they were only familiar with ‘context free’ research. I explained ‘[reporting QI work]’ is like writing up trial results and then also the experience of running that trial”.

Among participants who had experience with scholarly writing about healthcare improvement, or who had previous experience with scientific medical writing in general, the SQUIRE Guidelines were perceived as easier to use. Said participant 13, a physician researcher: “Much of the SQUIRE guidelines is commonsense, it is a transplantation of research methods into QI, making sure your QI project is written up as rigorously as traditional research”. Participant 17, physician graduate of an improvement fellowship, reflected on his experience since first using the SQUIRE Guidelines several years prior. He confirmed what others in his focus group noted about the challenges of using Guidelines and the impossibility of trying to include every item, saying “I whined to my advisor [when I was a fellow] about that very problem —[the document you create if you use SQUIRE exactly as written is unintelligible]. But the problem is I used it to write my very first paper. SQUIRE seems very simple to me now”. He had realised with experience and re-reading the Guidelines that SQUIRE did not require him to include every item in the manuscript, and that part of the challenge he had faced as a fellow was the work of being new to scientific writing in general.

Table 1 Selected Standards for Quality Improvement Reporting Excellence (SQUIRE) Guideline items that elicited specific comments from end users (quotes from focus groups and interviews that are illustrative of specific areas of concern in the document)

Item(s) eliciting comment	End user comments
Introduction	
Item 5. <i>Intended improvement</i> a) Describes the specific aim of the intervention b) Specifies who and what triggered the decision to make changes, and why now	<i>"Is this where the process of the project goes, or is it in the methods section?" Participant 23, nurse researcher</i> <i>"We worked with this item and item 6 in several loops, our intended improvement changed because things changed in our work." Participant 6, healthcare improvement fellow</i>
Item 6. <i>Study question</i> State the improvement—related question and any secondary questions that the study of the intervention was designed to answer	<i>"Primary and secondary improvement related question is confusing, what does that mean?... We had a hard time with the [difference between] improvement question and the study question." Participant 7, healthcare improvement fellow</i>
Methods section	
Item 9. <i>Planning the intervention</i> Intervention, component parts, factors contributing to choice of intervention, initial plans for intervention implementation	<i>"One of the things that permeates much of this writing here, the guidelines in this section [methods and results items], is this assumption of a very preconceived linear, extremely methodical approach to studying it. We know that lots of improvement work is messy and emergent... So the guideline, the way the guideline is framed, or written, gives you the sense that if you've been in this messy reality, you are off, you can't use this."</i>
Item 10. <i>Planning the study of the intervention</i> Plans for assessing how well intervention was implemented, mechanisms of change and plans for testing, plans for implementing chosen study design, internal and external validity	<i>"...the methods section in general lacks a place for qualitative work such as member checking or triangulation—various ways to measure validity—consider adding words like credibility or trustworthiness..."</i>
Item 11. <i>Methods of evaluation</i> Instruments and procedures to assess effectiveness of intervention and assessment instruments, methods to assure data quality and accuracy	<i>"...the initial intervention made us do [things] differently in the next cycle, is that a result, or is that a method?" Participants 2–4, improvement fellowship faculty members</i> <i>"More about the theory of the intervention should go here. This is something that I have learned more about as a QI professional lately—it has required more professional development for me." Participant 12, professional healthcare improvement consultant</i>
Results section	
Item 13. <i>Outcomes</i> The setting, the course of the intervention, degree of success, evolution of the plan, data on change in processes of care, benefits, harms, unexpected results, problems, failures, strength of associations, missing data	<i>"It is odd to describe the setting again here..." Participant 24, healthcare improvement fellow</i> <i>"I think the question of problems and failures is a different question from outcomes. And I would have a separate section describing my problems and failures. ...to specifically identify the problems that they encountered ...You know—were the problems related to data collection? Were the problems related to putting reports together that were meaningful? Were the problems reporting to the end users?..." Participant 12, professional healthcare improvement consultant</i> <i>"People don't know what process measures are."</i> <i>"This 'document degree of success' [item 13a] in implementing intervention components is confusing—do you mean grade the degree of success—and what is the instrument to do that?" Participants 8 and 11, improvement fellowship faculty members</i>
Discussion	
Item 14. <i>Summary</i> The most important successes and difficulties in implementing intervention components, main changes observed	<i>"Feels overall unnecessarily complicated—there is duplication." Participants 1–4, Improvement fellowship faculty members [regarding item #14] "Are there lists of barriers to implementation that could help with this?" Participant 13, physician researcher</i>
Item 15. <i>Relation to other evidence</i> Compare and contrast study results with relevant findings from other studies	<i>"...context is primarily mentioned in the limitations section. I don't really think that that's right. Context isn't—when we think about QI, context is so central to QI, it's not a limitation, it's a description of what the problem was, what the care setting was..." [regarding item #17] "Draw inferences—how do you measure the strength of the data?"</i>
Item 16. <i>Limitations</i> Possible sources of confounding, factors affecting external validity	<i>"...describe what contextual factors we learned may impact effectiveness—I would like clarity on how to measure that..."</i>
Item 17. <i>Interpretation</i> Reasons for differences observed and outcomes, draw inferences about causal mechanisms, review opportunity cost and financial cost, practical usefulness	<i>"Item 18b [implication for future studies] should be combined with [item] 14, and isn't this the thrust of the whole discussion?" Participants 1, 10 and 11, improvement fellowship faculty members</i>
Item 18. <i>Conclusions</i> Practical usefulness, implications for further studies	

Items that touched on evolving areas in healthcare improvement were interpreted differently across both end users and the advisory group of thinkers

The evolution of the healthcare improvement scholarly literature in the intervening years since the publication of the SQUIRE Guidelines has led to the development of concepts that were not fully anticipated at the time of initial release. Items 10b and 13a_{ii-iv} in particular revealed the areas where these changes are occurring. The specific items are shown in their entirety in the [box 1](#).

Item 10b states, “Describes mechanisms by which intervention components were expected to cause changes, and plans for testing whether those mechanisms were effective”. In the advisory group of thinkers, this item was interpreted to mean ‘the theory’ of an intervention. The word ‘theory’, however, meant different things to different people. For some, the word ‘theory’ meant ‘mechanism by which an intervention was expected to work’, for others it meant ‘lean or six sigma for example’, and for still others it meant ‘logic model’. These responses pointed to different aspects of study design and methods that would require clarification in the revised Guidelines.

Items 13a_{ii-iv} are “Explains the actual course of the intervention...”, “Documents degree of success in implementing...” and “Describes how and why the initial plan evolved...”. On a mechanical level, participants disagreed about whether these items belonged in the results section or the methods section. On a more global level, participants disagreed about whether one should include iterations and development of a project at all in a manuscript, even though the items called for their inclusion. The confusion about the interpretation of these items was shown well with one participant, who said: “...the important things you can’t really write about! ...for me what was really important was how I thought about what the project was...and how it failed because [what I proposed] wasn’t relevant to the people who were supposed to do the intervention. How do you write about that?” (participant 15, physician graduate of improvement fellowship). Some felt that failed iterations of the work, which might be included as part of item 13a_{iv}, should be presented because it could help others to learn: “...is the failure [of the intervention] unique? If it is not generalizable or useful, then it should not be included. If it informs the results, it

Box 1 Standards for Quality Improvement Reporting Excellence (SQUIRE) items that were interpreted differently across end users and the advisory group of thinkers.

Item 10. Planning the study of the intervention

- a. Outlines plans for assessing how well the intervention was implemented (dose or intensity of exposure)
- b. Describes mechanisms by which intervention components were expected to cause changes, and plans for testing whether those mechanisms were effective
- c. Identifies the study design (eg, observational, quasi-experimental, experimental) chosen for measuring impact of the intervention on primary and secondary outcomes, if applicable
- d. Explains plans for implementing essential aspects of the chosen study design, as described in publication guidelines for specific designs, if applicable (see, eg, <http://www.equator-network.org>)
- e. Describes aspects of the study design that specifically concerned internal validity (integrity of the data) and external validity (generalisability)

Item 13. Outcomes

- a. Nature of setting and improvement intervention
 - i. Characterises relevant elements of setting or settings (eg, geography, physical resources, organisational culture, history of change efforts), and structures and patterns of care (eg, staffing, leadership) that provided context for the intervention
 - ii. Explains the actual course of the intervention (eg, sequence of steps, events or phases; type and number of participants at key points), preferably using a timeline diagram or flow chart
 - iii. Documents degree of success in implementing intervention components
 - iv. Describes how and why the initial plan evolved, and the most important lessons learned from that evolution, particularly the effects of internal feedback from tests of change (reflexiveness)
- b. Changes in processes of care and patient outcomes associated with the intervention
 - i. Presents data on changes observed in the care delivery process
 - ii. Presents data on changes observed in measures of patient outcome (eg, morbidity, mortality, function, patient/staff satisfaction, service use, cost, care disparities)
 - iii. Considers benefits, harms, unexpected results, problems, failures
 - iv. Presents evidence regarding the strength of association between observed changes/improvements and intervention components/context factors
 - v. Includes summary of missing data for intervention and outcomes

For full guidelines, see http://www.squire-statement.org/assets/pdfs/SQUIRE_guidelines_table.pdf

should be included [in the paper]" (participant 46, head of an institutional improvement programme). Others felt that journal editors would likely not want this information in a manuscript saying "That experience [of getting buy-in from the participants and developing the intervention] may well have usefully formed a paper in itself, but may not be of interest to the sorts of journals, and the *caliber* of journal that I would publish the results in..." (participant 14, physician author). Still others felt it was incorrect to include anything except the most successful parts of the work, because to do anything else would be to 'drift into research' and ask too much of the authors: "If we put the onus on everybody out there who's trying to improve care to deal with that sophisticated question [of why and how the improvement occurred], I just think we are putting a barrier in place that is going to be a mountain" (participant 22, physician improvement consultant). These responses showed that if the development and iterations of a project and its failures were desired in reporting, then the Guidelines would need to make this explicit and provide guidance for how to incorporate it into a manuscript in SQUIRE 2.0.

DISCUSSION

As the first step in the revision of the SQUIRE Guidelines, we have evaluated their current status by working directly with both the people who have used SQUIRE and an advisory group of thinkers in the field of healthcare improvement. To our knowledge, an evaluation of this type has not been done before for a publication guideline. We used this approach because we believed end users in particular would provide unique and important insights into the changes and challenges of the field, informing and strengthening the revision process.

At the most basic level, certain items in SQUIRE were interpreted by end users to be redundant, unclear or simply not comprehensible. If this had been the only issue, careful copy editing of the document and broadening of the explanation and elaboration document would be all that was needed. However, many of the comments and discussions around particular items illustrated broader issues, which we hypothesise are related to skill in writing, knowledge of the science of improvement and the evolution of the field. The development of SQUIRE 2.0 will require attention to these matters.

End users' impressions of the SQUIRE Guidelines varied by their level of experience in both research and writing, and their skills and knowledge in the field of healthcare improvement in general. Those with more experience felt the Guidelines were easier to understand than those with less experience. We interpret these findings to reflect the hard work of learning to write for the scientific literature as well as the challenge of writing about healthcare

improvement. This type of writing has many dimensions to bring to the written form, such as time dependence and contextual issues, for which we lack well-defined scientific language and methods.

A potential response to end users' descriptions of the challenge of writing might be to say that the role of SQUIRE is to urge people to do high-quality healthcare improvement work and report it completely, not teach people to write. However, the relative newness of the field of scholarly writing about improvement, and the fact that one of the explicit goals of SQUIRE was to advance the quality of such writings³ suggests that attention should be paid to the work of teaching the skill of scholarly writing. We endorse the goals of the developers of the Equator Network, which is to provide students and new researchers with guidance on what constitutes good research practice and reporting, as educating the next generation of researchers will help move the field forward.¹⁶ Thus, one important goal of SQUIRE 2.0 will be to support the development of skills in scholarly writing about improvement.

Disagreement over key concepts in healthcare improvement reporting was identified in end users and the advisory group of thinkers. Two areas came up specifically, the use of 'theory' to guide improvement work (and what the word 'theory' means), and whether iterations of improvement work (and failed iterations, in particular) should be reported. These disagreements may be simply due to evolution in the field, but may also be due to a lack of knowledge of the science of improvement. Whether it is one or both of these reasons, SQUIRE 2.0 must address these directly so that authors understand the need to report them. Given that SQUIRE is intended for writing—but reporting theory and the iterations of a work require attention during the design and execution of a study—one might wonder how SQUIRE could hope to help. Moher *et al*¹⁷ have reported that there is indirect evidence that publication guidelines affect researchers' study design practices, a finding we confirmed in our study as participants reported SQUIRE being very useful for planning improvement work. Thus, there is reason to hope that a clear SQUIRE 2.0 would be able to support better reporting in these important areas.

The work we have presented here has limitations. Qualitative evaluations such as this one may not reflect the full range of relevant findings. To capture the widest range possible, we enrolled a wide range of participants until we reached thematic saturation—no additional findings were emerging from the data obtained in additional interviews or focus groups.¹³

Another challenge in qualitative evaluations relates to generalisability to other populations. To increase the chance that our findings would represent people writing in the field generally, we used maximum variation sampling.¹⁸ It is possible however that there are users of SQUIRE who have not yet published or are

not otherwise locatable by internet and personal contact searches who are different from our participants. Readers may also be concerned that the presence of authors of the prior Guidelines during two of the focus groups would inhibit responses, but we did not find the responses during these focus groups to be substantially different from those of other participants. Lastly, since interviews were conducted in English, subtle nuances of meaning may have been difficult to capture from those who were not native English speakers.

This unique evaluation of the SQUIRE Guidelines by end users and an advisory group of thinkers in the field has revealed the areas requiring clarification and identified the information needs of users. General item clarity, the use of theory in guiding an intervention, the concept of studying the intervention and the reporting of iterations of an intervention will require attention in SQUIRE 2.0. Further, SQUIRE 2.0 should seek to be a source of reliable definitions and a resource for the important concepts noted above as the field evolves. The findings from this study will support the initial revision of the SQUIRE Guidelines, so that they can be tested and further revised prior to the release of SQUIRE 2.0 in the fall of 2015.

Author affiliations

¹VA Outcomes Group, Department of Veterans Affairs Medical Center, White River Junction, VT

²Department of Surgery - Otolaryngology, Geisel School of Medicine at Dartmouth, Hanover, NH

³The Dartmouth Institute for Health Policy & Clinical Practice, Lebanon, NH

⁴The Geisel School of Medicine at Dartmouth, Hanover, NH, USA

⁵Department of Veterans Affairs Medical Center, White River Junction, VT, USA

Twitter Follow Louise Davies at @louisedaviesmd

Contributors LD: conception and design, data acquisition, analysis, drafting of the manuscript, obtaining funding; PB, FD, DS: interpretation of data, critical revision of manuscript, obtaining funding; GO: conception and design, interpretation of data, critical revision of manuscript, obtaining funding.

Funding The Robert Wood Johnson Foundation (#70024), The UK Health Foundation (#7099).

Disclaimer The views expressed do not necessarily represent the views of the Robert Wood Johnson Foundation, The UK Health Foundation, The Department of Veterans Affairs or the United States Government.

Competing interests None declared.

Patient consent End user interview and focus group participants were informed at the time of the invitation and at the start of their meetings that the project was being conducted as research, that their voices would be recorded and their confidentiality protected during the research project. Documentation of written informed consent was not required by the CPHS, agreement to participate served as the consent.

Ethics approval This study was approved by The Dartmouth Committee for the Protection of Human Subjects (CPHS) as research, exempt from human subjects review.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Raw data in de-identified form can be made available to requestors by direct request to the authors.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

- Davidoff F, Batalden PB, Stevens DP, *et al.* Development of the SQUIRE Publication Guidelines: evolution of the SQUIRE project. *BMJ Qual Saf* 2008;34:681–7.
- Davidoff F, Batalden P. Toward stronger evidence on quality improvement. Draft publication guidelines: the beginning of a consensus project. *Qual Saf Health Care* 2005;14:319–25.
- Huber AJ, Ogrinc G, Davidoff F. SQUIRE (Standards for Quality Improvement Reporting Excellence). In: Moher D, Altman D, Schulz K, *et al.*, eds. *Guidelines for Reporting Health Research: a user's manual*. London, UK: Wiley Blackwell, 2014.
- Cook JM, O'Donnell C, Dinnen S, *et al.* Measurement of a model of implementation for health care: toward a testable theory. *Implement Sci* 2012;7:59.
- Peters DH, Adam T, Alonge O, *et al.* Implementation research: what it is and how to do it. *BMJ* 2013;347:f6753.
- Kaplan HC, Provost LJ, Froehle CM, *et al.* The Model for Understanding Success in Quality (MUSIQ): building a theory of context in healthcare quality improvement. *BMJ Qual Saf* 2012;21:13–20.
- Bate P, Robert G, Fulop N, *et al.* Bamber J ed. *Perspectives on Context: A selection of essays considering the role of context in successful quality improvement*. London: The Health Foundation, 2014.
- Ogrinc G, Nelson WA, Adams SM, *et al.* An instrument to differentiate between clinical research and quality improvement. *IRB* 2013;35:1–8.
- Marshall M, Pronovost P, Dixon-Woods M. Promotion of improvement as a science. *Lancet* 2013;381:419–21.
- Research scan: improvement Science*. London, UK: Health Foundation, 2011.
- Moher D, Schulz KF, Simera I, *et al.* Guidance for developers of health research reporting guidelines. *PLoS Med* 2010;7:e1000217.
- Chevalier JB, Buckles DJ. *Participatory action research: theory and methods for engaged inquiry*. New York, NY: Routledge, 2013.
- Strauss AL, Corbin JM. *Basics of qualitative research: techniques and procedures for developing grounded theory*. 2nd edn. Thousand Oaks: Sage Publications, 1998.
- Kuper A, Lingard L, Levinson W. Critically appraising qualitative research. *BMJ* 2008;337:a1035.
- Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;19:349–57.
- Simera I, Moher D, Hirst A, *et al.* Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med* 2010;8:24.
- Moher D, Jones A, Lepage L, *et al.* Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001;285:1992–5.
- Lincoln YS, Guba EG. *Naturalistic inquiry*. Newbury Park, California: Sage, 1985.