

Publication Guidelines for Quality Improvement Studies in Health Care: Evolution of the SQUIRE Project

Frank Davidoff, MD¹, Paul Batalden, MD², David Stevens, MD², Greg Ogrinc, MD, MS^{2,3}, and Susan Mooney, MD, MS^{2,4} for the SQUIRE development group

¹Institute for Healthcare Improvement, Wethersfield, CT, USA; ²The Dartmouth Institute for Health Policy and Clinical Practice, Center for Leadership and Improvement, Lebanon, NH, USA; ³White River Junction VA, White River Junction, VT, USA; ⁴Alice Peck Day Memorial Hospital, Lebanon, NH, USA.

In 2005 we published draft guidelines for reporting studies of quality improvement interventions as the initial step in a consensus process for development of a more definitive version. The current article contains the revised version, which we refer to as SQUIRE (Standards for QUality Improvement Reporting Excellence). We describe the consensus process, which included informal feedback, formal written commentaries, input from publication guideline developers, review of the literature on the epistemology of improvement and on methods for evaluating complex social programs, and a meeting of stakeholders for critical review of the guidelines' content and wording, followed by commentary on sequential versions from an expert consultant group. Finally, we examine major differences between SQUIRE and the initial draft, and consider limitations of and unresolved questions about SQUIRE; we also describe ancillary supporting documents and alternative versions under development, and plans for dissemination, testing, and further development of SQUIRE.

KEY WORDS: quality improvement; publication; standards.

Editors Note: The SQUIRE Guidelines are intended to advance research in quality improvement. Quality of care and patient safety are at the heart of general internal medicine and consequently, the readers of the *Journal of General Internal Medicine*. A longer, more detailed explanation of the development of the SQUIRE consensus development effort appears in the October supplement to the journal *Quality and Safety in Health Care*. Because of the importance of the topic and its relevance to our readership, we are publishing this portion of that supplement.

J Gen Intern Med 23(12):2125-30
DOI: 10.1007/s11606-008-0797-4

INTRODUCTION

A great deal of meaningful and effective work is now done in clinical settings to improve the quality and safety of care.

Members of the SQUIRE development group who provided input during the development process and endorsed the SQUIRE guidelines are listed at the end of this article

Received July 17, 2008

Revised August 3, 2008

Accepted September 5, 2008

Published online October 2, 2008

Unfortunately, relatively little of that work is reported in the biomedical literature, and much of what is published could be more effectively presented. Failure to publish is potentially a serious barrier to the development of improvement science, since public sharing of concepts, methods, and findings is essential to the progress of all scientific work, both theoretical and applied. To help strengthen the evidence base for improvement in health care, in 2005 we proposed draft guidelines for reporting planned original studies of improvement interventions (1). Our aims were to stimulate the publication of high-caliber improvement studies and to increase the completeness, accuracy, and transparency of published reports of that work.

Our initial draft guidelines were based largely on the authors' personal experience with improvement work and were intended only as an initial step toward creation of an established publication standard. We have now refined and extended that draft. In the current article we present the resulting revised version, which we refer to as the Standards for QUality Improvement Reporting Excellence, or SQUIRE (Table 1). We also describe the SQUIRE consensus development process; examine the major differences between the current version of SQUIRE and the initial draft guidelines; consider the limitations of and questions about SQUIRE; describe ancillary supporting documents and variant versions that are under development; and explain plans for dissemination, testing, and further development of the SQUIRE guidelines.

THE CONSENSUS PROCESS

The SQUIRE development process proceeded along six lines. First, we obtained informal feedback on the utility, strengths, and limitations of the initial draft guidelines from potential authors in a series of seminars, as well as from experienced guideline developers at the organizational meeting of the EQUATOR network². Second, authors, peer reviewers, and journal editors "road tested" the draft guidelines as a working tool for editing and revising submitted manuscripts^{3,4}. Third, we solicited and published written commentaries on the initial version of the guidelines⁵⁻⁹. Fourth, we conducted an ongoing review of the relevant literature on epistemology, methodology, and evaluation of complex interventions, particularly in social sciences and the evaluation of social programs. Fifth, in April 2007 we subjected the draft guidelines to intensive analysis, comment, and recommendations for change at a 2-day meeting of 30 stakeholders. Finally, following that meeting, we obtained further critical appraisal of the guidelines through

three cycles of a Delphi process that involved an international group of over 50 consultants.

Informal Feedback

Informal input about the draft guidelines from authors and peer reviewers raised the following relevant issues: (1) uncertainty as to when (that is, to which studies) the guidelines apply, (2) the possibility their use might force quality improvement (QI) reports into a rigid, narrow format, (3) the concern that slavish application might result in lengthy and unreadable reports that were indiscriminately laden with detail, and (4) difficulty knowing if, when, and how other publication guidelines should be used in conjunction with publication guidelines for quality improvement studies.

Deciding When to Use the Guidelines. Publications on improvement in health care are emerging in four general categories: empirical studies on development and testing of quality improvement interventions; stories, theories, and frameworks; literature reviews and syntheses; and the development and testing of improvement-related tools (Rubenstein et al., unpublished). Our consensus process has made it clear that the SQUIRE guidelines can and should apply to reporting in the first category: original planned empirical studies on the development and testing of improvement interventions.

Forcing Articles into a Rigid Format. Publication guidelines are often referred to as checklists, since like other such documents they serve the function of “aides memoires,” which have proven valuable in managing information in complex systems¹⁰. Rigid or mechanical application of checklists can of course prevent users from making sense of complex information^{11,12}, but, paradoxically, checklists, like all constraints, can also serve as a crucial driver for creativity. The SQUIRE guidelines must therefore always be understood and used as signposts, not shackles¹³.

Creating Longer Articles. Improvement is a complex undertaking, and its evaluation can produce substantial amounts of qualitative and quantitative information. Of course, adding irrelevant information simply to “cover” guideline items would be counterproductive; on the other hand, added length that makes reports of improvement studies more complete, coherent, usable, and systematic helps to meet a principal aim of SQUIRE. Publishing portions of improvement studies electronically can make the content of long papers publicly available while preserving the scarce resource of print publication.

Conjoint Use with Other Publication Guidelines. Most other biomedical publication guidelines are designed to improve the reporting of studies that use specific experimental designs. The SQUIRE guidelines, in contrast, are concerned with reporting studies in a defined content area - improvement and safety. These two guideline types are therefore complementary; when appropriate, other specific design-related guidelines can and should therefore be used in conjunction with SQUIRE.

Formal Commentaries

Written commentaries suggested that the “pragmatic” focus of the initial draft guidelines was an important complement to traditional experimental clinical science⁵. The guidelines were also seen as a potentially valuable instrument for strengthening the design and conduct of improvement research, resulting in greater synergy with improvement practice⁸, and increasing the feasibility of combining improvement studies in systematic reviews. However, the commentaries also raised several concerns: (1) that the draft guidelines were inattentive to racial and ethnic disparities in care⁷; (2) that their IMRaD structure (Introduction, Methods, Results, and Discussion) could be incompatible with the reality that improvement interventions, by design, change over time⁶; (3) and that their use could result in a “dumbing down” of improvement science⁹.

Health Disparities. It would not be useful (even if it were possible) to address every relevant content issue in a concise set of guidelines for reporting improvement studies. We do agree, however, that disparities in care are not considered often enough in improvement work and that improvement initiatives should address this important issue whenever possible. We have therefore highlighted this issue in the SQUIRE guidelines (Table 1, Item 13.b.ii).

The IMRaD Structure. The study protocols traditionally described in the Methods sections of clinical trials are rigidly fixed, as required by the dictates of experimental design¹⁴. Improvement, in contrast, is a reflexive learning process; that is, improvement interventions are most effective when they are modified in response to outcome feedback. On these grounds, it has been suggested that reporting improvement interventions in the IMRaD format would require multiple sequential Methods sections, one for each iteration of the evolving intervention⁶. We maintain, however, that the reflexive nature of improvement does not exempt improvement studies from answering the four fundamental questions required in all scholarly inquiry — Why did you start? What did you do? What did you find? What does it mean? — the same questions that define the four elements of the IMRaD framework^{15,16}. In our view, this requirement justifies providing a single Methods section to describe the initial improvement plan and the theory (mechanism) on which it is based. The changes in interventions over time, and the learning that comes from making those changes, are themselves important improvement outcomes and therefore belong in the Results section rather than in a series of separate Methods sections¹.

“Dumbing Down” Improvement Reports. The declared purpose of all publication guidelines is to improve the completeness and transparency of reporting. Since it is precisely those characteristics of reporting that make it possible to detect weak, sloppy, or poorly designed studies, it is difficult to understand how use of the draft guidelines might have led to a “dumbing down” of improvement science. The underlying concern here appears to have less to do with transparency than with the inference that the draft guidelines failed to require sufficiently rigorous standards of evidence¹⁴. We recognize that those traditional experimental standards are

Table 1. SQUIRE Guidelines (Standards for Quality Improvement Reporting Excellence)

Text section; item number and name	Section or item description
Title and abstract	Did you provide clear and accurate information for finding, indexing, and scanning your paper?
1. Title	<ul style="list-style-type: none"> a. Indicates the article concerns the improvement of quality (broadly defined to include the safety, effectiveness, patient-centeredness, timeliness, efficiency, and equity of care) b. States the specific aim of the intervention c. Specifies the study method used (for example, “A qualitative study” or “A randomized cluster trial”)
2. Abstract	Summarizes precisely all key information from various sections of the text using the abstract format of the intended publication
Introduction	Why did you start?
3. Background knowledge	Provides a brief, non-selective summary of current knowledge of the care problem being addressed and characteristics of organizations in which it occurs
4. Local problem	Describes the nature and severity of the specific local problem or system dysfunction that was addressed
5. Intended improvement	<ul style="list-style-type: none"> a. Describes the specific aim (changes/improvements in care processes and patient outcomes) of the proposed intervention b. Specifies who (champions, supporters) and what (events, observations) triggered the decision to make changes, and why now (timing)
6. Study question	States precisely the primary improvement-related question and any secondary questions that the study of the intervention was designed to answer
Methods	What did you do?
7. Ethical issues	Describes ethical aspects of implementing and studying the improvement, such as privacy concerns, protection of participants’ physical well-being, and potential author conflicts of interest, and how ethical concerns were addressed
8. Setting	Specifies how elements of the local care environment considered most likely to influence change/improvement in the involved site or sites were identified and characterized
9. Planning the intervention	<ul style="list-style-type: none"> a. Describes the intervention and its component parts in sufficient detail that others could reproduce it b. Indicates main factors that contributed to choice of the specific intervention (for example, analysis of causes of dysfunction; matching relevant improvement experience of others with the local situation) c. Outlines initial plans for how the intervention was to be implemented: e.g., <i>what</i> was to be done (initial steps; functions to be accomplished by those steps; how tests of change would be used to modify intervention) and <i>by whom</i> (intended roles, qualifications, and training of staff)
10. Planning the study of the intervention	<ul style="list-style-type: none"> 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 (for example, 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, for example, www.EQUATOR-network.org) e. Describes aspects of the study design that specifically concerned internal validity (integrity of the data) and external validity (generalizability)
11. Methods of evaluation	<ul style="list-style-type: none"> a. Describes instruments and procedures (qualitative, quantitative, or mixed) used to assess (a) the effectiveness of implementation, (b) the contributions of intervention components and context factors to effectiveness of the intervention, and (c) primary and secondary outcomes b. Reports efforts to validate and test reliability of assessment instruments c. Explains methods used to assure data quality and adequacy (for example, blinding; repeating measurements and data extraction; training in data collection; collection of sufficient baseline measurements)
12. Analysis	<ul style="list-style-type: none"> a. Provides details of qualitative and quantitative (statistical) methods used to draw inferences from the data b. Aligns unit of analysis with level at which the intervention was implemented, if applicable c. Specifies degree of variability expected in implementation, change expected in primary outcome (effect size), and ability of study design (including size) to detect such effects d. Describes analytic methods used to demonstrate effects of time as a variable (for example, statistical process control)
Results	What did you find?
13. Outcomes	<ul style="list-style-type: none"> (a) Nature of setting and improvement intervention <ul style="list-style-type: none"> i. Characterizes relevant elements of setting or settings (for example, geography, physical resources, organizational culture, history of change efforts) and structures and patterns of care (for example, staffing, leadership) that provided context for the intervention ii. Explains the actual course of the intervention (for example, sequence of steps, events or phases; type and number of participants at key points), preferably using a time-line 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 <ul style="list-style-type: none"> i. Presents data on changes observed in the care delivery process ii. Presents data on changes observed in measures of patient outcome (for example, morbidity, mortality, function, patient/staff satisfaction, service utilization, 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

(continued on next page)

Table 1. (continued)

Text section; item number and name	Section or item description
Discussion	What do the findings mean?
14. Summary	a. Summarizes the most important successes and difficulties in implementing intervention components, and main changes observed in care delivery and clinical outcomes b. Highlights the study's particular strengths
15. Relation to other evidence	Compares and contrasts study results with relevant findings of others, drawing on broad review of the literature; use of a summary table may be helpful in building on existing evidence
16. Limitations	a. Considers possible sources of confounding, bias, or imprecision in design, measurement, and analysis that might have affected study outcomes (internal validity) b. Explores factors that could affect generalizability (external validity), for example: representativeness of participants, effectiveness of implementation, dose-response effects, features of local care setting c. Addresses likelihood that observed gains may weaken over time and describes plans, if any, for monitoring and maintaining improvement; explicitly states if such planning was not done d. Reviews efforts made to minimize and adjust for study limitations e. Assesses the effect of study limitations on interpretation and application of results
17. Interpretation	a. Explores possible reasons for differences between observed and expected outcomes b. Draws inferences consistent with the strength of the data about causal mechanisms and size of observed changes, paying particular attention to components of the intervention and context factors that helped determine the intervention's effectiveness (or lack thereof) and types of settings in which this intervention is most likely to be effective c. Suggests steps that might be modified to improve future performance d. Reviews issues of opportunity cost and actual financial cost of the intervention
18. Conclusions	a. Considers overall practical usefulness of the intervention b. Suggests implications of this report for further studies of improvement interventions
Other information	Were there other factors relevant to the conduct and interpretation of the study?
19. Funding	Describes funding sources, if any, and role of funding organization in design, implementation, interpretation, and publication of study

- These guidelines provide a framework for reporting formal, planned studies designed to assess the nature and effectiveness of interventions to improve the quality and safety of care
- It may not always be appropriate or even possible to include information about every numbered guideline item in reports of original studies, but authors should at least consider every item in writing their reports
- Although each major section (i.e., Introduction, Methods, Results, and Discussion) of a published original study generally contains some information about the numbered items within that section, information about items from one section (for example, the Introduction) is also often needed in other sections (for example, the Discussion).

powerful instruments for protecting the integrity of outcome measurements, largely by minimizing selection bias^{13,17}, and we fully accept their importance. But while those standards are relevant in improvement studies, they are not sufficient, since they fail to take into account the unique purpose and characteristics of the improvement process.

Unlike the “conceptually neat and procedurally unambiguous” interventions — drugs, tests, and procedures — whose efficacy is studied in clinical research, improvement is essentially a social process. Its immediate purpose is to change human performance, and it is driven primarily by experiential learning^{18,19}. Improvement is therefore inherently context dependent and, as noted, reflexive; both its interventions and outcomes are unstable, and it generally involves complex, multi-component interventions. Although traditional experimental and quasi-experimental methods are clearly important for learning *whether* improvement interventions change behavior, they do not address the crucial pragmatic (or “realist”) questions about improvement: *what* it is about the mechanism of a particular intervention that works, *for whom*, and *under what circumstances*^{20–22}? Using combinations of methods that simultaneously answer all these questions is not an easy task, since the experimental and pragmatic methodologies can work at cross purposes. We have attempted in the SQUIRE guidelines to maintain a balance between these two complementary epistemological approaches.

Consensus Meeting of Editors and Research Scholars

With support from the Robert Wood Johnson Foundation, we undertook a critical appraisal of the draft guidelines at a 2-day meeting in April 2007. Thirty participants attended, including clinicians, improvement professionals, epidemiologists, clinical researchers, and journal editors, several from outside the US. Prior to the meeting we sent participants a reading list and a concept paper on the epistemology of improvement. In plenary and small group sessions, participants critically discussed and debated the content and wording of every item in the draft guidelines, and recommended changes; they also provided input on plans for dissemination, adoption, and future uses of the guidelines. Working from transcribed audio-recordings of all meeting sessions, and flip charts listing the key discussion points, a coordinating group (the authors of this paper) then revised, refined, and expanded the draft guidelines.

Delphi Process. Following the consensus meeting, we circulated sequential revisions of the guidelines for further comments and suggestions in three cycles of a Delphi process. The group involved in that process included the meeting participants plus roughly 20 additional expert consultants. We then surveyed all participants as to their willingness to endorse the final consensus version (SQUIRE).

Several features of SQUIRE that are different from the initial draft guidelines are worth special mention. First, the SQUIRE guidelines distinguish clearly between improvement practice (that is, planning and implementing improvement interventions) and the evaluation of improvement interventions (that is, designing and executing studies to assess whether those interventions work, and why they do or do not work). Second, they highlight the essential and unique properties of improvement interventions, particularly their social nature, focus on changing performance, context-dependence, complexity, non-linearity, adaptation, and reflexiveness. Third, they specify elements of study design that assess both *whether* improvement interventions work (by minimizing bias and confounding) and *why* interventions are or are not effective (by marking out the effects of context and identifying mechanisms of change). Fourth, this version explicitly addresses the often confusing ethical dimensions of improvement and improvement studies²³.

LIMITATIONS AND QUESTIONS

The SQUIRE guidelines have been characterized as providing both too little and too much information: too little, because they fail to represent adequately the many unique and nuanced issues in the practice of improvement^{14,17,20–22,24}; too much, because the detail and density of the item descriptions can seem intimidating to authors. We recognize that the SQUIRE item descriptions are significantly more detailed than those of some other publication guidelines. In our view, however, the complexity of the improvement process, plus the relative unfamiliarity of improvement interventions and of the methods for evaluating them, justifies that level of detail, particularly in light of the diverse backgrounds of people working to improve health care. Moreover, the level of detail the SQUIRE guidelines is similar to that in guidelines for reporting observational studies, which also involve considerable complexities of study design²⁵. To increase their usability, we plan to make available a shortened electronic version of SQUIRE, accompanied by a glossary of terms used in the item descriptions that may be unfamiliar to potential users.

APPLYING SQUIRE

Authors' interest in using publication guidelines increases when journals make them part of the peer review and editorial process. We therefore encourage the widest possible use of the SQUIRE guidelines by editors. Unfortunately, little is known about the most effective ways to apply publication guidelines; editors have therefore learned by experience how to use them in practice, and the specifics of their use varies widely from journal to journal. We also lack systematic knowledge of how authors can use publication guidelines most productively; our experience suggests, however, that the SQUIRE guidelines will be most helpful to authors if they remain generally aware of the content of SQUIRE as they write initial drafts, then, as they make revisions, refer to individual guideline items in making detailed critical appraisal of what they have written. To increase our empirical knowledge about how publication guidelines can be used most effectively, we strongly encourage editors and authors to collect, examine, and report their experiences in using SQUIRE and other publication guidelines.

CURRENT AND FUTURE DIRECTIONS

In the October 2008 Supplement to *Quality and Safety in Health Care*, we present an explanation and elaboration (E&E) document²⁶. Like other such documents^{27–30}, the SQUIRE E&E document provides much necessary depth and detail that cannot be included in a set of concise guideline items. It presents the rationale for including each guideline item in SQUIRE, along with published examples of reporting for each item, and commentary on the strengths and weaknesses of those examples.

The SQUIRE website (<http://www.squire-statement.org>) will provide an authentic electronic home for the guidelines themselves and a medium for their progressive refinement. We also intend the site to serve as an interactive electronic community for authors, students, teachers, reviewers, and editors who are interested in the emerging body of scholarly and practical knowledge on improvement.

Although the primary purpose of SQUIRE is to improve the reporting of improvement studies, we believe the guidelines can also serve useful educational purposes, particularly for understanding the epistemology of improvement and the methodologies for evaluating improvement work. We believe, similarly, that SQUIRE can help in planning improvement interventions and studies of those interventions, as well as in developing skill in writing about improvement. We encourage those uses, as well as efforts to assess SQUIRE's impact on the completeness and transparency of published improvement studies^{31,32}, and to obtain empirical evidence that individual guideline items contribute materially to the value of published information in improvement science.

Acknowledgments: *Some of the work reported in this paper was done at the SQUIRE Advisory Committee Meeting, April 3–5, 2007, which was supported in part by a grant from the Robert Wood Johnson Foundation (RWJF grant number 58073).*

We are grateful to Rosemary Gibson and Laura Leviton for their unflinching support of this project, and to Joy McAvoy for her invaluable administrative work in coordinating the entire development process.

Contributors: *The following people contributed critical input to the guidelines during their development: Kay Dickersin, Donald Goldmann, Peter Goetzsche, Gordon Guyatt, Hal Luft, Kathryn McPherson, Victor Montori, Dale Needham, Duncan Neuhauser, Kaveh Shojania, Vincenza Snow, Ed Wagner, and Val Weber.*

Endorsement: *The following participants in the consensus process also provided critical input on the guidelines and endorsed the final version. Their endorsements are personal and do not imply endorsement by any group, organization, or agency: David Aron, Virginia Barbour, Jesse Berlin, Steven Berman, Donald Berwick, Maureen Bisognano, Andrew Booth, Isabelle Boutron, Peter Buerhaus, Marshall Chiu, Benjamin Crabtree, Linda Cronenwett, Mary Dixon-Woods, Brad Doebbling, Denise Dougherty, Martin Eccles, Susan Ellenberg, William Garrity, Lawrence Green, Trisha Greenhalgh, Linda Headrick, Susan Horn, Julie Johnson, Kate Koplan, David Korn, Uma Kotagal, Seth Landefeld, Elizabeth Loder, Joanne Lynn, Susan Mallett, Peter Margolis, Diana Mason, Don Minckler, Brian Mittman, Cynthia Mulrow, Eugene Nelson, Paul Plsek, Peter Pronovost, Lloyd Provost, Philippe Ravaud, Roger Resar, Jane Roessner, John-Arne Rottingen, Lisa Rubenstein, Harold Sox, Ted Speroff, Richard Thomson, Erik von Elm, Elizabeth Wager, Doug Wakefield, Bill Weeks, Hywel Williams, and Sankey Williams.*

Financial support: *The SQUIRE project was supported in part by a grant from the Robert Wood Johnson Foundation (RWJF grant number 58073).*

Conflicts of interest: None disclosed.

Corresponding Author: Frank Davidoff, MD; Institute for Health-care Improvement, 143 Garden Street, Wethersfield, CT 06109, USA (e-mail: fdavidoff@cox.net).

REFERENCES

1. **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.
2. EQUATOR Network. Enhancing the QQuality And Transparency Of health Research. Available at: <http://www.equator-network.org>. Accessed May 9, 2008.
3. **Janisse T.** A next step: reviewer feedback on quality improvement publication guidelines. *Permanente Journal.* 2007;11:1.
4. Guidelines for authors: guidelines for submitting more extensive quality research. *Quality and Safety in Health Care.* Available at: http://qshc.bmj.com/ifa/article_type.dtl#extensive
5. **Berwick D.** Broadening the view of evidence-based medicine. *Qual Saf Health Care.* 2005;14:315–6.
6. **Thomson RG.** Consensus publication guidelines: the next step in the science of quality improvement? *Qual Saf Health Care.* 2005;14:317–8.
7. **Chin MH, Chien AT.** Reducing racial and ethnic disparities in health care: an integral part of quality improvement scholarship. *Qual Saf Health Care.* 2006;15:79–80.
8. **Baker GR.** Strengthening the contribution of quality improvement research to evidence based health care. *Qual Saf Health Care.* 2006;15:150–1.
9. **Pronovost P, Wachter R.** Proposed standards for quality improvement research and publication: one step forward and two steps back. *Qual Saf Health Care.* 2006;15:152–3.
10. **Gawande A.** The checklist. *The New Yorker.* December 10, 2007, pp. 86–95.
11. **Rennie D.** Reporting randomized controlled trials. An experiment and a call for responses from readers. *JAMA.* 1995;273:1054–5.
12. **Williams JW, Holleman JR, Samsa GP, Simel DL.** Randomized controlled trial of 3 vs 10 days of timethoprim/sulfamethoxazole for acute maxillary sinusitis. *JAMA.* 1995;273:1015–21.
13. **Rutledge A. On Creativity.** Available at <http://www.alistapart/articles/oncreativity>. Accessed on May 9, 2008.
14. **Shadish WR, Cook TD, Campbell DT.** *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* New York: Houghton Mifflin Company; 2002.
15. **Day RA.** The origins of the scientific paper: the IMRaD format. *J Am Med Writ Assoc.* 1989;4:16–8.
16. **Huth EJ.** The research paper: general principles for structure and content. In *Writing and Publishing in Medicine*, 3rd edn. Philadelphia: Williams & Wilkins; 1999, pp. 63–73.
17. **Jadad AR, Enkin MW.** *Randomized Controlled Trials. Questions, Answers, and Musings.* 2 London: Blackwell Publishing/BMJ Books; 2007.
18. **Batalden P, Davidoff F.** Teaching quality improvement. The devil is in the details. *JAMA.* 2007;298:1059–61.
19. **Kolb DA.** *Experiential Learning. Experience as the Source of Learning and Development.* Englewood Cliffs, NJ: Prentice-Hall; 1984.
20. **Pawson R, Greenhalgh T, Harvey G, Walshe K.** Realist review - a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy.* 2005;10(Suppl 121–34).
21. **Pawson R, Tilley N.** *Realistic Evaluation.* Thousand Oaks, CA: SAGE Publications; 1997.
22. *Evaluation for the 21st Century.* In: E Chelmsky, William Shadish, eds. Thousand Oaks, CA: SAGE Publications; 1997.
23. *Health Care Quality Improvement: Ethical and Regulatory Issues.* In: Jennings B, Baily MA, Bottrell M, Lynn J, eds. Garrison, NY: The Hastings Center; 2007.
24. **Glouberman S, Zimmerman B.** Complicated and complex systems: What would successful reform of medicine look like? In: Forest PG, McKintosh K, Marchildon G, eds. *Health Care Services and the Process of Change.* Toronto: University of Toronto Press; 2004.
25. **von Elm E, Altman DG, Egger M, Pocock S, Gotsche PC, Vandenbroucke JP, for the STROBE initiative.** The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med.* 2007;147:573–7.
26. **Ogrinc G, Mooney SE, Estrada C, et al.** The SQUIRE (Standards for Quality Improvement Reporting Excellence) guidelines for quality improvement reporting: explanation and elaboration. *Qual Saf Health Care* 2008; Supplement, in press.
27. **Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Goetzche PC, Lang T, CONSORT GROUP (Consolidated Standards of Reporting Trials).** The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med.* 2001;134:663–94.
28. **Bossuyt PM, Reitsma JB, Bruns DE, Gastonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HCW, Lijmer JG.** The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem.* 2003;49:7–18.
29. **Vandenbroucke JP, von Elm E, Altman DG, Gotsche P, Mulrow CD, Pocock SJ, for the STROBE initiative, et al.** Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med.* 2007;147:W–164.
30. **Keech A, GebSKI V, Pike R.** *Interpreting and Reporting Clinical Trials. A Guide to the CONSORT Statement and the Principles of Randomized Trials.* Sydney: Australian Medical Publishing Company; 2007.
31. **Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al.** Does the CONSORT checklist improve the quality of reports of randomized controlled trials? A systematic review. *Med J Aust.* 2006;185:263–7.
32. **Egger M, Juni P, Bartlett C.** Value of flow diagrams in reports of randomized controlled trials. *JAMA.* 2001;285:1996–9.