



Presented at the Academic Surgical Congress 2019

Quality assessment of the literature on surgical quality improvement

Olivia A. Sacks, BA^a, Andrew J. Lambour, MD^b, Allison R. Wilcox, MD^b,
Shruthi Nammalwar, BS^a, Sandra L. Wong, MD, MS^{a,b,*}

^a Geisel School of Medicine at Dartmouth, Hanover, NH^b Department of Surgery, Dartmouth-Hitchcock Medical Center, Lebanon, NH

ARTICLE INFO

Article history:

Accepted 9 May 2019

Available online 26 June 2019

ABSTRACT

Background: A proliferation of work on surgical quality improvement has brought about an increase in quality improvement publications. We assessed the quality of surgical quality improvement publications using the Standards of Quality Improvement Reporting Excellence (SQUIRE) guidelines.

Methods: We conducted a comprehensive review of the surgical quality improvement literature from 2008 to 2018. Articles were reviewed for concordance with 18 SQUIRE statements and 40 subheadings using a dichotomous (yes or no) scale.

Results: Fifty-five articles were included. No publication adhered to all 18 SQUIRE statements. On average, quality improvement publications met 11 out of 18 (61%) of the main statements and 26 out of 40 (65%) of the subheadings. Articles were concordant with introductory components, such as problem description ($n = 55$, 100%) and rationale ($n = 52$, 95%), but were less adherent to statements describing methodology, results, and discussion sections including measures ($n = 7$, 13%), results ($n = 3$, 5.5%), interpretation ($n = 2$, 3.6%), and conclusions ($n = 2$, 3.6%). Only 4 articles cited the SQUIRE guidelines (7.3%). Articles that cited SQUIRE were not more concordant to the statements than those that did not cite SQUIRE.

Conclusion: Our analysis demonstrates that SQUIRE guidelines have not been adopted widely as a framework for the reporting of surgical quality improvement studies. Increased adherence to SQUIRE guidelines has the potential to improve the development and dissemination of surgical quality improvement projects.

© 2019 Elsevier Inc. All rights reserved.

Introduction

Surgeons have embraced efforts in quality improvement (QI), which has led to a proliferation of interventions aimed at advancing clinical practice and patient outcomes.^{1,2} Not unexpectedly, surgical QI projects most often focus on local change, and many of these projects go unpublished. Even when published, they may be presented without the detail necessary to be replicated at other hospitals. As a result, the impact and scope of surgical QI studies is limited, and many in the research community have called to increase the standards for publication of QI projects.^{1–7}

There is consensus that reporting guidelines improve the design and reporting of healthcare studies.^{6–9} The guidelines of the Standards for Quality Improvement Reporting Excellence (SQUIRE), established in 2008 and rereleased as SQUIRE 2.0 in 2015, provide a comprehensive scaffold for the planning, execution, and reporting of QI projects.^{10–14} The SQUIRE guidelines are the only reporting guideline for QI studies indexed in the Enhancing the Quality and Transparency Of Health Research (EQUATOR) network, an initiative dedicated to improving the quality of health-related research.¹⁵ SQUIRE 2.0 is also included in the instructions for manuscript preparation instructions in several major journals (eg, *Journal of the American Medical Association*) and has been copublished by journals such as *BMJ Quality and Safety*, *The Journal of Surgical Research*, *The Journal of the American College of Surgeons*, and *The Joint Commission Journal*.¹⁴

The extent to which the SQUIRE guidelines are used for the development and dissemination of surgical QI projects has not been evaluated formally. In this study, we performed a systematic

Presented at the 14th Annual Academic Surgical Congress, Houston, TX, February 5–7, 2019.

* Reprint requests: Sandra L. Wong, MD, MS, Department of Surgery, Dartmouth-Hitchcock Medical Center, 1 Medical Center Drive, Lebanon, NH 03756.

E-mail address: sandra.l.wong@dartmouth.edu (S.L. Wong).

<https://doi.org/10.1016/j.surg.2019.05.016>

0039-6060/© 2019 Elsevier Inc. All rights reserved.

review to assess the quality of the surgical QI literature using the SQUIRE 2.0 guidelines as the standard of reporting.

Methods

The MEDLINE database was accessed via PubMed, and comprehensive literature searches were completed to identify studies for potential inclusion in this literature review. Two health sciences librarians were consulted for assistance regarding search terms, databases, and standard methodology for this type of review. The database was queried using Quality Improvement and Specialties, Surgical [MESH] or Surgical Procedures, Operative [MESH] or surgery [subheading] in the top 30 surgery journals, as determined by the Journal Citation Reports impact factor, and the top 20 first and second tier journals that publish QI work were ranked by The Institute for Healthcare Improvement.^{15,16}

Articles were pooled from the search results, and the abstracts were screened to identify potentially relevant articles for full-text review. To meet our criteria for full-text review and inclusion in the study, the articles were required to (1) be published in an English language journal between November 2008 and January 2018, (2) report system level work that addressed the quality, safety or value of healthcare, (3) have an identifiable intervention, and (4) measure at least one outcome related to the intervention. The SQUIRE guidelines were consulted in the development of these inclusion criteria.¹⁴

SQUIRE 2.0 consists of 18 main statements and 40 subheadings. SQUIRE 2.0 was updated from the first version to eliminate redundancies, improve ease of use, and clarify best use practices.¹² The updated guidelines were used to evaluate all articles, including those published before the release of SQUIRE 2.0, for the following 2 reasons: (1) using a single evaluation rubric ensures consistency and allows comparison of the articles, and (2) SQUIRE 2.0 is considered an improved version of the original guidelines, which clarifies many of the statements from SQUIRE 1.0. SQUIRE 1.0 was not substantially different in content than SQUIRE 2.0, but it was updated to address “uncertainty about what was important to include and lack of clarity in items.”¹² Because the most current version of SQUIRE 2.0 was thought to be the best way to evaluate overall quality of the articles, SQUIRE 2.0 was the only tool used for assessment across the board.^{10–14}

One investigator (O.S.) reviewed the abstracts and titles for full-text review. Then, 3 investigators (O.S., A.W., S.N.) appraised the selected articles for the inclusion criteria. Articles that met these criteria were evaluated independently by all 3 of the investigators for concordance with the 40 SQUIRE 2.0 subheadings using a dichotomous (yes or no) scale. Adherence to each of the 18 main SQUIRE 2.0 statements was determined by adherence to all subheadings of that statement. The evaluators were blinded to each other's work, and inter-rater reliability was calculated using Fleiss's kappa. After this, the 3 authors came to consensus on any differences in their evaluation of the articles. Differences were adjudicated by reviewing the Explanation and Elaboration section of the SQUIRE website for clarification.^{10,13}

Primary outcomes were (1) complete adherence to all 18 main SQUIRE statements, (2) adherence to each individual main SQUIRE statement, and (3) adherence to each SQUIRE statement subheading. Secondary outcomes included adherence to the 40 subheadings stratified by citation of the SQUIRE 2.0 guidelines, type of journal (general interest versus surgical), study setting (academic versus community), and geographic area where the intervention took place. Comparison of the mean adherence to all 40 subheadings between the subgroups was evaluated using a paired *t* test. Data were analyzed using StataSE (StataCorp, TX).

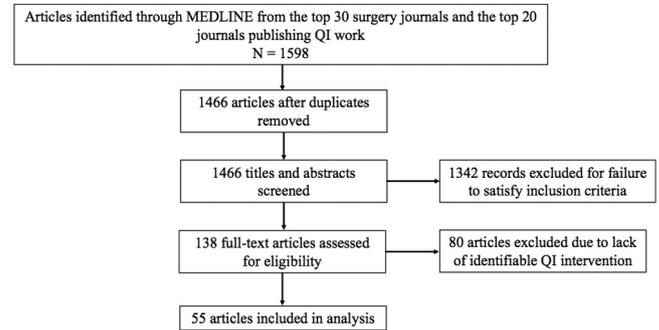


Fig 1. Flowchart of included surgical quality improvement reports.

Results

A total of 1,598 articles were identified by the MEDLINE search. Titles and abstracts of all 1,598 articles were reviewed for inclusion in the study (Fig 1). Articles were excluded most commonly for failing to include an identifiable, surgery-related QI intervention or failing to measure outcomes related to the intervention. In the end, 138 articles were selected for full-text review, after which a total of 55 articles were included in the study (Appendix 1). The level of agreement between the authors on adherence to the SQUIRE statements before group evaluation of the articles ranged from 82% to 99%.

Characteristics of the articles included in the analysis are presented in Table I. The proportion of articles from surgery journals was fairly similar to that of general interest journals ($n = 30$, 55%

Table I
Characteristics of studies included in the literature review

Type of journal	N (%)
Surgery journal	30 (55)
General interest journal	25 (45)
Journals most represented	
<i>Journal of the American College of Surgeons</i>	11 (20)
<i>Surgery</i>	4 (7)
<i>BMJ Quality and Safety</i>	4 (7)
<i>American Journal of Medical Quality</i>	4 (7)
<i>PLOS One</i>	4 (7)
Study setting	
Academic medical center	38 (69)
Community hospital	6 (11)
>1 hospital (mixed)	11 (20)
Specialty	
General surgery	36 (65)
Obstetrics and gynecology	8 (15)
Orthopedic surgery	5 (9)
Other	6 (11)
Type of intervention	
Multidisciplinary team approach	12 (22)
Training of hospital personnel	12 (22)
Checklist implementation	7 (13)
Electronic system implementation	5 (9)
Patient education	5 (9)
Other/mixed	14 (25)
Hospital area of intervention	
Perioperative	22 (40)
Operating room	11 (20)
Outpatient	10 (18)
Other/both	12 (22)
Geographic location	
North America	44 (80)
Europe	6 (11)
Asia	2 (3)
Africa	2 (3)
Australia	1 (2)

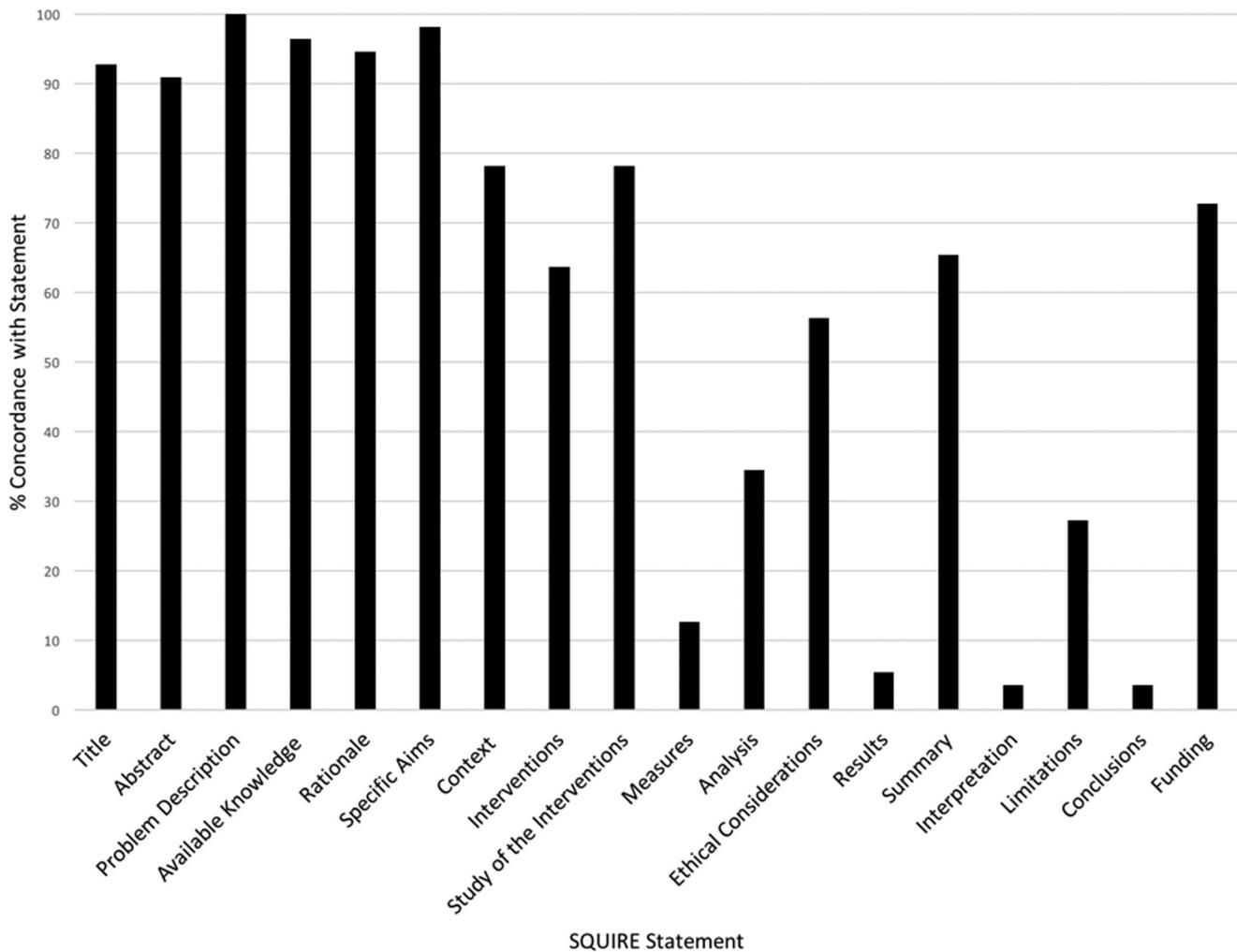


Fig 2. Concordance of surgical quality improvement with SQUIRE 2.0 guidelines.

and $n = 25$, 45%, respectively). The most represented journal in the analysis was *Journal of the American College of Surgeons* ($n = 11$, 20%). The QI studies that were described in the articles were executed most commonly in the United States ($n = 44$, 80%) and at academic medical centers ($n = 38$, 69%).

Twenty-three of the projects included in the study were published after the release of SQUIRE 2.0 (42%). The majority of the articles reflected QI work done by general surgery ($n = 36$, 65%), obstetrics and gynecology ($n = 8$, 15%), and orthopedics ($n = 5$, 9%). Most were perioperative QI projects ($n = 22$, 40%). Eleven projects were OR based (20%), and 10 were conducted in an outpatient setting (18%). The majority of projects were centered around multidisciplinary team training ($n = 12$, 22%) and education programs for hospital personnel ($n = 12$, 22%).

No publication adhered to all 18 statements of the SQUIRE guidelines. On average, articles satisfied 11 out of 18 (61%) of the main statements (Fig 2). Adherence was greatest for statements 1 to 7 of title ($n = 50$, 91%), abstract ($n = 50$, 91%), problem description ($n = 55$, 100%), available knowledge ($n = 53$, 96%), rationale ($n = 52$, 95%), specific aims ($n = 54$, 98%), and context ($n = 43$, 78%). Publications were less adherent to statements 8 to 12 describing the methodology of intervention(s) ($n = 35$, 64%), study of the intervention(s) ($n = 43$, 78%), measures ($n = 7$, 13%), analysis ($n = 19$, 35%), and ethics ($n = 31$, 56%). Three articles met all criteria for the results section ($n = 3$, 5.5%). For statements 14 to 18, concordance

was low overall for summary ($n = 36$, 65%), interpretation ($n = 2$, 3.6%), limitations ($n = 15$, 27%), conclusions ($n = 2$, 3.6%), and funding ($n = 40$, 73%).

On average, the publications adhered to 26 out of 40 (65%) of the subheading topics (Table II). Notable highlights from analysis of these subheadings found that 13 publications communicated the “methods employed for assessing completeness and accuracy of data” (24%). Six described “unintended consequences, such as unexpected benefits, problems, failures, or costs associated with the intervention(s),” and 2 included “details about missing data” (11% and 3.6%, respectively). Nineteen articles included “potential for spread to other contexts,” and 26 included “suggested next steps” (35 and 47%, respectively).

Four of 55 articles cited the SQUIRE guidelines (7.3%). Articles that cited SQUIRE were compliant with a mean of 29 out of 40 (73%) subheadings, compared to 26 out of 40 subheadings (65%) for those that did not cite SQUIRE ($P = .11$). These same articles were compliant with the same number of main statements (11 out of 18, 61%). The majority of articles that cited SQUIRE (3 out of 4, 75%) were published after the release of SQUIRE 2.0. In the subgroup analysis, there was no difference in the adherence to the 40 subheadings between articles from general interest versus surgery journals, between academic medical centers versus other hospitals, or between QI projects from US hospitals versus other countries ($P < .68$ each).

Table II
The SQUIRE 2.0¹¹ statements and subheadings

	Description of statement	Articles that satisfied each guideline, n (%)
1. Title	Indicate that the manuscript concerns an initiative to improve healthcare (broadly defined to include the quality, safety, effectiveness, patient centeredness, timeliness, cost, efficiency, and equity of healthcare)	51 (93)
2. Abstract	a. Provide adequate information to aid in searching and indexing b. Summarize all key information from various sections of the text using the abstract format of the intended publication or a structured summary such as: background, local problem, methods, interventions, results, conclusions	52 (95) 52 (95)
3. Problem description	Nature and significance of the local problem	55 (100)
4. Available knowledge	Summary of what is currently known about the problem, including relevant previous studies	53 (96)
5. Rationale	Informal or formal frameworks, models, concepts, and/or theories used to explain the problem, any reasons or assumptions that were used to develop the intervention(s), and reasons why the intervention(s) was expected to work	52 (95)
6. Specific aims	Purpose of the project and of this report	54 (98)
7. Context	Contextual elements considered important at the outset of introducing the intervention(s)	43 (78)
8. Intervention(s)	a. Description of the intervention(s) in sufficient detail that others could reproduce it b. Specifics of the team involved in the work	51 (92) 37 (67)
9. Study of the Intervention(s)	a. Approach chosen for assessing the impact of the intervention(s) b. Approach used to establish whether the observed outcomes were due to the intervention(s)	54 (98) 43 (78)
10. Measures	a. Measures chosen for studying processes and outcomes of the intervention(s), including rationale for choosing them, their operational definitions, and their validity and reliability b. Description of the approach to the ongoing assessment of contextual elements that contributed to the success, failure, efficiency, and cost c. Methods used for assessing completeness and accuracy of data	49 (89) 15 (27) 13 (24)
11. Analysis	a. Qualitative and quantitative methods used to draw inferences from the data b. Methods for understanding variation within the data, including the effects of time as a variable	50 (91) 19 (35)
12. Ethical Considerations	Ethical aspects of implementing and studying the intervention(s) and how they were addressed, including, but not limited to, formal ethics review and potential conflict(s) of interest	31 (56)
13. Results	a. Initial steps of the intervention(s) and their evolution over time (eg, timeline diagram, flow chart, or table), including modifications made to the intervention during the project b. Details of the process measures and outcome c. Contextual elements that interacted with the intervention(s) d. Observed associations between outcomes, interventions, and relevant contextual elements e. Unintended consequences such as unexpected benefits, problems, failures, or costs associated with the intervention(s) f. Details about missing data	25 (45) 54 (98) 24 (44) 29 (52) 6 (11) 2 (5)
14. Summary	a. Key findings, including relevance to the rationale and specific aims b. Particular strengths of the project	44 (80) 36 (65)
15. Interpretation	a. Nature of the association between the intervention(s) and the outcomes b. Comparison of results with findings from other publications c. Impact of the project on people and systems d. Reasons for any differences between observed and anticipated outcomes, including the influence of context e. Costs and strategic trade-offs, including opportunity costs	51 (93) 45 (82) 43 (78) 26 (47) 6 (11)
16. Limitations	a. Limits to the generalizability of the work b. Factors that might have limited internal validity such as confounding, bias, or imprecision in the design, methods, measurement, or analysis c. Efforts made to minimize and adjust for limitations	45 (82) 43 (78) 16 (29)
17. Conclusions	a. Usefulness of the work b. Sustainability c. Potential for spread to other contexts d. Implications for practice and for further study in the field e. Suggested next steps	53 (97) 21 (38) 19 (35) 36 (65) 26 (47)
18. Funding	Sources of funding that supported this work. Role, if any, of the funding organization in the design, implementation, interpretation, and reporting	40 (72)

Discussion

This study is the first to formally analyze the quality of surgical QI literature using the SQUIRE 2.0 guidelines as a reporting standard. Of the articles reviewed, none adhered to all of the guidelines which were developed to improve the planning, execution, and reporting of quality improvement work in surgery. Thus, the majority of surgical QI literature lacks adequate description in the methods, results, and discussion sections. These components are critical to the successful reporting and execution of QI work. In particular, half of the studies did not include a complete explanation of their intervention, which limits reproducibility across institutions.

One other published article sought to determine if the SQUIRE guidelines have improved QI literature. Howell et al reviewed

articles from the top QI journals written before and after the publication of the first version of the SQUIRE guidelines.¹⁷ Adherence was measured to the 50 SQUIRE 1.0 statements with the same dichotomous (yes or no) scale used in this review. The study was terminated early owing to the lack of any recognized improvement in the QI literature after the release of the SQUIRE guidelines. Taken together with the results of our study, these findings suggest that the SQUIRE guidelines do not seem to have made a substantial or meaningful impact on the way researchers design and report QI work.

Increasing the quality of QI reporting will help bridge the desire to improve care with the need for data to support such improvements.^{1–6} There is little disagreement about the importance of a framework for describing QI interventions and their results. SQUIRE 2.0 not only provides such a framework and should

help to define site-specific context for the QI work itself, but SQUIRE 2.0 will help to plan QI projects. Whether there is broad uptake of SQUIRE for this purpose is unknown and beyond the scope of this study, but it seems that there has yet to be wide dissemination.

Two main mechanisms may explain the low adherence to the SQUIRE guidelines. First, the surgical community may not know about the guidelines or may not have designed improvement work with the level of detail that SQUIRE suggests. Second, the guidelines themselves may be too difficult to follow completely. Of note, there are a large number of SQUIRE subheadings, and these may have varied importance or be interpreted differently depending on the project and the researcher. In our study, the small number of articles that cited the SQUIRE guidelines did not exhibit greater adherence than those that did not cite the guidelines. This observation suggests that both mechanisms may be at play in the failure to follow the guidelines. To support clear and complete reporting in the QI literature, guidelines should be easy to access, interpret, and follow.

It is notable that many journals which published the articles included in this analysis have co-published the SQUIRE 2.0 guidelines, possibly indicating complicit support of the guidelines. Explicit support from journals may be the key to integrating successfully the reporting guidelines into the design and execution of all research studies. Furthermore, journal endorsement of the EQUATOR network could serve as a streamlined way to point authors to the correct guideline for their type of study.⁸

In addition, many QI projects do not require approval by the local institutional review board, thereby reinforcing a cultural attitude that QI design and reporting does not have to be held to the same standard as other research projects. Having an institutional review process for QI projects similar to an institutional review board for research projects could help ensure adherence to many of the quality metrics essential for reporting and reproducing interventions at external institutions. As the distinction between QI work and research becomes increasingly blurred, greater focus on the SQUIRE guidelines for the design and reporting of QI projects may be needed.

This study has several limitations. The first limitation is the sample size. Although the literature search was designed to be inclusive, we may not have enough articles available to power subgroup analyses properly. Second, reviewers were not blinded to the study authors or journal which may have introduced bias into the evaluation of articles. Third, all 40 SQUIRE 2.0 statements were weighted equally in our analysis, but not all SQUIRE statements are necessarily of equal importance to conduct of high-quality improvement work and reporting of QI literature. Even the SQUIRE elaboration and instructions^{10,11} indicate that SQUIRE statements can be adapted for reporting, although no guidance is given in this regard. Recognizing that each QI report is different, we used a static rating system, possibly decreasing the external validity of our results.

In this study, our results demonstrate that the quality of the literature on surgical quality improvement is poor when assessed using the SQUIRE 2.0 guidelines. Complete reporting of QI design, processes, and outcomes has the potential to lead to increased efficiency and decreased redundancy while facilitating collaboration across hospital systems. Increased journal endorsement,

institutional review processes for QI projects, and use of the guidelines can increase the standards of surgical QI reporting.

Funding/Support

The authors have indicated that they have no conflicts of interest (or funding) regarding the content of this article.

Conflict of interest/Disclosure

All authors have no conflicts of interest to declare.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.surg.2019.05.016>.

References

1. Mouch CA, Regenbogen SE, Revels SL, Wong SL, Lemak CH, Morris AM. The quality of surgical care in safety net hospitals: A systematic review. *Surgery*. 2014;155:826–838.
2. Kao LS, Lally KP, Thomas EJ, Tyson JE. Improving quality improvement: A methodologic framework for evaluating effectiveness of surgical quality improvement. *JAMA*. 2009;208:621–626.
3. Berenholtz SM, Needham DM, Lubomski LH, Goeschel CA, Pronovost PJ. Improving the quality of quality improvement projects. *Jt Comm J Qual Patient Saf*. 2010;36:468–473.
4. McCulloch P. How to improve surgical research. *Br Med J*. 2011;343:d4121.
5. Cevasco M, Ashley SW. Quality measurement and improvement in general surgery. *Perm*. 2011;15:48–53.
6. Grady D, Redberg RF, O'Malley PG. Quality improvement for quality improvement studies. *JAMA Intern Med*. 2018;178:187.
7. Mueck KM, Putnam LR, Kao LS. Improving the quality of quality improvement reporting: Standards for Quality Improvement Reporting Excellence (SQUIRE) 2.0 Guidelines. *JAMA Surg*. 2016;151:311–312.
8. Simera I, Altman DG, Moher D, Schulz KF, Hoey J. Guidelines for reporting health research: The EQUATOR Network's survey of guideline authors. *PLoS Med*. 2008;5:e139.
9. Kamal AH, Agrawal S. Rigor in quality improvement studies and the role of time-series methodologies. *JAMA Intern Med*. 2018;178:725–725.
10. 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;17(Suppl 1):i13–i32.
11. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for Quality Improvement Reporting Excellence): Revised publication guidelines from a detailed consensus process. *J Perm*. 2015;19:65–70.
12. Davies L, Batalden P, Davidoff F, Stevens D, Ogrinc G. The SQUIRE Guidelines: An evaluation from the field, 5 years post release. *BMJ Qual Saf*. 2015;24:769–775.
13. Revised Standard for Quality Improvement Reporting Excellence SQUIRE 2.0. Available from: <http://squire-statement.org>. Accessed March 17, 2019.
14. Enhancing the QUALity and Transparency Of health Research: Equator Network. Available from: <http://equator-network.org>. Accessed March 17, 2019.
15. 2017 Journal Impact Factor, Journal Citation Reports Science Edition. Clarivate Analytics. Available from: <https://clarivate.com/products/journal-citation-reports>. Accessed January 17, 2019.
16. Where to Submit Your Writing: Journals Publishing Student Work. Available from: <http://www.ihl.org/education/IHIOpenSchool/resources/Pages/WhereToSubmitYourWritingQIFriendlyPeerReviewedJournals.aspx>. Accessed January 17, 2019.
17. Howell V, Schwartz AE, O'Leary JD, Mc Donnell C. The effect of the SQUIRE (Standards of Quality Improvement Reporting Excellence) guidelines on reporting standards in the quality improvement literature: A before-and-after study. *BMJ Qual Saf*. 2015;24:400–406.