Published in final edited form as:

J Hand Surg Am. 2013 August ; 38(8): 1584–9.e2. doi:10.1016/j.jhsa.2013.05.008.

Using the STROBE Statement to Assess Reporting of Observational Trials in Hand Surgery

Amelia A. Sorensen, MD, Robert D. Wojahn, BS, Mary Claire Manske, MD, and Ryan P. Calfee, MD, MSc

Washington University School of Medicine, Department of Orthopaedic Surgery, St. Louis, Missouri

Abstract

Purpose—To use the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement checklist to critically evaluate the change in quality of observational trial reporting in the *Journal of Hand Surgery American* between 2005 and 2011.

Methods—A cross-sectional analysis of observational studies published in *Journal of Hand Surgery American* was designed to sample 2 6-month periods of publication (March 2005 - August 2005 and June 2011 - November 2011). Fifty-one items were extracted from the STROBE statement for evaluation. Overall STROBE compliance rates for articles and specific checklist items were determined. Final compliance percentages from each period were compared by Student t-testing. Changes in item compliance over time were quantified.

Results—Overall compliance with the STROBE statement was 38% (range, 10-54%) in 2005 and 58% (range, 39%-85%) for 2011 manuscripts representing a significant improvement. Seventy-five percent or greater of articles (2005/2011) provided the explicit reporting of background (100%/97%), follow-up time (85%/94%), overall interpretation of data (100%/94%), and results of similar studies (95%/89%). Twenty-five percent or less of articles provided the study design in the abstract (10%/20%), a clear description of the study's setting (10%/23%), the handling of missing data (0%/6%), the potential directions of bias (5%/11%) and use of a power analysis (0%/17%). Eighty-six percent (44/51) of items were more frequently satisfied in 2011 articles compared to 2005 publications. Absolute increases in compliance rates of 40% were noted in 10 items (20%) with no worsening in compliance for an individual item over 6%.

Discussion—The overall quality of the reporting of observational trials in *Journal of Hand Surgery American* improved from 2005 to 2011. Current observational trials in hand surgery could still benefit from increased reporting of methodologic details including the use of power analyses, the handling of missing data, and consideration of potential bias.

© 2013 The American Society for Surgery of the Hand. Published by Elsevier Inc. All rights reserved.

Corresponding Author: Amelia Sorensen, M.D., Washington University School of Medicine, Department of Orthopaedic Surgery, 660 South Euclid Avenue, Campus Box 8233, St. Louis, MO 63110, 314-747-4705, sorensena@wudosis.wustl.edu.

Level of Evidence: III-Diagnostic review

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Key Terms

STROBE; Cohort Studies; Observational Studies; Reporting

Introduction

High quality research is necessary to provide a basis for the practice of evidence-based medicine. Surgical disciplines, compared to medical specialties, are less likely to produce randomized controlled trials due to feasibility and logistical and ethical concerns (1). As a result, hand surgical literature includes a preponderance of observational studies. At a time when translation of scientific findings into clinical practice is at a premium, investigators must conduct scientifically sound studies and produce detailed transparent manuscripts “so that readers can follow what was planned, what was done, what was found, and what conclusions were drawn (2).”

The Strengthening the Reporting of Observation Studies in Epidemiology (STROBE) statement was produced in 2007 to “improve the quality of observational study reporting, improve transparency in reporting, and allow for critical assessment by others of the strengths and weaknesses in study design, conduct, and analysis (2).” A team of 23 editors, epidemiologists, methodologists, statisticians, and practitioners from Europe and North America created this statement (3). The STROBE statement provides a 22 item checklist of items for inclusion in the reporting of all observational studies including components of the study design, interventions, data collection, analytic techniques, and potential bias. Available checklists are specific for each of the 3 observational study designs (i.e., cohort, case-control, and cross-sectional). Use of these checklists is intended to improve the reader's ability to assess, interpret, and generalize study findings.

The quality of general orthopedic and plastic surgery reporting has been analyzed using the STROBE statement (4,5). However, it is currently unknown how well reporting in hand literature would satisfy the STROBE checklist and how the quality of hand surgical research has changed since the introduction of the STROBE statement in 2007. The purpose of this study was to use the STROBE statement to critically evaluate the change in quality of observational study reporting in the Journal of Hand Surgery (JHS [Am]) between 2005 (prior to both JHS requiring level of evidence reporting and the publication of the STROBE statement) and 2011 (4 years after the introduction of the STROBE statement). We tested the null hypothesis that the quality of reporting for observational studies would be unchanged between the 2 time periods.

Methods

Identifying studies

A cross-sectional analysis of observational studies published in JHS [Am] was designed to sample 2 6-month periods of publication. This was chosen to result in precise estimation (95% CI +/- 4%) of mean overall compliance with STROBE items assuming 20 articles per period (assuming 50% compliance and 35 applicable checklist items per article) and

confirmed to produce a study comparable to similar prior investigations (6). The first study period was from March 2005 through August 2005. This preceded the decision detailed in the September 2005 JHS [Am] issue to publish assigned levels of evidence to clinical studies in order to “improve the quality of published articles and provide objective benchmarks to measure the scientific methodology of the clinical studies.” (7) The second study period was from June 2011 through November 2011. This period was selected to represent current reporting 6 years after introduction of level of evidence reporting and 4 years after publication of the STROBE statement.

The abstracts of all published articles in JHS during the selected time periods were reviewed by 2 investigators. For purposes of this investigation, all cohort, case-control and cross-sectional observational studies were selected for evaluation with the STROBE statement. Eighty-eight abstracts were reviewed from 2005. Twenty-one observational studies were identified consisting of 4 case-control studies, 17 cohort studies, and 1 cross-sectional study. Ninety-three abstracts were reviewed from 2011. Thirty-five observational studies were identified consisting of 4 case-control, 28 cohort, and 3 cross-sectional studies.

Checklist

Fifty-one items were extracted from the STROBE checklists for evaluation. These items were based upon previous published assessments and defined explicitly for each reviewer (8). To maximize inter-rater reliability, each investigator independently reviewed 1 observational study prior to scoring articles for this investigation. All investigators collectively reviewed this initial study, discussed any rating discrepancies, and clarified scoring of review items.

Each checklist item was categorized as “yes” (met the criteria), “no” (did not meet the criteria) or “not applicable.” Each manuscript was reviewed by 2 of 4 authors with 1 author (AAS) reading all articles. Each reviewer evaluated the article content independently. Discrepancies were resolved by consensus among the reviewers and the senior author(RPC).

Data Analysis

Reviewer datasets were collected on article grading sheets and entered into an electronic database for tabulation and descriptive statistics. If a study contained no statistical analysis, a “no” was entered for the first item concerning use of statistical methods, and subsequent items for statistics were marked not applicable. When determining both overall STROBE compliance rates for articles and compliance rates with specific checklist items, the number of yes ratings was divided by the sum of the yes and no responses. Ratings of not applicable did not impact the calculated percentages of compliance. The final compliance percentages were then compared between the early and current articles by the Student t-testing.

Results

Overall compliance with STROBE guidelines was 38% (95% CI 35%-42%: range, 10-54%) in 2005 and 58% (95% CI 55%-60%: range, 39%-85%) for 2011 articles. This increase in overall compliance represented a statistically significant improvement ($P<0.01$) for the current manuscripts.

2005 Articles

Twelve areas (24%) showed excellent compliance with greater than 75% of articles satisfactorily reporting these items. These items included providing a balanced summary in the abstract (95%), the study's background (100%), characteristics of study patients (90%), and an overall interpretation of findings (100%). (Table 1)

Thirteen areas (25%) had intermediate compliance with 26% to 74% of articles reporting these items. These areas more inconsistently documented included describing the objective of the study (57%), detailing the method of follow-up (45%), explaining the statistical methods for analysis (48%), and presenting limitations of the study (38%).

Twenty-six areas (51%) had 25% or less of articles reporting these items. These rarely satisfied items included providing a clear description of the study's setting (10%), matching criteria for case-control studies (0%), discussing how bias was addressed (0%), providing a power analysis (0%), handling of missing data (0%), and providing rates of loss to follow-up (6%).

2011 Articles

Twenty areas (39%) were reported consistently with greater than 75% papers satisfying these requirements. These items included providing a study objective (80%), the follow-up time (94%), outcome measures (100%), methods of assessment (94%), and an overall interpretation of the data (94%).

Fourteen areas (28%) were reported by 26%-74% of papers including providing the study design early in the paper (49%), source of the data (74%), description of statistical methods (71%), the number at each stage of the study (73%), reasons for nonparticipation (39%), and limitations of the study (69%).

Study compliance was less than 25% in only 14 areas (28%). This included presenting the study design in the abstract (20%), matching criteria for case-control studies (25%), addressing bias (3%), providing a power analysis (17%), and addressing missing data (6%).

Individual Item Change Over Time

Eighty-six percent (44/51) of items were more frequently satisfied in 2011 articles compared to 2005 publications. Absolute increases in compliance rates of 40% were noted in 10 items (20%). Meanwhile, there was not substantial evidence of decreased reporting of any item (maximum deterioration of 6%). (Table 2)

Discussion

Judged using the STROBE statement, overall quality of the reporting of observational trials improved in JHS [Am] between 2005 and 2011. Improvement was particularly remarkable in detailing methodology (matching criteria, patient selection, and eligibility). Statistical reporting saw improvement between 2005 and 2011 with a greater percentage of papers reporting on the statistical treatment of data and providing subgroup and sensitivity analyses in the later time period.

Despite improved reporting, the STROBE checklist identified several areas that provide room for further improvement in our literature. While statistical reporting has improved, only 17% of 2011 trials provided a power analysis. This was particularly relevant when judging null results that could either indicate a lack of true difference between study groups or indicate an underpowered statistical analysis that failed to detect a true difference. The setting for the studies were infrequently reported or only described as “our institution”. It is important to include a complete description of the setting to allow the reader to know if the results of a particular study would generalize to their population. Only 2 papers used a flow diagram (9, 10). Flow diagrams provide an effective way to convey information about participants (8). Missing data and losses to follow-up were also infrequently addressed in the Methods and Results sections at both time points. These are potential sources of information and determination of attritional bias. Additionally, while a majority of studies mentioned at least 1 limitation, few specifically addressed biases. No studies at either time point included within the body of the paper sources of funding, although this is not requested by JHS [Am] beyond the author disclosures provided in the print version of JHS [Am] studies.

The success and shortcomings of the observational studies we reviewed are comparable to reporting across multiple clinical specialties. Parson reviewed the quality of general orthopedic research reporting with STROBE guidelines in 2011 and found a 58% overall compliance rate, which was replicated in our review of hand literature (4). Three years after release of the STROBE statement 60 cohort studies across selective peer-reviewed journals were reviewed. Fifteen percent reported sources of bias, and 20% reported a power analysis (6). Failure to report on the treatment of missing data (6% in our current studies) was identified as a deficiency in the dermatology literature (6%), oncology (31%), ophthalmology (12%), and sexual health literature (35%) (11-14). Other commonly cited areas for improvement have included the use of flow diagrams, addressing patients lost to follow-up, and identification of the role of funding (11-14).

The present study has several limitations. The finite number of studies evaluated may increase the possibility of random error, and the limited time frame may increase the possibility of selection bias. Second, despite reporting overall STROBE compliance rates, we cannot assume that each item has the same impact on the quality of reporting. Thus, estimating a summary compliance percentage is an imperfect proxy for the quality of the article. However, we believe that the overall rates of compliance are useful as a gauge of change in reporting quality over time. Third, investigators reviewed articles without blinding, which introduced potential bias related to presumed quality associated with specific authors or institutions. Although we cannot exclude the possibility of rating variation between reviewers or as a result of reviewer fatigue, we attempted to limit these influences by having 2 graders for each study and reaching a consensus during data assimilation.

Excellence in reporting is a key component of research that improves the readability and reproducibility of investigations. Although the reporting of observational trials in the hand surgical literature is improving, it is not perfect. The primary limitations of manuscripts in our review appear to be areas of omission that may have been considered during study design and analysis but not reported. For this reason we would encourage authors to consult

the STROBE guidelines during manuscript preparation as 1 resource to improve the quality of their reporting. In our experience, use of the STROBE and other relevant checklists for scientific reporting have also proved invaluable as references during study inception when potential limitations can be considered and addressed prior to data collection. Alternatively, standardized reporting guidelines offer another avenue for reviewers to consider during the peer-review process as we all work toward the highest quality of scientific reporting in hand surgery.

References

1. Hoppe DJ, Schemitsch EH, Morshed S, Tornetta P 3rd, Bhandari M. Hierarchy of evidence: where observational studies fit in and why we need them. *J Bone Joint Surg Am*. 2009 May; 91(Suppl 3): 2–9. [PubMed: 19411493]
2. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007; 370:1453–1457. [PubMed: 18064739]
3. Vandenbroucke JP. The making of STROBE. *Epidemiology*. 2007 Nov; 18(6):797–799. [PubMed: 18049193]
4. Parsons NR, Hiskens R, Price CL, Achten J, Costa ML. A systematic survey of the quality of research reporting in general orthopaedic journals. *J Bone Joint Surg Br*. 2011 Sep; 93(9):1154–1159. [PubMed: 21911523]
5. Knobloch K, Yoon U, Vogt PM. Observational studies and STROBE reporting quality in plastic surgical conference abstracts. *Plast Reconstr Surg*. 2011 Aug; 128(2):108e–109e.
6. Poorolajal J, Cheraghi Z, Irani AD, Rezaeian S. Quality of Cohort Studies Reporting Post the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement. *Epidemiol Health*. 2011:33.
7. Hentz R, Meals R, Stern P. Levels of evidence and the Journal of Hand Surgery. *J Hand Surg Am*. 2005 Sep; 30A:5.
8. Sheffler LC, Yoo B, Bhandari M, Ferguson T. Observational Studies in Orthopaedic Surgery: The STROBE Statement as a Tool for Transparent Reporting. *J Bone Joint Surg Am*. 2013 Feb 6. 95(3):e14. 1–12. [PubMed: 23389794]
9. Wong K, von Schroeder HP. Delays and poor management of scaphoid fractures: factors contributing to nonunion. *J Hand Surg Am*. 2011 Sep; 36(9):1471–1474. [PubMed: 21803508]
10. Yu YR, Makhni MC, Tabrizi S, Rozental TD, Mundanthanam G, Day CS. Complications of low-profile dorsal versus volar locking plates in the distal radius: a comparative study. *J Hand Surg Am*. 2011 Jul; 36(7):1135–1141. [PubMed: 21712136]
11. Langan SM, Schmitt J, Coenraads PJ, Svensson A, von Elm E, Williams HC, European Dermato-Epidemiology Network (EDEN). STROBE and reporting observational studies in dermatology. *Br J Dermatol*. 2011 Jan; 164(1):1–3. [PubMed: 21175559]
12. Papathanasiou AA, Zintzaras E. Assessing the quality of reporting of observational studies in cancer. *Ann Epidemiol*. 2010 Jan; 20(1):67–73. [PubMed: 20006277]
13. Fung AE, Palanki R, Bakri SJ, Depperschmidt E, Gibson A. Applying the CONSORT and STROBE statements to evaluate the reporting quality of neovascular age-related macular degeneration studies. *Ophthalmology*. 2009 Feb; 116(2):286–296. [PubMed: 19091408]
14. Müller M, Egger M. Strengthening the reporting of observational epidemiology (STROBE) in sexual health. *Sex Transm Infect*. 2009 Jun; 85(3):162–164. [PubMed: 19478105]

Table 1

Percentage of articles addressing STROBE statement items according to publication date.

STROBE Statement Item	2005*	2011*
1a. Abstract: Study design with specific descriptive term	(2/21) 10%	(7/35) 20%
1b. Abstract: Balanced summary	(20/21) 95%	(32/35) 91%
2. Introduction: Background and rationale	(21/21) 100%	(34/35) 97%
3. Introduction: Objectives: include hypothesis, objective or purpose	(12/21) 57%	(28/35) 80%
4. Methods: Study design early in paper	(4/21) 19%	(17/35) 49%
5. Methods: Setting and location	(2/21) 10%	(8/35) 23%
5. Methods: Recruitment period	(14/21) 67%	(28/35) 80%
5. Methods: Follow-up time **	(17/20) 85%	(30/33) 94%
6a. Methods: Eligibility	(9/21) 43%	(32/35) 91%
6a. Methods: Sources, Methods of patient or case selection	(4/21) 19%	(29/35) 83%
6a. Methods: Cohort: method of follow-up	(9/17) 45%	(18/28) 62%
6b. Methods: Case Control: matching criteria	(0/4) 0%	(1/4) 25%
6b. Methods: Case Control: number of exposed	(4/4) 100%	(4/4) 100%
7. Methods: Define outcomes	(18/21) 86%	(35/35) 100%
7. Methods: Define exposures	(19/21) 90%	(34/35) 97%
7. Methods: Define predictors	(1/21) 5%	(5/35) 14%
7. Methods: Define potential confounders	(0/21) 0%	(2/35) 6%
7. Methods: Define effect modifiers	(1/21) 5%	(2/35) 2%
7. Methods: Diagnostic criteria	(16/21) 76%	(33/35) 94%
8. Methods: Source of data	(10/21) 48%	(26/35) 74%
8. Methods: Methods of assessment or measurement	(14/21) 67%	(33/35) 94%
9. Methods: How bias addressed	(0/21) 0%	(1/35) 3%
10. Methods: Power analysis	(0/21) 0%	(6/35) 17%
11. Methods: How quantitative variables addressed	(4/21) 19%	(21/35) 60%
12a. Methods: Statistical methods	(10/21) 48%	(25/35) 71%
12b. Methods: Statistical subgroups/interactions	(3/10) 30%	(13/25) 52%
12c. Methods: How missing data addressed	(0/21) 0%	(2/35) 6%
12d. Methods: Cohort: How loss to follow up addressed	(1/17) 6%	(6/28) 21%
12d. Methods: Case Control: How matched	(0/4) 0%	(1/4) 25%
12d. Methods: Cross-Sectional: sampling strategy	(0/1) 0%	(2/3) 67%
12e. Methods: Sensitivity analyses	(1/10) 10%	(12/25) 48%
13a. Results: Number at each stage of study	(5/21) 24%	(19/35) 73%
13a. Results: Reasons for nonparticipation	(0/21) 0%	(13/35) 39%
13a. Results: Use of a flow diagram	(0/21) 0%	(2/35) 6%
14a. Results: Characteristics of study participants	(19/21) 90%	(34/35) 97%
14b. Results: Number with missing data	(1/21) 5%	(3/35) 9%

STROBE Statement Item	2005*	2011*
14c. Results: Cohort: Follow-up time	(11/17) 69%	(24/28) 89%
15. Results: Number of events or exposures	(13/21) 62%	(32/32) 100%
16a. Results: Unadjusted estimates	(4/21) 19%	(23/35) 66%
16a. Results: Confounder adjusted estimates with reasoning	N/A	(2/2) 100%
16a. Results: 95% Confidence Interval	(0/21) 0%	(4/35) 11%
16b. Results: Category boundaries for continuous variables	(2/21) 10%	(21/35) 64%
16c. Results: Translate Relative Risk to Absolute Risk	N/A	N/A
17. Results: Other analyses (subgroups/interactions/sensitivity)	(3/10) 33%	(19/25) 76% /
18. Discussion: Key results with referenced objectives	(16/21) 76%	(31/35) 89%
19. Discussion: Limitations of study	(8/21) 38%	(23/35) 69%
19. Discussion: Directions or magnitude of bias	(1/21) 5%	(4/35) 11%
20. Discussion: Overall interpretation	(21/21) 100%	(33/35) 94%
20. Discussion: Results similar studies	(18/21) 86%	(30/35) 86%
21. Discussion: Generalizability	(20/21) 95%	(31/35) 89%
22. Funding: Source, role of funders, funding for original study	(0/21) 0%	(0/35) 0%
Overall Compliance	39%	58%

* (number of articles reporting/number of articles potentially reporting this item)

** Not applicable to cross sectional studies

Table 2

STROBE items demonstrating the greatest improvement since 2005.

STROBE Statement Item	2005	2011	Change
16a. Results: Confounder adjusted estimates with reasoning	N/A	100%	100%
14a. Results: Characteristics of study participants	9%	97%	88%
12d. Methods: Cross-sectional: sampling strategy	0%	67%	67%
6a. Methods: Sources, methods of patient and case selection	19%	83%	64%
16b. Results: Category boundaries for continuous variables	10%	64%	54%
13a. Results: Number at each stage of study	24%	73%	49%
6a. Methods: Eligibility	43%	91%	48%
16a. Results: Unadjusted estimates	19%	66%	47%
17. Results: Other analyses (subgroups/interactions/sensitivity)	33%	76%	43%
11. Methods: How quantitative variables addressed	19%	60%	41%