Basic Statistical Reporting for Articles Published in Biomedical Journals: The "Statistical Analyses and Methods in the Published Literature" or The SAMPL Guidelines"

Thomas A. Lang[a] and **Douglas G. Altman**[b]
[a] Principal, Tom Lang Communications and Training International
[b] Director, Centre for Statistics in Medicine, Oxford University

Have they reflected that the sciences founded on observation can only be promoted by statistics? . . . If medicine had not neglected this instrument, this means of progress, it would possess a greater number of positive truths, and stand less liable to the accusation of being a science of unfixed principles, vague and conjectural.

Jean-Etienne Dominique Esquirol, an early French psychiatrist,
quoted in The Lancet, 1838 [1]

Introduction

The first major study of the quality of statistical reporting in the biomedical literature was published in 1966 [2]. Since then, dozens of similar studies have been published, every one of which has found that large proportions of articles contain errors in the application, analysis, interpretation, or reporting of statistics or in the design or conduct of research. (See, for example, references 3 through 19.) Further, large proportions of these errors are serious enough to call the authors' conclusions into question [5,18,19]. The problem is made worse by the fact that most of these studies are of the world's leading peer-reviewed general medical and specialty journals.

Although errors have been found in more complex statistical procedures [20,21,22], paradoxically, many errors are in basic, not advanced, statistical methods [23]. Perhaps advanced methods are suggested by consulting statisticians, who then competently perform the analyses, but it is also true that authors are far more likely to use only elementary statistical methods, if they use any at all [23-26]. Still, articles with even major errors continue to pass editorial and peer review and to be published in leading journals.

The truth is that the problem of poor statistical reporting is long-standing, widespread, potentially serious, concerns mostly basic statistics, and yet is largely unsuspected by most readers of the biomedical literature [27].

More than 30 years ago, O'Fallon and colleagues recommended that "Standards governing the content and format of statistical aspects should be developed to guide authors in the preparation of manuscripts" [28]. Despite the fact that this call has since been echoed by several others (17,18,29-32), most journals have still not included in their Instructions for Authors more than a paragraph or two about reporting statistical methods [33]. However, given that many statistical errors concern basic statistics, a

Lang T, Altman D. Basic statistical reporting for articles published in clinical medical journals: the SAMPL Guidelines. In: Smart P, Maisonneuve H, Polderman A (eds). *Science Editors' Handbook,* European Association of Science Editors, 2013.

comprehensive—and comprehensible—set of reporting guidelines might improve how statistical analyses are documented.

In light of the above, we present here a set of statistical reporting guidelines suitable for medical journals to include in their Instructions for Authors. These guidelines tell authors, journal editors, and reviewers how to report basic statistical methods and results. Although these guidelines are limited to the most common statistical analyses, they are nevertheless sufficient to prevent most of the reporting deficiencies routinely found in scientific articles; they may also help to prevent some reporting errors by focusing attention on key points in the analyses.

Unlike many of other guidelines, the SAMPL guidelines were not developed by a formal consensus-building process, but they do draw considerably from published guidelines [27,34-37]. In addition, a comprehensive review of the literature on statistical reporting errors reveals near universal agreement on how to report the most common methods [27].

Statistical analyses are closely related to the design and activities of the research itself. However, our guidelines do not address the issues related to the design and conduct of research. Instead, we refer readers to the EQUATOR Network website (www.equator-network.org) where guidelines for reporting specific research designs can be found. (For example, see the CONSORT [38], TREND [39], STROBE [40]) These guidelines for reporting methodologies all include items on reporting statistics, but the guidelines presented here are more specific and complement, not duplicate, those in the methodology guidelines.

We welcome feedback and anticipate the need to update this guidance in due course.

Reporting Basic Statistical Analyses and Methods in the Published Literature: The SAMPL Guidelines for Biomedical Journals

Guiding Principles for Reporting Statistical Methods and Results

Our first guiding principle for statistical reporting comes from The International Committee of Medical Journal Editors, whose Uniform Requirements for Manuscripts Submitted to Biomedical Journals include the following excellent statement about reporting statistical analyses:

> "**Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results.** [Emphasis added.] When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid relying solely on statistical hypothesis testing, such as *P* values, which fail to convey important information about effect size. References for the design of the study and statistical methods should be to standard works when possible (with pages stated). Define statistical terms, abbreviations, and most symbols. Specify the computer software used" [33,41].

Our second guiding principle for statistical reporting is to **provide enough detail that the results can be incorporated into other analyses.** In general, this principle requires reporting the descriptive statistics from which other statistics are derived, such as the numerators and denominators of percentages, especially in risk, odds, and hazards ratios. Likewise, P values are not sufficient for re-analysis. Needed instead are descriptive statistics for the variable being compared, including sample size of the groups involved, the estimate (or "effect size") associated with the P value, and a measure of precision for the estimate, usually a 95% confidence interval.

General Principles for Reporting Statistical Methods

Preliminary analyses

- Identify any statistical procedures used to modify raw data before analysis. Examples include mathematically transforming continuous measurements to make distributions closer to the normal distribution, creating ratios or other derived variables, and collapsing continuous data into categorical data or combining categories.

Primary analyses

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- When possible, identify the smallest difference considered to be clinically important.
- Describe fully the main methods for analyzing the primary objectives of the study.
- Make clear which method was used for each analysis, rather than just listing in one place all the statistical methods used.
- Verify that that data conformed to assumptions of the test used to analyze them. In particular, specify that 1) skewed data were analyzed with non-parametric tests, 2) paired data were analyzed with paired tests, and 3) the underlying relationship analyzed with linear regression models was linear.
- Indicate whether and how any allowance or adjustments were made for multiple comparisons (performing multiple hypothesis tests on the same data).
- If relevant, report how any outlying data were treated in the analysis.

- Say whether tests were one- or two-tailed and justify the use of one-tailed tests.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- Name the statistical package or program used in the analysis.

Supplementary analyses

- Describe methods used for any ancillary analyses, such as sensitivity analyses, imputation of missing values, or testing of assumptions underlying methods of analysis.
- Identify post-hoc analyses, including unplanned subgroup analyses, as exploratory.

General Principles for Reporting Statistical Results

Reporting numbers and descriptive statistics

- Report numbers—especially measurements—with an appropriate degree of precision. For ease of comprehension and simplicity, round as much as is reasonable. For example, mean age can often be rounded to the nearest year without compromising either the clinical or the statistical analysis. If the smallest meaningful difference on a scale is 5 points, scores can be reported as whole numbers; decimals are not necessary.
- Report total sample and group sizes for each analysis.
- Report numerators and denominators for all percentages.
- Summarize data that are approximately normally distributed with means and standard deviations (SD). Use the form: mean (SD), not mean ± SD.
- Summarize data that are not normally distributed with medians and interpercentile ranges, ranges, or both. Report the upper and lower boundaries of interpercentile ranges and the minimum and maximum values of ranges, not just the size of the range
- Do NOT use the standard error of the mean (SE) to indicate the variability of a data set. Use standard deviations, inter-percentile ranges, or ranges instead.
- Display the data in tables or figures. Tables present exact values, and figures provide an overall assessment of the data.[42,43]

Reporting risk, rates, and ratios

- Identify the type of rate (incidence rates; survival rates), ratio (odds ratios; hazards ratios), or risk (absolute risks; relative risk differences), being reported.
- Identify the quantities represented in the numerator and denominator (e.g., the number of men with prostate cancer divided by the number of men capable of having prostate cancer).
- Identify the time period over with each rate applies.
- Identify any unit of population (that is, the unit multiplier: e.g., x 100; x 10,000) associated with the rate.
- Consider reporting a measure of precision (a confidence interval) for estimated risks, rates, and ratios.

Reporting hypothesis tests

- State the hypothesis being tested.
- Identify the variables in the analysis and summarize the data for each variable with the appropriate descriptive statistics.
- If possible, identify the minimum difference considered to be clinically important.
- For equivalence and non-inferiority studies, report the largest difference between groups that will still be accepted as indicating biological equivalence (the equivalence margin).
- Identify the name of the test used in the analysis. Report whether the test was one- or two-tailed and for paired or independent samples.
- Confirm that the assumptions of the test were met by the data.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- At least for primary outcomes, such as differences or agreement between groups, diagnostic sensitivity, and slopes of regression lines, report a measure of precision, such as the 95% confidence interval.
- Do NOT use the standard error of the mean (SE) to indicate the precision of an estimate. The SE is essentially a 68% confidence coefficient: use the 95% confidence coefficient instead.
- Although not preferred to confidence intervals, if desired, *P* values should be reported as equalities when possible and to one or two decimal places (e.g., $P = 0.03$ or 0.22 not as inequalities: e.g., $P < 0.05$). Do NOT report "NS"; give the actual P value. The smallest *P* value that need be reported is $P < 0.001$, save in studies of genetic associations.
- Report whether and how any adjustments were made for multiple statistical comparisons.
- Name the statistical software package used in the analysis.

Reporting association analyses

- Describe the association of interest.
- Identify the variables used and summarize each with descriptive statistics.
- Identify the test of association used.
- Indicate whether the test was one- or two-tailed. Justify the use of one-tailed tests.
- For *tests* of association (e.g., a *chi*-square test), report the *P* value of the test (because association is defined as a statistically significant result).
- For *measures* of association (i.e., the *phi* coefficient), report the value of the coefficient and a confidence interval. Do not describe the association as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories given their biological implications or realities.
- For primary comparisons, consider including the full contingency table for the analysis.
- Name the statistical package or program used in the analysis.

Reporting correlation analyses

- Describe the purpose of the analysis.
- Summarize each variable with the appropriate descriptive statistics.
- Identify the correlation coefficient used in the analysis (e.g., Pearson, Spearman).
- Confirm that the assumptions of the analysis were met.
- Report the alpha level (e.g., 0.05) that indicates whether the correlation coefficient is statistically significant.

- Report the value of the correlation coefficient. Do not describe correlation as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories given their biological implications or realities.

- For primary comparisons, report the (95%) confidence interval for the correlation coefficient, whether or not it is statistically significant.

- For primary comparisons, consider reporting the results as a scatter plot. The sample size, correlation coefficient (with its confidence interval), and *P* value can be included in the data field.

- Name the statistical package or program used in the analysis.

Reporting regression analyses

- Describe the purpose of the analysis.

- Identify the variables used in the analysis and summarize each with descriptive statistics.

- Confirm that the assumptions of the analysis were met. For example, in linear regression indicate whether an analysis of residuals confirmed the assumptions of linearity.

- If relevant, report how any outlying values were treated in the analysis.

- Report how any missing data were treated in the analyses.

- For either simple or multiple (multivariable) regression analyses, report the regression equation.

- For multiple regression analyses: 1) report the alpha level used in the univariate analysis; 2) report whether the variables were assessed for a) colinearity and b) interaction; and 3) describe the variable selection process by which the final model was developed (e.g., forward-stepwise; best subset).

- Report the regression coefficients (beta weights) of each explanatory variable and the associated confidence intervals and *P* values, preferably in a table.

- Provide a measure of the model's "goodness-of-fit" to the data (the coefficient of determination, r^2, for simple regression and the coefficient of multiple determination, R^2, for multiple regression).

- Specify whether and how the model was validated.

- For primary comparisons analyzed with simple linear regression analysis, consider reporting the results graphically, in a scatter plot showing the regression line and its confidence bounds. Do not extend the regression line (or the interpretation of the analysis) beyond the minimum and maximum values of the data.

- Name the statistical package or program used in the analysis.

Reporting analyses of variance (ANOVA) or of covariance (ANCOVA)

- Describe the purpose of the analysis.

- Identify the variables used in the analysis and summarize each with descriptive statistics.

- Confirm that the assumptions of the analysis were met. For example, indicate whether an analysis of residuals confirmed the assumptions of linearity.

- If relevant, report how any outlying data were treated in the analysis.

- Report how any missing data were treated in the analyses.

- Specify whether the explanatory variables were tested for interaction, and if so how these interactions were treated.

- If appropriate, in a table, report the *P* value for each explanatory variable, the test statistics and, where applicable, the degrees of freedom for the analysis.

- Provide an assessment of the goodness-of-fit of the model to the data, such as R^2.
- Specify whether and how the model was validated.
- Name the statistical package or program used in the analysis.

Reporting survival (time-to-event) analyses

- Describe the purpose of the analysis.
- Identify the dates or events that mark the beginning and the end of the time period analyzed.
- Specify the circumstances under which data were censored.
- Specify the statistical methods used to estimate the survival rate.
- Confirm that the assumptions of survival analysis were met.
- For each group, give the estimated survival probability at appropriate follow-up times, with confidence intervals, and the number of participants at risk for death at each time. It is often more helpful to plot the cumulative probability of not surviving, especially when events are not common.
- Reporting median survival times, with confidence intervals, is often useful to allow the results to be compared with those of other studies.
- Consider presenting the full results in a graph (e.g., a Kaplan-Meier plot) or table.
- Specify the statistical methods used to compare two or more survival curves.
- When comparing two or more survival curves with hypothesis tests, report the *P* value of the comparison
- Report the regression model used to assess the associations between the explanatory variables and survival or time-to-event.
- Report a measure of risk (e.g., a hazard ratio) for each explanatory variable, with a confidence interval.

Reporting Bayesian analyses

- Specify the pre-trial probabilities ("priors").
- Explain how the priors were selected.
- Describe the statistical model used.
- Describe the techniques used in the analysis.
- Identify the statistical software program used in the analysis.
- Summarize the posterior distribution with a measure of central tendency and a credibility interval
- Assess the sensitivity of the analysis to different priors.

References

1. Esquirol JED. Cited in: Pearl R. Introduction to Medical Biometry and Statistics. Philadelphia: WB Saunders, 1941.
2. Schor S, Karten I. Statistical evaluation of medical journal manuscripts. JAMA. 1966;195:1123-8.
3. Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal anti-inflammatory drugs in rheumatoid arthritis. Control Clin Trials. 1989;50:356.
4. Hall JC, Hill D, Watts, JM. Misuse of statistical methods in the Australasian surgical literature. Aust N Z J Surg. 1982;52:541-3.
5. Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. Circulation. 1980;61:1-7.
6. Lionel ND, Herxheimer A. Assessing reports of therapeutic trials. BMJ. 1970;3:637- 40.
7. Altman DG. Statistics in medical journals: developments in the 1980s. Stat Med. 1991;10:1897-913.

8. White S J. Statistical errors in papers in the British Journal of Psychiatry. Br J Psychiatr. 1979;135:336-42.
9. Gore SM, Jones IG, Rytter EC. Misuse of statistical methods: critical assessment of articles in BMJ from January to March 1976. BMJ. 1977;1:85-7.
10. Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 negative trials. N Engl J Med. 1978;299:690-4.
11. Reed JF, Slaichert W. Statistical proof in inconclusive "negative" trials. Arch Intern Med. 1981;141:1307-10.
12. Gardner MJ, Altman DG, Jones DR, Machin D. Is the statistical assessment of papers submitted to the British Medical Journal effective? BMJ. 1983;286:1485-8.
13. MacArthur RD, Jackson GG. An evaluation of the use of statistical methodology in the Journal of Infectious Diseases. J Infect Dis. 1984;149:349-54.
14. Avram MJ, Shanks CA, Dykes MH, et al. Statistical methods in anesthesia articles: an evaluation of two American journals during two six-month periods. Anesth Analg. 1985;64:607-11.
15. Godfrey K. Comparing the means of several groups. N Engl J Med.1985;313:1450-6.
16. A survey of three medical journals. N Engl J Med. 1987;317:426-32.
17. Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. N Engl J Med. 1987 Aug 13;317(7):426-32.
18. Murray GD. Statistical aspects of research methodology. Br J Surg. 1991;78:777-81.
19. Yancy JM. Ten rules for reading clinical research reports [Editorial]. Am J Surg. 1990;159:553-9.
20. Burton A, Altman DG. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. *Br J Cancer* 2004;91:4-8.
21. Mackinnon A. The use and reporting of multiple imputation in medical research – a review. *J Intern Med* 2010;268:586-93.
22. Schwarzer G, Vach W, Schumacher M. On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology. *Stat Med* 2000;19:541-61.
23 George SL. Statistics in medical journals: a survey of current policies and proposals for editors. Med Pediatr Oncol. 1985;13:109-12.
24. Emerson JD, Colditz GA. The statistical content of published medical research: some implications for biomedical education. Med Educ. 1985:19(3);248-255. DOI: 10.1111/j.1365-2923.1985.tb01315.x
25. Golden J, Zhu W, Sayre JW. A review of the statistical analysis used in papers published in Clinical Radiology and British Journal of Radiology. Clin Radiol. 1996;51(1):47- 50.
26. Lee CM, Soin HK, Einarson TR. Statistics in the Pharmacy Literature. Ann Pharmacother. 2004; 38(9):1412-1418. DOI 10.1345/aph.1D493
27. Lang T, Secic M. How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers, Second edition. Philadelphia: American College of Physicians, 2006.
28. O'Fallon JR, Duby SD, Salsburg DS, et al. Should there be statistical guidelines for medical research papers? Biometrics, 1978;34:687-95.
29. Shott S. Statistics in veterinary research. J Am Vet Med Assoc. 1985;187:138-41.
30. Hayden GF. Biostatistical trends in Pediatrics: implications for the future. Pediatrics. 1983;72:84-7.
31. Altman DG, Bland JM. Improving doctors' understanding of statistics. J R Statist Soc A. 1991;154:223-67.
32. Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. BMJ. 1983; 286:1489-93.
33. Bailar JC 3rd, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. Ann Intern Med. 1988 108(2):266-73.
34. Bond GR, Mintz J, McHugo GJ. Statistical guidelines for the Archives of PM&R. *Arch Phys Med Rehabil* 1995;76:784-7.
35. Wilkinson L and the Task Force on Statistical Inference. Statistical methods in psychology journals. Guidelines and explanations. *Am Psychologist* 1999;54:594-604.
36. Curran-Everett D, Benos DJ; American Physiological Society. Guidelines for reporting statistics in journals published by the American Physiological Society. *Am J Physiol Endocrinol Metab* 2004;287:E189-91. (plus other journals)
37. Curran-Everett D, Benos DJ. Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel. *Adv Physiol Educ* 2007;31:295-8.
38. Moher D, Schulz K, Altman DG, for the CONSORT Group. CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Ann Intern Med. 2001;134:657-62.
39. Des Jarlais DC, Lyles C, Crepaz N, Trend Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health. 2004; 94(3):361-6. PMID: 14998794
40. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Ann Intern Med. 2007; 147(8):573-577. PMID: 17938396
41. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication, 2011. www.icmje.org. Accessed December 12, 2012.
42. Schriger DL, Arora S, Altman DG. The content of medical journal instructions for authors. *Ann Emerg Med* 2006;48:743-749, 749.e1-4.
43. Lang T. How to Write, Publish, and Present in the Health Sciences: A Guide for Clinicians and Laboratory Researchers. Philadelphia: American College of Physicians, 2010.

International Journal for Quality in Health Care; **Volume 19, Number 6:** pp. 349–357 10.1093/intqhc/mzm042
Advance Access Publication: 14 September 2007

Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups

ALLISON TONG[1,2], PETER SAINSBURY[1,3] AND JONATHAN CRAIG[1,2]

[1]School of Public Health, University of Sydney, NSW 2006, Australia, [2]Centre for Kidney Research, The Children's Hospital at Westmead, NSW 2145, Australia, and [3]Population Health, Sydney South West Area Health Service, NSW 2170, Australia

Abstract

Background. Qualitative research explores complex phenomena encountered by clinicians, health care providers, policy makers and consumers. Although partial checklists are available, no consolidated reporting framework exists for any type of qualitative design.

Objective. To develop a checklist for explicit and comprehensive reporting of qualitative studies (indepth interviews and focus groups).

Methods. We performed a comprehensive search in Cochrane and Campbell Protocols, Medline, CINAHL, systematic reviews of qualitative studies, author or reviewer guidelines of major medical journals and reference lists of relevant publications for existing checklists used to assess qualitative studies. Seventy-six items from 22 checklists were compiled into a comprehensive list. All items were grouped into three domains: (i) research team and reflexivity, (ii) study design and (iii) data analysis and reporting. Duplicate items and those that were ambiguous, too broadly defined and impractical to assess were removed.

Results. Items most frequently included in the checklists related to sampling method, setting for data collection, method of data collection, respondent validation of findings, method of recording data, description of the derivation of themes and inclusion of supporting quotations. We grouped all items into three domains: (i) research team and reflexivity, (ii) study design and (iii) data analysis and reporting.

Conclusions. The criteria included in COREQ, a 32-item checklist, can help researchers to report important aspects of the research team, study methods, context of the study, findings, analysis and interpretations.

Keywords: focus groups, interviews, qualitative research, research design

Qualitative research explores complex phenomena encountered by clinicians, health care providers, policy makers and consumers in health care. Poorly designed studies and inadequate reporting can lead to inappropriate application of qualitative research in decision-making, health care, health policy and future research.

Formal reporting guidelines have been developed for randomized controlled trials (CONSORT) [1], diagnostic test studies (STARD), meta-analysis of RCTs (QUOROM) [2], observational studies (STROBE) [3] and meta-analyses of observational studies (MOOSE) [4]. These aim to improve the quality of reporting these study types and allow readers to better understand the design, conduct, analysis and findings of published studies. This process allows users of published research to be more fuller informed when they critically appraise studies relevant to each checklist and decide upon applicability of research findings to their local settings. Empiric studies have shown that the use of the CONSORT statement is associated with improvements in the quality of reports of randomized controlled trials [5]. Systematic reviews of qualitative research almost always show that key aspects of study design are not reported, and so there is a clear need for a CONSORT-equivalent for qualitative research [6].

The Uniform Requirements for Manuscripts Submitted to Biomedical Journals published by the International Committee of Medical Journal Editors (ICMJE) do not provide reporting guidelines for qualitative studies. Of all the mainstream biomedical journals (Fig. 1), only the British Medical Journal (BMJ) has criteria for reviewing qualitative research. However, the guidelines for authors specifically record that the checklist is not routinely used. In addition, the checklist is not comprehensive and does not provide specific guidance to assess some of the criteria. Although checklists for critical appraisal are available for qualitative research, there is no widely endorsed reporting framework for any type of qualitative research [7].

We have developed a formal reporting checklist for in-depth interviews and focus groups, the most common methods for data collection in qualitative health research.

Address reprint requests to: Allison Tong, Centre for Kidney Research, The Children's Hospital at Westmead, NSW 2145, Australia. Tel: +61-2-9845-1482; Fax: +61-2-9845-1491; E-mail: allisont@health.usyd.edu.au, allisont@chw.edu.au

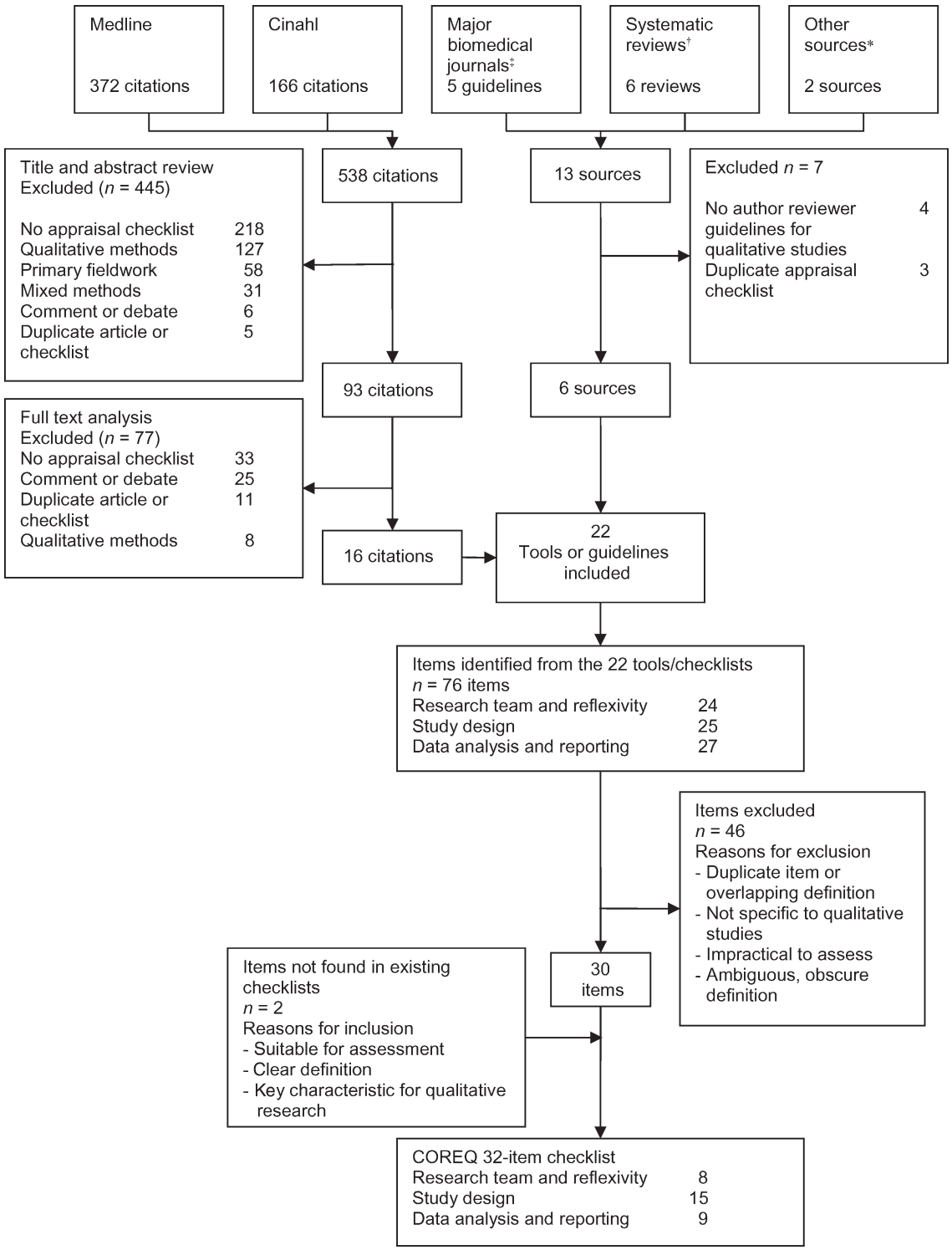


Figure 1 Development of the COREQ Checklist. *References [26, 27], †References [6, 28–32], ‡Author and reviewer guidelines provided by BMJ, JAMA, Lancet, Annals of Internal Medicine, NEJM.

These two methods are particularly useful for eliciting patient and consumer priorities and needs to improve the quality of health care [8]. The checklist aims to promote complete and transparent reporting among researchers and indirectly improve the rigor, comprehensiveness and credibility of interview and focus-group studies.

Basic definitions

Qualitative studies use non-quantitative methods to contribute new knowledge and to provide new perspectives in health care. Although qualitative research encompasses a broad range of study methods, most qualitative research

publications in health care describe the use of interviews and focus groups [8].

Interviews

In-depth and semi-structured interviews explore the experiences of participants and the meanings they attribute to them. Researchers encourage participants to talk about issues pertinent to the research question by asking open-ended questions, usually in one-to-one interviews. The interviewer might re-word, re-order or clarify the questions to further investigate topics introduced by the respondent. In qualitative health research, in-depth interviews are often used to study the experiences and meanings of disease, and to explore personal and sensitive themes. They can also help to identify potentially modifiable factors for improving health care [9].

Focus groups

Focus groups are semi-structured discussions with groups of 4–12 people that aim to explore a specific set of issues [10]. Moderators often commence the focus group by asking broad questions about the topic of interest, before asking the focal questions. Although participants individually answer the facilitator's questions, they are encouraged to talk and interact with each other [11]. This technique is built on the notion that the group interaction encourages respondents to explore and clarify individual and shared perspectives [12]. Focus groups are used to explore views on health issues, programs, interventions and research.

Methods

Development of a checklist

Search strategy. We performed a comprehensive search for published checklists used to assess or review qualitative studies, and guidelines for reporting qualitative studies in: Medline (1966—Week 1 April 2006), CINAHL (1982—Week 3 April 2006), Cochrane and Campbell protocols, systematic reviews of qualitative studies, author or reviewer guidelines of major medical journals and reference lists of relevant publications. We identified the terms used to index the relevant articles already in our possession and performed a broad search using those search terms. The electronic databases were searched using terms and text words for research (standards), health services research (standards) and qualitative studies (evaluation). Duplicate checklists and detailed instructions for conducting and analysing qualitative studies were excluded.

Data extraction. From each of the included publications, we extracted all criteria for assessing or reporting qualitative studies. Seventy-six items from 22 checklists were compiled into a comprehensive list. We recorded the frequency of each item across all the publications. Items most frequently included in the checklists related to sampling method, setting for data collection, method of data collection, respondent validation of findings, method of recording data, description of the derivation of themes and inclusion of supporting quotations. We grouped all items into three domains: (i) research team and reflexivity, (ii) study design and (iii) data analysis and reporting. (see Tables 2–4)

Within each domain we simplified all relevant items by removing duplicates and those that were ambiguous, too broadly defined, not specific to qualitative research, or impractical to assess. Where necessary, the remaining items were rephrased for clarity. Based upon consensus among the authors, two new items that were considered relevant for reporting qualitative research were added. The two new items were identifying the authors who conducted the interview or focus group and reporting the presence of non-participants during the interview or focus group. The COREQ checklist for explicit and comprehensive reporting of qualitative studies consists of 32 criteria, with a descriptor to supplement each item (Table 1).

COREQ: content and rationale (see Tables 1)

Domain 1: research team and reflexivity

(i) Personal characteristics: Qualitative researchers closely engage with the research process and participants and are therefore unable to completely avoid personal bias. Instead researchers should recognize and clarify for readers their identity, credentials, occupation, gender, experience and training. Subsequently this improves the credibility of the findings by giving readers the ability to assess how these factors might have influenced the researchers' observations and interpretations [13–15].

(ii) Relationship with participants: The relationship and extent of interaction between the researcher and their participants should be described as it can have an effect on the participants' responses and also on the researchers' understanding of the phenomena [16]. For example, a clinician–researcher may have a deep understanding of patients' issues but their involvement in patient care may inhibit frank discussion with patient–participants when patients believe that their responses will affect their treatment. For transparency, the investigator should identify and state their assumptions and personal interests in the research topic.

Domain 2: study design

(i) Theoretical framework: Researchers should clarify the theoretical frameworks underpinning their study so readers can understand how the researchers explored their research questions and aims. Theoretical frameworks in qualitative research include: grounded theory, to build theories from the data; ethnography, to understand the culture of groups with shared characteristics; phenomenology, to describe the meaning and significance of experiences; discourse analysis, to analyse linguistic expression; and content analysis, to systematically organize data into a structured format [10].

Table 1 Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist

No	Item	Guide questions/description
Domain 1: Research team and reflexivity		
Personal Characteristics		
1.	Interviewer/facilitator	Which author/s conducted the interview or focus group?
2.	Credentials	What were the researcher's credentials? *E.g. PhD, MD*
3.	Occupation	What was their occupation at the time of the study?
4.	Gender	Was the researcher male or female?
5.	Experience and training	What experience or training did the researcher have?
Relationship with participants		
6.	Relationship established	Was a relationship established prior to study commencement?
7.	Participant knowledge of the interviewer	What did the participants know about the researcher? e.g. *personal goals, reasons for doing the research*
8.	Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. *Bias, assumptions, reasons and interests in the research topic*
Domain 2: study design		
Theoretical framework		
9.	Methodological orientation and Theory	What methodological orientation was stated to underpin the study? *e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis*
Participant selection		
10.	Sampling	How were participants selected? *e.g. purposive, convenience, consecutive, snowball*
11.	Method of approach	How were participants approached? e.g. *face-to-face, telephone, mail, email*
12.	Sample size	How many participants were in the study?
13.	Non-participation	How many people refused to participate or dropped out? Reasons?
Setting		
14.	Setting of data collection	Where was the data collected? e.g. *home, clinic, workplace*
15.	Presence of non-participants	Was anyone else present besides the participants and researchers?
16.	Description of sample	What are the important characteristics of the sample? *e.g. demographic data, date*
Data collection		
17.	Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?
18.	Repeat interviews	Were repeat interviews carried out? If yes, how many?
19.	Audio/visual recording	Did the research use audio or visual recording to collect the data?
20.	Field notes	Were field notes made during and/or after the interview or focus group?
21.	Duration	What was the duration of the interviews or focus group?
22.	Data saturation	Was data saturation discussed?
23.	Transcripts returned	Were transcripts returned to participants for comment and/or correction?
Domain 3: analysis and findingsz		
Data analysis		
24.	Number of data coders	How many data coders coded the data?
25.	Description of the coding tree	Did authors provide a description of the coding tree?
26.	Derivation of themes	Were themes identified in advance or derived from the data?
27.	Software	What software, if applicable, was used to manage the data?
28.	Participant checking	Did participants provide feedback on the findings?
Reporting		
29.	Quotations presented	Were participant quotations presented to illustrate the themes / findings? Was each quotation identified? e.g. *participant number*
30.	Data and findings consistent	Was there consistency between the data presented and the findings?
31.	Clarity of major themes	Were major themes clearly presented in the findings?
32.	Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?

(ii) Participant selection: Researchers should report how participants were selected. Usually purposive sampling is used which involves selecting participants who share particular characteristics and have the potential to provide rich, relevant and diverse data pertinent to the research question [13, 17]. Convenience sampling is less optimal because it may fail to capture important perspectives from difficult-to-reach people [16]. Rigorous attempts to recruit participants and reasons for non-participation should be stated to reduce the likelihood of making unsupported statements [18].

Table 2 Items included in 22 published checklists: Research team and reflexivity domain

Item	References																						
	[26][a]	[27][a]	[6][b]	[28][b]	[32][b]	[13]	[15]	[14]	[17]	[33]	[34]	[35]	[16]	[19]	[36]	[7]	[37]	[23]	[38]	[39]	[22]	BMJ
Research team and reflexivity																						
Nature of relationship between the researcher and participants		•		•	•		•		•						•				•			
Examination of role, bias, influence	•	•			•	•	•	•							•						•	
Description of role		•		•					•	•				•	•					•	•	
Identity of the interviewer		•		•		•					•		•		•							
Continued and prolonged engagement		•				•							•	•					•	•		
Response to events	•	•				•	•	•														
Prior assumptions and experience		•						•									•			•		
Professional status		•					•								•							
Journal, record of personal experience		•								•				•								
Effects of research on researcher		•				•	•															
Qualifications		•													•							
Training of the interviewer/facilitator			•		•																	
Expertise demonstrated		•																	•			
Perception of research at inception								•						•								
Age							•															
Gender							•															
Social class							•															
Reasons for conducting study		•																				
Sufficient contact													•									
Too close to participants													•									
Empathy																	•					
Distance between researcher and participants							•															
Background								•														
Familiarity with setting																					•	

[a]Other publications, [b]Systematic review of qualitative studies; BMJ, British Medical Journal—editor's checklist for appraising qualitative research); •, item included in the checklist.

Table 3 Items included in 22 published checklists: Study design

Item	References																							
	[26][a]	[27][a]	[6][b]	[28] [b]	[32][b]	[13]	[15]	[14]	[17]	[33]	[34]	[35]	[16]	[19]	[36]	[7]	[37]	[23]	[38]	[39]	[22]	BMJ
Study design																						
Methodological orientation, ontological or epistemological basis		•		•				•	•					•					•	•	•	•
Sampling—convenience, purposive	•	•			•	•	•	•	•	•	•	•	•	•		•		•	•	•	•	•	•
Setting		•		•	•			•			•				•					•		
Characteristics and description of sample		•			•			•			•		•	•								
Reasons for participant selection	•	•				•		•			•											
Non-participation	•	•		•	•																	
Inclusion and exclusion, criteria		•			•	•													•			
Identity of the person responsible for recruitment				•	•						•				•							
Sample size		•		•	•						•										•	
Method of approach		•									•					•						
Description of explanation of research to participants	•				•										•							
Level and type of participation														•								
Method of data collection, e.g. focus group, in-depth interview	•	•	•	•	•	•		•	•		•	•	•	•		•				•	•	
Audio and visual recording	•	•	•	•	•	•				•	•		•						•		•	•
Transcripts			•	•	•	•			•		•		•						•			•
Setting and location	•	•		•	•		•		•		•				•						•	•
Saturation of data	•	•	•			•			•				•	•							•	
Use of a topic guide, tools, questions	•	•	•								•					•			•	•		
Field notes			•	•	•	•													•			•
Changes and modifications	•	•		•	•														•		•	
Duration of interview, focus group		•				•					•									•		
Sensitive to participant language and views		•										•		•								
Number of interviews, focus groups		•				•																
Time span																					•	
Time and resources available to the study		•																				

[a]Other publications, [b]Systematic review of qualitative studies; BMJ, British Medical Journal—editor's checklist for appraising qualitative research; •, item included in the checklist.

Table 4 Items included in 22 published checklists: Analysis and reporting

Item	References																							
	[26][a]	[27][a]	[6][b]	[28][b]	[32][b]	[13]	[15]	[14]	[17]	[33]	[34]	[35]	[16]	[19]	[36]	[7]	[37]	[23]	[38]	[39]	[22]	BMJ
Respondent validation	•	•	•		•		•		•	•			•	•			•	•	•	•		
Limitations and generalizability	•	•		•	•		•		•		•		•	•				•	•			
Triangulation	•	•		•	•	•	•	•	•					•			•		•			
Original data, quotation		•	•	•	•			•	•		•			•		•				•	•	•
Derivation of themes explicit	•	•	•	•	•		•	•			•								•			•
Contradictory, diverse, negative cases	•	•		•	•		•			•				•					•			•
Number of data analysts	•	•	•			•			•			•	•						•			•
In-depth description of analysis	•			•	•			•			•			•							•	•
Sufficient supporting data presented	•	•		•	•		•				•					•						
Data, interpretation and conclusions linked and integrated		•		•	•							•		•						•		
Retain context of data		•					•	•						•					•			
Explicit findings, presented clearly	•	•		•					•	•												
Outside checks													•	•				•	•			
Software used		•				•													•			•
Discussion both for and against the researchers' arguments	•	•		•	•																	
Development of theories, explanations		•					•			•		•										
Numerical data		•									•							•				•
Coding tree or coding system		•					•												•		•	
Inter-observer reliability		•									•										•	
Sufficient insight into meaning/perceptions of participants		•								•												
Reasons for selection of data to support findings		•			•																	
New insight		•						•														
Results interpreted in credible, innovative way									•													
Eliminate other theories													•									
Range of views														•								
Distinguish between researcher and participant voices								•														
Proportion of data taken into account														•								

[a]Other publications, [b]Systematic review of qualitative studies; BMJ, British Medical Journal—editor's checklist for appraising qualitative research, •, item included in the checklist.

Researchers should report the sample size of their study to enable readers to assess the diversity of perspectives included.

(iii) Setting: Researchers should describe the context in which the data were collected because it illuminates why participants responded in a particular way. For instance, participants might be more reserved and feel disempowered talking in a hospital setting. The presence of non-participants during interviews or focus groups should be reported as this can also affect the opinions expressed by participants. For example, parent interviewees might be reluctant to talk on sensitive topics if their children are present. Participant characteristics, such as basic demographic data, should be reported so readers can consider the relevance of the findings and interpretations to their own situation. This also allows readers to assess whether perspectives from different groups were explored and compared, such as patients and health care providers [13, 19].

(iv) Data collection: The questions and prompts used in data collection should be provided to enhance the readers' understanding of the researcher's focus and to give readers the ability to assess whether participants were encouraged to openly convey their viewpoints. Researchers should also report whether repeat interviews were conducted as this can influence the rapport developed between the researcher and participants and affect the richness of data obtained. The method of recording the participants' words should be reported. Generally, audio recording and transcription more accurately reflect the participants' views than contemporaneous researcher notes, more so if participants checked their own transcript for accuracy [19–21]. Reasons for not audio recording should be provided. In addition, field notes maintain contextual details and non-verbal expressions for data analysis and interpretation [19, 22]. Duration of the interview or focus group should be reported as this affects the amount of data obtained. Researchers should also clarify whether participants were recruited until no new relevant knowledge was being obtained from new participants (data saturation) [23, 24].

Domain 3: analysis and findings

(i) Data analysis: Specifying the use of multiple coders or other methods of researcher triangulation can indicate a broader and more complex understanding of the phenomenon. The credibility of the findings can be assessed if the process of coding (selecting significant sections from participant statements), and the derivation and identification of themes are made explicit. Descriptions of coding and memoing demonstrate how the researchers perceived, examined and developed their understanding of the data [17, 19]. Researchers sometimes use software packages to assist with storage, searching and coding of qualitative data. In addition, obtaining feedback from participants on the research findings adds validity to the researcher's interpretations by ensuring that the participants' own meanings and perspectives are represented and not curtailed by the researchers' own agenda and knowledge [23].

(ii) Reporting: If supporting quotations are provided, researchers should include quotations from different participants to add transparency and trustworthiness to their findings and interpretations of the data [17]. Readers should be able to assess the consistency between the data presented and the study findings, including the both major and minor themes. Summary findings, interpretations and theories generated should be clearly presented in qualitative research publications.

Discussion

The COREQ checklist was developed to promote explicit and comprehensive reporting of qualitative studies (interviews and focus groups). The checklist consists of items specific to reporting qualitative studies and precludes generic criteria that are applicable to all types of research reports. COREQ is a comprehensive checklist that covers necessary components of study design, which should be reported. The criteria included in the checklist can help researchers to report important aspects of the research team, study methods, context of the study, findings, analysis and interpretations.

At present, we acknowledge there is no empiric basis that shows that the introduction of COREQ will improve the quality of reporting of qualitative research. However this is no different than when CONSORT, QUOROM and other reporting checklists were introduced. Subsequent research has shown that these checklists have improved the quality of reporting of study types relevant to each checklist [5, 25], and we believe that the effect of COREQ is likely to be similar. Despite differences in the objectives and methods of quantitative and qualitative methods, the underlying aim of transparency in research methods and, at the least, the theoretical possibility of the reader being able to duplicate the study methods should be the aims of both methodological approaches. There is a perception among research funding agencies, clinicians and policy makers, that qualitative research is 'second class' research. Initiatives like COREQ are designed to encourage improvement in the quality of reporting of qualitative studies, which will indirectly lead to improved conduct, and greater recognition of qualitative research as inherently equal scientific endeavor compared with quantitative research that is used to assess the quality and safety of health care. We invite readers to comment on COREQ to improve the checklist.

References

1. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;**285**:1987–91.
2. Moher D, Cook DJ, Eastwood S *et al.* Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;**354**:1896–900.

3. STROBE Statement: Strengthening the reporting of observational studies in epidemiology. http://www.strobe-statement.org/Checkliste.html
4. Stroup DF, Berlin JA, Morton SC *et al*. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;**283**:2008–12.
5. Moher D, Jones A, Lepage L. Use of the CONSORT Statement and quality of reports of randomized trials. A comparative before-and-after evaluation. *JAMA* 2001;**285**:1992–5.
6. Mills E, Jadad AR, Ross C *et al*. Systematic review of qualitative studies exploring parental beliefs and attitudes toward childhood vaccination identified common barriers to vaccination. *J Clin Epidemiol* 2005;**58**:1081–8.
7. Knafl KA, Howard NJ. Interpreting and reporting qualitative research. *Res Nurs Health* 1984;**7**:7–14.
8. Sofaer S. Qualitative research methods. *Int J Qual Health Care* 2002;**14**:329–36.
9. Wright EB, Holcombe C, Salmon P. Doctor's communication of trust, care, and respect in breast cancer: qualitative study. *BMJ* 2004;**328**:864–8.
10. Liamputtong P, Ezzy D. *Qualitative Research Methods*. 2nd edn. Melbourne, Victoria: Oxford University Press, 2005.
11. Krueger RA, Casey MA. *Focus Groups. A Practical Guide for Applied Research*. Thousand Oaks CA: Sage Publications, 2000.
12. Morgan DL. *Focus Groups as Qualitative Research*. Newbury Park, California: Sage, 1988.
13. Giacomini MK, Cook DJ. Users' guides to the medical literature XXIII. Qualitative research in health care. A. Are the results of the study valid? *JAMA* 2000;**284**:357–62.
14. Malterud K. Qualitative research:standards challenges guidelines. *Lancet* 2001;**358**:483–8.
15. Mays N, Pope C. Qualitative research in health care: assessing quality in qualitative research. *BMJ* 2000;**320**:50–2.
16. Elder NC, William L. Reading and evaluating qualitative research studies. *J Fam Pract* 1995;**41**:279–85.
17. Cote L, Turgeon J. Appraising qualitative research articles in medicine and medical education. *Med Teach* 2005;**27**:71–5.
18. Altheide D, Johnson J. Criteria for assessing interpretive validity in qualitative research. In Denzin N, Lincoln Y (eds). *Handbook of Qualitative Research*. Thousand Oaks CA: Sage Publications, 1994.
19. Fossey E, Harvey C, McDermott F *et al*. Understanding and evaluating qualitative research. *Aust N Z J Psychiatry* 2002;**36**:717–32.
20. Seale C, Silverman S. Ensuring rigour in qualitative research. *Eur J Public Health* 1997;**7**:379–84.
21. Scheff T. Single case analysis in the health sciences. *Eur J Public Health* 1995;**5**:72–4.
22. Bluff R. Evaluating qualitative research. *Br J Midwifery* 1997;**5**:232–5.
23. Popay J, Rogers A, Williams G. Rationale and standards for the systematic review of qualitative literature in health services research. *Qual Health Res* 1998;**8**:341–51.
24. Blumer H. *Critiques of Research, in the Social Sciences*. New Brunswick, NJ: Transaction Books, 1979.
25. Delaney A, Bagshaw SM, Ferland A *et al*. A systematic evaluation of the quality of meta-anlyses in the critical care literature. *Crit Care* 2005;**9**:575–82.
26. Critical Skills Appraisal Programme (CASP) 10 Questions to help you make sense of qualitative research: Milton Keynes Primary Care Trust, 2002.
27. Spencer L, Ritchie J, Lewis J *et al*. *Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence*. London: Cabinet Office. Government Chief Social Researcher's Office, 2003.
28. Campbell R, Pound P, Pope C *et al*. Evaluating meta-ethnography: a synthesis of qualitative research on lay experience of diabetes and diabetes care. *Soc Sci Med* 2003;**56**:671–84.
29. Feder GS, Hutson M, Ramsay I *et al*. Women exposed to intimate partner violence: expectations and experiences when they encounter health care professionals: a meta-analysis of qualitative studies. *Arch Intern Med* 2006;**166**:22–37.
30. Pound P, Britten N, Morgan M *et al*. Resisting medicines: a synthesis of qualitative studies of medicine taking. *Soc Sci Med* 2005;**61**:133–55.
31. Smith LK, Pope C, Botha JL. Patients' help-seeking experiences and delay in cancer presentation: a qualitative synthesis. *Lancet* 2005;**366**:825–31.
32. Walter FM, Emery J, Braithwaite D *et al*. Lay understanding of familial risk of common chronic diseases: a systematic review and synthesis of qualitative research. *Ann Fam Med* 2004;**2**:583–94.
33. Inui TS, Frankel RM. Evaluating the quality of qualitative research: a proposal pro-term. *J Gen Intern Med* 1991;**6**:485–6.
34. Boulton M, Fitzpatrick R, Swinburn C. Qualitative research in health care: II A structured review and evaluation of studies. *J Eval Clin Pract* 1996;**2**:171–9.
35. Dixon-Woods M, Shaw RL, Agarwal S *et al*. The problem of appraising qualitative research. *Qual Saf Health Care* 2004;**13**:223–5.
36. Hoddinott P, Pill R. A review of recently published qualitative research in general practice. More methodological questions than answers? *Fam Pract* 1997;**14**:313–9.
37. Kuzel AJ, Engel JD, Addison RB *et al*. Desirable features of qualitative research. *Fam Pract Res J* 1994;**14**:369–78.
38. Treloar C, Champness S, Simpson PL *et al*. Critical appraisal checklist for qualitative research studies. *Indian J Pediatr* 2000;**67**:347–51.
39. Cesario S, Morin K, Santa-Donato A. Evaluating the level of evidence in qualitative research. *J Obstet Gynecol Neonatal Nurs* 2001;**31**:708–14.

Accepted for publication 7 July 2007

JOURNAL OF MEDICAL CASE REPORTS

RESEARCH ARTICLE Open Access

The CARE guidelines: consensus-based clinical case reporting guideline development

Joel J Gagnier[1,2], Gunver Kienle[3], Douglas G Altman[4], David Moher[5,6], Harold Sox[7], David Riley[8*] and the CARE Group[9]

Abstract

Background: A case report is a narrative that describes, for medical, scientific, or educational purposes, a medical problem experienced by one or more patients. Case reports written without guidance from reporting standards are insufficiently rigorous to guide clinical practice or to inform clinical study design.
Primary Objective. Develop, disseminate, and implement systematic reporting guidelines for case reports.

Methods: We used a three-phase consensus process consisting of (1) pre-meeting literature review and interviews to generate items for the reporting guidelines, (2) a face-to-face consensus meeting to draft the reporting guidelines, and (3) post-meeting feedback, review, and pilot testing, followed by finalization of the case report guidelines.

Results: This consensus process involved 27 participants and resulted in a 13-item checklist—a reporting guideline for case reports. The primary items of the checklist are title, key words, abstract, introduction, patient information, clinical findings, timeline, diagnostic assessment, therapeutic interventions, follow-up and outcomes, discussion, patient perspective, and informed consent.

Conclusions: We believe the implementation of the CARE (CAse REport) guidelines by medical journals will improve the completeness and transparency of published case reports and that the systematic aggregation of information from case reports will inform clinical study design, provide early signals of effectiveness and harms, and improve healthcare delivery.

Keywords: Case report, Case study, EQUATOR Network, Patient reports, Meaningful use, Health research reporting guidelines

Introduction

A case report is a detailed narrative that describes, for medical, scientific, or educational purposes, a medical problem experienced by one or several patients

Case reports present clinical observations customarily collected in healthcare delivery settings. They have proved helpful in the identification of adverse and beneficial effects, the recognition of new diseases, unusual forms of common diseases, and the presentation of rare diseases [1]. For example, our understanding of the relationship between thalidomide and congenital abnormalities [2] and the use of propranolol for the treatment of infantile hemangiomas began with case reports [3]. Case reports may generate hypotheses for future clinical studies, prove useful in the evaluation of global convergences of systems-oriented approaches, and guide the individualization and personalization of treatments in clinical practice [4,5]. Furthermore, case reports offer a structure for case-based learning in healthcare education and may facilitate the comparison of healthcare education and delivery across cultures.

Case reports are common and account for a growing number of articles in medical journals [6]; however their quality is uneven [7,8]. For example, one study evaluated 1316 case reports from four peer-reviewed emergency-medicine journals and found that more than half failed to provide information related to the primary treatment that would have increased transparency and replication [9]. Written without the benefit of reporting guidelines,

* Correspondence: driley@gahmllc.com
[8]Global Advances in Health and Medicine, Portland, OR, USA
Full list of author information is available at the end of the article

case reports often are insufficiently rigorous to be aggregated for data analysis, inform research design, or guide clinical practice [7,9].

Reporting guidelines exist for a variety of study designs including randomized controlled trials (Consolidated Standards of Reporting Trials, or CONSORT) [10], observational studies (Strengthening the Reporting of Observational studies in Epidemiology, or STROBE) [11], and systematic reviews and meta-analyses (Preferred Reporting Items for Systematic Reviews and Meta-Analyses, or PRISMA) [12]. Empirical evidence suggests that a journal's adoption of the CONSORT statement as a guide to authors is associated with an increase in the completeness of published randomized trials [13]. Guidelines have been developed for adverse-event case reports [14]; however, general reporting guidelines for case reports do not exist. Our primary objective was to develop reporting guidelines for case reports through a consensus-based process.

Methods

Research design

We followed the Guidance for Developers of Health Research Reporting Guidelines [15] and developed a three phase consensus process [16]. This consisted of (1) a pre-meeting literature review followed by interviews to generate items for a case report checklist, (2) a face-to-face consensus meeting for drafting a reporting guideline, and (3) post-meeting feedback and pilot testing followed by finalization of the case report guidelines.

Participants

We contacted 28 individuals who fulfilled at least one of four criteria [17-19]: (1) publication of articles related to case reports; (2) publication of a manual, handbook, or method guidelines related to case reports; (3) publication of a systematic review of methods or reporting related to case reports; and (4) publication of other reporting guidelines for clinical research.

Consensus process

Phase 1: Four of the authors, the steering committee (JG, GK, DM, and DR), searched the literature for publications on the role of case reports, recommendations for their publication, and surveys on reporting quality. A letter was sent to 28 potential participants explaining the purpose of the meeting, details of the consensus technique, and requesting their participation in generating specific recommendations for case reporting. Twenty-seven people agreed to participate and were scheduled for a telephone interview and sent a selection of key articles on case reports. During the telephone interview, participants were asked (1) what information was required to be included in case-reporting guidelines, (2) the rationale for their suggestions, and (3) for references that supported their reasoning.

Three of the authors (JG, GK, and DR) grouped the recommendations from the literature search and interviews by theme together with their rationale, references, and operational definitions. No quantitative scoring was done.

Phase 2: The face-to-face consensus meeting at the University of Michigan in Ann Arbor (October 2012) included 18 participants from Phase 1, one research assistant and two student observers. The meeting began with a review of the blinded recommendations elicited during the Phase 1 interviews, in whole group and small group sessions. On the second day, open discussion of each potential item continued, during which clarifications, opinions, justifications, operational definitions, and new ideas were expressed. By the end of the second day, the group had agreed upon a set of preliminary reporting recommendations.

Phase 3: The draft checklist was refined by the steering committee and sent for two rounds of review to the complete group (Phase 1 and 2 participants). The finalized reporting guidelines incorporated the feedback from the entire CARE group.

Results

The CAse REport (CARE) guidelines checklist is structured to correspond with key components of a case report and capture useful clinical information (including 'meaningful use' information mandated by some insurance plans).

The checklist begins with a statement that describes the narrative of a case report. The meeting CARE group felt that a case report should tell a story using prose that has a consistent style across all sections, including the rationale for any conclusions and take-away messages.

We recommend a timeline (item 7) in the form of a table or figure that gives the specific dates and times of important components of the case. This might include family and past medical history, genetic information, current symptoms, diagnostic test results, interventions, and events that occurred during follow-up. The timeline should show how the key events of the case unfolded.

We created separate checklist items for diagnostic assessments (item 8) and therapeutic interventions (item 9) with the recognition that both items will often be relevant in a case report.

The group discussed at length whether to include the patient's perspective on his or her experience. In the end, we advocated for patient-reported outcomes and experiences whenever possible (item 12). There was also discussion about the need for guidelines for patient-reported outcomes of their care. In a similar vein, a recent extension of the CONSORT statement was published for patient-reported outcomes in randomized trials; CONSORT-PRO [20].

Finally, we included an item on informed consent (item 13). We believe that authors have an ethical duty to obtain informed consent from the patient to publish patient information in a case report. Consent becomes informed when the patient or a relative reads the case report and approves its contents. If the patient cannot give consent and attempts to find a relative to give proxy consent have failed, the authors should seek permission to publish from an institutional committee. There may be other circumstances where an ethics committee or Institutional Review Board (IRB) approval may be necessary. The CARE guidelines are shown in the following Table 1.

Table 1 The CARE guidelines checklist

The Narrative: A case report tells a story in a narrative format that includes the presenting concerns, clinical findings, diagnoses, interventions, outcomes (including adverse events), and follow-up. The narrative should include a discussion of the rationale for any conclusions and any take-away messages.

Item name	Item no.	Brief description
Title	1	The words "case report" (or "case study") should appear in the title along with phenomenon of greatest interest (eg, symptom, diagnosis, test, intervention)
Keywords	2	The key elements of this case in 2-5 words
Abstract	3	a) Introduction—What does this case add?
		b) Case Presentation:
		- The main symptoms of the patient
		- The main clinical findings
		- The main diagnoses and interventions
		- The main outcomes
		c) Conclusion—What were the main "take-away" lessons from this case?
Introduction	4	Brief background summary of this case referencing the relevant medical literature
Patient Information	5	a) Demographic information (eg, age, gender, ethnicity, occupation)
		b) Main symptoms of the patient (his or her chief complaints)
		c) Medical, family, and psychosocial history—including diet, lifestyle, and genetic information whenever possible, and details about relevant comorbidities including past interventions and their outcomes
Clinical Findings	6	Describe the relevant physical examination (PE) findings
Timeline	7	Depict important dates and times in this case (table or figure)
Diagnostic Assessment	8	a) Diagnostic methods (eg, PE, laboratory testing, imaging, questionnaires)
		b) Diagnostic challenges (eg, financial, language/cultural)
		c) Diagnostic reasoning including other diagnoses considered
		d) Prognostic characteristics (eg, staging) where applicable
Therapeutic Intervention	9	a) Types of intervention (eg, pharmacologic, surgical, preventive, self-care)
		- Administration of intervention (eg, dosage, strength, duration)
		- Changes in intervention (with rationale)
Follow-up and Outcomes	10	a) Summarize the clinical course of all follow-up visits including
		- Clinician and patient-assessed outcomes
		- Important follow-up test results (positive or negative)
		- Intervention adherence and tolerability (and how this was assessed)
		- Adverse and unanticipated events
Discussion	11	a) The strengths and limitations of the management of this case
		b) The relevant medical literature
		c) The rationale for conclusions (including assessments of cause and effect)
		d) The main "take-away" lessons of this case report
Patient Perspective	12	The patient should share his or her perspective or experience whenever possible
Informed Consent	13	Did the patient give informed consent? Please provide if requested

Discussion

This 13-item checklist provides a framework to satisfy the need for completeness and transparency for published case reports. We attempted to strike a balance between adequate detail and the concise writing that is one of the appealing characteristics of a case report. Our consensus process resulted in a set of essential items for authors to consider when submitting a case report for publication.

While case reports have long been an important source of new ideas and information in medicine [21], it appears that case reports are likely to begin to play a role in the discovery of what works and for whom. BioMed Central launched the *Journal of Medical Case Reports* in 2007 [22] and a Cases Database in 2012 with more than 11,000 published case reports from 50 medical journals. In 6 months, it has grown to more than 26,000 case reports from 212 medical journals [23]. The CARE guidelines checklist is part of a growing effort to improve the reporting of case reports.

There is substantial empirical evidence that reporting guidelines improve the completeness of published scientific reports eg, see references [13,24,25]. A recent Cochrane review examining the influence of journal endorsement of the CONSORT statement on reporting included 53 publications assessing 16,604 randomized controlled trials and found that CONSORT-endorsing journals consistently have better overall reporting [13]. However, the potential impact of the CONSORT statement and related reporting guidelines has not been fully realized. A study examining the instructions to peer reviewers of 116 health research journals found that only 41 (35%) provided online instructions to peer reviewers. Of those, only 19 (46%) mentioned or referred to reporting guidelines as a useful resource [26]. In response, the authors provide several recommendations for editors to improve the peer review of submitted manuscripts, suggesting that journals have a responsibility to support peer reviewers [26].

The developers of reporting guidelines have a responsibility to plan a dissemination and implementation strategy that supports guidelines utilization [15]. Our efforts have several components:

1. The CARE guidelines will be presented at international conferences and workshops including the Peer Review and Biomedical Publication Congress in Chicago on September 10, 2013.
2. This article will be published simultaneously in multiple medical journals and outreach to the 212 journals depositing case reports into the BioMed Central Case Report Database.
3. We will develop a more detailed explanation and elaboration article to outline the rationale for each item and include empirical evidence and examples of good reporting from published case reports.
4. The CARE guidelines are being pilot tested, and preliminary results support the guidelines as currently written (personal communication with Helmut Kiene, Erika Oberg, Bill Manahan). Guidelines extensions for specialties are being developed.
5. The CARE guidelines and related documents will be available on a dedicated website (www.CARE-statement.org), the EQUATOR Network website (www.equator-network.org), and translated into multiple languages.
6. Authors, journal editors, peer reviewers and the wider medical community are encouraged to use the CARE checklist and provide feedback that can be incorporated into regular updates of the CARE guidelines.
7. We will conduct and support research into the impact of the CARE guidelines on the reporting of case reports.

Limitations

The CARE guidelines and their development have several possible limitations. First, these guidelines were developed through a consensus method and thus represent the opinions of the participants. However, consensus was easily reached during our meeting, we referred to the empirical evidence where available, and we received feedback from a wide selection of individuals, beyond those involved in our consensus meeting. Second, we recognize that causality determinations are a challenge for case reports even when following reporting guidelines [27,28]. The CARE guidelines emphasize information quality independent of causality assessments. Different specialties, practitioners, and patients are likely to require extensions of the CARE guidelines with specialty specific information. We welcome discussions with groups interested in using the CARE guidelines as the basis for their specific reporting needs.

Though not mentioned in our guidelines, medical journals often require authors to address three issues: (a) potential competing interests, (b) de-identification of patient-related data, and (c) ethics committee or Institutional Review Board (IRB) approval if obtained or necessary.

Conclusions

Anticipating a long future for case reports, we have provided guidance in the form of reporting standards for use by healthcare stakeholders around the world. The growth of case reports in an era in which clinical trials and systematic reviews dominate the tables of content of medical journals indicates that case reports have value, particularly with the increasing importance of individualized care.

Unlike randomized controlled trials, case reports are individual reports related to the care of individual patients where the sample size is one. When systematically collected and combined into larger datasets, they can be analyzed, enhancing the early discovery of effectiveness and harms.

We anticipate that the analysis of systematically aggregated information from patient encounters (now mandated by some insurance plans) will provide scalable, data-driven insights into what works for which patients in real time, facilitating comparisons across medical systems and cultures. Practitioners will soon be able to provide—and in some cases they are required to provide—patients with information from their encounters. This will transform how we think about "evidence" and revolutionize its creation, diffusion, and use—opening new opportunity landscapes. When it becomes clear how new data contributes to evidence, the stewardship needed to produce high-quality data will be more rewarding and our attitude toward "observation" will shift. The CARE guidelines provide a framework to satisfy the need for precision, completeness, and transparency.

Author contributions

JG, GK, DGA, DM, HS, and DR met the ICMJE criteria for authorship. JG and DR wrote the first draft of the article. DGA, JG, GK, DM, DR, and HS critically reviewed and edited drafts. The entire CARE group participated in parts or all of the guidelines development process and contributed to the editing and revision of the CARE guidelines and this article. All authors read and approved the final manuscript.

The CARE group

1. Alyshia Allaire, BA, Portland, OR, USA
2. Douglas G. Altman, DSc, Centre for Statistics in Medicine, University of Oxford, Oxford, UK
3. Jeffrey Aronson, MB, ChB, MA Dphil, FRCP, FB, PharmacolS, University of Oxford, Oxford, UK*
4. James Carpenter, MD, Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, MI, USA
5. Joel Gagnier, ND, MSc, PhD, Departments of Orthopaedic Surgery and Epidemiology, University of Michigan, Ann Arbor, MI, USA
6. Patrick Hanaway, MD, Director of Medical Education, Institute for Functional Medicine, Asheville, NC, USA*
7. Carolyn Hayes, PhD, RN, NEA-BC, Dana-Farber Brigham and Women's Cancer Center, Boston, MA, USA
8. David Jones, MD, President, Institute for Functional Medicine, Ashland, OR, USA
9. Marietta Kaszkin-Bettag, PhD, University of Frankfurt, Pharmalex GmbH, Mannheim, Germany
10. Michael Kidd, AM, MD, Editor-in-Chief *Journal of Medical Case Reports*, Faculty of Health Sciences, Flinders University, Adelaide, Australia*
11. Helmut Kiene, Dr med, Editor, *Global Advances in Health and Medicine,* Institute for Applied Epistemology and Research Methodology, University of Witten/Herdecke, Freiburg, Germany
12. Gunver Kienle, Dr med, Editor, *Global Advances in Health and Medicine,* Institute for Applied Epistemology and Research Methodology, University of Witten/Herdecke, Freiburg, Germany
13. Ben Kligler, MD, MPH, Co-Editor-in-Chief *Explore*, Beth Israel Medicine Center, New York, NY, USA*
14. Lori Knutson, RN, BSN, HN-BC, Integrative Healthcare Solutions, Minneapolis, MN, USA
15. Christian A. Koch, MD, PhD, FACP, FACE, Deputy Editor *Journal of Medical Case Reports,* University of Mississippi, Jackson, MS, USA*
16. Karen Milgate, MPP, Independent Health Policy Consultant, Washington, DC, USA*
17. Michele Mittelman, RN, MPH, Editor, *Global Advances in Health and Medicine*, Dover, MA, USA
18. David Moher, PhD, Ottawa Hospital Research Institute; Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, ON Canada*
19. Hanna Oltean, MPH, University of Michigan, Ann Arbor, MI, USA
20. Greg Plotnikoff, MD, MTS, FACP, Editor, *Global Advances in Health and Medicine,* Allina Center for Healthcare Innovations and the Penny George Institute for Health and Healing, Minneapolis, MN, USA
21. Richard Alan Rison, MD, FAANEM, Deputy Editor, *Journal of Medical Case Reports,* Section Editor, *BMC Research Notes*, PIH Health Hospital, Whittier, University of Southern California, Los Angeles, CA, USA*
22. David Riley MD, Editor-in-Chief, *Global Advances in Health and Medicine*, Portland, OR, USA
23. Anil Sethi, MS, Johns Hopkins School of Medicine - Information architecture and IT, Palo Alto, CA, USA*
24. Larissa Shamseer, MSc, Ottawa Hospital Research Institute, Ottawa, ON, Canada
25. Richard Smith, MB, ChB, MSc, United Healthcare Chronic Disease Initiative, London, UK
26. Harold Sox, MD, The Dartmouth Institute, Hannover, NH, USA
27. Peter Tugwell, MD, FRCP, University of Ottawa, Ottawa, ON, Canada

* Participated in the guidelines development process, the review and editing of the CARE guidelines and this article, but did not attend the face-to-face consensus meeting.

Acknowledgments

Joel Gagnier, University of Michigan, and David Riley, *Global Advances in Health and Medicine*, organized this consensus-based guideline-development project. The Department of Orthopaedic Surgery, the Office of the Vice-President of Research at the University of Michigan, and *Global Advances in Health and Medicine*® provided funding for this project. David Moher is funded through a University of Ottawa Research Chair. Funding support was used to reimburse the travel-related expenses of conference attendees. There were no honoraria. The volunteer steering committee consisted of Joel J. Gagnier, Gunver Kienle, David Moher, and David Riley. There are no conflicts of interest to report.

Author details

[1]Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, MI, USA. [2]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA. [3]Institute for Applied Epistemology and Medical Methodology, University of Witten/Herdecke, Freiburg, Germany. [4]Centre for Statistics in Medicine, University of Oxford, Oxford, UK. [5]Ottawa Hospital Research Institute, Ottawa, Canada. [6]Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, ON, Canada. [7]The Dartmouth Institute and Geisel School of Medicine at Dartmouth, Hannover, NH, USA. [8]Global Advances in Health and Medicine, Portland, OR, USA. [9]For a complete list of the CARE Group see the Author Contributions.

Received: 26 June 2013 Accepted: 3 September 2013
Published: 10 September 2013

References

1. Hauben M, Aronson JK: **Gold standards in pharmaco-vigilance: the use of definitive anecdotal reports of adverse drug reactions as pure gold and high-grade ore.** *Drug Saf* 2007, **30**(8):645–55.
2. Vandenbroucke JP: *Thalidomide: an unanticipated adverse effect. JLL Bulletin: Commentaries on the history of treatment evaluation.* http://www.jameslindlibrary.org/illustrating/articles/thalidomide-an-unanticipated-adverse-effect. Accessed January 10, 2013.
3. Levy M: **Propranolol for infantile hemangiomas.** *Global Adv Health Med* 2012, **1**(2):14–6. doi:10.7453/gahmj.2012.1.2.004.
4. Jenicek M: *Clinical case reporting in evidence-based medicine.* 2nd edition. New York: Oxford University Press; 2001.
5. Riley D: **Case reports in the era of clinical trials.** *Global Adv Health Med* 2013, **2**(2):10–11.
6. *"Case reports" search results.* Bethesda, MD: PubMed, US National Library of Medicine; 2013. http://www.ncbi.nlm.nih.gov/pubmed/?term=case+reports, Accessed June 10, 2013.

7. Kaszkin-Bettag M, Hildebrandt W: **Case report on cancer therapies: the urgent need to improve the reporting quality.** *Global Adv Health Med* 2012, **1**(2):8–10. doi:10.7453/gahmj.2012.1.2.002.
8. Kljakovic M: **Single cases in general practice and general medical journals.** *Aust Fam Physician* 2002, **31**(7):669–73.
9. Richason TP, Paulson SM, Lowenstein SR, Heard KJ: **Case reports describing treatments in the emergency medicine literature: missing and misleading information.** *BMC Emerg Med* 2009, **9**:10.
10. Schulz KF, Altman DG, Moher D: **CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials.** *Ann Intern Med* 2010, **152**(11):726–32.
11. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative: **The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies.** *BMJ* 2007, **335**(7624):806–8.
12. Moher D, Liberati A, Tetzlaff J, Altman DG, the PRISMA Group: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement.** *PLoS Med*, **6**(7):e1000097. doi:10.1371/journal.pmed.1000097.
13. Turner L, Shamseer L, Altman DG, Schulz KF, Moher D: **Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane Review.** *Syst Rev* 2012, **1**(1):60. doi:10.1186/2046-4053-1-60.
14. Kelly WN, Arellano FM, Barnes J, Bergman U, Edwards RI, Fernandez AM, *et al*: **Guidelines for submitting adverse event reports for publication.** *Drug Saf* 2007, **30**(5):367–73.
15. Moher D, Schulz KF, Simera I, Altman DG: **Guidance for developers of health research reporting guidelines.** *PLoS Med* 2010, **7**(2):e1000217. doi:10.1371/journal.pmed.100217.
16. Williams PL, Webb C: **The Delphi technique: a methodological discussion.** *J Adv Nurs* 1994, **19**(1):180–6.
17. Thomas B: **Using nominal group technique to identify researchable problems.** *J Nurs Educ* 1983, **22**(8):335–7.
18. Tully MP, Cantrill JA: **Exploring the domains of appropriateness of drug therapy using the nominal group technique.** *Pharm World Sci* 2002, **24**(4):128–31.
19. Horton JN: **Nominal group technique: a method of decision-making by committee.** *Anaesthesia* 1980, **35**(8):11–4.
20. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD, CONSORT PRO Group: **Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension.** *JAMA* 2013, **309**(8):814–22.
21. Vandenbroucke JP: **In defense of case reports and case series.** *Ann Intern Med* 2001, **134**(4):330–4.
22. *Journal of Medical Case Reports.* http://www.jmedicalcasereports.com/. Accessed June 17, 2013.
23. *Cases Database.* http://www.casesdatabase.com/. Accessed June 17, 2013.
24. Hopewell S, Clarke M, Moher D, *et al*: **CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration.** *PLoS Med* 2008, **5**(1):e20. doi:10.1371/journal.pmed.0050020.
25. Hopewell S, Ravaud P, Baron G, Boutron I: **Effect of editors implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis.** *BMJ* 2012, **344**:e4178. doi:10.1136/mj.e4178.
26. Hirst A, Altman DG: **Are peer reviewers encouraged to use reporting guidelines? A survey of 116 health research journals.** *PLoS ONE* 2012, **7**(4):e35621. doi:10.1371/journal.pone.0035621.
27. Hill AB: **The environment and disease: association or causation?** *Proc R Soc Med* 1965, **58**(5):295.
28. Kiene H, Hamre H, Kienle G: **In support of clinical case reports: a system of causality assessment.** *Global Adv Health Med* 2013, **2**(2):76–87. doi:10.7453/gahmj.2012.061.

doi:10.1186/1752-1947-7-223
Cite this article as: Gagnier *et al.*: The CARE guidelines: consensus-based clinical case reporting guideline development. *Journal of Medical Case Reports* 2013 **7**:223.

Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative

Patrick M Bossuyt, Johannes B Reitsma, David E Bruns, Constantine A Gatsonis, Paul P Glasziou, Les M Irwig, Jeroen G Lijmer, David Moher, Drummond Rennie, Henrica C W de Vet for the STARD steering group

The Standards for Reporting of Diagnostic Accuracy (STARD) steering group aims to improve the accuracy and completeness of reporting of studies of diagnostic accuracy. The group describes and explains the development of a checklist and flow diagram for authors of reports

Abstract

Objective To improve the accuracy and completeness of reporting of studies of diagnostic accuracy, to allow readers to assess the potential for bias in a study, and to evaluate a study's generalisability.

Methods The Standards for Reporting of Diagnostic Accuracy (STARD) steering committee searched the literature to identify publications on the appropriate conduct and reporting of diagnostic studies and extracted potential items into an extensive list. Researchers, editors, and members of professional organisations shortened this list during a two day consensus meeting, with the goal of developing a checklist and a generic flow diagram for studies of diagnostic accuracy.

Results The search for published guidelines about diagnostic research yielded 33 previously published checklists, from which we extracted a list of 75 potential items. At the consensus meeting, participants shortened the list to a 25 item checklist, by using evidence, whenever available. A prototype of a flow diagram provides information about the method of patient recruitment, the order of test execution, and the numbers of patients undergoing the test under evaluation and the reference standard, or both.

Conclusions Evaluation of research depends on complete and accurate reporting. If medical journals adopt the STARD checklist and flow diagram, the quality of reporting of studies of diagnostic accuracy should improve to the advantage of clinicians, researchers, reviewers, journals, and the public.

Introduction

The world of diagnostic tests is highly dynamic. New tests are developed at a fast rate, and the technology of existing tests is continuously being improved. Exaggerated and biased results from poorly designed and reported diagnostic studies can trigger their premature dissemination and lead physicians into making incorrect treatment decisions. A rigorous evaluation of diagnostic tests before introduction into clinical practice could not only reduce the number of unwanted clinical consequences related to misleading estimates of test accuracy but also limit healthcare costs by preventing unnecessary testing. Studies to determine the diagnostic accuracy of a test are a vital part of this evaluation process.[1–3]

In studies of diagnostic accuracy, the outcomes from one or more tests under evaluation are compared with outcomes from the reference standard—both measured in subjects who are suspected of having the condition of interest. The term test refers to any method for obtaining additional information on a patient's health status. It includes information from history and physical examination, laboratory tests, imaging tests, function tests, and histopathology. The condition of interest or target condition can refer to a particular disease or to any other identifiable condition that may prompt clinical actions, such as further diagnostic testing, or the initiation, modification, or termination of treatment. In this framework, the reference standard is considered to be the best available method for establishing the presence or absence of the condition of interest. The reference standard can be a single method, or a combination of methods, to establish the presence of the target condition. It can include laboratory tests, imaging tests, and pathology, as well as dedicated clinical follow up of subjects. The term accuracy refers to the amount of agreement between the information from the test under evaluation, referred to as the index test, and the reference standard. Diagnostic accuracy can be expressed in many ways, including sensitivity and specificity, likelihood ratios, diagnostic odds ratio, and the area under a receiver-operator characteristic curve.[4–6]

Several potential threats to the internal and external validity of a study on diagnostic accuracy exist. A survey of studies of diagnostic accuracy published in four major medical journals between 1978 and 1993 revealed that the quality of methods was mediocre at best.[7] However, evaluations were hampered because many reports lacked information on key elements of

Editorial by Straus

Department of Clinical Epidemiology and Biostatistics, Academic Medical Center, University of Amsterdam, PO Box 22700, 1100 DE Amsterdam, Netherlands
Patrick M Bossuyt *professor of clinical epidemiology*
Johannes B Reitsma *clinical epidemiologist*
Jeroen G Lijmer *clinical epidemiologist*

Clinical Chemistry, University of Virginia, Charlottesville, VA 22903-0757, USA
David E Bruns *editor*

Center for Statistical Sciences, Brown University, Providence, RI 02912, USA
Constantine A Gatsonis *professor of medical science (biostatistics) and applied mathematics*

continued over

BMJ 2003;326:41–4

A list of members of the STARD steering committee and the STARD group appears on bmj.com

School of Population Health, University of Queensland, Brisbane, Queensland 4006, Australia
Paul P Glasziou *professor of evidence based practice*
Department of Public Health and Community Medicine, University of Sydney, Sydney, NSW 2006, Australia
Les M Irwig *professor of epidemiology*
Thomas C Chalmer's Center for Systematic Reviews, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON K1H 8LI, Canada
David Moher *director*
JAMA, 515 N State St, Chicago, IL 60610, USA
Drummond Rennie *deputy editor*
Institute for Research in Extramural Medicine, VU University Medical Center, 1081 BT Amsterdam, Netherlands
Henrica C W de Vet *professor of epidemiology*
Correspondence to: P Bossuyt stard@amc.uva.nl

design, conduct, and analysis of diagnostic studies.[7] The absence of critical information about the design and conduct of diagnostic studies has been confirmed by authors of meta-analyses.[8 9] As in any other type of research, flaws in study design can lead to biased results. One report showed that diagnostic studies with specific design features are associated with biased, optimistic estimates of diagnostic accuracy compared with studies without such features.[10]

At the 1999 Cochrane colloquium meeting in Rome, the Cochrane diagnostic and screening test methods working group discussed the low methodological quality and substandard reporting of diagnostic test evaluations. The working group felt that the first step towards correcting these problems was to improve the quality of reporting of diagnostic studies. Following the successful CONSORT initiative,[11–13] the working group aimed to develop a checklist of items that should be included in the report of a study on diagnostic accuracy.

The objective of the Standards for Reporting of Diagnostic Accuracy (STARD) initiative is to improve the quality of reporting of studies of diagnostic accuracy. Complete and accurate reporting allows readers to detect the potential for bias in a study (internal validity) and to assess the generalisability and applicability of results (external validity).

Methods

The STARD steering committee (see bmj.com) started with an extensive search to identify publications on the conduct and reporting of diagnostic studies. This search included Medline, Embase, BIOSIS, and the methodological database from the Cochrane Collaboration up to July 2000. In addition, the members of the steering committee examined reference lists of retrieved articles, searched personal files, and contacted other experts in the field of diagnostic research. They reviewed all relevant publications and extracted an extended list of potential checklist items.

Subsequently, the STARD steering committee convened a two day consensus meeting for invited experts from the following interest groups: researchers, editors, methodologists, and professional organisations. The aim of the conference was to reduce the extended list of potential items, where appropriate, and to discuss the optimal format and phrasing of the checklist. The selection of items to retain was based on evidence whenever possible.

The meeting format consisted of a mixture of small group sessions and plenary sessions. Each small group focused on a group of related items in the list. The suggestions of the small groups then were discussed in plenary sessions. Overnight, a first draft of the STARD checklist was assembled on the basis of suggestions from the small group and additional remarks from the plenary sessions. All meeting attendees discussed this version the next day and made additional changes. The members of the STARD group could suggest further changes through a later round of comments by email.

Potential users field tested the conference version of the checklist and flow diagram, and additional comments were collected. This version was placed on the CONSORT website, with a call for comments. The STARD steering committee discussed all comments and assembled the final checklist.

Results

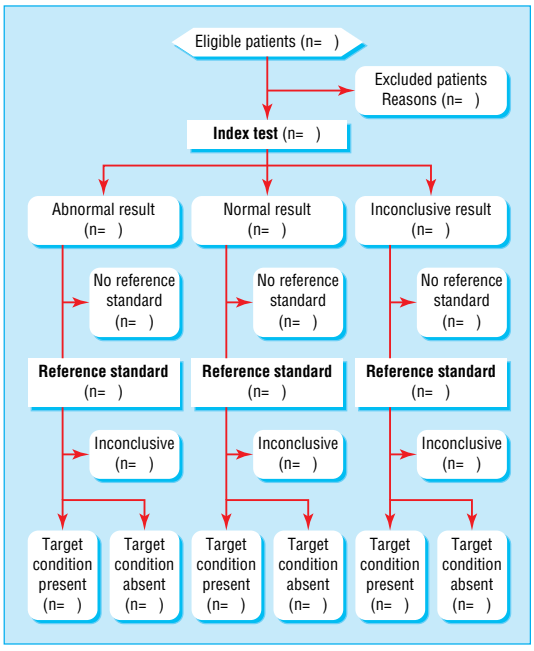
The search for published guidelines for diagnostic research yielded 33 lists. Based on these published guidelines and on input of steering and STARD group members, the steering committee assembled a list of 75 items. During the consensus meeting on 16–17 September 2000, participants consolidated and eliminated items to form the 25 item checklist. Conference members made major revisions to the phrasing and format of the checklist.

The STARD group received valuable comments and remarks during the various stages of evaluation after the conference, which resulted in the version of the STARD checklist in the table.

A flow diagram provides information about the method of patient recruitment (for example, enrolment of a consecutive series of patients with specific symptoms or of cases and controls), the order of test execution, and the number of patients undergoing the test under evaluation (index test) and the reference test. The figure shows a prototype flowchart that reflects the most commonly employed design in diagnostic research. Examples that reflect other designs appear on the STARD website (www.consort-statement.org\stardstatement.htm).

Discussion

The purpose of the STARD initiative is to improve the quality of reporting of diagnostic studies. The items in the checklist and flowchart can help authors to describe essential elements of the design and conduct of the study, the execution of tests, and the results. We arranged the items under the usual headings of a medical research article, but this is not intended to dictate the order in which they have to appear within an article.



Prototype of a flow diagram for a study on diagnostic accuracy

STARD checklist for reporting diagnostic accuracy studies

Section and topic	Item	Description
Title, abstract, and keywords	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading "sensitivity and specificity")
Introduction	2	State the research questions or aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups
Methods:		
Participants	3	Describe the study population: the inclusion and exclusion criteria and the settings and locations where the data were collected
	4	Describe participant recruitment: was this based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?
	5	Describe participant sampling: was this a consecutive series of participants defined by selection criteria in items 3 and 4? If not, specify how participants were further selected
	6	Describe data collection: was data collection planned before the index tests and reference standard were performed (prospective study) or after (retrospective study)?
Test methods	7	Describe the reference standard and its rationale
	8	Describe technical specifications of material and methods involved, including how and when measurements were taken, or cite references for index tests or reference standard, or both
	9	Describe definition of and rationale for the units, cut-off points, or categories of the results of the index tests and the reference standard
	10	Describe the number, training, and expertise of the persons executing and reading the index tests and the reference standard
	11	Were the readers of the index tests and the reference standard blind (masked) to the results of the other test? Describe any other clinical information available to the readers.
Statistical methods	12	Describe methods for calculating or comparing measures of diagnostic accuracy and the statistical methods used to quantify uncertainty (eg 95% confidence intervals)
	13	Describe methods for calculating test reproducibility, if done
Results:		
Participants	14	Report when study was done, including beginning and ending dates of recruitment
	15	Report clinical and demographic characteristics (eg age, sex, spectrum of presenting symptoms, comorbidity, current treatments, and recruitment centre)
	16	Report how many participants satisfying the criteria for inclusion did or did not undergo the index tests or the reference standard, or both; describe why participants failed to receive either test (a flow diagram is strongly recommended)
Test results	17	Report time interval from index tests to reference standard, and any treatment administered between
	18	Report distribution of severity of disease (define criteria) in those with the target condition and other diagnoses in participants without the target condition
	19	Report a cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, report the distribution of the test results by the results of the reference standard
	20	Report any adverse events from performing the index test or the reference standard
Estimates	21	Report estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals)
	22	Report how indeterminate results, missing responses, and outliers of index tests were handled
	23	Report estimates of variability of diagnostic accuracy between readers, centres, or subgroups of participants, if done
	24	Report estimates of test reproducibility, if done
Discussion	25	Discuss the clinical applicability of the study findings

The guiding principle in the development of the STARD checklist was to select items that would help readers judge the potential for bias in the study and to appraise the applicability of the findings. Two other general considerations shaped the content and format of the checklist. Firstly, the STARD group believes that one general checklist for studies of diagnostic accuracy, rather than different checklists for each field, is likely to be more widely disseminated and perhaps accepted by authors, peer reviewers, and journal editors. Although the evaluation of imaging tests differs from that of tests in the laboratory, we felt that these differences were more in degree than in kind. The second consideration was the development of a checklist specifically aimed at studies of diagnostic accuracy. We did not include general issues in the reporting of research findings, such as the recommendations contained in the uniform requirements for manuscripts submitted to biomedical journals.[14]

Wherever possible, the STARD group based the decision to include an item on evidence linking the item to biased estimates (internal validity) or to variations in measures of diagnostic accuracy (external validity). The evidence varied from narrative articles that explained theoretical principles and papers that presented the results of statistical modelling to empirical evidence derived from diagnostic studies. For several items, the evidence was rather limited.

A separate background document explains the meaning and rationale of each item and briefly summarises the type and amount of evidence.[15] This background document should enhance the use, understanding, and dissemination of the STARD checklist.

The STARD group put considerable effort into the development of a flow diagram for diagnostic studies. A flow diagram has the potential to communicate vital information about the design of a study and the flow of participants in a transparent manner.[16] A comparable flow diagram has become an essential element in the CONSORT standards for reporting of randomised trials.[12 16] The flow diagram could be even more essential in diagnostic studies, given the variety of designs employed in diagnostic research. Flow diagrams in the reports of studies of diagnostic accuracy indicate the process of sampling and selecting participants (external validity); the flow of participants in relation to the timing and outcomes of tests; the number of subjects who fail to receive the index test or the reference standard, or

both (potential for verification bias[17–19]); and the number of patients at each stage of the study, which provides the correct denominator for proportions (internal consistency).

The STARD group plans to measure the impact of the statement on the quality of published reports on diagnostic accuracy with a before and after evaluation.[13] Updates of the STARD initiative's documents will be provided when new evidence on sources of bias or variability becomes available. We welcome any comments, whether on content or form, to improve the current version.

This initiative to improve the reporting of studies was supported by a large number of people around the globe who commented on earlier versions. This paper is also being published in the first issues in 2003 of *Annals of Internal Medicine*, *Clinical Chemistry*, *Journal of Clinical Microbiology*, *Lancet*, and *Radiology*. *Clinical Chemistry* is also publishing the background document.

Contributors: PMB and JGL are the initiators of the STARD project. Rijk van Ginkel did the initial search for published guidelines on the design and conduct of diagnostic studies. All authors contributed to the list of potential items for the checklist. PMB, JBR, and JGL prepared the consensus meeting. All authors discussed the comments received during the various stages of the evaluation process. All authors were involved in assembling the final checklist. JBR wrote the first draft of the article, and all authors contributed to the final manuscript. PMB, JBR, and JGL are the guarantors. A list of the members of the STARD steering committee and the STARD group appears on bmj.com

Funding: Financial support to convene the STARD group was provided in part by the Dutch Health Care Insurance Board, Amstelveen, Netherlands; the International Federation of Clinical Chemistry, Milan, Italy; the Medical Research Council's Health Services Research Collaboration, Bristol; and the Academic Medical Center, Amsterdam, Netherlands.

Competing interests: None declared.

1 Guyatt GH, Tugwell PX, Feeny DH, Haynes RB, Drummond M. A framework for clinical evaluation of diagnostic technologies. *Can Med Assoc J* 1986;134:587-94.
2 Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11:88-94.
3 Kent DL, Larson EB. Disease, level of impact, and quality of research methods. Three dimensions of clinical efficacy assessment applied to magnetic resonance imaging. *Invest Radiol* 1992;27:245-54.
4 Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures. Principles and applications. *Ann Intern Med* 1981;94:557-92.
5 Sackett DL, Haynes RB, Guyatt GH, Tugwell P. The selection of diagnostic tests. In: Sackett D, ed. *Clinical epidemiology*. Boston, Toronto, and London: Little, Brown and Company, 1991:47-57.
6 Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8:283-98.
7 Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. *JAMA* 1995;274:645-51.
8 Nelemans PJ, Leiner T, de Vet HCW, van Engelshoven JMA. Peripheral arterial disease: meta-analysis of the diagnostic performance of MR angiography. *Radiology* 2000;217:105-14.
9 Devries SO, Hunink MGM, Polak JF. Summary receiver operating characteristic curves as a technique for meta-analysis of the diagnostic performance of duplex ultrasonography in peripheral arterial disease. *Acad Radiol* 1996;3:361-9.
10 Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061-6.
11 Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;276:637-9.
12 Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285:1987-91.
13 Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials. A comparative before-and-after evaluation. *JAMA* 2001;285:1992-5.
14 International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *JAMA* 1997;277:927-34. http://www.acponline.org/journals/annals/01jan97/unifreqr.htm (accessed 7 Nov 2002).
15 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem* 2003;49:7-18.
16 Egger M, Jüni P, Barlett C. Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001;285:1996-9.
17 Knottnerus JA. The effects of disease verification and referral on the relationship between symptoms and diseases. *Med Decis Making* 1987;7:139-48.
18 Panzer RJ, Suchman AL, Griner PF. Workup bias in prediction research. *Med Decis Making* 1987;7:115-9.
19 Begg CB. Biases in the assessment of diagnostic tests. *Stat Med* 1987;6:411-23.

(Accepted 15 August 2002)

One hundred years ago

Art and nature

Mr. T. A. Cook's *Spirals in Nature and Art* is a book which will appeal to artists and men of science alike. The author describes it on the title page as "A Study of Spiral Formations, based on the manuscripts of Leonardo da Vinci, with special reference to the architecture of the open staircase at Blois, in Touraine, now, for the first time, shown to be from his designs." The book will be found extremely interesting, not only because its subject centres in Leonardo da Vinci, that wonderful painter, man of science, engineer, biologist, mathematician, and architect, but also because, as Professor Ray Lankester in his preface points out, "the training which he (Mr. Cook) received in Paris has emboldened him to enter upon a course of speculative generalization which a more restricted method of study might have prevented. He looks, in fact, upon the results of others' labours with a mind that is more ready to perceive its general value than are those intellects which have concentrated a unique energy upon a single set of problems." When Mr. Cook compares certain architectural beauties with certain natural forms—for example, the spiral staircase at Blois (attributed to Leonardo da Vinci) with the spiral structure of the shell of a mollusc—the resemblance is seen to be obvious, and the beauty and fitness of each is perceived at once. This suggests that the artist, in striking out this spiral form, has been moved or inspired by some deeply underlying natural law, the coincidence implying that there is a rational basis for aesthetics to be discovered; the artist or architect should endeavour, as did the best minds of da Vinci's day, to grasp the problems of proportion in architecture, reflecting the laws of construction and growth exemplified throughout organic life. They should go to Nature and study the ways in which she has solved problems of an allied if not directly comparable kind, and solved them always in a way which gratifies the aesthetic sense of man.

If this be true, then the human aesthetic sense is shown to have its place in the true order of Nature—to be a reflex of, or part of, that order. Da Vinci evolved his theory of spirals not only from shell forms, but also from climbing plants; in the dressing of women's hair, as in the study for the "Leda," he closely follows the coils of the ammonite; he noted that the spiral formation of a screw suggested the movements of a flying bird; and among his drawings are studies of the curves of waves and of the effects of currents upon the banks of the mainland and of islands.

(*BMJ* 1903;i:377)

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies

Erik von Elm, Douglas G Altman, Matthias Egger, Stuart J Pocock, Peter C Gøtzsche, Jan P Vandenbroucke, for the STROBE initiative

Much biomedical research is observational. The reporting of such research is often inadequate, which hampers the assessment of its strengths and weaknesses and of a study's generalisability. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) initiative developed recommendations on what should be included in an accurate and complete report of an observational study. We defined the scope of the recommendations to cover three main study designs: cohort, case-control, and cross-sectional studies. We convened a 2-day workshop in September, 2004, with methodologists, researchers, and journal editors to draft a checklist of items. This list was subsequently revised during several meetings of the coordinating group and in e-mail discussions with the larger group of STROBE contributors, taking into account empirical evidence and methodological considerations. The workshop and the subsequent iterative process of consultation and revision resulted in a checklist of 22 items (the STROBE statement) that relate to the title, abstract, introduction, methods, results, and discussion sections of articles. 18 items are common to all three study designs and four are specific for cohort, case-control, or cross-sectional studies. A detailed explanation and elaboration document is published separately and is freely available on the websites of *PLoS Medicine*, *Annals of Internal Medicine*, and *Epidemiology*. We hope that the STROBE statement will contribute to improving the quality of reporting of observational studies.

Lancet 2007; 370: 1453–57

Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland (E von Elm MD, Prof M Egger MD); **Centre for Statistics in Medicine, University of Oxford, Oxford, UK** (Prof D G Altman DSc); **Department of Social Medicine, University of Bristol, Bristol, UK** (M Egger); **London School of Hygiene and Tropical Medicine, University of London, London, UK** (Prof S J Pocock PhD); **Nordic Cochrane Centre, Copenhagen, Denmark** (P C Gøtzsche MD); **and Department of Clinical Epidemiology, Leiden University Hospital, Leiden, Netherlands** (Prof J P Vandenbroucke MD)

Correspondence to:
Dr Erik von Elm, Institute of Social and Preventive Medicine (ISPM), University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland
strobe@ispm.unibe.ch

Introduction

Many questions in medical research are investigated in observational studies.[1] Much of the research into the cause of diseases relies on cohort, case-control, or cross-sectional studies. Observational studies also have a role in research into the benefits and harms of medical interventions.[2] Randomised trials cannot answer all important questions about a given intervention. For example, observational studies are more suitable to detect rare or late adverse effects of treatments, and are more likely to provide an indication of what is achieved in daily medical practice.[3]

Research should be reported transparently so that readers can follow what was planned, what was done, what was found, and what conclusions were drawn. The credibility of research depends on a critical assessment by others of the strengths and weaknesses in study design, conduct, and analysis. Transparent reporting is also needed to judge whether and how results can be included in systematic reviews.[4,5] However, in published observational research important information is often missing or unclear. An analysis of epidemiological studies published in general medical and specialist journals found that the rationale behind the choice of potential confounding variables was often not reported.[6] Only a few reports of case-control studies in psychiatry explained the methods used to identify cases and controls.[7] In a survey of longitudinal studies in stroke research, 17 of 49 articles (35%) did not specify the eligibility criteria.[8] Others have argued that without sufficient clarity of reporting, the benefits of research might be achieved more slowly,[9] and that there is a need for guidance in reporting observational studies.[10,11]

Recommendations on the reporting of research can improve reporting quality. The Consolidated Standards of Reporting Trials (CONSORT) statement was developed in 1996 and revised 5 years later.[12] Many medical journals supported this initiative,[13] which has helped to improve the quality of reports of randomised trials.[14,15] Similar initiatives have followed for other research areas—eg, for the reporting of meta-analyses of randomised trials[16] or diagnostic studies.[17] We established a network of methodologists, researchers, and journal editors to develop recommendations for the reporting of observational research: the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement.

Aims and use of the STROBE statement

The STROBE statement is a checklist of items that should be addressed in articles reporting on the three main study designs of analytical epidemiology: cohort, case-control, and cross-sectional studies. The intention is solely to provide guidance on how to report observational research well: these recommendations are not prescriptions for designing or conducting studies. Also, while clarity of reporting is a prerequisite to evaluation, the checklist is not an instrument to evaluate the quality of observational research.

Here we present the STROBE statement and explain how it was developed. In a detailed companion paper, the explanation and elaboration article,[18–20] we justify the inclusion of the different checklist items and give methodological background and published examples of what we consider transparent reporting. We strongly recommend using the STROBE checklist in conjunction with the explanatory article, which is available freely on the websites of *PLoS Medicine* (www.plosmedicine.org), *Annals of Internal Medicine* (www.annals.org), and *Epidemiology* (www.epidem.com).

	Item	Recommendation	Reported on manuscript page
Title and abstract			
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
Methods			
Study design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) *Cohort study*—give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—give the eligibility criteria, and the sources and methods of selection of participants	
		(b) *Cohort study*—for matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—for matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	
Data sources/ measurement	8*	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) *Cohort study*—if applicable, explain how loss to follow-up was addressed *Case-control study*—if applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—if applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report the numbers of individuals at each stage of the study—eg, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg, demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate the number of participants with missing data for each variable of interest	
		(c) *Cohort study*—summarise follow-up time (eg, average and total amount)	
Outcome data	15*	*Cohort study*—report numbers of outcome events or summary measures over time *Case-control study*—report numbers in each exposure category, or summary measures of exposure *Cross-sectional study*—report numbers of outcome events or summary measures	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(b) Report category boundaries when continuous variables were categorised	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg, analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

*Give such information separately for cases and controls in case-control studies, and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies. An explanation and elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the websites of *PLoS Medicine*, *Annals of Internal Medicine*, and *Epidemiology*). Separate versions of the checklist for cohort, case-control, and cross-sectional studies are available on the STROBE website.

Table: **The STROBE statement—checklist of items that should be addressed in reports of observational studies**

Development of the STROBE statement

We established the STROBE initiative in 2004, obtained funding for a workshop, and set up a website (www.strobe-statement.org). We searched textbooks, bibliographic databases, reference lists, and personal files for relevant material, including previous recommendations, empirical studies of reporting, and articles describing relevant methodological research. Because observational research makes use of many different study designs, we felt that the scope of STROBE had to be clearly defined early on. We decided to focus on the three study designs that are used most widely in analytical observational research: cohort, case-control, and cross-sectional studies.

We organised a 2-day workshop in Bristol, UK, in September, 2004. 23 individuals attended this meeting, including editorial staff from *Annals of Internal Medicine*, *BMJ*, *Bulletin of the World Health Organization*, *International Journal of Epidemiology*, *JAMA*, *Preventive Medicine*, and *The Lancet*, as well as epidemiologists, methodologists, statisticians, and practitioners from Europe and North America. Written contributions were sought from ten other individuals who declared an interest in contributing to STROBE, but could not attend. Three working groups identified items deemed to be important to include in checklists for each type of study. A provisional list of items prepared in advance (available from our website) was used to facilitate discussions. The three draft checklists were then discussed by all participants and, where possible, items were revised to make them applicable to all three study designs. In a final plenary session, the group decided on the strategy for finalising and disseminating the STROBE statement.

After the workshop we drafted a combined checklist including all three designs and made it available on our website. We invited participants and additional scientists and editors to comment on this draft checklist. We subsequently published three revisions on the website, and two summaries of comments received and changes made. During this process the coordinating group (ie, the authors of the present paper) met on eight occasions for 1 or 2 days and held several telephone conferences to revise the checklist and to prepare the present paper and the explanation and elaboration paper.[18–20] The coordinating group invited three additional co-authors with methodological and editorial expertise to help write the explanation and elaboration paper, and sought feedback from more than 30 people, who are listed at the end of this paper. We allowed several weeks for comments on subsequent drafts of the paper and reminded collaborators about deadlines by e-mail.

STROBE components

The STROBE statement is a checklist of 22 items that we consider essential for good reporting of observational studies (table). These items relate to the article's title and abstract (item 1), the introduction (items 2 and 3), methods (items 4–12), results (items 13–17), and discussion sections (items 18–21), and other information (item 22 on funding). 18 items are common to all three designs, while four (items 6, 12, 14, and 15) are design-specific, with different versions for all or part of the item. For some items (indicated by asterisks), information should be given separately for cases and controls in case-control studies, or exposed and unexposed groups in cohort and cross-sectional studies. Although presented here as a single checklist, separate checklists are available for each of the three study designs on the STROBE website.

For more on **the STROBE initiative** see www.strobe-statement.org

Implications and limitations

The STROBE statement was developed to assist authors when writing up analytical observational studies, to support editors and reviewers when considering such articles for publication, and to help readers when critically appraising published articles. We developed the checklist through an open process, taking into account the experience gained with previous initiatives, in particular CONSORT. We reviewed the relevant empirical evidence as well as methodological work, and subjected consecutive drafts to an extensive iterative process of consultation. The checklist presented here is thus based on input from a large number of individuals with diverse backgrounds and perspectives. The comprehensive explanatory article,[18–20] which is intended for use alongside the checklist, also benefited greatly from this consultation process.

Observational studies serve a wide range of purposes, on a continuum from the discovery of new findings to the confirmation or refutation of previous findings.[18–20] Some studies are essentially exploratory and raise interesting hypotheses. Others pursue clearly defined hypotheses in available data. In yet another type of studies, the collection of new data is planned carefully on the basis of an existing hypothesis. We believe the present checklist can be useful for all these studies, since the readers always need to know what was planned (and what was not), what was done, what was found, and what the results mean. We acknowledge that STROBE is currently limited to three main observational study designs. We would welcome extensions that adapt the checklist to other designs—eg, case-crossover studies or ecological studies—and also to specific topic areas. Four extensions are now available for the CONSORT statement.[21–24] A first extension to STROBE is underway for gene-disease association studies: the STROBE Extension to Genetic Association studies (STREGA) initiative.[25] We ask those who aim to develop extensions of the STROBE statement to contact the coordinating group first to avoid duplication of effort.

The STROBE statement should not be interpreted as an attempt to prescribe the reporting of observational research in a rigid format. The checklist items should be addressed in sufficient detail and with clarity somewhere in an article, but the order and format for presenting information

depends on author preferences, journal style, and the traditions of the research field. For instance, we discuss the reporting of results under a number of separate items, while recognising that authors might address several items within a single section of text or in a table. Also, item 22, on the source of funding and the role of funders, could be addressed in an appendix or in the methods section of the article. We do not aim at standardising reporting. Authors of randomised clinical trials were asked by an editor of a specialist medical journal to "CONSORT" their manuscripts on submission.[26] We believe that manuscripts should not be "STROBEd", in the sense of regulating style or terminology. We encourage authors to use narrative elements, including the description of illustrative cases, to complement the essential information about their study, and to make their articles an interesting read.[27]

We emphasise that the STROBE statement was not developed as a tool for assessing the quality of published observational research. Such instruments have been developed by other groups and were the subject of a recent systematic review.[28] In the explanation and elaboration paper, we used several examples of good reporting from studies whose results were not confirmed in further research—the important feature was the good reporting, not whether the research was of good quality. However, if STROBE is adopted by authors and journals, issues such as confounding, bias, and generalisability could become more transparent, which might help temper the over-enthusiastic reporting of new findings in the scientific community and popular media,[29] and improve the methodology of studies in the long term. Better reporting may also help to have more informed decisions about when new studies are needed, and what they should address.

We did not undertake a comprehensive systematic review for each of the checklist items and subitems, or do our own research to fill gaps in the evidence base. Further, although no one was excluded from the process, the composition of the group of contributors was influenced by existing networks and was not representative in terms of geography (it was dominated by contributors from Europe and North America) and probably was not representative in terms of research interests and disciplines. We stress that STROBE and other recommendations on the reporting of research should be seen as evolving documents that require continual assessment, refinement, and, if necessary, change. We welcome suggestions for the further dissemination of STROBE—eg, by re-publication of the present article in specialist journals and in journals published in other languages. Groups or individuals who intend to translate the checklist to other languages should consult the coordinating group beforehand. We will revise the checklist in the future, taking into account comments, criticism, new evidence, and experience from its use. We invite readers to submit their comments via the STROBE website.

Contributors

The following individuals have contributed to the content and elaboration of the STROBE statement: Douglas G Altman, Maria Blettner, Paolo Boffetta, Hermann Brenner, Geneviève Chêne, Cyrus Cooper, George Davey-Smith, Erik von Elm, Matthias Egger, France Gagnon, Peter C Gøtzsche, Philip Greenland, Sander Greenland, Claire Infante-Rivard, John Ioannidis, Astrid James, Giselle Jones, Bruno Ledergerber, Julian Little, Margaret May, David Moher, Hooman Momen, Alfredo Morabia, Hal Morgenstern, Cynthia D Mulrow, Fred Paccaud, Stuart J Pocock, Charles Poole, Martin Röösli, Dietrich Rothenbacher, Kenneth Rothman, Caroline Sabin, Willi Sauerbrei, Lale Say, James J Schlesselman, Jonathan Sterne, Holly Syddall, Jan P Vandenbroucke, Ian White, Susan Wieland, Hywel Williams, Guang Yong Zou.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgments

The workshop was funded by the European Science Foundation (ESF). Additional funding was received from the Medical Research Council Health Services Research Collaboration, and the National Health Services Research and Development Methodology Programme. We are grateful to Gerd Antes, Kay Dickersin, Shah Ebrahim, and Richard Lilford for supporting the STROBE initiative. We are grateful to the following institutions that have hosted working meetings of the coordinating group: Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland; Department of Social Medicine, University of Bristol, Bristol, UK; London School of Hygiene and Tropical Medicine, London, UK; Nordic Cochrane Centre, Copenhagen, Denmark; and Centre for Statistics in Medicine, Oxford, UK. We are grateful to six reviewers who provided helpful comments on a previous draft of this paper.

References

1 Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. *BMJ* 2004; **328:** 39–41.
2 Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996; **312:** 1215–18.
3 Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 2006; **174:** 635–41.
4 Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ* 2001; **323:** 42–46.
5 Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ* 1998; **316:** 140–44.
6 Pocock SJ, Collier TJ, Dandreo KJ, et al. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ* 2004; **329:** 883.
7 Lee W, Bindman J, Ford T, et al. Bias in psychiatric case-control studies: literature survey. *Br J Psychiatry* 2007; **190:** 204–09.
8 Tooth L, Ware R, Bain C, Purdie DM, Dobson A. Quality of reporting of observational longitudinal research. *Am J Epidemiol* 2005; **161:** 280–88.
9 Bogardus ST Jr, Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research: the need for methodological standards. *JAMA* 1999; **281:** 1919–26.
10 Anon. Guidelines for documentation of epidemiologic studies. Epidemiology Work Group of the Interagency Regulatory Liaison Group. *Am J Epidemiol* 1981; **114:** 609–13.
11 Rennie D. CONSORT revised – improving the reporting of randomized trials. *JAMA* 2001; **285:** 2006–07.
12 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001; **357:** 1191–94.
13 Moher D, Altman DG, Schulz KF, Elbourne DR. Opportunities and challenges for improving the quality of reporting clinical research: CONSORT and beyond. *CMAJ* 2004; **171:** 349–50.
14 Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006; **185:** 263–67.
15 Egger M, Jüni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001; **285:** 1996–99.

16 Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; **354:** 1896–900.

17 Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* 2003; **138:** 40–44.

18 Vandenbroucke JP, von Elm E, Altman DG, et al, for the STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med* 2007; **4:** e297.

19 Vandenbroucke JP, von Elm E, Altman DG, et al, for the STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Ann Intern Med* (in press).

20 Vandenbroucke JP von Elm E, Altman DG, et al, for the STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology* (in press).

21 Ioannidis JP, Evans SJ, Gotzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; **141:** 781–88.

22 Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004; **328:** 702–08.

23 Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006; **295:** 1152–60.

24 Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann Intern Med* 2006; **144:** 364–67.

25 Ioannidis JP, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. *Nat Genet* 2006; **38:** 3–5.

26 Ormerod AD. CONSORT your submissions: an update for authors. *Br J Dermatol* 2001; **145:** 378–79.

27 Schriger DL. Suggestions for improving the reporting of clinical research: the role of narrative. *Ann Emerg Med* 2005; **45:** 437–43.

28 Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007; **36:** 666–76.

29 Bartlett C, Sterne J, Egger M. What is newsworthy? Longitudinal study of the reporting of medical research in two British newspapers. *BMJ* 2002; **325:** 81–84.

SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials

An-Wen Chan, MD, DPhil; Jennifer M. Tetzlaff, MSc; Douglas G. Altman, DSc; Andreas Laupacis, MD; Peter C. Gøtzsche, MD, DrMedSci; Karmela Krleža-Jerić, MD, DSc; Asbjørn Hróbjartsson, PhD; Howard Mann, MD; Kay Dickersin, PhD; Jesse A. Berlin, ScD; Caroline J. Doré, BSc; Wendy R. Parulekar, MD; William S.M. Summerskill, MBBS; Trish Groves, MBBS; Kenneth F. Schulz, PhD; Harold C. Sox, MD; Frank W. Rockhold, PhD; Drummond Rennie, MD; and David Moher, PhD

The protocol of a clinical trial serves as the foundation for study planning, conduct, reporting, and appraisal. However, trial protocols and existing protocol guidelines vary greatly in content and quality. This article describes the systematic development and scope of SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013, a guideline for the minimum content of a clinical trial protocol.

The 33-item SPIRIT checklist applies to protocols for all clinical trials and focuses on content rather than format. The checklist recommends a full description of what is planned; it does not prescribe how to design or conduct a trial. By providing guidance for key content, the SPIRIT recommendations aim to facilitate the drafting of high-quality protocols. Adherence to SPIRIT would also enhance the transparency and completeness of trial protocols for the benefit of investigators, trial participants, patients, sponsors, funders, research ethics committees or institutional review boards, peer reviewers, journals, trial registries, policymakers, regulators, and other key stakeholders.

Ann Intern Med. 2013;158:200-207. www.annals.org
For author affiliations, see end of text.
This article was published at www.annals.org on 8 January 2013.

The protocol of a clinical trial plays a key role in study planning, conduct, interpretation, oversight, and external review by detailing the plans from ethics approval to dissemination of results. A well-written protocol facilitates an appropriate assessment of scientific, ethical, and safety issues before a trial begins; consistency and rigor of trial conduct; and full appraisal of the conduct and results after trial completion. The importance of protocols has been emphasized by journal editors (1–6), peer reviewers (7–10), researchers (11–15), and public advocates (16).

Despite the central role of protocols, a systematic review revealed that existing guidelines for protocol content vary greatly in their scope and recommendations, seldom describe how the guidelines were developed, and rarely cite broad stakeholder involvement or empirical evidence to support their recommendations (17). These limitations may partly explain why an opportunity exists to improve the quality of protocols. Many protocols for randomized trials do not adequately describe the primary outcomes (inadequate for 25% of trials) (18, 19), treatment allocation methods (inadequate for 54% to 79%) (20, 21), use of blinding (inadequate for 9% to 34%) (21, 22), methods for reporting adverse events (inadequate for 41%) (23), components of sample size calculations (inadequate for 4% to 40%) (21, 24), data analysis plans (inadequate for 20% to 77%) (21, 24–26), publication policies (inadequate for 7%) (27), and roles of sponsors and investigators in study design or data access (inadequate for 89% to 100%) (28, 29). The problems that underlie these protocol deficiencies may in turn lead to avoidable protocol amendments, poor trial conduct, and inadequate reporting in trial publications (15, 30).

In response to these gaps in protocol content and guidance, we launched the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) initiative in 2007. This international project aims to improve the completeness of trial protocols by producing evidence-based recommendations for a minimum set of items to be addressed in protocols. The SPIRIT 2013 Statement includes a 33-item checklist (**Table 1**) and diagram (**Figure**). An associated explanatory paper (SPIRIT 2013 Explanation and Elaboration) (31) details the rationale and supporting evidence for each checklist item, along with guidance and model examples from actual protocols.

Development of the SPIRIT 2013 Statement

The SPIRIT 2013 Statement was developed in broad consultation with 115 key stakeholders, including trial investigators ($n = 30$); health care professionals ($n = 31$); methodologists ($n = 34$); statisticians ($n = 16$); trial coordinators ($n = 14$); journal editors ($n = 15$); and representatives from the research ethics community ($n = 17$), industry and nonindustry funders ($n = 7$), and regulatory agencies ($n = 3$), whose roles are not mutually exclusive. As detailed later, the SPIRIT guideline was developed through 2 systematic reviews, a formal Delphi consensus process, 2 face-to-face consensus meetings, and pilot-testing (32).

The SPIRIT checklist evolved through several iterations. The process began with a preliminary checklist of 59 items derived from a systematic review of existing protocol guidelines (17). In 2007, 96 expert panelists from 17 low- ($n = 1$), middle- ($n = 6$), and high-income ($n = 10$) countries refined this initial checklist over 3 iterative Delphi consensus survey rounds by e-mail (33). Panelists rated each item on a scale of 1 (not important) to 10 (very important), suggested new items, and provided comments that were circulated in subsequent rounds. Items with a median score of 8 or higher in the final round were included, whereas those rated 5 or lower were excluded.

Table 1. **SPIRIT 2013 Checklist: Recommended Items to Address in a Clinical Trial Protocol and Related Documents***

Section/Item	Item Number	Description
Administrative information		
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry.
	2b	All items from the World Health Organization Trial Registration Data Set (**Appendix Table**, available at www.annals.org)
Protocol version	3	Date and version identifier
Funding	4	Sources and types of financial, material, and other support
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors
	5b	Name and contact information for the trial sponsor
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities
	5d	Composition, roles, and responsibilities of the coordinating center, steering committee, end point adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see item 21a for DMC)
Introduction		
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention
	6b	Explanation for choice of comparators
Objectives	7	Specific objectives or hypotheses
Trial design	8	Description of trial design, including type of trial (e.g., parallel group, crossover, factorial, single group), allocation ratio, and framework (e.g., superiority, equivalence, noninferiority, exploratory)
Methods		
Participants, interventions, and outcomes		
Study setting	9	Description of study settings (e.g., community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centers and individuals who will perform the interventions (e.g., surgeons, psychotherapists)
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (e.g., drug dose change in response to harms, participant request, or improving/worsening disease)
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (e.g., drug tablet return, laboratory tests)
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (e.g., systolic blood pressure), analysis metric (e.g., change from baseline, final value, time to event), method of aggregation (e.g., median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended
Participant timeline	13	Time schedule of enrollment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (**Figure**).
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations
Recruitment	15	Strategies for achieving adequate participant enrollment to reach target sample size
Assignment of interventions (for controlled trials)		
Allocation		
Sequence generation	16a	Method of generating the allocation sequence (e.g., computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (e.g., blocking) should be provided in a separate document that is unavailable to those who enroll participants or assign interventions.
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (e.g., central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned
Implementation	16c	Who will generate the allocation sequence, who will enroll participants, and who will assign participants to interventions
Blinding (masking)	17a	Who will be blinded after assignment to interventions (e.g., trial participants, care providers, outcome assessors, data analysts), and how
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial

Continued on following page

Table 1—**Continued**

Section/Item	Item Number	Description
Data collection, management, and analysis		
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (e.g., duplicate measurements, training of assessors) and a description of study instruments (e.g., questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol.
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (e.g., double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol.
Statistical methods	20a	Statistical methods for analyzing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol.
	20b	Methods for any additional analyses (e.g., subgroup and adjusted analyses)
	20c	Definition of analysis population relating to protocol nonadherence (e.g., as-randomized analysis), and any statistical methods to handle missing data (e.g., multiple imputation)
Monitoring		
Data monitoring	21a	Composition of DMC; summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed.
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor
Ethics and dissemination		
Research ethics approval	24	Plans for seeking REC/IRB approval
Protocol amendments	25	Plans for communicating important protocol modifications (e.g., changes to eligibility criteria, outcomes, analyses) to relevant parties (e.g., investigators, RECs/IRBs, trial participants, trial registries, journals, regulators)
Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorized surrogates, and how (see item 32)
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site
Access to data	29	Statement of who will have access to the final trial data set, and disclosure of contractual agreements that limit such access for investigators
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, health care professionals, the public, and other relevant groups (e.g., via publication, reporting in results databases, or other data-sharing arrangements), including any publication restrictions
	31b	Authorship eligibility guidelines and any intended use of professional writers
	31c	Plans, if any, for granting public access to the full protocol, participant-level data set, and statistical code
Appendices		
Informed consent materials	32	Model consent form and other related documentation given to participants and authorized surrogates
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable

DMC = data monitoring committee; IRB = institutional review board; REC = research ethics committee; SPIRIT = Standard Protocol Items: Recommendations for Interventional Trials.

* It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation and Elaboration (31) for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group and is reproduced with permission.

Items rated between 5 and 8 were retained for further discussion at the consensus meetings.

After the Delphi survey, 16 members of the SPIRIT Group (named as authors of this paper) met in December 2007 in Ottawa, Ontario, Canada, and 14 members met in September 2009 in Toronto, Ontario, Canada, to review the survey results, discuss controversial items, and refine the draft checklist. After each meeting, the revised checklist was recirculated to the SPIRIT Group for additional feedback.

A second systematic review identified empirical evidence about the relevance of specific protocol items to trial conduct or risk of bias. The results of this review informed the decision to include or exclude items on the SPIRIT

checklist. This review also provided the evidence base of studies cited in the SPIRIT 2013 Explanation and Elaboration paper (31). Some items had little or no identified empirical evidence (for example, the title) and are included in the checklist on the basis of a strong pragmatic or ethical rationale.

Finally, we pilot-tested the draft checklist in 2010 and 2011 with University of Toronto graduate students who used the document to develop trial protocols as part of a master's-level course on clinical trial methods. Their feedback on the content, format, and usefulness of the checklist was obtained through an anonymous survey and incorporated into the final SPIRIT checklist.

Definition of a Clinical Trial Protocol

Although every study requires a protocol, the precise definition of a protocol varies among individual investigators, sponsors, and other stakeholders. For the SPIRIT initiative, the protocol is defined as a document that provides sufficient detail to enable understanding of the background, rationale, objectives, study population, interventions, methods, statistical analyses, ethical considerations, dissemination plans, and administration of the trial; replication of key aspects of trial methods and conduct; and appraisal of the trial's scientific and ethical rigor from ethics approval to dissemination of results.

The protocol is more than a list of items. It should be a cohesive document that provides appropriate context and narrative to fully understand the elements of the trial. For example, the description of a complex intervention may need to include training materials and figures to enable replication by persons with appropriate expertise.

The full protocol must be submitted for approval by an institutional review board (IRB) or research ethics committee (34). It is recommended that trial investigators or sponsors address the SPIRIT checklist items in the protocol before submission. If the details for certain items have not yet been finalized, then this should be stated in the protocol and the items updated as they evolve.

The protocol is a "living" document that is often modified during the trial. A transparent audit trail with dates of important changes in trial design and conduct is an essential part of the scientific record. Trial investigators and sponsors are expected to adhere to the protocol as approved by the IRB and to document amendments made in the most recent protocol version. Important protocol amendments should be reported to IRBs and trial registries as they occur and subsequently be described in trial reports.

Scope of the SPIRIT 2013 Statement

The SPIRIT 2013 Statement applies to the content of a clinical trial protocol, including its appendices. A clinical trial is a prospective study in which 1 or more interventions are assigned to human participants to assess the effects on health-related outcomes. The primary scope of SPIRIT 2013 relates to randomized trials, but the same considerations substantially apply to all types of clinical trials, regardless of study design, intervention, or topic.

The SPIRIT 2013 Statement provides guidance for minimum protocol content. Certain circumstances may warrant additional protocol items. For example, a factorial study design may require specific justification; crossover trials have unique statistical considerations, such as carryover effects; and industry-sponsored trials may have additional regulatory requirements.

Figure. **Example template of recommended content for the schedule of enrollment, interventions, and assessments.**

	Study Period							
	Enrollment	Allocation	Postallocation					Closeout
Time point*	$-t_1$	0	t_1	t_2	t_3	t_4	etc.	t_x
Enrollment:								
Eligibility screen	X							
Informed consent	X							
[List other procedures]	X							
Allocation		X						
Interventions:								
[Intervention A]			◆	—	◆			
[Intervention B]			X		X			
[List other study groups]			◆	—	—	◆		
Assessments:								
[List baseline variables]	X	X						
[List outcome variables]				X		X	etc.	X
[List other data variables]			X	X	X	X	etc.	X

Recommended content can be displayed using various schematic formats. See SPIRIT 2013 Explanation and Elaboration (31) for examples. This template is copyrighted by the SPIRIT Group and is reproduced with permission. SPIRIT = Standard Protocol Items: Recommendations for Interventional Trials.
* List specific time points in this row.

The protocol and its appendices are often the sole repository of detailed information relevant to every SPIRIT checklist item. Using existing trial protocols, we have been able to identify model examples of every item (31), which illustrates the feasibility of addressing all checklist items in a single protocol document. For some trials, relevant details may appear in related documents, such as statistical analysis plans, case record forms, operations manuals, or investigator contracts (35, 36). In these instances, the protocol should outline the key principles and refer to the separate documents so that their existence is known.

The SPIRIT 2013 Statement primarily relates to the content of the protocol rather than its format, which is often subject to local regulations, traditions, or standard operating procedures. Nevertheless, adherence to certain formatting conventions, such as a table of contents; section headings; glossary; list of abbreviations; list of references; and a schematic schedule of enrollment, interventions, and assessments, will facilitate protocol review (**Figure**).

Finally, the intent of SPIRIT 2013 is to promote transparency and a full description of what is planned—not to prescribe how a trial should be designed or conducted. The checklist should not be used to judge trial quality, because the protocol of a poorly designed trial may address all checklist items by fully describing its inadequate design features. Nevertheless, the use of SPIRIT 2013 may improve the validity and success of trials by reminding investigators about important issues to consider during the planning stages.

Relation to Existing Clinical Trial Guidance

With its systematic development process, consultation with international stakeholders, and explanatory paper citing relevant empirical evidence (31), SPIRIT 2013 builds on other international guidance applicable to clinical trial protocols. It adheres to the ethical principles mandated by the 2008 Declaration of Helsinki, particularly the requirement that the protocol address specific ethical considerations, such as competing interests (34).

In addition, SPIRIT 2013 encompasses the protocol items recommended by the International Conference on Harmonisation Good Clinical Practice E6 guidance, written in 1996 for clinical trials whose data are intended for submission to regulatory authorities (37). The SPIRIT Statement builds on the Good Clinical Practice guidance by providing additional recommendations on specific key protocol items (for example, allocation concealment, trial registration, and consent processes). In contrast to SPIRIT, the Good Clinical Practice guidance used informal consensus methods, has unclear contributorship, and lacks citation of supporting empirical evidence (38).

The SPIRIT 2013 Statement also supports trial registration requirements from the World Health Organization (39), the International Committee of Medical Journal Editors (40), legislation pertaining to ClinicalTrials.gov (41), the European Commission (42), and others. For example, item 2b of the SPIRIT checklist recommends that the protocol list the World Health Organization Trial Registration Data Set (**Appendix Table**, available at www.annals.org), which is the minimum amount of information that the International Committee of Medical Journal Editors mandates for trial registries. Having this data set in its own protocol section is intended not only to serve as a form of trial summary but also to help improve the quality of information in registry entries. Registration-specific data could be easily identified in the protocol section and copied into the registry fields. In addition, protocol amendments applicable to this section could prompt investigators to update their registry data.

The SPIRIT 2013 Statement mirrors applicable items from CONSORT 2010 (Consolidated Standards of Reporting Trials) (43). Consistent wording and structure used for items common to both checklists will facilitate the transition from a SPIRIT-based protocol to a final report based on CONSORT. The SPIRIT Group has also engaged leaders of other initiatives relevant to protocol standards, such as trial registries, the Clinical Data Interchange Standards Consortium Protocol Representation Group, and Pragmatic Randomized Controlled Trials in HealthCare, to align international efforts in promoting transparency and high-quality protocol content.

Potential Effect

An extensive range of stakeholders could benefit from widespread use of the SPIRIT 2013 Statement and its explanatory paper (**Table 2**). Pilot-testing and informal feedback have shown that it is particularly valuable for trial investigators when they draft their protocols. It can also serve as an informational resource for new investigators, peer reviewers, and IRB members.

There is also potential benefit for trial implementation. The excessive delay from the time of protocol development to ethics approval and the start of participant recruitment remains a major concern for clinical trials (44). Improved completeness of protocols could help increase the efficiency of protocol review by reducing avoidable queries to investigators about incomplete or unclear information. With full documentation of key information and increased awareness of important considerations before the trial begins, the use of SPIRIT may also help to reduce the number and burden of subsequent protocol amendments—many of which can be avoided with careful protocol drafting and development (15). Widespread adoption of SPIRIT 2013 as a single standard by IRBs, funding agencies, regulatory agencies, and journals could simplify the work of trial investigators and sponsors, who could fulfill the common application requirements of multiple stakeholders with a single SPIRIT-based protocol. Better protocols would also help trial personnel to implement the study as the protocol authors intended.

Table 2. **Potential Benefits and Proposed Stakeholder Actions for Supporting Adherence to SPIRIT 2013**

Stakeholder	Proposed Actions	Potential Benefits
Clinical trial groups, investigators, sponsors	Adopt SPIRIT as standard guidance Use as tool for writing protocols	Improved quality, completeness, and consistency of protocol content Enhanced understanding of rationale and issues to consider for key protocol items Increased efficiency of protocol review
Research ethics committees/institutional review boards, funding agencies, regulatory agencies	Mandate or encourage adherence to SPIRIT for submitted protocols Use as training tool	Improved quality, completeness, and consistency of protocol submissions Increased efficiency of review and reduction in queries about protocol requirements
Educators	Use SPIRIT checklist and explanatory paper as a training tool	Enhanced understanding of the rationale and issues to consider for key protocol items
Patients, trial participants, policymakers	Advocate use of SPIRIT by trial investigators and sponsors	Improved protocol content relevant to transparency, accountability, critical appraisal, and oversight
Trial registries	Encourage SPIRIT-based protocols Register full protocols to accompany results disclosure	Improved quality of registry records Prompt for trialists to update registry record when SPIRIT checklist item 2b (Registration Data Set) is updated Improved quality, completeness, and consistency of protocol content for registries that house full protocols and results
Journal editors and publishers	Endorse SPIRIT as standard guidance for published and unpublished protocols Include reference to SPIRIT in instructions for authors Ask that protocols be submitted with manuscripts, circulate them to peer reviewers, and encourage authors to make them available as Web appendices	Improved quality, completeness, and consistency of protocol content Enhanced peer review of trial manuscripts through improved protocol content, which can be used to assess protocol adherence and selective reporting Improved transparency and interpretation of trials by readers

SPIRIT = Standard Protocol Items: Recommendations for Interventional Trials.

Furthermore, adherence to SPIRIT 2013 could help ensure that protocols contain the requisite information for critical appraisal and trial interpretation. High-quality protocols can provide important information about trial methods and conduct that is not available from journals or trial registries (45–47). As a transparent record of the researchers' original intent, comparisons of protocols with final trial reports can help to identify selective reporting of results and undisclosed amendments (48), such as changes to primary outcomes (19, 49). However, clinical trial protocols are not generally accessible to the public (45). The SPIRIT 2013 Statement will have a greater effect when protocols are publicly available to facilitate full evaluation of trial validity and applicability (11, 12, 14, 50).

The SPIRIT 2013 guideline needs the support of key stakeholders to achieve its greatest impact (**Table 2**), as seen with widely adopted reporting guidelines, such as CONSORT (51). We will post the names of organizations that have endorsed SPIRIT 2013 on the SPIRIT Web site (www.spirit-statement.org) and provide resources to facilitate implementation. Widespread adoption of the SPIRIT recommendations can help improve protocol drafting, content, and implementation; facilitate registration, efficiency, and appraisal of trials; and ultimately enhance transparency for the benefit of patient care.

From Women's College Research Institute, Women's College Hospital, and Keenan Research Centre at the Li Ka Shing Knowledge Institute of St. Michael's Hospital, University of Toronto, Toronto, Ontario, Canada; Ottawa Methods Centre, Ottawa Hospital Research Institute, Ethics Office, Canadian Institutes of Health Research, and University of Ottawa, Ottawa, Ontario, Canada; Centre for Statistics in Medicine, University of Oxford, Oxford, United Kingdom; Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark; University of Utah School of Medicine, Salt Lake City, Utah; Center for Clinical Trials, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; Janssen Research & Development, Janssen Pharmaceutical Companies of Johnson & Johnson, Titusville, New Jersey; UK Medical Research Council Clinical Trials Unit, *The Lancet*, and *BMJ*, London, United Kingdom; NCIC Clinical Trials Group, Cancer Research Institute, Queen's University, Kingston, Ontario, Canada; Quantitative Sciences, FHI 360, and GlaxoSmithKline, Research Triangle Park, North Carolina; The Dartmouth Institute for Health Policy and Clinical Practice, The Geisel School of Medicine at Dartmouth, Hanover, New Hampshire; The PR Lee Institute for Health Policy Studies, University of California, San Francisco, San Francisco, California.

Disclaimer: Dr. Krleža-Jerić was formerly employed by the Canadian Institutes of Health Research (Knowledge Translation Branch), and Dr. Parulekar is affiliated with the NCIC Clinical Trials Group. The funders otherwise had no input into the design and conduct of the project; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. Dr. Berlin is employed by the Janssen Pharmaceutical Companies of Johnson & Johnson, Dr. Sox is Editor Emeritus of *Annals of Internal Medicine*, and Dr. Rockhold is employed by GlaxoSmithKline.

Acknowledgment: The authors thank Drs. Mona Loufty and Patricia Parkin for pilot-testing the SPIRIT checklist with their graduate course students. The authors also acknowledge the participation of Dr. Geneviève Dubois-Flynn in the 2009 SPIRIT meeting.

Grant Support: Financial support for the SPIRIT meetings was provided by the Canadian Institutes of Health Research (grant DET-106068), National Cancer Institute of Canada (now Canadian Cancer Society Research Institute), and Canadian Agency for Drugs and Technologies in Health. The Canadian Institutes of Health Research has also funded ongoing dissemination activities (grant MET-117434).

Potential Conflicts of Interest: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M12-1905.

Requests for Single Reprints: An-Wen Chan, MD, DPhil, Women's College Research Institute, Women's College Hospital, University of Toronto, 790 Bay Street, Toronto, Ontario M5G 1N8, Canada; e-mail, anwen.chan@utoronto.ca.

Current author addresses and author contributions are available at www.annals.org.

References

1. **Rennie D.** Trial registration: a great idea switches from ignored to irresistible. JAMA. 2004;292:1359-62. [PMID: 15355937]
2. Strengthening the credibility of clinical research [Editorial]. Lancet. 2010;375:1225. [PMID: 20382309]
3. **Summerskill W, Collingridge D, Frankish H.** Protocols, probity, and publication. Lancet. 2009;373:992. [PMID: 19304003]
4. **Jones G, Abbasi K.** Trial protocols at the BMJ [Editorial]. BMJ. 2004;329:1360. [PMID: 15591544]
5. **Groves T.** Let SPIRIT take you . . . towards much clearer trial protocols. BMJ Group Blogs. 25 September 2009. Accessed at http://blogs.bmj.com/bmj/2009/09/25/trish-groves-let-spirit-take-you-towards-much-clearer-trial-protocols/ on 1 October 2012.
6. **Altman DG, Furberg CD, Grimshaw JM, Rothwell PM.** Lead editorial: trials-using the opportunities of electronic publishing to improve the reporting of randomised trials [Editorial]. Trials. 2006;7:6. [PMID: 16556322]
7. **Turner EH.** A taxpayer-funded clinical trials registry and results database [Editorial]. PLoS Med. 2004;1:e60. [PMID: 15562322]
8. **Coultas D.** Ethical considerations in the interpretation and communication of clinical trial results. Proc Am Thorac Soc. 2007;4:194-8. [PMID: 17494731]
9. **Siegel JP.** Editorial review of protocols for clinical trials [Letter]. N Engl J Med. 1990;323:1355. [PMID: 2215630]
10. **Murray GD.** Research governance must focus on research training. BMJ. 2001;322:1461-2.
11. **Chan A-W.** Access to clinical trial data [Editorial]. BMJ. 2011;342:d80. [PMID: 21228020]
12. **Miller JD.** Registering clinical trial results: the next step. JAMA. 2010;303:773-4. [PMID: 20179288]
13. **Krleža-Jerić K, Chan A-W, Dickersin K, Sim I, Grimshaw J, Gluud C.** Principles for international registration of protocol information and results from human trials of health related interventions: Ottawa statement (part 1). BMJ. 2005;330:956-8. [PMID: 15845980]
14. **Lassere M, Johnson K.** The power of the protocol. Lancet. 2002;360:1620-2. [PMID: 12457782]
15. **Getz KA, Zuckerman R, Cropp AB, Hindle AL, Krauss R, Kaitin KI.** Measuring the incidence, causes, and repercussions of protocol amendments. Drug Inf J. 2011;45:265-75.
16. Public Citizen Health Research Group v. Food and Drug Administration, 964 F Supp. 413 (DDC 1997).
17. **Tetzlaff JM, Chan A-W, Kitchen J, Sampson M, Tricco A, Moher D.** Guidelines for randomized clinical trial protocol content: a systematic review. Syst Rev. 2012;1:43. [PMID: 23006870]
18. **Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG.** Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA. 2004;291:2457-65. [PMID: 15161896]
19. **Smyth RM, Kirkham JJ, Jacoby A, Altman DG, Gamble C, Williamson PR.** Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists. BMJ. 2011;342:c7153. [PMID: 21212122]
20. **Pildal J, Chan A-W, Hróbjartsson A, Forfang E, Altman DG, Gøtzsche PC.** Comparison of descriptions of allocation concealment in trial protocols and the published reports: cohort study. BMJ. 2005;330:1049. [PMID: 15817527]
21. **Mhaskar R, Djulbegovic B, Magazin A, Soares HP, Kumar A.** Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. J Clin Epidemiol. 2012;65:602-9. [PMID: 22424985]
22. **Hróbjartsson A, Pildal J, Chan A-W, Haahr MT, Altman DG, Gøtzsche PC.** Reporting on blinding in trial protocols and corresponding publications was often inadequate but rarely contradictory. J Clin Epidemiol. 2009;62:967-73. [PMID: 19635403]
23. **Scharf O, Colevas AD.** Adverse event reporting in publications compared with sponsor database for cancer clinical trials. J Clin Oncol. 2006;24:3933-8. [PMID: 16921045]
24. **Chan A-W, Hróbjartsson A, Jørgensen KJ, Gøtzsche PC, Altman DG.** Discrepancies in sample size calculations and data analyses reported in randomised trials: comparison of publications with protocols. BMJ. 2008;337:a2299. [PMID: 19056791]
25. **Al-Marzouki S, Roberts I, Evans S, Marshall T.** Selective reporting in clinical trials: analysis of trial protocols accepted by *The Lancet* [Letter]. Lancet. 2008;372:201. [PMID: 18640445]
26. **Hernández AV, Steyerberg EW, Taylor GS, Marmarou A, Habbema JD, Maas AI.** Subgroup analysis and covariate adjustment in randomized clinical trials of traumatic brain injury: a systematic review. Neurosurgery. 2005;57:1244-53. [PMID: 16331173]
27. **Gøtzsche PC, Hróbjartsson A, Johansen HK, Haahr MT, Altman DG, Chan A-W.** Constraints on publication rights in industry-initiated clinical trials [Letter]. JAMA. 2006;295:1645-6. [PMID: 16609085]
28. **Gøtzsche PC, Hróbjartsson A, Johansen HK, Haahr MT, Altman DG, Chan A-W.** Ghost authorship in industry-initiated randomised trials. PLoS Med. 2007;4:e19. [PMID: 17227134]
29. **Lundh A, Krogsbøll LT, Gøtzsche PC.** Access to data in industry-sponsored trials [Letter]. Lancet. 2011;378:1995-6. [PMID: 22153200]
30. **Hopewell S, Dutton S, Yu LM, Chan A-W, Altman DG.** The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. BMJ. 2010;340:c723. [PMID: 20332510]
31. **Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al.** SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ. 2013;346:e7586.
32. **Moher D, Schulz KF, Simera I, Altman DG.** Guidance for developers of health research reporting guidelines. PLoS Med. 2010;7:e1000217. [PMID: 20169112]
33. **Tetzlaff JM, Moher D, Chan A-W.** Developing a guideline for clinical trial protocol content: Delphi consensus survey. Trials. 2012;13:176. [PMID: 23006145]
34. **World Medical Association.** WMA Declaration of Helsinki-Ethical Principles for Medical Research Involving Human Subjects. Accessed at www.wma.net/en/30publications/10policies/b3/index.html on 1 October 2012.
35. **International Conference on Harmonisation.** ICH Harmonised Tripartite Guideline: General Considerations for Clinical Trials: E8. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. 17 July 1997. Accessed at www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E8/Step4/E8_Guideline.pdf on 1 October 2012.
36. **International Conference on Harmonisation.** ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials: E9. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. 5 February 1998. Accessed at www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf on 1 October 2012.
37. **International Conference on Harmonisation.** ICH Harmonised Tripartite Guideline: Guideline for Good Clinical Practice: E6. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. 10 June 1996. Accessed at www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6_R1/Step4/E6_R1__Guideline.pdf on 1 October 2012.
38. **Grimes DA, Hubacher D, Nanda K, Schulz KF, Moher D, Altman DG.** The Good Clinical Practice guideline: a bronze standard for clinical research. Lancet. 2005;366:172-4. [PMID: 16005342]
39. **Sim I, Chan A-W, Gülmezoglu AM, Evans T, Pang T.** Clinical trial registration: transparency is the watchword. Lancet. 2006;367:1631-3. [PMID: 16714166]
40. **Laine C, De Angelis C, Delamothe T, Drazen JM, Frizelle FA, Haug C, et al.** Clinical trial registration: looking back and moving ahead [Editorial]. Ann Intern Med. 2007;147:275-7. [PMID: 17548404]
41. Food and Drug Administration Amendments Act of 2007, HR 2580, 110th Congress, 1st Sess, Title VIII, §801 (2007). Expanded Clinical Trial Registry

Data Bank. Accessed at www.govtrack.us/congress/billtext.xpd?bill=h110-3580 on 1 October 2012.
42. **European Commission.** Communication from the Commission regarding the guideline on the data fields contained in the clinical trials database provided for in Article 11 of Directive 2001/20/EC to be included in the database on medicinal products provided for in Article 57 of Regulation (EC) No 726/2004 (2008/C 168/02). Official Journal of the European Union. 2008;51:3-4.
43. **Schulz KF, Altman DG, Moher D; CONSORT Group.** CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med. 2010;152:726-32. [PMID: 20335313]
44. **National Research Council.** A National Cancer Clinical Trials System for the 21st Century: Reinvigorating the NCI Cooperative Group Program. Washington, DC: National Academies Pr; 2010.
45. **Chan A-W.** Out of sight but not out of mind: how to search for unpublished clinical trial evidence. BMJ. 2012;344:d8013. [PMID: 22214892]
46. **Wieseler B, Kerekes MF, Vervoelgyi V, McGauran N, Kaiser T.** Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. BMJ. 2012;344:d8141. [PMID: 22214759]
47. **Reveiz L, Chan A-W, Krleža-Jerić K, Granados CE, Pinart M, Etxeandia I, et al.** Reporting of methodologic information on trial registries for quality assessment: a study of trial records retrieved from the WHO search portal. PLoS One. 2010;5:e12484. [PMID: 20824212]
48. **Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR.** Comparison of protocols and registry entries to published reports for randomised controlled trials. Cochrane Database Syst Rev. 2011:MR000031. [PMID: 21249714]
49. **Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan A-W, Cronin E, et al.** Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS One. 2008;3:e3081. [PMID: 18769481]
50. **GlaxoSmithKline.** Public disclosure of clinical research. Global Public Policy Issues. October 2011. Accessed at www.gsk.com/policies/GSK-on-disclosure-of-clinical-trial-information.pdf on 1 October 2012.
51. **Turner L, Shamseer L, Altman DG, Schulz KF, Moher D.** Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. Syst Rev. 2012;1:60. [PMID: 23194585]

Current Author Addresses: Dr. Chan: Women's College Research Institute, Women's College Hospital, University of Toronto, 790 Bay Street, Toronto, Ontario M5G 1N8, Canada.
Ms. Tetzlaff: Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, 501 Smyth Road, Ottawa, Ontario K1H 8L6, Canada.
Dr. Altman: Centre for Statistics in Medicine, University of Oxford, Wolfson College Annexe, Linton Road, Oxford OX2 6UD, United Kingdom.
Dr. Laupacis: Keenan Research Centre at the Li Ka Shing Knowledge Institute of St. Michael's Hospital, 30 Bond Street, Toronto, Ontario M5B 1W8, Canada.
Drs. Gøtzsche and Hróbjartsson: Nordic Cochrane Centre, Rigshospitalet Department 3343, Blegdamsvej 9, 2100 Copenhagen Ø, Denmark.
Dr. Krleža-Jerić: Department of Epidemiology and Community Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Ontario K1H 8M5, Canada.
Dr. Mann: Division of Medical Ethics and Humanities, University of Utah School of Medicine, 75 South 2000 East #108, Salt Lake City, UT 84132.
Dr. Dickersin: Center for Clinical Trials, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Mail Room W5010, Baltimore, MD 21205.
Dr. Berlin: Janssen Research & Development, Janssen Pharmaceutical Companies of Johnson & Johnson, 1125 Trenton Harbourton Road, Titusville, NJ 08560.
Ms. Doré: UK Medical Research Council Clinical Trials Unit, 125 Kingsway, London WC2B 6NH, United Kingdom.
Dr. Parulekar: NCIC Clinical Trials Group, Cancer Research Institute, Queen's University, 10 Stuart Street, Kingston, Ontario K7L 3N6, Canada.
Dr. Summerskill: *The Lancet*, 32 Jamestown Road, London NW1 7BY, United Kingdom.
Dr. Groves: *BMJ*, BMA House, Tavistock Square, London WC1H 9JP, United Kingdom.
Dr. Schulz: Quantitative Sciences, FHI 360, Research Triangle Park, 2224 East NC Highway 54, Durham, NC 27713.
Dr. Sox: The Dartmouth Institute for Health Policy and Clinical Practice, The Geisel School of Medicine at Dartmouth, HB 7500, Hanover, NH 03755.
Dr. Rockhold: GlaxoSmithKline, 5 Moore Drive, PO Box 13398, Research Triangle Park, NC 27709.
Dr. Rennie: The Philip R. Lee Institute for Health Policy Studies, University of California, San Francisco, 3333 California Street, Laurel Heights 265, San Francisco, CA 94143-0936.
Dr. Moher: Clinical Epidemiology Program, Ottawa Hospital Research Institute, Department of Epidemiology and Community Medicine, University of Ottawa, 501 Smyth Road, Ottawa, Ontario K1H 8L6, Canada.

Author Contributions: Conception and design: A.-W. Chan, J.M. Tetzlaff, D.G. Altman, A. Laupacis, P.C. Gøtzsche, K. Krleža-Jerić, A. Hróbjartsson, H. Mann, K. Dickersin, J.A. Berlin, W.R. Parulekar, K.F. Schulz, H.C. Sox, D. Rennie, D. Moher.
Analysis and interpretation of the data: A.-W. Chan, J.M. Tetzlaff, D.G. Altman, A. Laupacis, P.C. Gøtzsche, K. Krleža-Jerić, A. Hróbjartsson, K. Dickersin, C.J. Doré, W.R. Parulekar, T. Groves, K.F. Schulz, F.W. Rockhold, D. Rennie, D. Moher.
Drafting of the article: A.-W. Chan, J.M. Tetzlaff, P.C. Gøtzsche, H. Mann, K. Dickersin, J.A. Berlin, C.J. Doré, W.R. Parulekar, K.F. Schulz.
Critical revision of the article for important intellectual content: A.-W. Chan, J.M. Tetzlaff, D.G. Altman, A. Laupacis, K. Krleža-Jerić, A. Hróbjartsson, H. Mann, K. Dickersin, J.A. Berlin, C.J. Doré, W.R. Parulekar, W.S.M. Summerskill, T. Groves, K.F. Schulz, H.C. Sox, F.W. Rockhold, D. Rennie, D. Moher.
Final approval of the article: A.-W. Chan, J.M. Tetzlaff, D.G. Altman, A. Laupacis, P.C. Gøtzsche, K. Krleža-Jerić, A. Hróbjartsson, H. Mann, K. Dickersin, J.A. Berlin, C.J. Doré, W.R. Parulekar, W.S.M. Summerskill, T. Groves, K.F. Schulz, H.C. Sox, F.W. Rockhold, D. Rennie, D. Moher.
Provision of study materials or patients: K. Krleža-Jerić, K. Dickersin.
Statistical expertise: D.G. Altman, P.C. Gøtzsche, C.J. Doré, K.F. Schulz, F.W. Rockhold.
Obtaining of funding: A.-W. Chan, A. Laupacis, D. Moher.
Administrative, technical, or logistic support: A.-W. Chan, P.C. Gøtzsche, K. Krleža-Jerić.
Collection and assembly of data: A.-W. Chan, J.M. Tetzlaff, P.C. Gøtzsche, A. Hróbjartsson, K. Dickersin, C.J. Doré, W.R. Parulekar, K.F. Schulz.

Appendix Table. **World Health Organization Trial Registration Data Set***

Item	Description
1. Primary registry and trial-identifying number	Name of primary registry and the unique identifier assigned by the primary registry
2. Date of registration in primary registry	Date when the trial was officially registered in the primary registry
3. Secondary identifying numbers	Other identifiers, if any Universal Trial Number Identifiers assigned by the sponsor Other trial registration numbers issued by other registries Identifiers issued by funding bodies, collaborative research groups, regulatory authorities, ethics committees/ institutional review boards, etc.
4. Sources of monetary or material support	Major sources of monetary or material support for the trial (e.g., funding agency, foundation, company, institution)
5. Primary sponsor	Person, organization, group, or other legal entity that takes responsibility for initiating and managing a study
6. Secondary sponsor(s)	Additional persons, organizations, or other legal persons, if any, who have agreed with the primary sponsor to take on responsibilities of sponsorship
7. Contact for public queries	E-mail address, telephone number, and postal address of the contact who will respond to general queries, including information about current recruitment status
8. Contact for scientific queries	Name and title, e-mail address, telephone number, postal address, and affiliation of the principal investigator and e-mail address, telephone number, postal address, and affiliation of the contact for scientific queries about the trial (if applicable)
9. Public title	Title intended for the lay public in easily understood language
10. Scientific title	Scientific title of the study as it appears in the protocol submitted for funding and ethical review; include trial acronym, if available
11. Countries of recruitment	Countries from which participants will be recruited
12. Health condition(s) or problem(s) studied	Primary health condition(s) or problem(s) studied (e.g., depression, breast cancer, medication error)
13. Intervention(s)	For each group of the trial, record a brief intervention name plus an intervention description Intervention name: For drugs, use the generic name; for other types of interventions, provide a brief descriptive name Intervention description: Must be sufficiently detailed for it to be possible to distinguish between the groups of a study; for example, interventions involving drugs may include dosage form, dosage, frequency, and duration
14. Key inclusion and exclusion criteria	Inclusion and exclusion criteria for participant selection, including age and sex
15. Study type	Method of allocation (randomized/nonrandomized) Blinding/masking (identify who is blinded) Assignment (e.g., single group, parallel, crossover, factorial) Purpose Phase (if applicable) For randomized trials: Method of sequence generation and allocation concealment
16. Date of first enrollment	Anticipated or actual date of enrollment of the first participant
17. Target sample size	Total number of participants to enroll
18. Recruitment status	Pending: Participants are not yet being recruited or enrolled at any site Recruiting Suspended: Temporary halt in recruitment and enrollment Complete: Participants are no longer being recruited or enrolled Other
19. Primary outcome(s)	The primary outcome should be the outcome used in sample size calculations or the main outcome used to determine the effects of the intervention For each primary outcome provide: Name of the outcome (do not use abbreviations) Metric or method of measurement used (be as specific as possible) Time point of primary interest
20. Key secondary outcome(s)	As for primary outcomes, for each secondary outcome provide: Name of the outcome (do not use abbreviations) Metric or method of measurement used (be as specific as possible) Time point of interest

* Adapted from www.who.int/ictrp/network/trds/en/index.html.

CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials

Kenneth F Schulz,[1] Douglas G Altman,[2] David Moher,[3] for the CONSORT Group

EDITORIAL by Antes
RESEARCH, p 697

[1]Family Health International, Research Triangle Park, NC 27709, USA
[2]Centre for Statistics in Medicine, University of Oxford, Wolfson College, Oxford
[3]Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada
Correspondence to: K F Schulz
kschulz@fhi.org

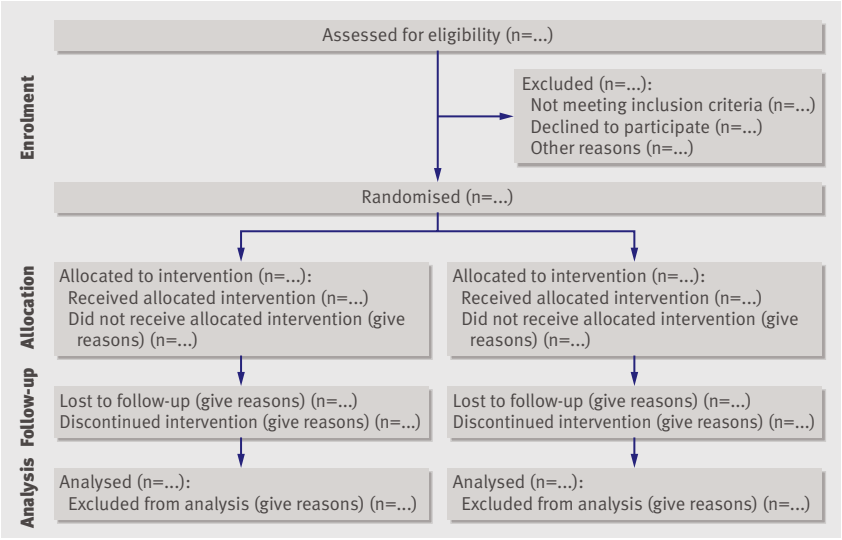
Accepted: 9 December 2009

Cite this as: *BMJ* 2010;340:c332
doi: 10.1136/bmj.c332

The CONSORT statement is used worldwide to improve the reporting of randomised controlled trials. **Kenneth Schulz and colleagues** describe the latest version, CONSORT 2010, which updates the reporting guideline based on new methodological evidence and accumulating experience

Randomised controlled trials, when appropriately designed, conducted, and reported, represent the gold standard in evaluating healthcare interventions. However, randomised trials can yield biased results if they lack methodological rigour.[1] To assess a trial accurately, readers of a published report need complete, clear, and transparent information on its methodology and findings. Unfortunately, attempted assessments frequently fail because authors of many trial reports neglect to provide lucid and complete descriptions of that critical information.[2-4]

That lack of adequate reporting fuelled the development of the original CONSORT (Consolidated Standards of Reporting Trials) statement in 1996[5] and its revision five years later.[6-8] While those statements improved the reporting quality for some randomised controlled trials,[9 10] many trial reports still remain inadequate.[2] Furthermore, new methodological evidence and additional experience has accumulated since the last revision in 2001. Consequently, we organised a CONSORT Group meeting to update the 2001 statement.[6-8] We introduce here the result of that process, CONSORT 2010.



Flow diagram of the progress through the phases of a parallel randomised trial of two groups (that is, enrolment, intervention allocation, follow-up, and data analysis)

Intent of CONSORT 2010

The CONSORT 2010 Statement is this paper including the 25 item checklist in the table and the flow diagram. It provides guidance for reporting all randomised controlled trials, but focuses on the most common design type—individually randomised, two group, parallel trials. Other trial designs, such as cluster randomised trials and non-inferiority trials, require varying amounts of additional information. CONSORT extensions for these designs,[11 12] and other CONSORT products, can be found through the CONSORT website (www.consort-statement.org). Along with the CONSORT statement, we have updated the explanation and elaboration article,[13] which explains the inclusion of each checklist item, provides methodological background, and gives published examples of transparent reporting.

Diligent adherence by authors to the checklist items facilitates clarity, completeness, and transparency of reporting. Explicit descriptions, not ambiguity or omission, best serve the interests of all readers. Note that the CONSORT 2010 Statement does not include recommendations for designing, conducting, and analysing trials. It solely addresses the reporting of what was done and what was found.

Nevertheless, CONSORT does indirectly affect design and conduct. Transparent reporting reveals deficiencies in research if they exist. Thus, investigators who conduct inadequate trials, but who must transparently report, should not be able to pass through the publication process without revelation of their trial's inadequacies. That emerging reality should provide impetus to improved trial design and conduct in the future, a secondary indirect goal of our work. Moreover, CONSORT can help researchers in designing their trial.

Background to CONSORT

Efforts to improve the reporting of randomised controlled trials accelerated in the mid-1990s, spurred partly by methodological research. Researchers had shown for many years that authors reported such trials poorly, and empirical evidence began to accumulate that some poorly conducted or poorly reported aspects of trials were associated with bias.[14] Two initiatives aimed at developing reporting guidelines culminated in one of us (DM) and Drummond Rennie organising the first CONSORT statement in 1996.[5]

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item
Title and abstract		
	1a	Identification as a randomised trial in the title
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts[21 31])
Introduction		
Background and objectives	2a	Scientific background and explanation of rationale
	2b	Specific objectives or hypotheses
Methods		
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons
Participants	4a	Eligibility criteria for participants
	4b	Settings and locations where the data were collected
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed
	6b	Any changes to trial outcomes after the trial commenced, with reasons
Sample size	7a	How sample size was determined
	7b	When applicable, explanation of any interim analyses and stopping guidelines
Randomisation:		
Sequence generation	8a	Method used to generate the random allocation sequence
	8b	Type of randomisation; details of any restriction (such as blocking and block size)
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how
	11b	If relevant, description of the similarity of interventions
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses
Results		
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome
	13b	For each group, losses and exclusions after randomisation, together with reasons
Recruitment	14a	Dates defining the periods of recruitment and follow-up
	14b	Why the trial ended or was stopped
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[28])
Discussion		
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses
Generalisability	21	Generalisability (external validity, applicability) of the trial findings
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence
Other information		
Registration	23	Registration number and name of trial registry
Protocol	24	Where the full trial protocol can be accessed, if available
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration[13] for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials,[11] non-inferiority and equivalence trials,[12] non-pharmacological treatments,[32] herbal interventions,[33] and pragmatic trials.[34] Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

Further methodologi cal research on similar topics reinforced earlier findings[15] and fed into the revision of 2001.[6-8] Subsequently, the expanding body of methodological research informed the refinement of CONSORT 2010. More than 700 studies comprise the CONSORT database (located on the CONSORT website), which provides the empirical evidence to underpin the CONSORT initiative.

Indeed, CONSORT Group members continually monitor the literature. Information gleaned from these efforts provides an evidence base on which to update the CONSORT statement. We add, drop, or modify items based on that evidence and the recommendations of the CONSORT Group, an international and eclectic group of clinical trialists, statisticians, epidemiologists, and biomedical editors. The CONSORT Executive (KFS, DGA, DM) strives for a balance of established and emerging researchers. The membership of the group is dynamic. As our work expands in response to emerging projects and needed expertise, we invite new members to contribute. As such, CONSORT continually assimilates new ideas and perspectives. That process informs the continually evolving CONSORT statement.

Over time, CONSORT has garnered much support. More than 400 journals, published around the world and in many languages, have explicitly supported the CONSORT

statement. Many other healthcare journals support it without our knowledge. Moreover, thousands more have implicitly supported it with the endorsement of the CONSORT statement by the International Committee of Medical Journal Editors (www.icmje.org). Other prominent editorial groups, the Council of Science Editors and the World Association of Medical Editors, officially support CONSORT. That support seems warranted: when used by authors and journals, CONSORT seems to improve reporting.[9]

Development of CONSORT 2010

Thirty one members of the CONSORT 2010 Group met in Montebello, Canada, in January 2007 to update the 2001 CONSORT statement. In addition to the accumulating evidence relating to existing checklist items, several new issues had come to prominence since 2001. Some participants were given primary responsibility for aggregating and synthesising the relevant evidence on a particular checklist item of interest. Based on that evidence, the group deliberated the value of each item. As in prior CONSORT versions, we kept only those items deemed absolutely fundamental to reporting a randomised controlled trial. Moreover, an item may be fundamental to a trial but not included, such as approval by an institutional ethical review board, because funding bodies strictly enforce ethical review and medical journals usually address reporting ethical review in their instructions for authors. Other items may seem desirable, such as reporting on whether on-site monitoring was done, but a lack of empirical evidence or any consensus on their value cautions against inclusion at this point. The CONSORT 2010 Statement thus addresses the minimum criteria, although that should not deter authors from including other information if they consider it important.

After the meeting, the CONSORT Executive convened teleconferences and meetings to revise the checklist. After seven major iterations, a revised checklist was distributed to the larger group for feedback. With that feedback, the executive met twice in person to consider all the comments and to produce a penultimate version. That served as the basis for writing the first draft of this paper, which was then distributed to the group for feedback. After consideration of their comments, the executive finalised the statement.

The CONSORT Executive then drafted an updated explanation and elaboration manuscript, with assistance from other members of the larger group. The substance of the 2007 CONSORT meeting provided the material for the update. The updated explanation and elaboration manuscript was distributed to the entire group for additions, deletions, and changes. That final iterative process converged to the CONSORT 2010 Explanation and Elaboration.[13]

Box 1 | Noteworthy general changes in CONSORT 2010 Statement

- We simplified and clarified the wording, such as in items 1, 8, 10, 13, 15, 16, 18, 19, and 21
- We improved consistency of style across the items by removing the imperative verbs that were in the 2001 version
- We enhanced specificity of appraisal by breaking some items into sub-items. Many journals expect authors to complete a CONSORT checklist indicating where in the manuscript the items have been addressed. Experience with the checklist noted pragmatic difficulties when an item comprised multiple elements. For example, item 4 addresses eligibility of participants and the settings and locations of data collection. With the 2001 version, an author could provide a page number for that item on the checklist, but might have reported only eligibility in the paper, for example, and not reported the settings and locations. CONSORT 2010 relieves obfuscations and forces authors to provide page numbers in the checklist for both eligibility and settings

Changes in CONSORT 2010

The revision process resulted in evolutionary, not revolutionary, changes to the checklist (table), and the flow diagram was not modified except for one word (figure). Moreover, because other reporting guidelines augmenting the checklist refer to item numbers, we kept the existing items under their previous item numbers except for some renumbering of items 2 to 5. We added additional items either as a sub-item under an existing item, an entirely new item number at the end of the checklist, or (with item 3) an interjected item into a renumbered segment. We have summarised the noteworthy general changes in box 1 and specific changes in box 2. The CONSORT website contains a side by side comparison of the 2001 and 2010 versions.

Implications and limitations

We developed CONSORT 2010 to assist authors in writing reports of randomised controlled trials, editors and peer reviewers in reviewing manuscripts for publication, and readers in critically appraising published articles. The CONSORT 2010 Explanation and Elaboration provides elucidation and context to the checklist items. We strongly recommend using the explanation and elaboration in conjunction with the checklist to foster complete, clear, and transparent reporting and aid appraisal of published trial reports.

CONSORT 2010 focuses predominantly on the two group, parallel randomised controlled trial, which accounts for over half of trials in the literature.[2] Most of the items from the CONSORT 2010 Statement, however, pertain to all types of randomised trials. Nevertheless, some types of trials or trial situations dictate the need for additional information in the trial report. When in doubt, authors, editors, and readers should consult the CONSORT website for any CONSORT extensions, expansions (amplifications), implementations, or other guidance that may be relevant.

The evidence based approach we have used for CONSORT also served as a model for development of other reporting guidelines, such as for reporting systematic reviews and meta-analyses of studies evaluating interventions,[16] diagnostic studies,[17] and observational studies.[18] The explicit goal of all these initiatives is to improve reporting. The Enhancing the Quality and Transparency of Health Research (EQUATOR) Network will facilitate development of reporting guidelines and help disseminate the guidelines: www.equator-network.org provides information on all reporting guidelines in health research.

With CONSORT 2010, we again intentionally declined to produce a rigid structure for the reporting of randomised trials. Indeed, SORT[19] tried a rigid format, and it failed in a pilot run with an editor and authors.[20] Consequently, the format of articles should abide by journal style, editorial directions, the traditions of the research field addressed, and, where possible, author preferences. We do not wish to standardise the structure of reporting. Authors should simply address checklist items somewhere in the article, with ample detail and lucidity. That stated, we think that manuscripts benefit from frequent subheadings within the major sections, especially the methods and results sections.

Box 2 | Noteworthy specific changes in CONSORT 2010 Statement

- *Item 1b (title and abstract)*—We added a sub-item on providing a structured summary of trial design, methods, results, and conclusions and referenced the CONSORT for abstracts article[21]
- *Item 2b (introduction)*—We added a new sub-item (formerly item 5 in CONSORT 2001) on "Specific objectives or hypotheses"
- *Item 3a (trial design)*—We added a new item including this sub-item to clarify the basic trial design (such as parallel group, crossover, cluster) and the allocation ratio
- *Item 3b (trial design)*—We added a new sub-item that addresses any important changes to methods after trial commencement, with a discussion of reasons
- *Item 4 (participants)*—Formerly item 3 in CONSORT 2001
- *Item 5 (interventions)*—Formerly item 4 in CONSORT 2001. We encouraged greater specificity by stating that descriptions of interventions should include "sufficient details to allow replication"[3]
- *Item 6 (outcomes)*—We added a sub-item on identifying any changes to the primary and secondary outcome (endpoint) measures after the trial started. This followed from empirical evidence that authors frequently provide analyses of outcomes in their published papers that were not the prespecified primary and secondary outcomes in their protocols, while ignoring their prespecified outcomes (that is, selective outcome reporting).[4 22] We eliminated text on any methods used to enhance the quality of measurements
- *Item 9 (allocation concealment mechanism)*—We reworded this to include mechanism in both the report topic and the descriptor to reinforce that authors should report the actual steps taken to ensure allocation concealment rather than simply report imprecise, perhaps banal, assurances of concealment
- *Item 11 (blinding)*—We added the specification of how blinding was done and, if relevant, a description of the similarity of interventions and procedures. We also eliminated text on "how the success of blinding (masking) was assessed" because of a lack of empirical evidence supporting the practice as well as theoretical concerns about the validity of any such assessment[23 24]
- *Item 12a (statistical methods)*—We added that statistical methods should also be provided for analysis of secondary outcomes
- *Sub-item 14b (recruitment)*—Based on empirical research, we added a sub-item on "Why the trial ended or was stopped"[25]
- *Item 15 (baseline data)*—We specified "A table" to clarify that baseline and clinical characteristics of each group are most clearly expressed in a table
- *Item 16 (numbers analysed)*—We replaced mention of "intention to treat" analysis, a widely misused term, by a more explicit request for information about retaining participants in their original assigned groups[26]
- *Sub-item 17b (outcomes and estimation)*—For appropriate clinical interpretability, prevailing experience suggested the addition of "For binary outcomes, presentation of both relative and absolute effect sizes is recommended"[27]
- *Item 19 (harms)*—We included a reference to the CONSORT paper on harms[28]
- *Item 20 (limitations)*—We changed the topic from "Interpretation" and supplanted the prior text with a sentence focusing on the reporting of sources of potential bias and imprecision
- *Item 22 (interpretation)*—We changed the topic from "Overall evidence." Indeed, we understand that authors should be allowed leeway for interpretation under this nebulous heading. However, the CONSORT Group expressed concerns that conclusions in papers frequently misrepresented the actual analytical results and that harms were ignored or marginalised. Therefore, we changed the checklist item to include the concepts of results matching interpretations and of benefits being balanced with harms
- *Item 23 (registration)*—We added a new item on trial registration. Empirical evidence supports the need for trial registration, and recent requirements by journal editors have fostered compliance[29]
- *Item 24 (protocol)*—We added a new item on availability of the trial protocol. Empirical evidence suggests that authors often ignore, in the conduct and reporting of their trial, what they stated in the protocol.[4 22] Hence, availability of the protocol can instigate adherence to the protocol before publication and facilitate assessment of adherence after publication
- *Item 25 (funding)*—We added a new item on funding. Empirical evidence points toward funding source sometimes being associated with estimated treatment effects[30]

CONSORT urges completeness, clarity, and transparency of reporting, which simply reflects the actual trial design and conduct. However, as a potential drawback, a reporting guideline might encourage some authors to report fictitiously the information suggested by the guidance rather than what was actually done. Authors, peer reviewers, and editors should vigilantly guard against that potential drawback and refer, for example, to trial protocols, to information on trial registers, and to regulatory agency websites. Moreover, the CONSORT 2010 Statement does not include recommendations for designing and conducting randomised trials. The items should elicit clear pronouncements of how and what the authors did, but do not contain any judgments on how and what the authors should have done. Thus, CONSORT 2010 is not intended as an instrument to evaluate the quality of a trial. Nor is it appropriate to use the checklist to construct a "quality score."

Nevertheless, we suggest that researchers begin trials with their end publication in mind. Poor reporting allows authors, intentionally or inadvertently, to escape scrutiny of any weak aspects of their trials. However, with wide adoption of CONSORT by journals and editorial groups, most authors should have to report transparently all important aspects of their trial. The ensuing scrutiny rewards well conducted trials and penalises poorly conducted trials. Thus, investigators should understand the CONSORT 2010 reporting guidelines before starting a trial as a further incentive to design and conduct their trials according to rigorous standards.

CONSORT 2010 supplants the prior version published in 2001. Any support for the earlier version accumulated from journals or editorial groups will automatically extend to this newer version, unless specifically requested otherwise. Journals that do not currently support CONSORT may do so by registering on the CONSORT website. If a journal supports or endorses CONSORT 2010, it should cite one of the original versions of CONSORT 2010, the CONSORT 2010 Explanation and Elaboration, and the CONSORT website in their "Instructions to authors." We suggest that authors who wish to cite CONSORT should cite this or another of the original journal versions of CONSORT 2010 Statement, and, if appropriate, the CONSORT 2010 Explanation and Elaboration.[13] All CONSORT material can be accessed through the original publishing journals or the CONSORT website. Groups or individuals who desire to translate the CONSORT 2010 Statement into

bmj.com: recent Research Methods & Reporting articles

- Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design (2010;340:c1066)
- Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes (2010;340:b5087)
- Economic impact of disease and injury: counting what matters (2010;340:c924)
- The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews (2010;340:c365)
- Meta-analysis of individual participant data: rationale, conduct, and reporting (2010;340:c221)

other languages should first consult the CONSORT policy statement on the website.

We emphasise that CONSORT 2010 represents an evolving guideline. It requires perpetual reappraisal and, if necessary, modifications. In the future we will further revise the CONSORT material considering comments, criticisms, experiences, and accumulating new evidence. We invite readers to submit recommendations via the CONSORT website.

Author contributions: KFS, DM, and DGA participated in meetings and regular conference calls, planned the CONSORT 2007 meeting at Montebello, developed the agenda, prepared background research, identified and invited participants, contributed to the CONSORT meeting, drafted the manuscript, and, after critical review by the CONSORT Group, finalised the text of the manuscript. Members of the CONSORT Group attended the meeting, except for those noted below, and provided input on and review of the revised checklist and text of this article. Some members also prepared background material.

The CONSORT Group contributors to CONSORT 2010: DG Altman, Centre for Statistics in Medicine, University of Oxford; Virginia Barbour, *PLoS Medicine*; Jesse A Berlin, Johnson & Johnson Pharmaceutical Research and Development, USA; Isabelle Boutron, University Paris 7 Denis Diderot, Assistance Publique des Hôpitaux de Paris, INSERM, France; PJ Devereaux, McMaster University, Canada; Kay Dickersin, Johns Hopkins Bloomberg School of Public Health, USA; Diana Elbourne, London School of Hygiene & Tropical Medicine; Susan Ellenberg, University of Pennsylvania School of Medicine, USA; Val Gebski, University of Sydney, Australia; Steven Goodman, *Journal of the Society for Clinical Trials*, USA; Peter C Gøtzsche, Nordic Cochrane Centre, Denmark; Trish Groves, *BMJ*; Steven Grunberg, American Society of Clinical Oncology, USA; Brian Haynes, McMaster University, Canada; Sally Hopewell, Centre for Statistics in Medicine, University of Oxford; Astrid James, *Lancet*; Peter Juhn, Johnson & Johnson, USA; Philippa Middleton, University of Adelaide, Australia; Don Minckler, University of California Irvine, USA; D Moher, Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Canada; Victor M Montori, Knowledge and Encounter Research Unit, Mayo Clinic College of Medicine, USA; Cynthia Mulrow, *Annals of Internal Medicine*, USA; Stuart Pocock, London School of Hygiene & Tropical Medicine; Drummond Rennie, *JAMA*, USA; David L Schriger, *Annals of Emergency Medicine*, USA; KF Schulz, Family Health International, USA; Iveta Simera, EQUATOR Network; Elizabeth Wager, Sideview.

Contributors to CONSORT 2010 who did not attend the Montebello meeting: Mike Clarke, UK Cochrane Centre; Gordon Guyatt, McMaster University, Canada.

Funding: We received financial support from United Kingdom National Institute for Health Research and the Medical Research Council; Canadian Institutes of Health Research; Presidents Fund, Canadian Institutes of Health Research; Johnson & Johnson; *BMJ*; and the American Society for Clinical Oncology. DGA is supported by Cancer Research UK, DM by a University of Ottawa Research Chair, and KFS by Family Health International. None of the sponsors had any involvement in the planning, execution, or writing of the CONSORT documents. Additionally, no funder played a role in drafting the manuscript.

Competing interests: Uniform disclosure of potential conflicts of interest: all authors have completed the ICMJE unified competing interest form at www.icmje.org/coi_disclosure.pdf (available from the corresponding author) and declare (1) DM received grants for this work from Johnson & Johnson, *BMJ*, and American Society for Clinical Oncology; KFS and DGA received support for travel to meetings for this work from Johnson & Johnson, *BMJ*, and American Society for Clinical Oncology; (2) KFS and DA had travel expenses reimbursed by the EQUATOR Network; KFS has received honoraria for delivering educational presentations for the American Board of Obstetrics and Gynecology Foundation for Excellence in Women's Health Care, Ortho-McNeil Janssen Scientific Affairs, and the American College of Obstetrics and Gynecology; and has done consultancy for Wyeth. All authors also declare (3) no spouses, partners, or children with relationships with commercial entities that might have an interest in the submitted work; (4) no non-financial interests that may be relevant to the submitted work.

1 Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6.
2 Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005;365:1159-62.
3 Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008;336:1472-4.
4 Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE* 2008;3:e3081.
5 Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;276:637-9.
6 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191-4.
7 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med* 2001;134:657-62.
8 Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285:1987-91.
9 Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006;185:263-7.
10 Hopewell S, Dutton S, Yu L-M, Chan A-W, Altman DG. The quality of reports of randomised trials in 2000 and 2006: a comparative study of articles indexed by PubMed. *BMJ* 2010;340:c723.
11 Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004;328:702-8.
12 Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006;295:1152-60.
13 Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869.
14 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
15 Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;352:609-13.
16 Moher D, Liberati A, Tetzlaff J, Altman DG, for the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
17 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;326:41-4.
18 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, for the STROBE Initiative. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;335:806-8.
19 Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. *JAMA* 1994;272:1926-31.
20 Rennie D. Reporting randomized controlled trials. An experiment and a call for responses from readers. *JAMA* 1995;273:1054-5.
21 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008;371:281-3.
22 Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457-65.
23 Sackett DL. Commentary: Measuring the success of blinding in RCTs: don't, must, can't or needn't? *Int J Epidemiol* 2007;36:664-5.
24 Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002;359:696-700.
25 Montori VM, Devereaux PJ, Adhikari NK, Burns KE, Eggert CH, Briel M, et al. Randomized trials stopped early for benefit: a systematic review. *JAMA* 2005;294:2203-9.
26 Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;319:670-4.
27 Nuovo J, Melnikow J, Chang D. Reporting number needed to treat and absolute risk reduction in randomized controlled trials. *JAMA* 2002;287:2813-4.
28 Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;141:781-8.
29 De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *Lancet* 2004;364:911-2.
30 Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;326:1167-70.
31 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 2008;5:e20.
32 Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008;148:295-309.
33 Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann Intern Med* 2006;144:364-7.
34 Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;337:a2390.

BMJ Quality & Safety Online First, published on 14 September 2015 as 10.1136/bmjqs-2015-004411

OPEN ACCESS

SQUIRE 2.0 (*Standards for QUality Improvement Reporting Excellence*): revised publication guidelines from a detailed consensus process

Greg Ogrinc,[1,2,3] Louise Davies,[1,2,3] Daisy Goodman,[1,2] Paul Batalden,[2,3] Frank Davidoff,[3] David Stevens[3,4]

For numbered affiliations see end of article.

Correspondence to
Dr Greg Ogrinc, Associate Chief of Staff for Education, White River Junction VA, 215 North Main St (111), White River Junction, VT, 05009, USA; greg.ogrinc@med.va.gov

This paper is being co-published in: *American Journal of Critical Care, American Journal of Medical Quality, Canadian Journal of Diabetes, Journal of Continuing Education in Nursing, Journal of American College of Surgeons, Journal of Surgical Research, Joint Commission Journal on Quality and Patient Safety, Journal of Nursing Care Quality, The Permanente Journal, GIMBE Evidence for Health (Italy), MedWave (Chile), and Scandinavian Journal of Trauma, Resuscitation, and Emergency Medicine.*

Received 19 May 2015
Revised 11 July 2015
Accepted 17 July 2015

To cite: Ogrinc G, Davies L, Goodman D, *et al*. *BMJ Qual Saf* Published Online First: [*please include* Day Month Year] doi:10.1136/bmjqs-2015-004411

ABSTRACT

Since the publication of Standards for QUality Improvement Reporting Excellence (SQUIRE 1.0) guidelines in 2008, the science of the field has advanced considerably. In this manuscript, we describe the development of SQUIRE 2.0 and its key components. We undertook the revision between 2012 and 2015 using (1) semistructured interviews and focus groups to evaluate SQUIRE 1.0 plus feedback from an international steering group, (2) two face-to-face consensus meetings to develop interim drafts and (3) pilot testing with authors and a public comment period. SQUIRE 2.0 emphasises the reporting of three key components of systematic efforts to improve the quality, value and safety of healthcare: the use of formal and informal theory in planning, implementing and evaluating improvement work; the context in which the work is done and the study of the intervention(s). SQUIRE 2.0 is intended for reporting the range of methods used to improve healthcare, recognising that they can be complex and multidimensional. It provides common ground to share these discoveries in the scholarly literature (http://www.squire-statement.org).

In 2005, draft publication guidelines for quality improvement reporting debuted in *Quality and Safety in Health Care*.[1] At that time, publications of scholarly work about healthcare improvement were often confusing and of limited value. Leaders in the field were working to consolidate the evidence for a science of improvement[2 3] and without guidance on how to write their findings, authors struggled to report their improvement work in a reliable and consistent way.[4 5] These factors influenced the initial publication in 2008 of the Standards for QUality Improvement Reporting Excellence (SQUIRE),[6] which we will refer to as SQUIRE 1.0. The guidelines were developed in an effort to reduce uncertainty about the information deemed to be important in scholarly reports of healthcare improvement and to increase the completeness, precision and transparency of those reports.

In the intervening years, the reach of systematic efforts to improve the quality, safety and value of healthcare has grown. Health professionals' education worldwide now includes improvement as a standard competency.[7–11] The science of the field also continues to advance through guidance on applying formal and informal theory in the development and interpretation of improvement programmes;[12] stronger ways to identify, assess and describe context;[13–16] recommendations for clearer, more complete descriptions of interventions[17] and development of initial guidance on how to study an intervention.[18]

In this setting, we have undertaken a revision of SQUIRE 1.0. When we began, it rapidly became apparent that a wide variety of approaches had developed for improving healthcare, ranging from formative to experimental to evaluative. Rather than limiting the revised guidelines to only a few of these, we fashioned them to be applicable across the many methods that are used. We aimed to reflect the dynamic nature of the field and support its further development. This article describes the development and content of SQUIRE 2.0 (table 1).

SQUIRE 2.0 DEVELOPMENTAL PATH

We developed SQUIRE 2.0 between 2012 and 2015 in three overlapping phases: (1) evaluation of the initial

Table 1 Revised Standards for QUality Improvement Reporting Excellence (SQUIRE 2.0) publication guidelines

Text section and item name	Section or item description
Notes to authors	▶ The SQUIRE guidelines provide a framework for reporting new knowledge about how to improve healthcare. ▶ The SQUIRE guidelines are intended for reports that describe system level work to improve the quality, safety and value of healthcare, and used methods to establish that observed outcomes were due to the intervention(s). ▶ A range of approaches exists for improving healthcare. SQUIRE may be adapted for reporting any of these. ▶ Authors should consider every SQUIRE item, but it may be inappropriate or unnecessary to include every SQUIRE element in a particular manuscript. ▶ The SQUIRE glossary contains definitions of many of the key words in SQUIRE. ▶ The explanation and elaboration document provides specific examples of well-written SQUIRE items and an in-depth explanation of each item. ▶ Please cite SQUIRE when it is used to write a manuscript.
Title and abstract	
1. Title	Indicate that the manuscript concerns an initiative to improve healthcare (broadly defined to include the quality, safety, effectiveness, patient-centredness, timeliness, cost, efficiency and equity of healthcare).
2. Abstract	a. Provide adequate information to aid in searching and indexing. b. Summarise 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.
Introduction	*Why did you start?*
3. Problem description	Nature and significance of the local problem.
4. Available knowledge	Summary of what is currently known about the problem, including relevant previous studies.
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
6. Specific aims	Purpose of the project and of this report.
Methods	*What did you do?*
7. Context	Contextual elements considered important at the outset of introducing the intervention(s).
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.
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).
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 employed for assessing completeness and accuracy of data.
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.
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.
Results	*What did you find?*
13. Results	a. Initial steps of the intervention(s) and their evolution over time (eg, time-line diagram, flow chart or table), including modifications made to the intervention during the project. b. Details of the process measures and outcomes. 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.

Continued

Table 1 Continued

Text section and item name	Section or item description
Discussion	*What does it mean?*
14. Summary	a. Key findings, including relevance to the rationale and specific aims. b. Particular strengths of the project.
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.
16. Limitations	a. Limits to the generalisability 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 minimise and adjust for limitations.
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.
Other information	
18. Funding	Sources of funding that supported this work. Role, if any, of the funding organisation in the design, implementation, interpretation and reporting.

SQUIRE guidelines, (2) early revisions and (3) pilot testing with late revisions.

We began the evaluation of SQUIRE 1.0 by collecting data to assess its clarity and usability.[19] Semistructured interviews and focus groups with 29 end users of SQUIRE 1.0 revealed that many found SQUIRE 1.0 helpful in planning and doing improvement work, but less so in the writing process. This issue was especially apparent in the efforts to write about the cyclic, iterative process that often occurs with improvement interventions. SQUIRE 1.0 was seen by many as unnecessarily complex with too much redundancy and lacking a clear distinction between 'doing improvement' and 'studying the improvement'. A recent independent study and editorial also documented and addressed some of these challenges.[20 21]

In the second phase, we convened an international advisory group of 18 experts that included editors, authors, researchers and improvement professionals. This group met through three conference calls, reviewed SQUIRE 1.0 and the results of the end-user evaluation, and provided detailed feedback on successive revisions. This advisory group and additional participants attended two consensus conferences in 2013 and 2014 where they engaged in intensive analysis and made recommendations that further guided the revision process.

In the third phase, 44 authors used an interim draft version of the updated SQUIRE guidelines to write sections of a manuscript. Each author then provided comments on the utility and understandability of the draft guidelines, and in their submitted section, identified the portions of their writing samples that fulfilled the items of that section.[22] We also obtained detailed feedback about this draft version through semistructured interviews with 11 biomedical journal editors. The data from this phase revealed areas needing further clarification, and which specific items were prone to misinterpretation. Finally, a penultimate draft was emailed to over 450 individuals around the world, including the advisory group, consensus meeting participants, authors, reviewers, editors, faculty in fellowship programmes and trainees. This version was also posted on the SQUIRE website with an invitation for public feedback. We used the information from this process to write SQUIRE 2.0 (table 1).

SQUIRE 2.0

Many publication guidelines, including CONSORT (randomised trials), STROBE (observational studies) and PRISMA (systematic reviews) focus on a particular study methodology (http://www.equator-network.org). In contrast, SQUIRE 2.0 is designed to apply across the many approaches used for systematically improving the quality, safety and value of healthcare. Methods range from iterative changes using plan–do–study–act cycles in single settings to retrospective analyses of large-scale programmes to multisite randomised trials. We encourage authors to apply other

publication guidelines—particularly those that focus on specific study methods—along with SQUIRE, as appropriate. Authors should carefully consider the relevance of each SQUIRE item, but recognise that it is sometimes not necessary, nor even possible, to include each item in a particular manuscript.

SQUIRE 2.0 retains the IMRaD (introduction, methods, results and discussion) structure.[23] Although used primarily for reporting research within a spectrum of study designs, this structure expresses the underlying logic of most systematic investigations, and is familiar to authors, editors, reviewers and readers. We continue to use A. Bradford Hill's four fundamental questions for writing: Why did you start? What did you do? What did you find? What does it mean?[24] In our evaluation of SQUIRE 1.0, novice authors found these questions to be straightforward, clear and useful.

SQUIRE 2.0 contains 18 items, but omits the multiple subitems that were a source of confusion for SQUIRE 1.0 users.[19] A range of approaches exists for improving healthcare, and SQUIRE may be adapted for reporting any of these. As stated above, authors should consider every SQUIRE item, but it may be inappropriate or unnecessary to include every SQUIRE item in a particular manuscript. In addition, authors need not use items in the order in which they appear. Major changes between SQUIRE 1.0 and 2.0 are concentrated in four areas: (1) terminology, (2) theory, (3) context and (4) studying the intervention(s).

Terminology

The elaborate detail in SQUIRE 1.0 was seen by users as both a blessing and a curse[19]: helpful in designing and executing quality improvement work, but less useful in the writing process. The level of detail sometimes led to confusion about what to include or not include in a manuscript. Consequently, we made the items in SQUIRE 2.0 shorter and more direct.

A major challenge in the reporting of systematic efforts to improve healthcare is the multiplicity of terms used to describe the work, which is challenging for novices and experts alike. Improvement work draws on the epistemology of a variety of fields, and depending on one's field of study, the same words can carry different connotations, a particularly undesirable state of affairs. Terms such as 'quality improvement', 'implementation science' and 'improvement science' refer to approaches that have many similarities, but can also connote important (and often-debated) differences. Other terms such as 'healthcare delivery science', 'patient safety' and even simply 'improvement' are also subject to surprising variation in interpretation. To address this problem in semantics, we created a glossary of terms used in SQUIRE 2.0 (box 1). The glossary provides the intended meaning of certain key terms as we have used them in SQUIRE 2.0 (table 1). These definitions may be helpful in other endeavours, but are not necessarily intended to be adopted for use in other contexts. Overall, we sought terms and definitions that would be useful to the largest possible audience. For example, we chose 'intervention(s)' to refer to the changes that are made. We decided not to use the word 'improvement' in the individual items (although it remains in the SQUIRE acronym) to encourage authors to report efforts that did not lead to changes for the better. Reporting well done, negative studies is vital for the learning in this discipline.

Theory

SQUIRE 2.0 includes a new item titled 'rationale'. Biomedical and clinical research is driven by iterative cycles of theory building and hypothesis testing. Healthcare improvement work has not consistently based the planning, design and execution of its programmes solidly in theory, to the detriment of the work. For this reason, SQUIRE 2.0 explicitly includes an item devoted to theory, although we chose to use the broader and less technical label 'rationale' to encourage authors to be explicit in reporting formal and informal theories, models, concepts or even hunches as to why they expected a particular intervention to work in a particular context. A plain language interpretation of 'rationale' might be, 'why did you think this would work?' A recent narrative review of the nature of theory and its use in improvement describes the many types and applications of theory, and considers pitfalls in using and not using theory.[12]

The addition of the 'rationale' item is intended to encourage clarity around assumptions about the nature of the intervention, the context and the expected outcomes. The presence of a well thought out rationale will align with appropriate measures and with the study of the intervention; it may also be the starting point for the next round of work. The 'summary' item in the discussion section encourages authors to revisit the original rationale in the light of its findings and in the larger context of similar projects.

Context

SQUIRE 2.0 accepts 'context' as the key features of the environment in which the work is immersed and which are interpreted as meaningful to the success, failure and unexpected consequences of the intervention(s), as well as the relationship of these to the stakeholders (eg, improvement team, clinicians, patients, families, etc).[13–16] Systematic efforts to improve healthcare should contain clear descriptions and acknowledgement of context, rather than efforts to control it or explain it away. SQUIRE 1.0 included context with items in all sections of the manuscript, but context did not rise to the level of a distinct item itself. SQUIRE 2.0 recognises context as a fundamental item in the methods section, but its relevance is not limited to this section. In addition to affecting the

Box 1 Glossary of key terms used in Standards for QUality Improvement Reporting Excellence (SQUIRE) 2.0. This glossary provides the intended meaning of selected words and phrases as they are used in the SQUIRE 2.0 guidelines. They may, and often do, have different meanings in other disciplines, situations and settings

Assumptions
Reasons for choosing the activities and tools used to bring about changes in healthcare services at the system level.
Context
Physical and sociocultural make-up of the local environment (eg, external environmental factors, organisational dynamics, collaboration, resources, leadership and the like), and the interpretation of these factors ('sense-making') by the healthcare delivery professionals, patients and caregivers that can affect the effectiveness and generalisability of intervention(s).
Ethical aspects
The value of system-level initiatives relative to their potential for harm, burden and cost to the stakeholders. Potential harms particularly associated with efforts to improve the quality, safety and value of healthcare services include opportunity costs, invasion of privacy and staff distress resulting from disclosure of poor performance.[25]
Generalisability
The likelihood that the intervention(s) in a particular report would produce similar results in other settings, situations or environments (also referred to as external validity).
Healthcare improvement
Any systematic effort intended to raise the quality, safety and value of healthcare services, usually done at the system level. We encourage the use of this phrase rather than 'quality improvement', which often refers to more narrowly defined approaches.
Inferences
The meaning of findings or data, as interpreted by the stakeholders in healthcare services—improvers, healthcare delivery professionals and/or patients and families.
Initiative
A broad term that can refer to organisation-wide programmes, narrowly focused projects or the details of specific interventions (eg, planning, execution and assessment).
Internal validity
Demonstrable, credible evidence for efficacy (meaningful impact or change) resulting from introduction of a specific intervention into a particular healthcare system.
Intervention(s)
The specific activities and tools introduced into a healthcare system with the aim of changing its performance for the better. Complete description of an intervention includes its inputs, internal activities and outputs (eg, in the form of a logic model) and the mechanism(s) by which these components are expected to produce changes in a system's performance.[17]
Opportunity costs
Loss of the ability to perform other tasks or meet other responsibilities resulting from the diversion of resources needed to introduce, test or sustain a particular improvement initiative.
Problem
Meaningful disruption, failure, inadequacy, distress, confusion or other dysfunction in a healthcare service delivery system that adversely affects patients, staff or the system as a whole, or that prevents care from reaching its full potential.
Process
The routines and other activities through which healthcare services are delivered.
Rationale
Explanation of why particular intervention(s) were chosen, and why it was expected to work, be sustainable and be replicable elsewhere.
Systems
The interrelated structures, people, processes and activities that together create healthcare services for and with individual patients and populations. For example, systems exist from the personal self-care system of a patient to the individual provider–patient dyad system, to the microsystem, to the macrosystem and all the way to the market/social/insurance system. These levels are nested within each other.
Theory or theories
Any 'reason-giving' account that asserts causal relationships between variables (causal theory) or that makes sense of an otherwise obscure process or situation (explanatory theory). Theories come in many forms, and serve different purposes in the phases of improvement work. It is important to be explicit and well founded about any informal and formal theory (or theories) that are used.

development of the rationale and subsequent design of the intervention(s), context plays a key role in the iterations of intervention(s) and the outcomes. While it is often not simple to capture or describe context, understanding its impact on the design, implementation, measurement and results make it a vital contributor in identifying and reporting the factors and mechanisms responsible for the success or failure of the intervention(s).

Studying the intervention(s)

The study of the intervention is, perhaps, the most challenging item in SQUIRE. In the evaluation of SQUIRE 1.0[19] and in the pilot testing,[22] many were perplexed by this item and its subelements. This item was intended to encourage a more formal assessment of the intervention and its associated outcomes. In SQUIRE 2.0, this section is called 'study of the intervention(s)' (table 1).

'Doing' an improvement project is fundamentally different from 'studying' it. The primary purpose of 'doing' improvement is to produce better local processes and outcomes rather than contribute to new generalisable knowledge. In contrast, the reason for 'studying' the intervention is mainly to contribute to the body of knowledge about the efficacy and generalisability of efforts for improving healthcare. Both 'doing' and 'studying' are required for a deep understanding of the nature and impact of the intervention(s) as well as the possible underlying mechanisms. 'Study of the intervention(s)' focuses mainly on whether and why an intervention 'works'. It should align with the rationale and may include, but is not limited to, preplanned formal testing of the proposed theory that the intervention(s) actually produced the observed changes, as well as the impact of the intervention(s) on the context in which the work was done.

SQUIRE 2.0 asks authors to be as transparent, complete and as accurate as possible about reporting 'doing' and 'studying' improvement work as both aspects of the work are key to scholarly reporting. The 'summary' and 'interpretation' items in the discussion encourage authors to explain potential mechanisms by which the intervention(s) resulted (or failed to result) in change, thereby developing explanatory theories that can be subsequently tested.

CONCLUSIONS

The development of SQUIRE 2.0 consisted of a detailed analysis of SQUIRE 1.0, input from experts in the field and thorough pilot testing. Many methods and philosophical approaches to improve the quality, safety and value of healthcare are available. The systematic efforts to improve healthcare are often complex and multidimensional, and their effectiveness is inherently context dependent. SQUIRE 2.0 provides common ground on which the discoveries contributed by the various approaches can advance the field by sharing them in the published literature.

At the same time, we recognise that simply publishing SQUIRE 2.0 will not effect this change; additional efforts and resources are required. For example, we have created an explanation and elaboration (E&E) document (Goodman D, Ogrinc G, Davies L; personal communication, 2015) to accompany this article. For each item in SQUIRE 2.0, the E&E provides one or more examples from the published literature and a commentary on how the example(s) meets or does not meet the item's standards; this information brings the content of each item to life. The SQUIRE website (http://www.squire-statement.org) contains a number of resources in addition to the guidelines themselves, including interactive E&E pages and video commentaries. The website supports an emerging online community for the continuous use, conversation about and evaluation of the guidelines.

Writing about improvement can be challenging. Sharing successes, failures and developments through scholarly literature is an essential component of the complex work required in order to improve healthcare services for patients, professionals and the public.

Author affiliations
[1]White River Junction VA Medical Center, White River Junction, Vermont, USA
[2]Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA
[3]The Dartmouth Institute for Health Policy and Clinical Practice, Hanover, New Hampshire, USA
[4]Institute for Healthcare Improvement, Cambridge, Massachusetts, USA

Acknowledgements This material is based on work supported by the Health Foundation and the Robert Wood Johnson Foundation, and included the use of facilities and material at the White River Junction VA in White River Junction, Vermont, USA.

Collaborators Name and affiliation of members of the SQUIRE 2.0 advisory group: Davina Allen, Cardiff University, UK. Ross Baker, University of Toronto, Canada. Helen Crisp, Health Foundation, UK. Mary Dixon-Woods, University of Leicester, UK. Don Goldmann, Institute for Healthcare Improvement, USA. Steve Goodman, Stanford University, USA. Leora Horwitz, New York University, USA. Pam Ironside, Indiana University, USA. Peter Margolis, University of Cincinnati, USA. Paul Miles, American Board of Pediatrics, USA. Shirley Moore, Case Western Reserve University, USA. Peter Pronovost, Johns Hopkins University, USA. Lisa Rubenstein, University of California Los Angeles, USA. Gwen Sherwood, University of North Carolina, USA. Kaveh Shojania, University of Toronto, Canada. Richard Thomson, Newcastle University, UK. Charles Vincent, Imperial College London, UK. Hub Wollersheim, Radboud University Medical Center, the Netherlands.

Contributors Each member of the author list contributed to the design, execution, writing, editing and approval of this work.

Funding Robert Wood Johnson Foundation (grant number 70024) and Health Foundation (grant number 7099).

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

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 DeVinney B, ed. *Expanding Research and Evaluation Designs to Improve the Science Base for Health Care and Public Health Quality Improvement Symposium*. Washington DC: Agency for Health Care Research and Quality. Conference Summary, 15 Sep 2005.
3 Shojania KG, Grimshaw JM. Evidence-based quality improvement: the state of the science. *Health Aff (Millwood)* 2005;24:138–50.
4 Grol RP, Bosch MC, Hulscher ME, *et al*. Planning and studying improvement in patient care: the use of theoretical perspectives. *Milbank Q* 2007;85:93–138.
5 Rubenstein LV, Hempel S, Farmer MM, *et al*. Finding order in heterogeneity: types of quality-improvement intervention publications. *Qual Saf Health Care* 2008;17:403–8.
6 Davidoff F, Batalden P, Stevens D, *et al*. Publication guidelines for quality improvement in health care: evolution of the SQUIRE project. *Qual Saf Health Care* 2008;17(Suppl 1):i3–9.
7 Batalden P, Leach D, Swing S, *et al*. General competencies and accreditation in graduate medical education. *Health Aff (Millwood)* 2002;21:103–11.
8 Interprofessional Education Collaborative Expert Panel. *Core Competencies for Interprofessional Collaborative Practice: Report of an Expert Panel*. Washington DC: Interprofessional Education Collaborative, 2011.
9 Teaching for Quality. 2013. https://www.aamc.org/initiatives/cei/te4q/ (accessed 21 Jun 2013).
10 Cronenwett L, Sherwood G, Barnsteiner J, *et al*. Quality and safety education for nurses. *Nurs Outlook* 2007;55:122–31.
11 Nasca TJ, Philibert I, Brigham T, *et al*. The next GME accreditation system—rationale and benefits. *N Engl J Med* 2012;366:1051–6.
12 Davidoff F, Dixon-Woods M, Leviton L, *et al*. Demystifying theory and its use in improvement. *BMJ Qual Saf* 2015;24:228–38.
13 Bate P, Robert G, Fulop N, *et al*. *Perspectives on context*. London: The Health Foundation, 2014.
14 Kaplan HC, Provost LP, 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.
15 Øvretveit J. Understanding the conditions for improvement: research to discover which context influences affect improvement success. *BMJ Qual Saf* 2011;20(Suppl 1):i18–23.
16 Taylor SL, Dy S, Foy R, *et al*. What context features might be important determinants of the effectiveness of patient safety practice interventions? *BMJ Qual Saf* 2011;20:611–17.
17 Hoffmann TC, Glasziou PP, Boutron I, *et al*. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:1–12.
18 Portela MC, Pronovost PJ, Woodcock T, *et al*. How to study improvement interventions: a brief overview of possible study types. *BMJ Qual Saf* 2015;24:325–36.
19 Davies L, Batalden P, Davidoff F, *et al*. The SQUIRE guidelines: an evaluation from the field, 5 years post release. *BMJ Qual Saf* 2015. Published Online First: 18 Jun 2015. doi:10.1136/bmjqs-2015-004116
20 Howell V, Schwartz AE, O'Leary JD, *et al*. 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;24:400–6.
21 Stevens D. SQUIRE and the evolving science of healthcare improvement. *BMJ Qual Saf* 2015;24:349–51.
22 Davies L, Donnelly K, Goodman D, *et al*. Findings from a novel approach to publication guideline revision: User road testing of a draft version of SQUIRE 2.0. *BMJ Qual Saf* 2015 (in press). doi:10.1136/bmjqs-2015-004117
23 Day RA. The origins of the scientific paper: the IMRaD format. *J Am Med Writers Assoc* 1989;4:16–18.
24 Huth E. *Writing and publishing in medicine*. 3rd edn. Baltimore: Williams and Wilkins, 1999.
25 Baily M, Bottrell M, Lynn J, *et al*. Special report: the ethics of using QI methods to improve health care quality and safety. *Hastings Center Rep* 2006;34:s1–40.

BMJ Quality & Safety

SQUIRE 2.0 (*Standards for QUality Improvement Reporting Excellence)* : revised publication guidelines from a detailed consensus process

Greg Ogrinc, Louise Davies, Daisy Goodman, Paul Batalden, Frank Davidoff and David Stevens

BMJ Qual Saf published online September 14, 2015

BMC Medical Research Methodology

CORRESPONDENCE Open Access

Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ

Allison Tong[1,2*†], Kate Flemming[3†], Elizabeth McInnes[4†], Sandy Oliver[5] and Jonathan Craig[1,2]

Abstract

Background: The syntheses of multiple qualitative studies can pull together data across different contexts, generate new theoretical or conceptual models, identify research gaps, and provide evidence for the development, implementation and evaluation of health interventions. This study aims to develop a framework for reporting the synthesis of qualitative health research.

Methods: We conducted a comprehensive search for guidance and reviews relevant to the synthesis of qualitative research, methodology papers, and published syntheses of qualitative health research in MEDLINE, Embase, CINAHL and relevant organisational websites to May 2011. Initial items were generated inductively from guides to synthesizing qualitative health research. The preliminary checklist was piloted against forty published syntheses of qualitative research, purposively selected to capture a range of year of publication, methods and methodologies, and health topics. We removed items that were duplicated, impractical to assess, and rephrased items for clarity.

Results: The Enhancing transparency in reporting the synthesis of qualitative research (ENTREQ) statement consists of 21 items grouped into five main domains: introduction, methods and methodology, literature search and selection, appraisal, and synthesis of findings.

Conclusions: The ENTREQ statement can help researchers to report the stages most commonly associated with the synthesis of qualitative health research: searching and selecting qualitative research, quality appraisal, and methods for synthesising qualitative findings. The synthesis of qualitative research is an expanding and evolving methodological area and we would value feedback from all stakeholders for the continued development and extension of the ENTREQ statement.

Keywords: Thematic synthesis, Standards, Qualitative health research, Reporting

Background

Methods to synthesise qualitative research began with the recognition that providing evidence-based healthcare and health policy requires a range of evidence beyond that provided by the 'rationalist' model of systematic reviewing of quantitative research [1]. Qualitative research aims to provide an in-depth understanding into human behaviour, emotion, attitudes and experiences. The synthesis of findings from multiple qualitative studies can provide a range and depth of meanings, experiences, and perspectives of participants across health-care contexts. Syntheses of qualitative research can pull together data across different contexts, generate new theoretical or conceptual models, identify research gaps, inform the development of primary studies, and provide evidence for the development, implementation and evaluation of health interventions [2-9]. The synthesis, or "bringing together" of the findings of primary qualitative studies is emerging as an important source of evidence for healthcare and policy [10]. Many aspects of the methods for synthesising qualitative research are in the early stages of development.

The number of published syntheses of qualitative health research is increasing (Figure 1). There are a wide range of qualitative synthesis methods with many common features, but also key differences [1]. The

* Correspondence: allison.tong@sydney.edu.au
†Equal contributors
[1]Sydney School of Public Health, University of Sydney, Sydney, NSW 2006, Australia
[2]Centre for Kidney Research, The Children's Hospital at Westmead, Westmead, NSW 2145, Australia
Full list of author information is available at the end of the article

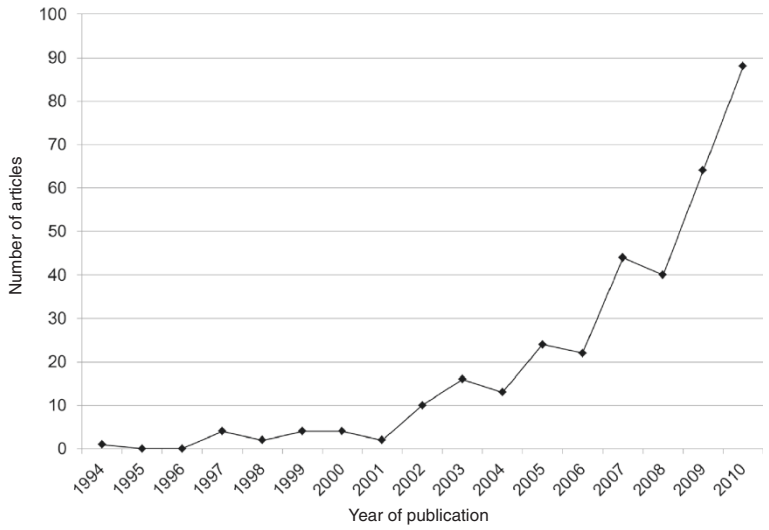


Figure 1 Number of published synthesis of qualitative health research.

main methods of qualitative synthesis include: meta-ethnography [11]; thematic synthesis [12]; critical interpretive synthesis [4]; narrative synthesis [13]; and meta-study [14-16]. One of the first methods identified for synthesising qualitative research - meta-ethnography - has subsequently influenced the development of other methods such as thematic analysis and critical interpretive synthesis through the use of its terminology and concepts, as well as extending and adapting its methods. Figure 2 provides examples of the wide-ranging terms used to describe different qualitative synthesis methods. Some of the adaptations of qualitative syntheses have, however, resulted in inconsistent use of terms for describing key stages of synthesis [17]. For users of qualitative syntheses the different labels used to describe similar qualitative synthesis methods and the inconsistent use of terms to describe the different stages within qualitative reviews can be confusing [1,18]. While there are differences in approaches and rationale for some qualitative synthesis methods (for example, Critical Interpretive Synthesis may be better suited for large diverse bodies of literature while meta-ethnography may be better for analysing a smaller number of papers) [4] there is a core set of techniques common to most qualitative synthesis methods.

While there are reporting guidelines for qualitative research [19], there are no published guidelines for *reporting* the synthesis of qualitative research. Reporting guidelines

Figure 2 Word cloud of the methodological terms used in published synthesis of qualitative health research. Word clouds give more prominence to words that appear more frequently in the source text. The methodological terms were extracted from the title/abstract/full text of 381 published synthesis of qualitative health research (to 31st May 2011) and entered into Tagxedo, an online tool which generated a word cloud.

for quantitative systematic reviews exist and these are helping to set standards for both the conduct and reporting of these reviews [20]. Currently, most synthesises of qualitative research have been undertaken by those with an interest in methodological development, and therefore reviews appear to be well-reported. Increasingly, the methodologies associated with synthesis of qualitative research are being used by researchers and students new to the process. It is important at this time to begin to establish reporting standards. Developing reporting guidelines for qualitative synthesis may assist researchers to improve both the conduct and reporting of qualitative syntheses and enable the end-user to better understand the processes involved in developing a qualitative synthesis.

The aim of this paper is to report on the first phase of the development of guidelines to encourage transparency in reporting syntheses of qualitative research; to assist end-users to identify the core steps involved and to provide a tool to help clarify to the various concepts and terms used to describe similar processes in qualitative syntheses.

Methods

ENTREQ Development

It is acknowledged that there is no single best or correct approach to developing guidelines [21]. Where feasible, we have reported the development of our guideline drawing from steps provided in 'Guidance for developers of health research reporting guidelines' by David Moher and colleagues [21], available at www.equator-network.org (an international initiative that seeks to improve reliability and value of medical research by promoting transparent reporting).

Identify the need for a guideline

We identified the need for a reporting guideline for syntheses of qualitative research as a result of our collective experiences in using, publishing, reviewing and teaching syntheses of qualitative health research, debriefing notes taken after an international conference symposium on the synthesis of qualitative health research (Qualitative Health Research Conference in Vancouver, Canada, 2010, KF/AT) and a seminar at the Qualitative Health Research Collaboration in Sydney, Australia, 2011, (AT/EM).

To further establish a need for a reporting guideline, we conducted a comprehensive search for guidance and reviews relevant to the synthesis of qualitative research, methodology papers, and published syntheses of qualitative health research using the terms for "qualitative research" combined with terms relating to synthesis (systematic review, synthesis, thematic synthesis, meta-ethnography, meta-study, meta-analysis) (Additional file 1). The searches were conducted in electronic medical literature databases including MEDLINE, EMBASE, and CINAHL from inception to 20th May 2011, and in Google Scholar. Relevant organisational websites including the EQUATOR Network database of reporting guidelines (www.equator-network.org) and reference lists of relevant articles were also searched. We identified 381 syntheses of qualitative research, with the number of publications exponentially increasing from 1994 to May 2011 (Additional file 2, Figure 1).

Generating items for inclusion in the checklist

The initial items for inclusion in the preliminary "Enhancing transparency in reporting the synthesis of qualitative research (ENTREQ) Statement" were generated inductively from guides to synthesising qualitative health research [1,10], seminal methodology papers [4,11,12,22-24] and the authors' experience in conducting and appraising qualitative syntheses (AT, KF, EM,). The items were compiled and grouped into five categories: introduction; methods and methodology; literature search and selection; appraisal; and synthesis of findings.

Pilot testing the checklist

In order to test our preliminary framework and to reach consensus for the inclusion of each item, the reporting framework was pilot tested against forty published syntheses of qualitative research, which were purposively selected from our search results to capture a range of year of publication, methods and methodologies, and health topics (Additional file 3). Three members of the research team (AT/KF/EM) independently piloted the guidance initially against 32 of these reviews, by extracting relevant data for each guidance item. During this time we met via teleconferences to discuss the results of the testing and made a series of revisions to the ENTREQ Statement. We removed items that were duplicated. Items were also rephrased for clarity where there was ambiguity. The revised statement was tested against the eight remaining reviews and no further changes were made. On average, it took 5 to 20 minutes to assess each review using the ENTREQ Statement. The results are provided in Additional file 3.

Results and discussion

ENTREQ Statement: content and rationale

The ENTREQ statement consists of 21 items grouped into five main domains: introduction, methods and methodology, literature search and selection, appraisal, and synthesis of findings (Table 1). For each item, a descriptor and examples are provided. Below we present a rationale for each domain and its associated items.

Introduction, methods and methodology (Domains 1 and 2)

The methodology and approaches selected are usually influenced by the research question (outlined in the

Table 1 Enhancing transparency in reporting the synthesis of qualitative research: the ENTREQ statement

No	Item	Guide and description
1	Aim	State the research question the synthesis addresses.
2	Synthesis methodology	Identify the synthesis methodology or theoretical framework which underpins the synthesis, and describe the rationale for choice of methodology *(e.g. meta-ethnography, thematic synthesis, critical interpretive synthesis, grounded theory synthesis, realist synthesis, meta-aggregation, meta-study, framework synthesis).*
3	Approach to searching	Indicate whether the search was pre-planned (*comprehensive search strategies to seek all available studies)* or iterative (*to seek all available concepts until they theoretical saturation is achieved).*
4	Inclusion criteria	Specify the inclusion/exclusion criteria *(e.g. in terms of population, language, year limits, type of publication, study type).*
5	Data sources	Describe the information sources used (e.g. *electronic databases (MEDLINE, EMBASE, CINAHL, psycINFO, Econlit), grey literature databases (digital thesis, policy reports), relevant organisational websites, experts, information specialists, generic web searches (Google Scholar) hand searching, reference lists)* and when the searches conducted; provide the rationale for using the data sources.
6	Electronic Search strategy	Describe the literature search *(e.g. provide electronic search strategies with population terms, clinical or health topic terms, experiential or social phenomena related terms, filters for qualitative research, and search limits).*
7	Study screening methods	Describe the process of study screening and sifting *(e.g. title, abstract and full text review, number of independent reviewers who screened studies).*
8	Study characteristics	Present the characteristics of the included studies *(e.g. year of publication, country, population, number of participants, data collection, methodology, analysis, research questions).*
9	Study selection results	Identify the number of studies screened and provide reasons for study exclusion *(e,g, for comprehensive searching, provide numbers of studies screened and reasons for exclusion indicated in a figure/flowchart; for iterative searching describe reasons for study exclusion and inclusion based on modifications t the research question and/or contribution to theory development).*
10	Rationale for appraisal	Describe the rationale and approach used to appraise the included studies or selected findings *(e.g. assessment of conduct (validity and robustness), assessment of reporting (transparency), assessment of content and utility of the findings).*
11	Appraisal items	State the tools, frameworks and criteria used to appraise the studies or selected findings *(e.g. Existing tools: CASP, QARI, COREQ, Mays and Pope* [25]*; reviewer developed tools; describe the domains assessed: research team, study design, data analysis and interpretations, reporting).*
12	Appraisal process	Indicate whether the appraisal was conducted independently by more than one reviewer and if consensus was required.
13	Appraisal results	Present results of the quality assessment and indicate which articles, if any, were weighted/excluded based on the assessment and give the rationale.
14	Data extraction	Indicate which sections of the primary studies were analysed and how were the data extracted from the primary studies? *(e.g. all text under the headings "results /conclusions" were extracted electronically and entered into a computer software).*
15	Software	State the computer software used, if any.
16	Number of reviewers	Identify who was involved in coding and analysis.
17	Coding	Describe the process for coding of data *(e.g. line by line coding to search for concepts).*
18	Study comparison	Describe how were comparisons made within and across studies *(e.g. subsequent studies were coded into pre-existing concepts, and new concepts were created when deemed necessary).*
19	Derivation of themes	Explain whether the process of deriving the themes or constructs was inductive or deductive.
20	Quotations	Provide quotations from the primary studies to illustrate themes/constructs, and identify whether the quotations were participant quotations of the author's interpretation.
21	Synthesis output	Present rich, compelling and useful results that go beyond a summary of the primary studies (e.g. *new interpretation, models of evidence, conceptual models, analytical framework, development of a new theory or construct).*

introduction), intended synthesis output, reviewer's philosophical position, context, and target audience. Also, reviewers may choose their approach according to the type of data available. For example, meta-ethnography works well with primary qualitative studies offering "thick descriptions" and in-depth analysis. Thematic synthesis is possible with "thinner" studies. A recent review of qualitative syntheses found that nine main approaches were used to synthesise qualitative research including: critical interpretive synthesis, grounded theory synthesis, meta-ethnography, meta-study, thematic synthesis, meta-narrative synthesis, textual narrative synthesis, framework synthesis, and ecological triangulation [1]. A summary of commonly used approaches for synthesising qualitative health research is provided in Table 2.

Literature search and selection (Domain 3)

Conducting a systematic search which is reproducible and comprehensive is a distinguishing characteristic of a

Table 2 Summary of common methodologies for the synthesis of qualitative health research*

Methodology	Critical interpretive synthesis	Grounded theory synthesis	Meta-ethnography	Meta-study	Thematic synthesis
Key seminal methodology references	Dixon-woods et al. 2006 [4]	Kearney 2001 [23], Eaves 2001 [22]	Noblit and Hare 1988 [11], Britten et al. 2002 [2]	Paterson et al. 2001 [24]	Thomas and Harden 2008 [12]
Philosophical positioning**	*Subjective idealism* – no single shared reality independent of multiple alterative human constructions	*Objective idealism* – a world of collectively shared understandings exists	*Objective idealism* – a world of collectively shared understandings exists	*Subjective idealism* – no single shared reality independent of multiple alterative human constructions	*Critical realism* – knowledge of reality is medicated by one's beliefs and perspectives
Literature search	Theoretical sampling	Theoretical sampling	Non-specified	Not-specified	Systematic, comprehensive
Quality appraisal	The degree to which the research findings can inform theory development	Implicit judgement about the context, quality and usefulness of the study	Judgement based on relevance; CASP	Focuses on rigour and the epistemological soundness of the research methods	Criteria related to aims, context, rationale, methods and findings, reliability, validity, appropriateness of methods for ensuring findings are grounded in participant perspectives
Analysis techniques and concepts	• Concurrent iteration of the research questions	• Concurrent data collection and analysis	• Reciprocal translational analysis (translation of concepts from individual studies – $1^{st}/2^{nd}$ order constructs)	• Analyse findings – meta-data-analysis	• Line by line coding of text from primary studies
	• Extract data and summarise papers	• Theory is derived inductively from the data	• Refutational synthesis (explore and explain contradictions between studies – $1^{st}/2^{nd}$ order constructs)	• Analyse methods – meta-method)	• Free codes organised into descriptive themes
	• Define and apply codes	• Constant comparison of data	• Lines of argument (grounded theorising based on synthesising translations)	• Analyse theory – meta-theory	• Further interpretation to develop analytical themes
	• Develop a critique, generate themes			• Bring together all three components of the analysis	
Synthesis output	• New theoretical conceptualisation – synthetic construct	• Generation of a new, higher-level grounded theory	• New insights – 3^{rd} order constructs	• Account for differences in research findings	• Analytical themes that offer a new interpretation that goes beyond the primary studies
				• New interpretation of phenomena studied	
Topic areas and study references†	Access to healthcare by vulnerable groups [4], pain management [26]	Domestic violence [23], caregiving [22]	Medicine-taking [3], patients' help-seeking experiences in cancer presentation [6], palliative care [27]	Chronic illness experience [14], influences on shared decisions making [15], adolescent health [16]	Children's experiences of health eating [12], chronic kidney disease [28], people's understanding of cancer risk [29], organ transplantation [7], patient-physician relationships [30]

*This is not a complete list of methodologies as methodologies for the synthesis of qualitative health research are wide ranging; **Adapted from Barnett-Page and Thomas [1] and Spencer et al. [31]. †References selected to reflect a range of topic areas in health research.

systematic review; however there are few developed and tested methods for locating qualitative research, and lack of consensus as to whether systematic searching is required [32]. Some argue that exhaustive searching is not necessary. Instead, reviewers may adopt an iterative approach where all the available concepts rather than studies are sought until saturation is reached [1].

A pre-planned sensitive search strategy may combine search terms relating to the population and context, with those relating to the health or clinical topic, and terms relating to experiential and social phenomena (such as knowledge, attitudes, beliefs, understanding, preferences, perspectives). These can then be combined with terms for qualitative methods and methodology. Methodological filters for qualitative research have been developed but have undergone little replication and validation [32]. There are also differences in the indexing of qualitative research within electronic databases such as

MEDLINE, EMBASE, PsycINFO and CINAHL. Within published syntheses of qualitative research there is often a lack of transparency about the search processes employed, with neither the search strategy nor databases detailed [33]. For a comprehensive approach, the PRISMA flowchart is recommended for reporting the different phases of searching, screening and identifying studies for inclusion in the qualitative synthesis [20]. Qualitative research can often be found in the grey literature (e.g. technical reports, working papers, thesis publications). To locate relevant studies, reviewers can search relevant organisational websites, Google Scholar, thesis databases, specialist journals, and consult with experts (researchers, providers, policy makers) in the relevant fields and librarians.

The inclusion and exclusion of studies may be defined by factors including population characteristics, health or clinical topic, methods and methodology (philosophical approach), language, time frame, or type of publication; and this should be justified. For readers to make an assessment about the transferability of the findings to their own setting, a description of the study characteristics, screening process, and reasons for excluding studies is needed.

Appraisal (Domain 4)

Quality assessment of qualitative research is challenging and contentious [25]. Just as there are no standardised criteria for assessing the quality of all quantitative research, standardising criteria for assessing the standard of conduct in all qualitative research which embraces a range of designs, is not possible or appropriate [31,34]. Also, there is little evidence on how the quality of reporting reflects the robustness, trustworthiness and transferability of the findings of qualitative studies [35]. Nevertheless, most published syntheses of qualitative research include a quality appraisal of the primary studies. The rationale underpinning quality assessment and the methods used to appraise quality vary widely but can be broadly characterised into three approaches: assessment of study conduct, appraisal of study reporting, and implicit judgement of the content and utility of the findings for theory development. Some syntheses exclude low quality studies, while others comment on or weight study findings according to their quality [36].

Several appraisal tools have been used including the Critical Appraisal Skills Program (CASP) [37] which addresses the principles and assumptions underpinning qualitative research but does not claim to be a definitive guide; the Qualitative Assessment Review Instrument Tool (QARI) [38], which suggests general questions that require the reader to make a judgement for example about the "congruency" of the research methodology with the state philosophical perspective, research questions, data collection, interpretation of the results; and Consolidated Criteria for Reporting Qualitative Research (COREQ) [19] which is the only framework developed explicitly for assessing reporting. Some reviewers have developed their own appraisal framework selecting items from existing criteria [25,39-41], augmented with additional criteria they deemed were specifically relevant to the research topic. These were usually identified by discussion and consensus among the research group. For example, Brunton et al. [42] conducted a systematic review of qualitative research on children and physical activity and used existing criteria proposed for assessing quality of qualitative research but included an additional item, "actively involved children to an appropriate degree in the design and conduct of the study," which they deemed relevant to their review [42].

Systematic reviewers of qualitative studies have found that many primary qualitative studies are poorly reported [3]. Also, some reviewers have found that studies with sparse detail about the conduct of the research tend to contribute less to the synthesis [28]. An assessment of the quality of reporting can allow readers to make an informed judgement about the credibility (can the research findings be trusted?), dependability (is the process of research logical, traceable and clearly documented?), transferability (are the research findings relevant to other settings?) and confirmability (are the research findings and interpretations linked to the data?). A reporting framework can also function as a screening tool for systematic reviewers to determine study eligibility and inform the development of future qualitative studies on the topic of interest. For example, it can highlight qualitative methods and methodologies that have been effective in gaining in-depth insight into participants' perspectives, beliefs and attitudes, and identify those which could generate more understanding about a phenomena, but have not been tried and tested. Also, the process of appraisal can facilitate a deeper understanding of included papers.

Existing frameworks for reporting qualitative research may be considered and used as a starting point and adapted to suit the synthesis topic. The framework should capture the range of methods and methodologies of the included studies. In some instances, multiple reviewers have independently assessed quality and discussed quality appraisal to achieve consensus. Also, the rationale for weighting or excluding studies based on quality appraisal should be explicit.

Synthesis of findings (Domain 5)

For clarity of reporting the analysis process, reviewers should define which sections of the included articles were actually analysed; and describe the process of

coding, comparing and interpreting the data. Specific analysis techniques and concepts are provided in Table 2. Details about use of software and number of reviewers involved in coding and analysis can allow readers to assess the dependability of the findings. It enables readers to assess whether data are managed in a systemic way. Quotations from the articles may be included to illustrate the themes or constructs identified. The target audience should also be considered when reporting and presenting the synthesis output. Ultimately, the synthesis should generate rich, compelling and new insights that go beyond a summary of the primary studies; however some "implicit judgment" and team discussion may be required to assess this.

Conclusions

The ENTREQ statement was developed to promote explicit and comprehensive reporting of the synthesis of qualitative studies. We acknowledge it is unlikely that a standardised set of procedures will ever be developed, more probably, a 'methodological palette' will be created from which reviewers can draw methods relevant to the focus of their review [9]. The proposed guidelines covers reporting items relating to methodology and methods, literature searching and selection, appraisal and the synthesis of findings.

The purpose of the ENTREQ statement is to offer guidance for researchers and reviewers to improve the reporting of synthesis of qualitative health research. We believe this document can be a useful resource and reference for those learning how to conduct a synthesis of qualitative research and readers of syntheses of qualitative health research. But we emphasise that this is not an absolute, definitive framework. Also, we acknowledge that we did not complete a Delphi exercise as recommended by the "Guidance for developers of health research reporting guidelines" [21] due to resource limitations. However, we believe that this initial development of the ENTREQ Statement is a crucial step for the development of a Delphi exercise.

We encourage authors to evaluate the checklist to assess whether it is useful for improving the completeness of reporting the synthesis of qualitative research. The synthesis of qualitative research is an expanding and evolving methodological area and we would value feedback from all stakeholders for the continued development and extension of the ENTREQ statement in terms of content, clarity and feasibility.

Additional files

Additional file 1: Search strategy.

Additional file 2: Search results.

Additional file 3: Pilot test: assessment of 40 published synthesis of qualitative research using the ENTREQ Statement.

Abbreviations

ENTREQ: Enhancing transparency in the reporting of qualitative health research.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

AT, KF, EM collected and analysed the data. AT/KF/EM drafted the manuscript. All authors critically reviewed and provided intellectual input on the manuscript. All authors read and approved the final manuscript.

Acknowledgments

AT is supported by the National Health and Medical Research Council (NHMRC) Fellowship [ID1037162].

Author details

[1]Sydney School of Public Health, University of Sydney, Sydney, NSW 2006, Australia. [2]Centre for Kidney Research, The Children's Hospital at Westmead, Westmead, NSW 2145, Australia. [3]Department of Health Sciences, University of York, Heslington, York Y010 5DD, UK. [4]National Centre for Clinical Outcomes Research, Australian Catholic University, St Vincent's Hospital, Darlinghurst, NSW 2010, Australia. [5]Institute of Education, University of London, London WC1H 0AL, UK.

Received: 25 June 2012 Accepted: 12 November 2012
Published: 27 November 2012

References

1. Barnett-Page E, Thomas J: *Methods for the synthesis of qualitative research: a critical review*. 01/09th edition. London: ESRC National Centre for Research Methods, Social Science Research Unit, Institute of Education, University of London; 2009.
2. Britten N, Campbell R, Pope C, Donovan J, Morgan M, Pill R: **Using meta ethnography to synthesise qualitative research: a worked example.** *J Health Serv Res Policy* 2002, **7**(4):209–215.
3. Campbell R, Pound P, Pope C, Britten N, Pill R, Morgan M, Donovan J: **Evaluating meta-ethnography: a synthesis of qualitative research on lay experience of diabetes and diabetes care.** *Soc Sci Med* 2003, **56**(4):671–684.
4. Dixon-Woods M, Cavers D, Agarwal S, Annandale E, Arthur A, Harvey J, Hsu R, Katbamna S, Olsen R, Smith L, *et al*: **Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups.** *BMC Med Res Methodol* 2006, **6:**35–47.
5. Pound P, Britten N, Morgan M, Yardley L, Pope C, Daker-White G, Campbell R: **Resisting medicines: a synthesis of qualitative studies of medicine taking.** *Soc Sci Med* 2005, **61**(1):133–155.
6. Smith LK, Pope C, Botha JL: **Patients' help-seeking experiences and delay in cancer presentation: a qualitative synthesis.** *Lancet* 2005, **366:**825–831.
7. Tong A, Morton RL, Howard K, Craig JC: **Adolescent experiences following organ transplantation: a systematic review of qualitative studies.** *J Pediatr* 2010, **155**(4):542–549.
8. McInnes RJ, Chambers JA: **Supporting breastfeeding mothers: qualitative synthesis.** *J Adv Nurs* 2008, **62**(4):407–427.
9. Flemming K: **The use of morphine to treat cancer-related pain: a synthesis of quantitative and qualitative research.** *J Pain Symptom Manag* 2010, **39**(1):139–154.
10. Ring N, Ritchie K, Mandava L, Jepson R: *A guide to synthesising qualitative research for researchers undertaking health technology assessments and systematic reviews*. 2010. Available from: http://www.nhshealthquality.org/nhsqis/8837.html.
11. Noblit GW, Hare HD: *Meta-ethnography: synthesising qualitative studies*. Thousand Oaks: Sage; 1988.
12. Thomas J, Harden A: **Methods for the thematic synthesis of qualitative research in systematic reviews.** *BMC Med Res Methodol* 2008, **8:**45.
13. Arai L, Britten N, Popay J, Roberts H, Petticrew M, Rodgers M, Sowden A: **Testing methodological developments in the conduct of narrative**

synthesis: a demonstration review of research on the implementation of smoke alarm interventions. *Evid Policy* 2007, **3**(3):361–383.
14. Thorne S, Paterson B, Acorn S, Canam C, Joachim G, Jillings C: **Chronic illness experience: insights from a meta-study.** *Qual Health Res* 2002, **12:**437–452.
15. Edwards M, Davies M, Edwards A: **What are the external influences on information exchange and shared decision-making in healthcare consultations: a meta-synthesis of the literature.** *Patient Educ Counsel* 2009, **75:**37–52.
16. Timminen KA, Holt NL: **A meta-study of qualitative research examining stressor appraisals and coping among adolescents in sport.** *J Sports Sci* 2010, **28**(14):1563–1580.
17. Atkins S, Lewin S, Smith H, Engel M, Fretheim A, Volminck J: **Conducting a meta-ethnography of qualitative literature: lessons learnt.** *BMC Med Res Methodol* 2008, **8:**21.
18. Dixon-Woods M: **Using framework-based synthesis for conducting reviews of qualitative studies.** *BMC Med* 2011, **9:**39.
19. Tong A, Sainsbury P, Craig JC: **Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups.** *Int J Qual Health Care* 2007, **19**(6):349–357.
20. Moher D, Liberati A, Tetzlaff J, Altman D, The PRISMA Group: **Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement.** *PLOS Med* 2009, **6**(7):e1000097.
21. Moher D, Schulz KF, Simera I, Altman D: **Guidance for developers of health research reporting guidelines.** *PLOS Med* 2010, **7**(2):e1000217. doi:1000210.1001371/journal.pmed.1000217.
22. Eaves YD: **A synthesis technique for grounded theory data analysis.** *J Adv Nurs* 2001, **35**(5):654–663.
23. Kearney MH: **Enduring love: a grounded formal theory of women's experience of domestic violence.** *Res Nurs Health* 2001, **24:**270–282.
24. Paterson BL, Thorne SE, Canam C, Jillings C: *Meta-study of qualitative health research. A practical guide to meta-analysis and meta-synthesis*. Thousand Oaks: Sage; 2001.
25. Mays N, Pope C: **Qualitative research in health care: assessing quality in qualitative research.** *Br Med J* 2000, **320:**50–52.
26. Flemming K: **The use of morphine to treat cancer-related pain: a synthesis of quantitative and qualitative research.** *J Pain Symptom Manage* 2010, **39**(1):139–154.
27. Hubbard G, McLachlan K, Forbat L, Munday D: **Recognition by family members that relatives with neurodegenerative disease are likely to die within a year: A meta-ethnography.** *Palliat Med* 2011, Online first.
28. Morton RL, Tong A, Howard K, Snelling P, Webster AC: **The views of patients and carers in treatment decision-making for chronic kidney disease: systematic review and thematic synthesis of qualitative studies.** *Br Med J* 2010, **340:**c112.
29. Lipworth WL, Davey H, Carter S, Hooker C: **Beliefs and beyond: what can we learn from qualitative studies of lay people's understanding of cancer risk.** *Health Expect* 2010, **13**(2):113–124.
30. Ridd M, Shaw A, Lewis G, Salisbury C: **The patient-doctor relationship: a synthesis of the qualitative literature on patients' perspectives.** *Br J Gen Prac* 2009, **29**(561):e116–e133.
31. Spencer L, Ritchie J, Lewis J, Dillon L: *Quality in qualitative evaluation: a framework for assessing research evidence*. London: Government Chief Social Researcher's Office; 2003.
32. Flemming K, Briggs M: **Electronic searching to locate qualitative research: evaluation of three strategies.** *J Adv Nurs* 2007, **57**(1):95–100.
33. Dixon-Woods M, Booth A, Sutton AJ: **Synthesizing qualitative research: a review of published reports.** *Qual Res* 2007, **7**(3):375–422.
34. Barbour R: **Checklists for improving rigour in qualitative research: a case of the tail wagging the dog?** *Br Med J* 2001, **332:**1115–1117.
35. Dixon-Woods M, Sutton A, Shaw R, Miller T, Smith J, Young B, Bonas S, Booth A, Jones D: **Appraising qualitative research for inclusion in systematic reviews: a quantitative and qualitative comparison of three methods.** *J Health Serv Res Policy* 2007, **12**(1):42–47.
36. Harden A, Gough D: **Quality and relevance appraisal**. In *An introduction to systematic reviews*. Edited by Gouhg D, Oliver S, Thomas J. London: Sage; 2012.
37. Questions to help you make sense of qualitative research; 2002. Available at http://www.casp-uk.net/wp-content/uploads/2011/11/CASP_Qualitative_Appraisal_Checklist_14oct10.pdf.
38. The Joanna Briggs Institute: *Joanna Briggs Institute Reviewers*. 2008th edition. The Joanna Briggs Institute; 2008.
39. Dixon-Woods M, Shaw RL, Agarwal S, Smith JA: **The problem of appraising qualitative research.** *Qual Saf Health Care* 2004, **13:**223–225.
40. Malterud K: **Qualitative research: standards, challenges, and guidelines.** *Lancet* 2001, **358:**483–488.
41. Popay J, Rogers A, Williams G: **Rationale and standards for the systematic review of qualitative literature in health services research.** *Qual Health Res* 1998, **8:**341–351.
42. Brunton G, Harden A, Rees R, Kavanagh J, Oliver S, Oakley A: *Children and physical activity: a systematic review of barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London; 2003.

doi:10.1186/1471-2288-12-181
Cite this article as: Tong *et al.*: Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Medical Research Methodology* 2012 **12**:181.

Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement

David Moher, PhD; Alessandro Liberati, MD, DrPH; Jennifer Tetzlaff, BSc; Douglas G. Altman, DSc; and the PRISMA Group*

Editor's Note: In order to encourage dissemination of the PRISMA Statement, this article is freely accessible on the Annals of Internal Medicine *Web site* (*www.annals.org*) *and will be also published in* PLOS Medicine, BMJ, Journal of Clinical Epidemiology, *and* Open Medicine. *The authors jointly hold the copyright of this article. For details on further use, see the PRISMA Web site* (*www.prisma-statement.org*).

Systematic reviews and meta-analyses have become increasingly important in health care. Clinicians read them to keep up to date with their field (1, 2), and they are often used as a starting point for developing clinical practice guidelines. Granting agencies may require a systematic review to ensure there is justification for further research (3), and some health care journals are moving in this direction (4). As with all research, the value of a systematic review depends on what was done, what was found, and the clarity of reporting. As with other publications, the reporting quality of systematic reviews varies, limiting readers' ability to assess the strengths and weaknesses of those reviews.

Several early studies evaluated the quality of review reports. In 1987, Mulrow examined 50 review articles published in four leading medical journals in 1985 and 1986 and found that none met all eight explicit scientific criteria, such as a quality assessment of included studies (5). In 1987, Sacks and colleagues (6) evaluated the adequacy of reporting of 83 meta-analyses on 23 characteristics in six domains. Reporting was generally poor; between one and 14 characteristics were adequately reported (mean = 7.7; standard deviation = 2.7). A 1996 update of this study found little improvement (7).

In 1996, to address the suboptimal reporting of meta-analyses, an international group developed a guidance called the QUOROM Statement (*QU*ality *O*f *R*eporting *O*f *M*eta-analyses), which focused on the reporting of meta-analyses of randomized, controlled trials (8). In this article, we summarize a revision of these guidelines, renamed PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses), which have been updated to address several conceptual and practical advances in the science of systematic reviews (**Box 1**).

See also:

Web-Only
Related Explanation and Elaboration article
Downloadable templates Table S1 and Figure S1
Conversion of graphics into slides

Terminology

The terminology used to describe a systematic review and meta-analysis has evolved over time. One reason for changing the name from QUOROM to PRISMA was the desire to encompass both systematic reviews and meta-analyses. We have adopted the definitions used by the Cochrane Collaboration (9). A systematic review is a review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyze data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyze and summarize the results of the included studies. Meta-analysis refers to the use of statistical techniques in a systematic review to integrate the results of included studies.

Developing the PRISMA Statement

A three-day meeting was held in Ottawa, Ontario, Canada, in June 2005 with 29 participants, including review authors, methodologists, clinicians, medical editors, and a consumer. The objective of the Ottawa meeting was to revise and expand the QUOROM checklist and flow diagram, as needed.

The executive committee completed the following tasks, prior to the meeting: a systematic review of studies examining the quality of reporting of systematic reviews, and a comprehensive literature search to identify methodological and other articles that might inform the meeting, especially in relation to modifying checklist items. An international survey of review authors, consumers, and groups commissioning or using systematic reviews and meta-analyses was completed, including the International Network of Agencies for Health Technology Assessment (INAHTA) and the Guidelines International Network (GIN). The survey aimed to ascertain views of QUOROM, including the merits of the existing checklist items. The results of these activities were presented during the meeting and are summarized on the PRISMA Web site (www.prisma-statement.org).

Ann Intern Med. 2009;151:264-269.
For author affiliations, see end of text.
* Membership of the PRISMA Group is provided in the Acknowledgment.
This article was published at www.annals.org on 21 July 2009.

Only items deemed essential were retained or added to the checklist. Some additional items are nevertheless desirable, and review authors should include these, if relevant (10). For example, it is useful to indicate whether the systematic review is an update (11) of a previous review, and to describe any changes in procedures from those described in the original protocol.

Shortly after the meeting a draft of the PRISMA checklist was circulated to the group, including those invited to the meeting but unable to attend. A disposition file was created containing comments and revisions from each respondent, and the checklist was subsequently revised 11 times. The group approved the checklist, flow diagram, and this summary paper.

Although no direct evidence was found to support retaining or adding some items, evidence from other domains was believed to be relevant. For example, Item 5 asks authors to provide registration information about the systematic review, including a registration number, if available. Although systematic review registration is not yet widely available (12, 13), the participating journals of the International Committee of Medical Journal Editors (ICMJE) (14) now require all clinical trials to be registered in an effort to increase transparency and accountability (15). Those aspects are also likely to benefit systematic reviewers, possibly reducing the risk of an excessive number of reviews addressing the same question (16, 17) and providing greater transparency when updating systematic reviews.

The PRISMA Statement

The PRISMA Statement consists of a 27-item checklist (**Table 1**; see also **Table S1**, available at www.annals.org, for a downloadable Word template for researchers to re-use) and a four-phase flow diagram (**Figure 1**; see also **Figure S1**, available at www.annals.org, for a downloadable Word template for researchers to re-use). The aim of the PRISMA Statement is to help authors improve the reporting of systematic reviews and meta-analyses. We have focused on randomized trials, but PRISMA can also be used as a basis for reporting systematic reviews of other types of research, particularly evaluations of interventions. PRISMA may also be useful for critical appraisal of published systematic reviews. However, the PRISMA checklist is not a quality assessment instrument to gauge the quality of a systematic review.

From QUOROM to PRISMA

The new PRISMA checklist differs in several respects from the QUOROM checklist, and the substantive specific changes are highlighted in **Table 2**. Generally, the PRISMA checklist "decouples" several items present in the QUOROM checklist and, where applicable, several checklist items are linked to improve consistency across the systematic review report.

Box 1. **Conceptual Issues in the Evolution From QUOROM to PRISMA**

Completing a Systematic Review Is an Iterative Process
The conduct of a systematic review depends heavily on the scope and quality of included studies: thus systematic reviewers may need to modify their original review protocol during its conduct. Any systematic review reporting guideline should recommend that such changes can be reported and explained without suggesting that they are inappropriate. The PRISMA Statement (Items 5, 11, 16, and 23) acknowledges this iterative process. Aside from Cochrane reviews, all of which should have a protocol, only about 10% of systematic reviewers report working from a protocol (22). Without a protocol that is publicly accessible, it is difficult to judge between appropriate and inappropriate modifications.

Conduct and Reporting Research Are Distinct Concepts
This distinction is, however, less straightforward for systematic reviews than for assessments of the reporting of an individual study, because the reporting and conduct of systematic reviews are, by nature, closely intertwined. For example, the failure of a systematic review to report the assessment of the risk of bias in included studies may be seen as a marker of poor conduct, given the importance of this activity in the systematic review process (37).

Study-Level Versus Outcome-Level Assessment of Risk of Bias
For studies included in a systematic review, a thorough assessment of the risk of bias requires both a "study-level" assessment (e.g., adequacy of allocation concealment) and, for some features, a newer approach called "outcome-level" assessment. An outcome-level assessment involves evaluating the reliability and validity of the data for each important outcome by determining the methods used to assess them in each individual study (38). The quality of evidence may differ across outcomes, even within a study, such as between a primary efficacy outcome, which is likely to be very carefully and systematically measured, and the assessment of serious harms (39), which may rely on spontaneous reports by investigators. This information should be reported to allow an explicit assessment of the extent to which an estimate of effect is correct (38).

Importance of Reporting Biases
Different types of reporting biases may hamper the conduct and interpretation of systematic reviews. Selective reporting of complete studies (e.g., publication bias) (28) as well as the more recently empirically demonstrated "outcome reporting bias" within individual studies (40, 41) should be considered by authors when conducting a systematic review and reporting its results. Though the implications of these biases on the conduct and reporting of systematic reviews themselves are unclear, some previous research has identified that selective outcome reporting may occur also in the context of systematic reviews (42).

The flow diagram has also been modified. Before including studies and providing reasons for excluding others, the review team must first search the literature. This search results in records. Once these records have been screened and eligibility criteria applied, a smaller number of articles will remain. The number of included articles might be smaller (or larger) than the number of studies, because articles may report on multiple studies and results from a particular study may be published in several articles. To capture this information, the PRISMA flow diagram now requests information on these phases of the review process.

Table 1. **Checklist of Items to Include When Reporting a Systematic Review or Meta-Analysis**

Section/Topic	Item #	Checklist Item	Reported on Page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see Item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot.	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, users, and policy makers).	
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias).	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	

ENDORSEMENT

The PRISMA Statement should replace the QUOROM Statement for those journals that have endorsed QUOROM. We hope that other journals will support PRISMA; they can do so by registering on the PRISMA Web site. To underscore to authors, and others, the importance of transparent reporting of systematic reviews, we encourage supporting journals to reference the PRISMA Statement and include the PRISMA Web address in their instructions to authors. We also invite editorial organizations to consider endorsing PRISMA and encourage authors to adhere to its principles.

THE PRISMA EXPLANATION AND ELABORATION PAPER

In addition to the PRISMA Statement, a supporting Explanation and Elaboration document has been produced (18) following the style used for other reporting guidelines (19–21). The process of completing this document included developing a large database of exemplars to highlight how best to report each checklist item, and identifying a comprehensive evidence base to support the inclusion of each checklist item. The Explanation and Elaboration document was completed after several face-to-face meetings and numerous iterations among several meeting participants, after which it was shared with the whole group for additional revisions and final approval. Finally, the group formed a dissemination subcommittee to help disseminate and implement PRISMA.

DISCUSSION

The quality of reporting of systematic reviews is still not optimal (22–27). In a recent review of 300 systematic

Figure 1. **Flow of information through the different phases of a systematic review.**

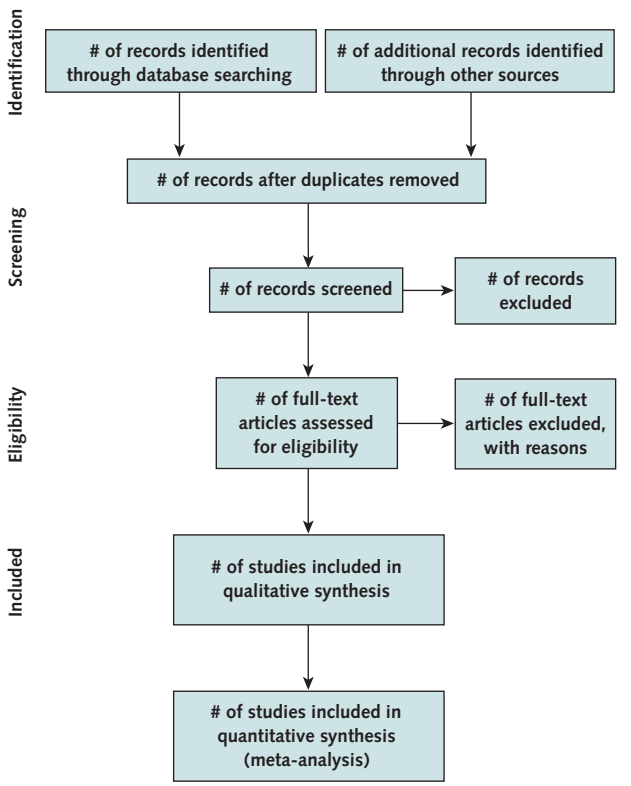


Table 2. **Substantive Specific Changes Between the QUOROM Checklist and the PRISMA Checklist***

Section/Topic	Item	QUOROM	PRISMA	Comment
Abstract		√	√	QUOROM and PRISMA ask authors to report an abstract. However, PRISMA is not specific about format.
Introduction	Objective		√	This new item (4) addresses the explicit question the review addresses using the PICO reporting system (which describes the participants, interventions, comparisons, and outcome[s] of the systematic review), together with the specification of the type of study design (PICOS); the item is linked to Items 6, 11, and 18 of the checklist.
Methods	Protocol		√	This new item (5) asks authors to report whether the review has a protocol and if so how it can be accessed.
Methods	Search	√	√	Although reporting the search is present in both QUOROM and PRISMA checklists, PRISMA asks authors to provide a full description of at least one electronic search strategy (Item 8). Without such information it is impossible to repeat the authors' search.
Methods	Assessment of risk of bias in included studies	√	√	Renamed from "quality assessment" in QUOROM. This item (12) is linked with reporting this information in the results (Item 19). The new concept of "outcome-level" assessment has been introduced.
Methods	Assessment of risk of bias across studies		√	This new item (15) asks authors to describe any assessments of risk of bias in the review, such as selective reporting within the included studies. This item is linked with reporting this information in the results (Item 22).
Discussion		√	√	Although both QUOROM and PRISMA checklists address the discussion section, PRISMA devotes three items (24–26) to the discussion. In PRISMA the main types of limitations are explicitly stated and their discussion required.
Funding			√	This new item (27) asks authors to provide information on any sources of funding for the systematic review.

* A tick indicates the presence of the topic in QUOROM or PRISMA.

reviews, few authors reported assessing possible publication bias (22), even though there is overwhelming evidence both for its existence (28) and its impact on the results of systematic reviews (29). Even when the possibility of publication bias is assessed, there is no guarantee that systematic reviewers have assessed or interpreted it appropriately (30). Although the absence of reporting such an assessment does not necessarily indicate that it was not done, reporting an assessment of possible publication bias is likely to be a marker of the thoroughness of the conduct of the systematic review.

Several approaches have been developed to conduct systematic reviews on a broader array of questions. For example, systematic reviews are now conducted to investigate cost-effectiveness (31), diagnostic (32) or prognostic questions (33), genetic associations (34), and policy making (35). The general concepts and topics covered by PRISMA are all relevant to any systematic review, not just those whose objective is to summarize the benefits and harms of a health care intervention. However, some modifications of the checklist items or flow diagram will be necessary in particular circumstances. For example, assessing the risk of bias is a key concept, but the items used to assess this in a diagnostic review are likely to focus on issues such as the spectrum of patients and the verification of disease status, which differ from reviews of interventions. The flow diagram will also need adjustments when reporting individual patient data meta-analysis (36).

We have developed an explanatory document (18) to increase the usefulness of PRISMA. For each checklist item, this document contains an example of good reporting, a rationale for its inclusion, and supporting evidence, including references, whenever possible. We believe this document will also serve as a useful resource for those teaching systematic review methodology. We encourage journals to include reference to the explanatory document in their Instructions to Authors.

Like any evidence-based endeavor, PRISMA is a living document. To this end we invite readers to comment on the revised version, particularly the new checklist and flow diagram, through the PRISMA Web site. We will use such information to inform PRISMA's continued development.

From Ottawa Methods Centre, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada; Università di Modena e Reggio Emilia, Modena, Italy; Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri, Milan, Italy; and Centre for Statistics in Medicine, University of Oxford, Oxford, United Kingdom.

Acknowledgment: The following people contributed to the PRISMA Statement: Doug Altman, DSc, Centre for Statistics in Medicine (Oxford, United Kingdom); Gerd Antes, PhD, University Hospital Freiburg (Freiburg, Germany); David Atkins, MD, MPH, Health Services Research & Development Service, Veterans Health Administration (Washington, DC); Virginia Barbour, MRCP, DPhil, *PLoS Medicine* (Cambridge, United Kingdom); Nick Barrowman, PhD, Children's Hospital of Eastern Ontario (Ottawa, Ontario, Canada); Jesse A. Berlin, ScD, Johnson & Johnson Pharmaceutical Research and Development (Titusville, New Jersey); Jocalyn Clark, PhD, *PLoS Medicine* (at the time of writing, *BMJ*; London, United Kingdom); Mike Clarke, PhD, UK Cochrane Centre (Oxford, United Kingdom) and School of Nursing and Midwifery, Trinity College (Dublin, Ireland); Deborah Cook, MD, Departments of Medicine, Clinical Epidemiology and Biostatistics, McMaster University (Hamilton, Ontario, Canada); Roberto D'Amico, PhD, Università di Modena e Reggio Emilia (Modena, Italy) and Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri (Milan, Italy); Jonathan J. Deeks, PhD, University of Birmingham (Birmingham, United Kingdom); P.J. Devereaux, MD, PhD, Departments of Medicine, Clinical Epidemiology and Biostatistics, McMaster University (Hamilton, Ontario, Canada); Kay Dickersin, PhD, Johns Hopkins Bloomberg School of Public Health (Baltimore, Maryland); Matthias Egger, MD, Department of Social and Preventive Medicine, University of Bern (Bern, Switzerland); Edzard Ernst, MD, PhD, FRCP, FRCP(Edin), Peninsula Medical School (Exeter, United Kingdom); Peter C. Gøtzsche, MD, MSc, The Nordic Cochrane Centre (Copenhagen, Denmark); Jeremy Grimshaw, MBChB, PhD, FRCFP, Ottawa Hospital Research Institute (Ottawa, Ontario, Canada); Gordon Guyatt, MD, Departments of Medicine, Clinical Epidemiology and Biostatistics, McMaster University (Hamilton, Ontario, Canada); Julian Higgins, PhD, MRC Biostatistics Unit (Cambridge, United Kingdom); John P.A. Ioannidis, MD, University of Ioannina Campus (Ioannina, Greece); Jos Kleijnen, MD, PhD, Kleijnen Systematic Reviews Ltd (York, United Kingdom) and School for Public Health and Primary Care (CAPHRI), University of Maastricht (Maastricht, the Netherlands); Tom Lang, MA, Tom Lang Communications and Training (Davis, California); Alessandro Liberati, MD, Università di Modena e Reggio Emilia (Modena, Italy) and Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri (Milan, Italy); Nicola Magrini, MD, NHS Centre for the Evaluation of the Effectiveness of Health Care–CeVEAS (Modena, Italy); David McNamee, PhD, *The Lancet* (London, United Kingdom); Lorenzo Moja, MD, MSc, Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri (Milan, Italy); David Moher, PhD, Ottawa Methods Centre, Ottawa Hospital Research Institute (Ottawa, Ontario, Canada); Cynthia Mulrow, MD, MSc, *Annals of Internal Medicine* (Philadelphia, Pennsylvania); Maryann Napoli, Center for Medical Consumers (New York, New York); Andy Oxman, MD, Norwegian Health Services Research Centre (Oslo, Norway); Ba' Pham, MMath, Toronto Health Economics and Technology Assessment Collaborative (Toronto, Ontario, Canada; at the time of the first meeting of the group, GlaxoSmithKline Canada [Mississauga, Ontario, Canada]); Drummond Rennie, MD, FRCP, FACP, University of California, San Francisco (San Francisco, California); Margaret Sampson, MLIS, Children's Hospital of Eastern Ontario (Ottawa, Ontario, Canada); Kenneth F. Schulz, PhD, MBA, Family Health International (Durham, North Carolina); Paul G. Shekelle, MD, PhD, Southern California Evidence-Based Practice Center (Santa Monica, California); Jennifer Tetzlaff, BSc, Ottawa Methods Centre, Ottawa Hospital Research Institute (Ottawa, Ontario, Canada); David Tovey, FRCGP, The Cochrane Library, Cochrane Collaboration (Oxford, United Kingdom; at the time of the first meeting of the group, *BMJ* [London, United Kingdom]); and Peter Tugwell, MD, MSc, FRCPC, Institute of Population Health, University of Ottawa (Ottawa, Ontario, Canada).

Grant Support: PRISMA was funded by the Canadian Institutes of Health Research; Università di Modena e Reggio Emilia, Italy; Cancer Research UK; Clinical Evidence BMJ Knowledge; The Cochrane Collaboration; and GlaxoSmithKline, Canada. Dr. Liberati is funded, in part, through grants of the Italian Ministry of University (COFIN-PRIN 2002 prot. 2002061749 and COFIN-PRIN 2006 prot. 2006062298). Dr. Altman is funded by Cancer Research UK. Dr. Moher is funded by a University of Ottawa Research Chair. None of the sponsors had any involvement in the planning, execution, or write-up of the PRISMA documents. Additionally, no funder played a role in drafting the manuscript.

Potential Financial Conflicts of Interest: None disclosed.

Corresponding Author: David Moher, PhD, Ottawa Methods Centre, Ottawa Hospital Research Institute, The Ottawa Hospital, General Campus, Critical Care Wing (Eye Institute), 6th Floor, 501 Smyth Road, Ottawa, Ontario K1H 8L6, Canada; e-mail, dmoher@ohri.ca.

Current author addresses are available at www.annals.org.

References

1. **Oxman AD, Cook DJ, Guyatt GH.** Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. JAMA. 1994;272:1367-71. [PMID: 7933399]
2. **Swingler GH, Volmink J, Ioannidis JP.** Number of published systematic reviews and global burden of disease: database analysis. BMJ. 2003;327:1083-4. [PMID: 14604930]
3. **Canadian Institutes of Health Research.** Randomized controlled trials registration/application checklist. December 2006. Accessed at www.cihr-irsc.gc.ca/e/documents/rct_reg_e.pdf on 19 May 2009.
4. **Young C, Horton R.** Putting clinical trials into context. Lancet. 2005;366:107-8. [PMID: 16005318]
5. **Mulrow CD.** The medical review article: state of the science. Ann Intern Med. 1987;106:485-8. [PMID: 3813259]
6. **Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC.** Meta-analyses of randomized controlled trials. N Engl J Med. 1987;316:450-5. [PMID: 3807986]
7. **Sacks HS, Reitman D, Pagano D, Kupelnick B.** Meta-analysis: an update. Mt Sinai J Med. 1996;63:216-24. [PMID: 8692168]
8. **Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF.** Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet. 1999;354:1896-900. [PMID: 10584742]
9. **Green S, Higgins J, eds.** Glossary. In: Cochrane Handbook for Systematic Reviews of Interventions 4.2.5. The Cochrane Collaboration; 2005. Accessed at www.cochrane.org/resources/glossary.htm on 19 May 2009.
10. **Strech D, Tilburt J.** Value judgments in the analysis and synthesis of evidence. J Clin Epidemiol. 2008;61:521-4. [PMID: 18471654]
11. **Moher D, Tsertsvadze A.** Systematic reviews: when is an update an update? Lancet. 2006;367:881-3. [PMID: 16546523]
12. University of York Centre for Reviews and Dissemination. 2009. Accessed at www.york.ac.uk/inst/crd/ on 19 May 2009.
13. The Joanna Briggs Institute protocols & work in progress. 2009. Accessed at www.joannabriggs.edu.au/pubs/systematic_reviews_prot.php on 19 May 2009.
14. **De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al; International Committee of Medical Journal Editors.** Clinical trial registration: a statement from the International Committee of Medical Journal Editors [Editorial]. CMAJ. 2004;171:606-7. [PMID: 15367465]
15. **Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A, Boddington E.** Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. Lancet. 2004;363:1341-5. [PMID: 15110490]
16. **Bagshaw SM, McAlister FA, Manns BJ, Ghali WA.** Acetylcysteine in the prevention of contrast-induced nephropathy: a case study of the pitfalls in the evolution of evidence. Arch Intern Med. 2006;166:161-6. [PMID: 16432083]
17. **Biondi-Zoccai GG, Lotrionte M, Abbate A, Testa L, Remigi E, Burzotta F, et al.** Compliance with QUOROM and quality of reporting of overlapping meta-analyses on the role of acetylcysteine in the prevention of contrast associated nephropathy: case study. BMJ. 2006;332:202-9. [PMID: 16415336]
18. **Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche P, Ioannidis JP, et al.** The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. 2009;151:W-65-94.
19. **Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al; 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. [PMID: 11304107]
20. **Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al; Standards for Reporting of Diagnostic Accuracy.** The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Ann Intern Med. 2003;138:W1-12. [PMID: 12513067]
21. **Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al; STROBE Initiative.** Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Ann Intern Med. 2007;147:W163-94. [PMID: 17938389]
22. **Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG.** Epidemiology and reporting characteristics of systematic reviews. PLoS Med. 2007;4:e78. [PMID: 17388659]
23. **Bhandari M, Morrow F, Kulkarni AV, Tornetta P 3rd.** Meta-analyses in orthopaedic surgery. A systematic review of their methodologies. J Bone Joint Surg Am. 2001;83-A:15-24. [PMID: 11205853]
24. **Kelly KD, Travers A, Dorgan M, Slater L, Rowe BH.** Evaluating the quality of systematic reviews in the emergency medicine literature. Ann Emerg Med. 2001;38:518-26. [PMID: 11679863]
25. **Richards D.** The quality of systematic reviews in dentistry. Evid Based Dent. 2004;5:17. [PMID: 15238972]
26. **Choi PT, Halpern SH, Malik N, Jadad AR, Tramèr MR, Walder B.** Examining the evidence in anesthesia literature: a critical appraisal of systematic reviews. Anesth Analg. 2001;92:700-9. [PMID: 11226105]
27. **Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ.** A systematic evaluation of the quality of meta-analyses in the critical care literature. Crit Care. 2005;9:R575-82. [PMID: 16277721]
28. **Dickersin K.** Publication bias: recognizing the problem, understanding its origins and scope, and preventing harm. In: Rothstein HR, Sutton AJ, Borenstein M, eds. Publication Bias in Meta-Analysis—Prevention, Assessment and Adjustments. Chichester, UK: J Wiley; 2005:11-33.
29. **Sutton AJ.** Evidence concerning the consequences of publication and related biases. In: Rothstein HR, Sutton AJ, Borenstein M, eds. Publication Bias in Meta-Analysis—Prevention, Assessment and Adjustments. Chichester, UK: J Wiley; 2005:175-92.
30. **Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I.** The case of the misleading funnel plot. BMJ. 2006;333:597-600. [PMID: 16974018]
31. **Ladabaum U, Chopra CL, Huang G, Scheiman JM, Chernew ME, Fendrick AM.** Aspirin as an adjunct to screening for prevention of sporadic colorectal cancer. A cost-effectiveness analysis. Ann Intern Med. 2001;135:769-81. [PMID: 11694102]
32. **Deeks JJ.** Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. BMJ. 2001;323:157-62. [PMID: 11463691]
33. **Altman DG.** Systematic reviews of evaluations of prognostic variables. BMJ. 2001;323:224-8. [PMID: 11473921]
34. **Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG.** Replication validity of genetic association studies. Nat Genet. 2001;29:306-9. [PMID: 11600885]
35. **Lavis J, Davies H, Oxman A, Denis JL, Golden-Biddle K, Ferlie E.** Towards systematic reviews that inform health care management and policy-making. J Health Serv Res Policy. 2005;10 Suppl 1:35-48. [PMID: 16053582]
36. **Stewart LA, Clarke MJ.** Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. Stat Med. 1995;14:2057-79. [PMID: 8552887]
37. **Moja LP, Telaro E, D'Amico R, Moschetti I, Coe L, Liberati A.** Assessment of methodological quality of primary studies by systematic reviews: results of the metaquality cross sectional study. BMJ. 2005;330:1053. [PMID: 15817526]
38. **Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al; GRADE Working Group.** GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336:924-6. [PMID: 18436948]
39. **Schünemann HJ, Jaeschke R, Cook DJ, Bria WF, El-Solh AA, Ernst A, et al; ATS Documents Development and Implementation Committee.** An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. Am J Respir Crit Care Med. 2006;174:605-14. [PMID: 16931644]
40. **Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG.** Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA. 2004;291:2457-65. [PMID: 15161896]
41. **Chan AW, Krleza-Jerić K, Schmid I, Altman DG.** Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. CMAJ. 2004;171:735-40. [PMID: 15451835]
42. **Silagy CA, Middleton P, Hopewell S.** Publishing protocols of systematic reviews: comparing what was done to what was planned. JAMA. 2002;287:2831-4. [PMID: 12038926]

Current Author Addresses: Dr. Moher and Ms. Tetzlaff: Ottawa Methods Centre, Ottawa Hospital Research Institute, The Ottawa Hospital, General Campus, Critical Care Wing (Eye Institute), 6th Floor, 501 Smyth Road, Ottawa, Ontario K1H 8L6, Canada.

Dr. Liberati: Università di Modena e Reggio Emilia and Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri, Via La Masa 19, 20156 Milan, Italy.

Dr. Altman: Centre for Statistics in Medicine, University of Oxford, Wolfson College Annexe, Linton Road, Oxford OX2 6UD, United Kingdom.