

Checklists to improve the quality of the orthopaedic literature

Raman Mundi, Harman Chaudhry, Ishu Singh¹, Mohit Bhandari

ABSTRACT

Several checklists have been developed in an effort to help journals and researchers improve the quality of reporting in research. The CONSORT statement and the CLEAR NPT evaluate randomized trials. The MOOSE and QUOROM checklists evaluate meta-analyses. The STROBE checklists assists readers in evaluating observational studies and the STARD checklist was developed for diagnostic test evaluation. The checklists presented here provide an invaluable source of guidance to authors, journal editors and readers who are seeking to prepare and evaluate reports. As evidence-based medicine continues to establish itself as the new paradigm by which medicine is practiced, the need for good reporting for all research designs must also become commonplace as opposed to the exception.

Key words: Critical appraisal, checklists, meta-analysis, diagnostic tests, quality assessment, randomized trials

INTRODUCTION

The current quality of reporting in both the medical and surgical literature is poor and in need of immediate improvement.¹⁻⁵ The orthopaedic literature is no exception. In fact, many studies have highlighted the substandard quality of reporting in even 'high quality' study designs such as randomized controlled trials and systematic reviews.⁶⁻⁹

The true benefits of well conducted studies with valid results can only be realized if they are presented to readers in a comprehensive and transparent manner. Owing to the need for improved reporting, several checklists have been developed to guide authors preparing manuscripts for various types of study designs. Although some journals have endorsed certain checklists with marginal improvements in reporting, the quality is still often poor as authors do not adhere to many of the checklist recommendations.^{5,10}

There is a dire need to promote awareness and understanding of these available checklists so that authors can begin taking

advantage of these invaluable guides. The objective of this report is to introduce several of the existing checklists for various study designs. In particular, the study designs focused on in this article include randomized controlled trials, systematic reviews, observational studies, diagnostic trials and qualitative studies.

CHECKLISTS FOR RANDOMIZED CONTROLLED TRIALS

Although randomized controlled trials (RCTs) are considered the 'gold standard' study designs for evaluating treatment effectiveness, they nevertheless remain subject to bias unless methodological and statistical safeguards are implemented into the trial. Briefly, such safeguards include allocation concealment, blinding, ensuring complete patient follow-up and analyses according to the intention-to-treat principle. This is just a sample of many safeguards and other components of randomized trials that should be reported in trial manuscripts. For instance, there are methodological issues unique to nonpharmacological trials (NPTs) that also merit reporting, such as standardization of the intervention and ensuring adequate care provider skill.

To aid investigators in preparing comprehensive and high quality manuscripts for randomized trials, two checklists have gained much attention: 1) the Consolidated Standards of Reporting Trials (CONSORT) statement and 2) the Checklist to Evaluate a Report of a Nonpharmacological Trial (CLEAR NPT).

The CONSORT statement was first published in The Journal of the American Medical Association in 1996 and revised in 2001. This statement consists of a 22-item checklist and flow diagram that serve as a detailed set of recommendations on

Department of Surgery, McMaster University, Hamilton, Ontario, Canada
¹Division of Orthopedic Surgery, Millcreek Community Hospital, Erie, PA, USA

Correspondence: Raman Mundi,
 Department of Surgery, Division of Orthopaedic Surgery, McMaster University,
 Hamilton General Hospital, 293 Wellington Street North, Suite 110, Hamilton,
 Ontario, L8L 8E7, Canada.
 E-mail: mundir2@mcmaster.ca

Disclaimer: Dr. Bhandari was supported in part by a Canada Research Chair. Raman Mundi is the recipient of the 'Studentships in Musculoskeletal Research' from Canadian Institutes of Health Research. Raman Mundi and Harman Chaudhry are recipients of the Bachelor of Health Sciences Programme '2007 Research Scholarships'.

how to prepare a report for a randomized trial or conversely, aid in critically appraising the reports of others. In particular, the checklist provides recommendations on how to report the design, analysis and interpretation of the study, whereas the flow diagram offers guidance on how to report the progress of participants through the trial.¹¹ Realizing that randomized trials of nonpharmacological interventions have unique challenges—such as complex interventions and difficulty blinding patients—not fully addressed by the revised CONSORT statement, an extension to the CONSORT was recently made to specifically address the issues facing these trials.¹² Both the standard CONSORT checklist and the modified version for NPTs can be seen in Figure 1. Figure 2 illustrates the standard CONSORT flow diagram and Figure 3 illustrates the modified flow diagram for NPTs. Adherence in the orthopaedic literature to the CONSORT recommendations has been demonstrated to be poor.^{8,9} For instance, Bhandari and colleagues⁸ evaluated the reports of 196 randomized trials investigating fracture care across 32 journals and found that the average report adhered to only 32% ± 29% of the CONSORT criteria.

Developed in 2005, the CLEAR NPT is a 15-item checklist (10 main items and 5 sub-items), that serves to critically appraise the reports of randomized trials of nonpharmacological interventions [Figure 4].¹³ As implied by its name, this checklist is useful for assessing nonpharmacological trials due to its focus on key methodological issues surrounding NPTs. Each item on the checklist can be answered with a quick Yes, No or Unclear, making it an efficient tool to evaluate the literature with. Further supporting the claims of studies utilizing the CONSORT statement, Chan and Bhandari¹⁴ have demonstrated that the quality of reporting for randomized trials in the orthopaedic literature as assessed by the CLEAR NPT is also suboptimal. Although this checklist's primary utility is to evaluate reports, it still serves as a useful guide to authors preparing manuscripts of randomized trials.

CHECKLISTS FOR SYSTEMATIC REVIEWS AND META-ANALYSES

With the plethora of studies being published constantly, summarizing the results of primary articles on a given topic is a useful and helpful practice for health care providers and policy makers.¹⁵ Systematic reviews are considered high quality evidence due to their systematic approach at collecting, critically appraising and synthesizing data from original articles on a specific topic. If a quantitative analysis is performed to arrive at a single best estimate of the treatment effect, these reviews are better known as meta-analyses. Due to their systematic nature and ability to put forth a single best estimate of the treatment

effect, meta-analyses can have a significant impact on patient care. However, meta-analyses may vary in their methodological rigor and produce results of varying credibility. For instance, the quality of a systematic review or meta-analyses is directly dependent on the quality of the studies included. Thus, systematic reviews and meta-analyses that consider only RCTs would provide stronger evidence than those which consider non-randomized studies as well. However, between 1996 and 2001, it was found that the majority of orthopaedic systematic reviews published in peer-reviewed journals outside the Cochrane collaboration included non-randomized trials.¹⁶ Clinicians reading the reports of systematic reviews and meta-analyses must be able to appraise the methods and validity of the study in order to confidently interpret their results.

The Quality of Reporting of Meta-Analyses (QUOROM) statement, consisting of an 18-item checklist and flow diagram, was developed to aid authors preparing reports of meta-analyses of RCTs [Figure 5 and 6].¹⁵ The QUOROM checklist outlines a set of recommendations on how to prepare the abstract, introduction, methods, results and discussion sections of a meta-analysis. The ultimate goal of these reporting guidelines is to provide readers with transparency regarding the search, selection, validity assessments, data abstraction, study characteristics, quantitative data synthesis and trial flow of the study.¹⁵ For instance, under the “methods” section of the checklist, authors are encouraged to report the criteria used to assess the quality of the included RCTs and the outcome of such quality assessments. This is imperative, as RCTs with deficiencies in certain methodological safeguards have been shown to produce biased results.⁵ Incorporating these ‘biased’ studies without caution into a meta-analysis would also result in a biased estimate of the treatment effect in the meta-analysis. The purpose of the flow diagram is to help authors on reporting details of the inclusion and exclusion of RCTs. Although the QUOROM statement is designed specifically to guide reporting of meta-analyses, authors of systematic reviews of RCTs can also benefit from these recommendations—with the exception of the reporting recommendations geared towards the quantitative analysis, as this step is only carried out in a meta-analysis.¹⁵

Not all meta-analyses can rely solely on RCTs to answer a question of interest. First and foremost, RCTs are relatively scarce in the orthopaedic literature, making it impractical to always exclusively use data from RCTs.⁶ Secondly, for issues surrounding risk factors for disease and harm, it would be unethical to randomize patients to groups in which they would be subject to any potentially harmful risks.¹⁷ For instance, if the question of interest was, “What is the

Section	Item	Standard CONSORT Description	Extension for Nonpharmacologic Trials
Title and abstract	1	How participants were allocated to interventions (e.g., "random allocation," "randomized," or "randomly assigned")	In the abstract, description of the experimental treatment, comparator, care providers, centers, and blinding status
Introduction			
Background	2	Scientific background and explanation of rationale	
Methods			
Participants	3	Eligibility criteria for participants and the settings and locations where the data were collected	When applicable, eligibility criteria for centers and those performing the interventions
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered	Precise details of both the experimental treatment and comparator
	4A		Description of the different components of the interventions and, when applicable, descriptions of the procedure for tailoring the interventions to individual participants
	4B		Details of how the interventions were standardized
	4C		Details of how adherence of care providers with the protocol was assessed or enhanced
Objectives	5	Specific objectives and hypotheses	
Outcomes	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors)	
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules	When applicable, details of whether and how the clustering by care providers or centers was addressed
Randomization-sequence generation	8	Method used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification)	When applicable, how care providers were allocated to each trial group
Allocation concealment	9	Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned	
Implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups	
Blinding (masking)	11A	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment	Whether or not those administering co-interventions were blinded to group assignment
	11B†		If blinded, method of blinding and description of the similarity of interventions†
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s); methods for additional analyses, such as subgroup analyses and adjusted analyses	When applicable, details of whether and how the clustering by care providers or centers was addressed
Results			
Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended)—specifically, for each group, report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome; describe protocol deviations from study as planned, together with reasons	The number of care providers or centers performing the intervention in each group and the number of patients treated by each care provider or in each center
Implementation of intervention	New item		Details of the experimental treatment and comparator as they were implemented
Recruitment	14	Dates defining the periods of recruitment and follow-up	
Baseline data	15	Baseline demographic and clinical characteristics of each group	When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centers (volume) in each group
Numbers analyzed	16	Number of participants (denominator) in each group included in each analysis and whether analysis was by "intention-to-treat"; state the results in absolute numbers when feasible (e.g., 10/20, not 50%)	
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group and the estimated effect size and its precision (e.g., 95% confidence interval)	
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those prespecified and those exploratory	
Adverse events	19	All important adverse events or side effects in each intervention group	
Discussion			
Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes	In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centers in each group
Generalizability	21	Generalizability (external validity) of the trial findings	Generalizability (external validity) of the trial findings according to the intervention, comparators, patients, and care providers and centers involved in the trial
Overall evidence	22	General interpretation of the results in the context of current evidence	

Figure 1: CONSORT checklist¹²

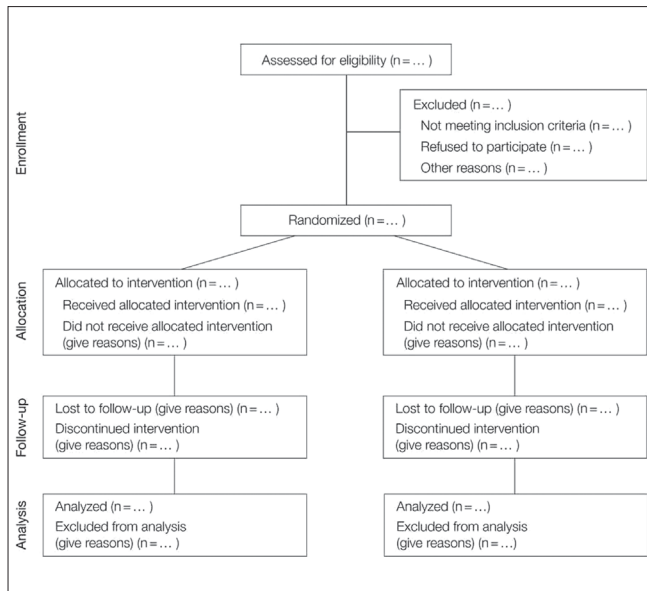


Figure 2: Standard CONSORT flow diagram¹¹

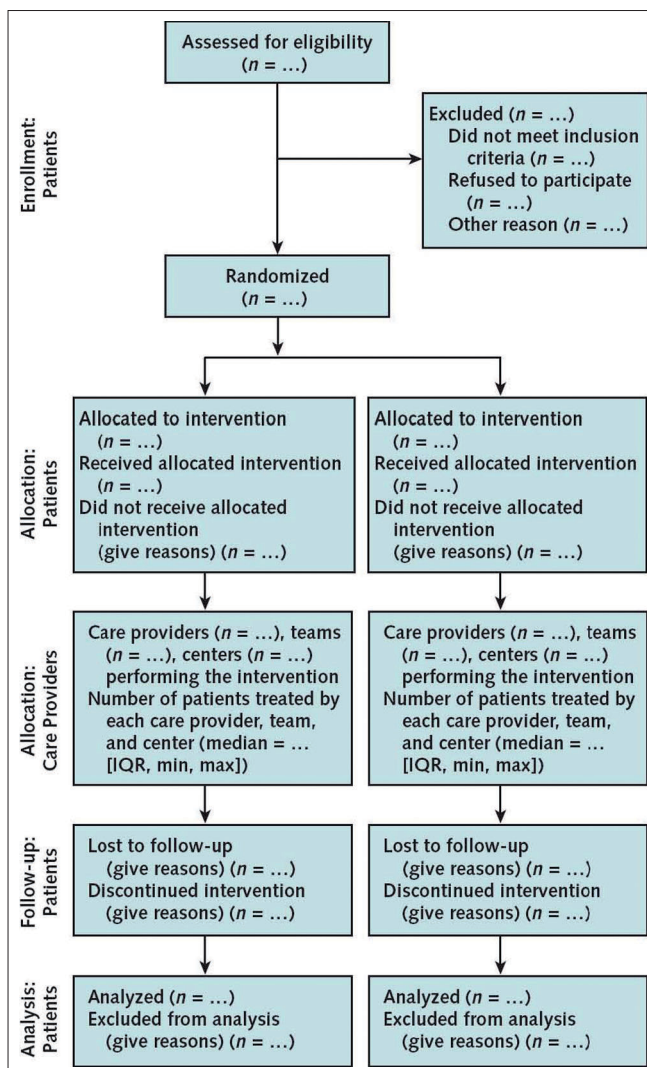


Figure 3: CONSORT flow diagram for Nonpharmacological Trials¹²

effect of smoking on fracture nonunion rates?”, it would not be feasible to carry out a randomized study in which patients were randomly allocated to either a smoking or non-smoking group.¹⁸ As a result, several meta-analyses rely upon observational studies - those in which patients are naturally exposed to risk factors or in which physician or patient preference determines allocation to a treatment or control intervention.¹⁹ The importance of comprehensive reporting for such meta-analyses can not be overstated, as observational studies are more prone to biased results than RCTs. In fact, arriving at a single estimate of the effect of a treatment or exposure when pooling data from observational studies requires extreme caution, as these results can often be misleading.¹⁶

In 2000, the Meta-Analyses of Observational Studies in Epidemiology (MOOSE) group produced a 35-item checklist that details how meta-analyses of observational studies should be reported. Specifically, the checklist provides recommendations on how to report background information, the search strategy and the methods, results, discussion and conclusion sections of the paper [Figure 7]. For each of these six categories, there are several corresponding items that are listed by the MOOSE groups as essential for reporting. For instance, under search strategy, reports should include the qualifications of the searchers, a detailed description of the search strategy and the method by which articles in foreign languages were utilized, among other details regarding the search.¹⁷

CHECKLISTS FOR OBSERVATIONAL STUDIES

As mentioned, not all research questions can be answered through RCTs. Therefore, observational studies have an important role in answering questions of treatment effectiveness and disease etiology. Three primary observational study designs include the cohort, case-control and cross-sectional studies. Briefly, cohort studies usually follow two groups of patients; one group in which everyone has been exposed to a risk factor or treatment and the other in which no exposure has occurred. The groups are then compared for the rate of development of disease or outcome of interest. In case-control studies, a group that has already developed an outcome of interest is compared to a group without the outcome for factors that may be associated with the outcome. Cross-sectional surveys are carried out at a single time point, at which both the outcome and factors of interest are measured. Due to the lack of randomization, observational studies are inherently more prone to potential biases. Even if investigators attempt to match groups for known prognostic factors, there may be underlying imbalances in unknown prognostic factors that may produce misleading and biased results.²⁰ Furthermore, case-control studies are

Item	Possible answers
1. Was the generation of allocation sequences adequate?	Yes; No; Unclear
2. Was the treatment allocation concealed?	Yes; No; Unclear
3. Were details of the intervention administered to each group made available? ^a	Yes; No; Unclear
4. Were care providers' experience or skill ^b in each arm appropriate? ^c	Yes; No; Unclear
5. Was participant (i.e., patients) adherence assessed quantitatively? ^d	Yes; No; Unclear
6. Were participants adequately blinded?	Yes; No, because blinding is not feasible; No, although blinding is feasible; Unclear
6.1. If participants were not adequately blinded	
6.1.1. Were all other treatments and care (i.e., cointerventions) the same in each randomized group?	Yes; No; Unclear
6.1.2. Were withdrawals and lost to follow-up the same in each randomized group?	Yes; No; Unclear
7. Were care providers or persons caring for the participants adequately blinded?	Yes; No, because blinding is not feasible; No, although blinding is feasible; Unclear
7.1. If care providers were not adequately blinded	
7.1.1. Were all other treatments and care (i.e., cointerventions) the same in each randomized group?	Yes; No; Unclear
7.1.2. Were withdrawals and lost to follow-up the same in each randomized group?	Yes; No; Unclear
8. Were outcome assessors adequately blinded to assess the primary outcomes?	Yes; No, because blinding is not feasible; No, although blinding is feasible; Unclear
8.1. If outcome assessors were not adequately blinded, were specific methods used to avoid ascertainment bias (systematic differences in outcome assessment)? ^e	Yes; No; Unclear
9. Was the follow-up schedule the same in each group? ^f	Yes; No; Unclear
10. Were the main outcomes analyzed according to the intention-to-treat principle?	Yes; No; Unclear

^a The answer should be "yes" for this item if these data were either described in the report or made available for each arm (reference to a preliminary report, online addendum etc.)

^b Care provider experience or skill will be assessed only for therapist-dependent interventions (i.e., interventions where the success of the treatment are directly linked to care providers' technical skill). For other treatment, this item is not relevant and should be removed from the checklist or answered "unclear."

^c Appropriate experience or skill should be determined according to published data, preliminary studies, guidelines, run-in period, or a group of experts and should be specified in the protocol for each study arm before the beginning of the survey.

^d Treatment adherence will be assessed only for treatments necessitating iterative interventions (e.g., physiotherapy that supposes several sessions, in contrast to a one-shot treatment such as surgery). For one-shot treatments, this item is not relevant and should be removed from the checklist or answered "unclear."

^e The answer should be "yes" for this item, if the main outcome is objective or hard, or if outcomes were assessed by a blinded or at least an independent endpoint review committee, or if outcomes were assessed by an independent outcome assessor trained to perform the measurements in a standardized manner, or if the outcome assessor was blinded to the study purpose and hypothesis.

^f This item is not relevant for trials in which follow-up is part of the question. For example, this item is not relevant for a trial assessing frequent vs. less frequent follow-up for cancer recurrence. In these situations, this item should be removed from the checklist or answered "unclear."

Figure 4: CLEAR NPT checklist¹³

always retrospective in nature (cohort studies may also be retrospective) which increases the potential for incomplete and biased data collection.¹⁸ Despite these limitations, observational studies have a crucial role to play in medical research and as such, satisfactory reporting to allow readers to evaluate these studies is of utmost importance.

The Strengthening the Reporting of Observational Research in Epidemiology (STROBE) statement, outlines how to prepare good manuscripts for these three observational study designs. It consists of a 22-item checklist which provides reporting recommendations for all sections of the paper, as well as on funding sources [Figure 8]. Of the 22-items, 18 are general to all three study designs and 4 are design-specific. In particular, information in the methods section regarding participants (item 6) and statistical methods (item 12), as well as in the results section regarding descriptive data (item 14) and outcome data (item 15) are design-specific. Although the STROBE group emphasizes

that reporting of all 22-items in this checklist is essential, they encourage authors to utilize their preferences and creativity when selecting the order and format of presenting such details.²¹

CHECKLISTS FOR STUDIES OF DIAGNOSTIC ACCURACY

Diagnostic tests are widely used by clinicians to diagnose health states and subsequently initiate, alter or terminate various treatment options.²² Diagnostic studies evaluate the accuracy of a diagnostic test (by its level of agreement to the current 'gold standard' for diagnosis) in predicting a disease, stage of a disease, health status or any health condition that could prompt clinical action.²² The 'gold standard' is typically impractical to use in regular clinical encounters and therefore the study is attempting to offer a more practical alternative.²³ As such, studies of diagnostic test accuracy have the potential to directly impact treatment decisions and, therefore, patient care. Unfortunately, it has

Heading	Subheading	Descriptor	Reported? (Y/N)	Page number
Title		Identify the report as a meta-analysis [or systematic review] of RCTs ²⁶		
Abstract		Use a structured format ²⁷		
		Describe		
	Objectives	The clinical question explicitly		
	Data sources	The databases (ie, list) and other information sources		
	Review methods	The selection criteria (ie, population, intervention, outcome, and study design); methods for validity assessment, data abstraction, and study characteristics, and quantitative data synthesis in sufficient detail to permit replication		
	Results	Characteristics of the RCTs included and excluded; qualitative and quantitative findings (ie, point estimates and confidence intervals); and subgroup analyses		
	Conclusion	The main results		
		Describe		
Introduction		The explicit clinical problem, biological rationale for the intervention, and rationale for review		
Methods	Searching	The information sources, in detail ²⁸ (eg, databases, registers, personal files, expert informants, agencies, hand-searching), and any restrictions (years considered, publication status, ²⁹ language of publication ^{30,31})		
	Selection	The inclusion and exclusion criteria (defining population, intervention, principal outcomes, and study design ³²)		
	Validity assessment	The criteria and process used (eg, masked conditions, quality assessment, and their findings ³³⁻³⁶)		
	Data abstraction	The process or processes used (eg, completed independently, in duplicate) ^{35,36}		
	Study characteristics	The type of study design, participants' characteristics, details of intervention, outcome definitions, &c, ³⁷ and how clinical heterogeneity was assessed		
	Quantitative data synthesis	The principal measures of effect (eg, relative risk), method of combining results (statistical testing and confidence intervals), handling of missing data; how statistical heterogeneity was assessed; ³⁸ a rationale for any a-priori sensitivity and subgroup analyses; and any assessment of publication bias ³⁹		
Results	Trial flow	Provide a meta-analysis profile summarising trial flow (see figure)		
	Study characteristics	Present descriptive data for each trial (eg, age, sample size, intervention, dose, duration, follow-up period)		
	Quantitative data synthesis	Report agreement on the selection and validity assessment; present simple summary results (for each treatment group in each trial, for each primary outcome); present data needed to calculate effect sizes and confidence intervals in intention-to-treat analyses (eg 2x2 tables of counts, means and SDs, proportions)		
Discussion		Summarise key findings; discuss clinical inferences based on internal and external validity; interpret the results in light of the totality of available evidence; describe potential biases in the review process (eg, publication bias); and suggest a future research agenda		
Quality of reporting of meta-analyses				

Figure 5: QUOROM checklist¹⁵

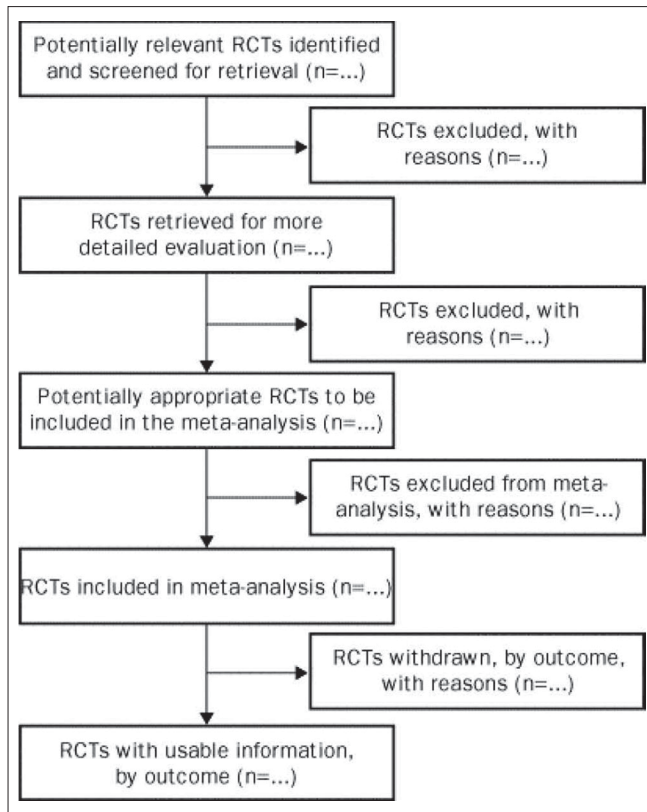


Figure 6: QUOROM flow diagram¹⁵

been demonstrated that methodologically compromised studies are more likely than methodologically rigorous studies to overestimate the accuracy of diagnostic tests.²⁴ In the hands of the uncritical clinician, such poor studies may lead to the unwarranted use and interpretation of a diagnostic test, ultimately to the detriment of high quality patient care.

Recognizing the importance of studies evaluating diagnostic accuracy, Bossuyt and colleagues²⁵ developed the Standards for the Reporting of Diagnostic Accuracy studies (STARD) statement. The STARD statement includes a 25-item checklist which outlines crucial information that authors should include in the abstract, introduction, methods, results and discussion sections of a report to enable an adequate assessment of both external validity (i.e. how generalizable study results are) and internal validity (i.e. the potential for bias) [Figure 9].²⁶ In addition to some relatively common elements, such as inclusion/exclusion criteria, method of data collection and method of data analysis, the STARD checklist also includes some unique items. For instance, it asks for a description and rationale of the gold standard to which the diagnostic test (referred to as the index test) is being compared. This is because even positive study results will be limited by the effectiveness of the 'gold standard' as a diagnostic tool.²³ The STARD statement also includes and encourages

authors to use a flow diagram to report the number of patients included and excluded in the diagnostic and/or 'gold standard' tests [Figure 10].²⁶

Rama and colleagues recently published an investigation of diagnostic accuracy studies using the STARD criteria in three orthopaedic journals.²⁷ They found that the majority of studies had deficiencies in reporting of methodology and, overall, reported less than two-thirds of the STARD criteria. Currently, no major orthopaedic journals have adopted the STARD statement;²⁸ this may be partly attributable to the scarcity of diagnostic accuracy studies, which constitute only 1% of the orthopaedic literature.²⁷ However, owing to the enormous implications that a newly implemented diagnostic test can have on patient care, we believe that the STARD statement must be endorsed to enable readers to adequately interpret study results and prevent untenable treatment decisions.

CHECKLISTS FOR QUALITATIVE STUDIES

Qualitative studies are useful in the surgical literature for comprehensively describing phenomena (from social, emotional and experiential perspectives) as well as for generating hypotheses that can subsequently be quantitatively verified or disproven.^{29,30} Giacomini, Cook and Guyatt have stated that qualitative studies can provide a "rigorous alternative to armchair hypothesizing".²⁹ For instance, one study published in the *Journal of Bone and Joint Surgery (American Volume)* explored the reasons why (from the patient's perspective) many elderly arthritic patients are unwilling to undergo a total joint replacement procedure.³¹ The principles of evidence-based medicine demand that articles should be critically appraised before results are implemented into clinical practice. However, the methodological rigour of qualitative studies has come under criticism; commentators have stated that there is a need for rigorous methodological standards in order to minimize the effect of bias on study results.^{32,33}

The most comprehensive available checklist for qualitative studies is the RATS guidelines developed by Clark and adopted as a 28-item checklist by BioMed Central in the instructions to authors section [Figure 11].^{34,35} RATS is an acronym which describes four components of a rigorously reported qualitative study: 1) Relevance of the study question; 2) Appropriateness of qualitative method; 3) Transparency of procedures; and 4) Soundness of interpretive approach.³⁵ In addition to the 28-item checklist, the RATS guideline offers a section on possible "red flags" authors should avoid.³⁴ Unfortunately, there does not appear to be an overwhelming consensus on the effectiveness of this particular checklist to ensure all

Reporting of background should include

- Problem definition
- Hypothesis statement
- Description of study outcome(s)
- Type of exposure or intervention used
- Type of study designs used
- Study population

Reporting of search strategy should include

- Qualifications of searchers (eg, librarians and investigators)
- Search strategy, including time period included in the synthesis and keywords
- Effort to include all available studies, including contact with authors
- Databases and registries searched
- Search software used, name and version, including special features used (eg, explosion)
- Use of hand searching (eg, reference lists of obtained articles)
- List of citations located and those excluded, including justification
- Method of addressing articles published in languages other than English
- Method of handling abstracts and unpublished studies
- Description of any contact with authors

Reporting of methods should include

- Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested
- Rationale for the selection and coding of data (eg, sound clinical principles or convenience)
- Documentation of how data were classified and coded (eg, multiple raters, blinding, and interrater reliability)
- Assessment of confounding (eg, comparability of cases and controls in studies where appropriate)
- Assessment of study quality, including blinding of quality assessors; stratification or regression on possible predictors of study results
- Assessment of heterogeneity
- Description of statistical methods (eg, complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated
- Provision of appropriate tables and graphics

Reporting of results should include

- Graphic summarizing individual study estimates and overall estimate
- Table giving descriptive information for each study included
- Results of sensitivity testing (eg, subgroup analysis)
- Indication of statistical uncertainty of findings

Reporting of discussion should include

- Quantitative assessment of bias (eg, publication bias)
- Justification for exclusion (eg, exclusion of non-English-language citations)
- Assessment of quality of included studies

Reporting of conclusions should include

- Consideration of alternative explanations for observed results
- Generalization of the conclusions (ie, appropriate for the data presented and within the domain of the literature review)
- Guidelines for future research
- Disclosure of funding source

Figure 7: MOOSE checklist¹⁷

Item	Item Number	Recommendation
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—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed. (b) Give reasons for nonparticipation at each stage. (c) Consider use of a flow diagram.
Descriptive data	14*	(a) Give characteristics of study participants (e.g., 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: Summarize follow-up time—e.g., 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 (e.g., 95% confidence intervals). Make clear which confounders were adjusted for and why they were included. (b) Report category boundaries when continuous variables were categorized. (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period.
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions and sensitivity analyses.
Discussion		
Key results	18	Summarize 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.
Generalizability	21	Discuss the generalizability (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.

Figure 8: STROBE checklist²²

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	
METHODS			
<i>Participants</i>	3	Describe the study population: The inclusion and exclusion criteria, setting and locations where the data were collected.	
	4	Describe participant recruitment: Was recruitment 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 the study population a consecutive series of participants defined by the 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 test and reference standard were performed (prospective study) or after (retrospective study)?	
<i>Test methods</i>	7	Describe the reference standard and its rationale.	
	8	Describe technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	
	9	Describe definition of and rationale for the units, cutoffs and/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	Describe whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	
<i>Statistical methods</i>	12	Describe methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	
	13	Describe methods for calculating test reproducibility, if done.	
RESULTS			
<i>Participants</i>	14	Report when study was done, including beginning and ending dates of recruitment.	
	15	Report clinical and demographic characteristics of the study population (e.g. age, sex, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centers).	
	16	Report the number of participants satisfying the criteria for inclusion that did or did not undergo the index tests and/or the reference standard; describe why participants failed to receive either test (a flow diagram is strongly recommended).	
<i>Test results</i>	17	Report time interval from the index tests to the reference standard, and any treatment administered between.	
	18	Report distribution of severity of disease (define criteria) in those with the target condition; 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, the distribution of the test results by the results of the reference standard.	
	20	Report any adverse events from performing the index tests or the reference standard.	
<i>Estimates</i>	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 the index tests were handled.	
	23	Report estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	
	24	Report estimates of test reproducibility, if done.	
DISCUSSION	25	Discuss the clinical applicability of the study findings.	

Figure 9: STARD checklist²⁷

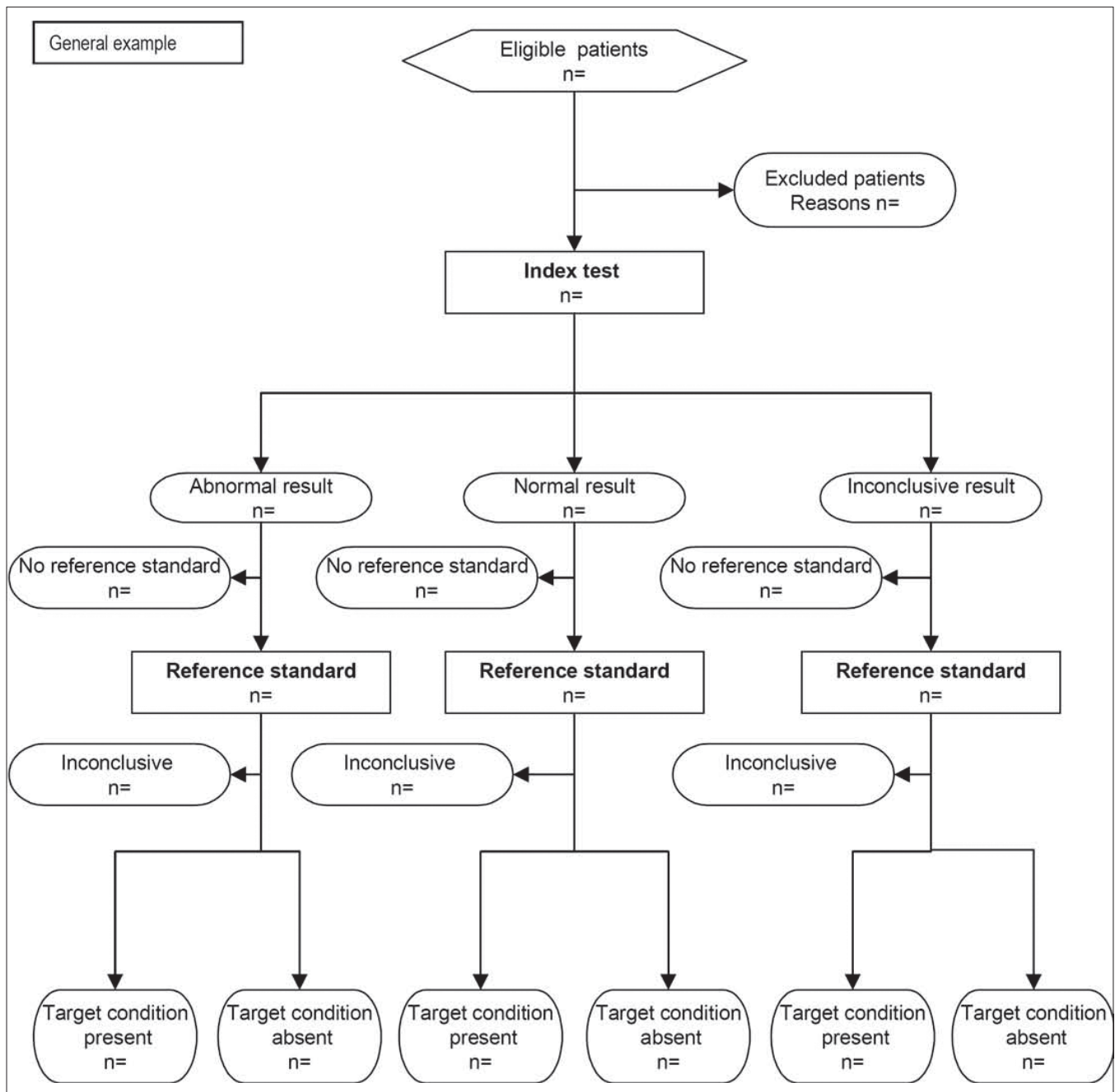


Figure 10: STARD flow diagram²⁷

pertinent methodological criteria have been met, as there is for the CONSORT criteria. However, authors of qualitative studies are advised to consider and incorporate the checklist criteria in reporting findings.

CONCLUSION

For over a decade now, several checklists have been developed in an effort to help investigators prepare reports for a variety of different study designs. If the time, effort and resources put forth in carrying out medical research

is to make its impact on patient care and policy decisions, then the importance of complete and comprehensive reporting can not be overstated. The checklists presented here provide an invaluable source of guidance to authors, journal editors and readers who are seeking to prepare and evaluate reports. To gain greater information on these and other available reporting checklists, we encourage readers to locate the original articles in which these checklists are published. Furthermore, certain checklists, such as the CONSORT statement, have corresponding explanation and elaboration papers which are informative and aid in

ASK THIS OF THE MANUSCRIPT	THIS SHOULD BE INCLUDED IN THE MANUSCRIPT	✓
<p>R Relevance of study question</p> <p>Is the research question interesting?</p> <p>Is the research question relevant to clinical practice, public health, or policy?</p>	<p>Research question explicitly stated</p> <p>Research question justified and linked to the existing knowledge base (empirical research, theory, policy)</p>	
<p>A Appropriateness of qualitative method</p> <p>Is qualitative methodology the best approach for the study aims?</p> <p><i>Interviews:</i> experience, perceptions, behaviour, practice, process</p> <p><i>Focus groups:</i> group dynamics, convenience, non-sensitive topics</p> <p><i>Ethnography:</i> culture, organizational behaviour, interaction</p> <p><i>Textual analysis:</i> documents, art, representations, conversations</p>	<p>Study design described and justified e.g., why was a particular method (i.e., interviews) chosen?</p>	
<p>T Transparency of procedures</p> <p><i>Sampling</i></p> <p>Are the participants selected the most appropriate to provide access to type of knowledge sought by the study?</p> <p>Is the sampling strategy appropriate?</p>	<p>Criteria for selecting the study sample justified and explained</p> <p><i>theoretical:</i> based on pre conceived or emergent theory</p> <p><i>purposive:</i> diversity of opinion</p> <p><i>volunteer:</i> feasibility, hard-to-reach groups</p>	
<p><i>Recruitment</i></p> <p>Was recruitment conducted using appropriate methods?</p> <p>Is the sampling strategy appropriate?</p> <p>Could there be selection bias?</p>	<p>Details of how recruitment was conducted and by whom</p> <p>Details of who chose not to participate and why</p>	
<p><i>Data collection</i></p> <p>Was collection of data systematic and comprehensive?</p> <p>Are characteristics of the study group and setting clear?</p> <p>Why and when was data collection stopped, and is this reasonable?</p>	<p>Method (s) outlined and examples given (e.g., interview questions)</p> <p>Study group and setting clearly described</p> <p>End of data collection justified and described</p>	
<p><i>Role of researchers</i></p> <p>Is the researcher (s) appropriate? How might they bias (good and bad) the conduct of the study and results?</p>	<p>Do the researchers occupy dual roles (clinician and researcher)?</p>	

Figure 11: RATS checklist³⁵

	Are the ethics of this discussed? Do the researcher(s) critically examine their own influence on the formulation of the research question, data collection, and interpretation?	
<p><i>Ethics</i></p> <p>Was informed consent sought and granted?</p> <p>Were participants' anonymity and confidentiality ensured?</p> <p>Was approval from an appropriate ethics committee received?</p>	<p>Informed consent process explicitly and clearly detailed</p> <p>Anonymity and confidentiality discussed</p> <p>Ethics approval cited</p>	
<p>S Soundness of interpretive approach</p> <p><i>Analysis</i></p> <p>Is the type of analysis appropriate for the type of study? <i>thematic:</i> exploratory, descriptive, hypothesis generating <i>framework:</i> e.g., policy <i>constant comparison/grounded theory:</i> theory generating, analytical</p> <p>Are the interpretations clearly presented and adequately supported by the evidence?</p> <p>Are quotes used and are these appropriate and effective?</p> <p>Was trustworthiness/reliability of the data and interpretations checked?</p>	<p>Analytic approach described in depth and justified</p> <p><i>Indicators of quality:</i> Description of how themes were derived from the data (inductive or deductive)</p> <p>Evidence of alternative explanations being sought</p> <p>Analysis and presentation of negative or deviant cases</p> <p>Description of the basis on which quotes were chosen</p> <p>Semi-quantification when appropriate</p> <p>Illumination of context and/or meaning, richly detailed</p> <p>Method of reliability check described and justified e.g., was an audit trail, triangulation, or member checking employed? Did an independent analyst review data and contest themes? How were disagreements resolved?</p>	
<p><i>Discussion and presentation</i></p> <p>Are findings sufficiently grounded in a theoretical or conceptual framework?</p> <p>Is adequate account taken of previous knowledge and how the findings add?</p> <p>Are the limitations thoughtfully considered?</p> <p>Is the manuscript well written and accessible?</p>	<p>Findings presented with reference to existing theoretical and empirical literature, and how they contribute</p> <p>Strengths and limitations explicitly described and discussed</p> <p>Evidence of following guidelines (format, word count)</p> <p>Detail of methods or additional quotes contained in appendix</p> <p>Written for a health sciences audience</p>	

Figure 11: (Continue)

<p>Are red flags present? these are common features of ill conceived or poorly executed qualitative studies, are a cause for concern, and must be viewed critically. They might be fatal flaws, or they may result from lack of detail or clarity.</p>	<p><i>Grounded theory</i>: not a simple content analysis but a complex, sociological, theory generating approach <i>Jargon</i>: descriptions that are trite, pat, or jargon filled should be viewed sceptically <i>Over interpretation</i>: interpretation must be grounded in "accounts" and semi-quantified if possible or appropriate <i>Seems anecdotal, self evident</i>: may be a superficial analysis, not rooted in conceptual framework or linked to previous knowledge, and lacking depth <i>Consent process thinly discussed</i>: may not have met ethics requirements <i>Doctor-researcher</i>: consider the ethical implications for patients and the bias in data collection and interpretation</p>	<p>?</p>
---	--	----------

Figure 11: (Continue)

promoting understanding of the checklists. As evidence-based medicine continues to establish itself as the new paradigm by which medicine is practiced, the need for good reporting for all research designs must also become commonplace as opposed to the exception.

REFERENCES

1. Tiruvoipati R, Balasubramanian SP, Atturu G, Peek GJ, Elbourne D. Improving the quality of reporting randomized controlled trials in cardiothoracic surgery: The way forward. *J Thorac Cardiovasc Surg* 2006;132:233-40.
2. Balasubramanian SP, Wiener M, Alshameeri Z, Tiruvoipati R, Elbourne D, Reed MW. Standards of reporting of randomized controlled trials in general surgery: Can we do better? *Ann Surg* 2006;244:663-7.
3. Agha R, Cooper D, Muir G. The reporting quality of randomised controlled trials in surgery: A systematic review. *Int J Surg* 2007;5:413-22.
4. Jacquier I, Boutron I, Moher D, Roy C, Ravaud P. The reporting of randomized clinical trials using a surgical intervention is in need of immediate improvement: A systematic review. *Ann Surg* 2006;244:677-83.
5. Mills E, Wu P, Gagnier J, Heels-Ansdell D, Montori V. An analysis of general medical and specialist journals that endorse CONSORT found that reporting was not enforced consistently. *J Clin Epidemiol* 2005;58:662-7.
6. Bhandari M, Richards RR, Sprague S, Schemitsch EH. The quality of reporting randomized trials in the Journal of Bone and Joint Surgery from 1988 through 2000. *J Bone Joint Surg Am* 2002;84:388-96.
7. Poolman R, Struijs PA, Krips R, Sierevelt IN, Lutz KH, Bhandari M. Does a "Level I Evidence" rating imply high quality of reporting in orthopaedic randomised controlled trials? *BMC Med Res Methodol* 2006;6:44-51.
8. Bhandari M, Guyatt GH, Lochner H, Sprague S, Tornetta P 3rd. Application of the consolidated standards of reporting trials (CONSORT) in the fracture care literature. *J Bone Joint Surg Am*

- 2002;84:485-9.
9. Cowan J, Lozano-Calderon S, Ring D. Quality of prospective controlled randomized trials: Analysis of trials of treatment for lateral epicondylitis as an example. *J Bone Joint Surg Am* 2007;89:1693-9.
10. Devereaux PJ, Manns BJ, Ghali WA, Quan H, Guyatt GH. The reporting of methodological factors in randomized controlled trials and the association with a journal policy to promote adherence to the Consolidated Standards of Reporting Trials (CONSORT) checklist. *Control Clin Trials* 2002;23:380-8.
11. The CONSORT group. CONSORT 2001 checklist and CONSORT 2001 flow diagram. [cited on 2008 Mar 1]. Available from: <http://www.consort-statement.org/index.aspx?o = 1031>.
12. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacological treatments: Explanation and elaboration. *Ann Intern Med* 2008;148:205-9.
13. Boutron I, Moher D, Tugwell P, Giraudeau B, Poiraudau S, Nizard R, et al. A checklist to evaluate a report of a nonpharmacological trial (CLEAR NPT) was developed using consensus. *J Clin Epidemiol* 2005;58:1233-40.
14. Chan S, Bhandari M. The quality of reporting of orthopaedic randomized trials with use of a checklist for nonpharmacological therapies. *J Bone Joint Surg Am* 2007;89:1970-8.
15. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. *Lancet* 1999;354:1896-900.
16. Audige L, Bhandari M, Griffin D, Middleton P, Reeves BC. Systematic reviews of nonrandomized clinical studies in the orthopaedic literature. *Clin Orthop Relat Res* 2004;427:249-57.
17. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: A proposal for reporting. *JAMA* 2000;283:2008-12.
18. Petrisor BA, Bhandari M. The hierarchy of evidence: levels and grades of recommendation. *Indian J Orthop* 2007;41:11-5.
19. Guyatt GH. Therapy and harm: An introduction. In: Guyatt GH, Rennie D, editors. *Users' guides to the medical literature: A manual for evidence-based clinical practice/The Evidence-*

- Based Medicine Working Group. Chicago: AMA Press; 2002. p. 49-53.
20. Levine M, Haslam D, Walter S, Cumming R, Lee H, Haines T, *et al.* Harm. *In: Guyatt GH, Rennie D, editors. Users' guides to the medical literature: A manual for evidence-based clinical practice/The Evidence-Based Medicine Working Group.* Chicago: AMA Press; 2002. p. 81-100.
 21. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Report of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *Ann Intern Med* 2007;147:573-7.
 22. 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. *Ann Intern Med* 2003;138:W1-12.
 23. Furukawa TA, Guyatt GH. Sources of bias in diagnostic accuracy studies and the diagnostic process. *CMAJ* 2006;174:481-2.
 24. Lijmer JG, Mol BW, Heisterkamp S, Bossel 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.
 25. 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. *Clin Chem* 2003;49:1-6.
 26. The STARD group. STARD checklist and STARD flow diagram. [cited on 2008 Mar 1]. Available from: <http://www.stard-statement.org>.
 27. Rama S, Boddu KR, Sharmila P, Apsingi S. Quality of reporting of orthopaedic diagnostic accuracy studies is suboptimal. *Clin Ortho Rel Res* 2006;446:247.
 28. The STARD group. Adopters of STARD. [cited on 2008 Mar 1]. Available from: <http://www.stard-statement.org>.
 29. Giacomini M, Cook D, Guyatt GH. Therapy and applying the results: Qualitative research. *In: Guyatt GH, Rennie D, editors. Users' guides to the medical literature: a manual for evidence-based clinical practice/The Evidence-Based Medicine Working Group.* Chicago: AMA Press; 2002. p. 433-48.
 30. Rich JA, Grey CM. Qualitative research on trauma surgery: getting beyond the numbers. *World J Surg* 2003;27:957-61.
 31. Clark JP, Hudak PL, Hawker GA, Coyte PC, Mohamed NN, Kreder HJ, *et al.* The moving target: A qualitative study of elderly patients' decision-making regarding total joint replacement surgery. *J Bone Joint Surg Am* 2004;86:1366-74.
 32. Poses RM, Isen AM. Qualitative research in medicine and health care. *J Gen Intern Med* 1998;13:32-8.
 33. Poses RM, Isen AM. In defense of qualitative research: responses to Poses and Isen perspective article. *J Gen Intern Med* 1998;13:64-72.
 34. Clark JP. Qualitative research review guidelines: RATS. [cited on 2008 Mar 1]. Available from: <http://www.biomedcentral.com/info/ifora/rats>.
 35. Clark JP. How to peer review a qualitative manuscript. *In: Godlee F, Jefferson T, editors. Peer review in health sciences. Second edition.* London: BMJ Books; 2003. p. 219-35.

Source of Support: Nil, Conflict of Interest: None.
--