Enhancing primary reports of randomized controlled trials: Three most common challenges and suggested solutions

Guowei Li\textsuperscript{a,b,c}, Meha Bhatt\textsuperscript{a}, Mei Wang\textsuperscript{a,c}, Lawrence Mbuagbawa\textsuperscript{a,c}, Zainab Samaan\textsuperscript{d}, and Lehana Thabane\textsuperscript{a,b,c,1}

\textsuperscript{a}Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada L8S 4L8; \textsuperscript{b}Centre for Evaluation of Medicines, Programs for Assessment of Technology in Health Research Institute, McMaster University, Hamilton, ON, Canada L8N 1Y3; \textsuperscript{c}Biostatistics Unit, Research Institute at St. Joseph’s Healthcare Hamilton, Hamilton, ON, Canada L8N 4A6; and \textsuperscript{d}Department of Psychiatry and Behavioral Neurosciences, McMaster University, Hamilton, ON, Canada L8S 4L8

Edited by David B. Allison, Indiana University, Bloomington, IN, and accepted by Editorial Board Member Susan T. Fiske July 14, 2017 (received for review May 30, 2017)

Evidence from a well-designed randomized controlled trial (RCT) is generally considered to be the gold standard that can inform clinical practice and guide decision-making. However, several deficiencies in the reporting of RCTs have frequently been identified, including incomplete, selective, and biased or inconsistent reporting. Such suboptimal reporting may lead to irreproducible results, substantial waste of resources, impaired study validity, erosion of public trust in science, and a high risk of research misconduct. In this article, we present an overview of the reporting of RCTs in the biomedical literature with a focus on the three most common reporting problems: (i) lack of adherence to reporting guidelines, (ii) inconsistencies between trial protocols or registrations and full reports, and (iii) inconsistencies between abstracts and their corresponding full reports. Unsatisfactory levels of adherence to guidelines and frequent inconsistencies between protocols or registrations and full reports, and between abstracts and full reports, were consistently found in various biomedical research fields. A variety of factors were found to be associated with these reporting challenges. Improved reporting can build public trust and credibility of science, save resources, and enhance the ethical integrity of research. Therefore, joint efforts from the various sectors of the biomedical community (researchers, journal editors and reviewers, educators, healthcare providers, and other research consumers) are needed to reduce and reverse the current suboptimal state of RCT reporting in the literature.

In evidence-based medicine, evidence from well-designed randomized controlled trials (RCTs) is generally considered to be the gold standard to inform clinical practice and guide decision-making (1). However, several deficiencies in the reporting of RCTs have frequently been identified, including incomplete, selective, and biased or inconsistent reporting. For example, it was found that only about 60% of RCT reports provided adequate information on interventions (2). This suboptimal reporting of clinical trials may lead to irreproducible results, substantial waste of resources, impaired study validity, erosion of public trust in science, and a high risk of research misconduct mainly due to the potential consequence of patients suffering and dying unnecessarily (3). Moreover, enhancing the quality and transparency of reporting is an ethical and scientific obligation in order to avoid wasting research resources, minimize risk of harms and maximize benefits of therapies, and enhance research integrity (4–6). Therefore, improving reporting of health research is an imperative that requires a variety of efforts by the biomedical community (researchers, journal editors and reviewers, educators, healthcare providers, and other research consumers). In this article, we highlight and provide some insights into the three most common reporting issues of RCTs: (i) lack of adherence to reporting guidelines, (ii) inconsistencies between protocols or registrations and full reports, and (iii) inconsistencies between abstracts and their corresponding full reports.

Lack of Adherence to Reporting Guidelines for RCTs

To improve reporting of clinical trials, the Consolidated Standards of Reporting Trials (CONSORT) statement was published in 1996 as the first guideline for reporting biomedical research (7). The CONSORT statement is evidence-based guidance that is regularly updated and extended, with the latest version published in 2010 (8). Since its appearance, there have been several extensions to CONSORT intended to adapt their application to (i) specific trial designs, such as cluster trials, N-of-1 trials, pragmatic trials, and pilot and feasibility trials, among others; (ii) specific types of interventions, such as herbal medicinal, nonpharmacological, and acupuncture; (iii) different types of outcomes, such as patient-reported outcomes and harms; and (iv) specific formats of trial reports such as abstracts and RCT protocols, all of which can be found in the Enhancing Quality and Transparency of Health Research (EQUATOR) network (9) and on the CONSORT website (10). Some studies have revealed that the use of the CONSORT statement is associated with enhanced quality of reporting of clinical trials and, in particular, that checking submitted manuscripts for missing items from the CONSORT list in peer-review processes can improve the quality of peer reviews and the final publications (11–14). For example, it was found that peer reviewers failed to detect important deficiencies in reporting of the methods and results of RCTs (15). The study by Cobo et al. (13) was the first RCT evaluating the effect of using the reporting guideline checklists during the review process on study quality. Cobo et al. (13) observed an improvement of manuscript quality in the group that used conventional peer reviews plus additional review looking for missing items from CONSORT and Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines in the peer-review process, compared with the group using conventional reviews alone.
Nevertheless, the levels of adherence to guidelines of RCTs remain suboptimal, unsatisfactory, inadequate, or poor. One study exploring the reporting quality in 150 RCTs of surgical interventions found a low level of guideline adherence, with only 55% of the CONSORT checklist items adequately reported (16). Most included trials did not satisfactorily report sequence generation for randomization (57%), allocation concealment (55%), blinding (65%), or sample-size calculation (55%). Similarly, a survey evaluating the reporting of subfertility trials found that 15% of trials (24 of 164) reported randomization inadequately, while only 10 trials (6%) provided transparent and complete details of sequence generation and allocation concealment (17). Unsatisfactory compliance with CONSORT was also observed in trials of herbal medicine interventions (18), oncology (19), cardiology (20), anesthesiology (21), dentistry (22), ophthalmology (23), and others (24). A recent Cochrane review also revealed the suboptimal CONSORT adherence levels and common incompleteness of reporting in 16,604 RCTs, which could provide a more generalizable picture of unsatisfactory reporting guideline adherence (25).

Adherence of RCTs to reporting guidelines remains an open area for further research with the potential for substantial improvement. Reasons for the lack of adherence to CONSORT may be (i) the unclear reporting requirements in the instructions to authors; (ii) the multiplicity of the CONSORT extensions, making the use of them difficult; and (iii) the lack of training and education of authors, making the active implementations and interventions warranted, among others.

Some studies found that adherence was improved in recently published RCTs. For example, in the study by Alvarez et al. (26) comparing the reporting quality of dermatology RCTs published in 1997 and 2006, it was observed that the reporting had significantly improved, with a guideline adherence level of 28% with CONSORT criteria in 2006, compared with 11% in 1997 (P value = 0.03). Nevertheless, another study assessing the reporting of RCTs of pediatric dentistry found negligible improvement in reporting quality after comparing the RCTs published between 1985 and 1997 with those published between 1998 and 2006 (27). Similar small changes in reporting quality were also observed in the RCTs of cerebral palsy published between 1990 and 1997 and between 1998 and 2002, with adherence levels of 41 and 46%, respectively (28). The generalizability of these findings may be limited, due to the nature of the observational study. For example, the nature of reporting in journals or diseases selected. Moreover, the cross-sectional designs of studies, the potential risk of confounding, and the phenomenon of regression toward the mean would impair the validity of their findings. Another study found that the reporting of RCTs had been improved over time especially for sequence generation and allocation concealment in 20,920 RCTs included in Cochrane reviews published between March 2011 and September 2014 (29). Nevertheless, there is clear evidence that the current state of reporting of RCTs requires more improvement, especially for the trials published in journals with a low impact factor (29).

To improve the reporting of RCTs, there have been some studies to investigate what factors are related to increased or decreased adherence to CONSORT. For example, Eithgen et al. (30) found that the journal Impact Factor and being published in CONSORT-endorsing journals were positively and significantly associated with improved reporting quality of RCTs evaluating stents used in percutaneous coronary interventions. Other factors related to improved reporting quality included trials with negative findings (31), a more recent year of publication (32), a large sample size (33), and trials receiving industrial funding (34), among others. Some modifiable factors—including endorsement of CONSORT by journals, a requirement by a journal to upload the CONSORT checklist at the time of submission, and considering the use of CONSORT guidelines in the study design and writing process—may represent a feasible option to assist in improved quality and transparency of trial reports (24).

However, the relationship between these factors and reporting quality was identified from cross-sectional designs, which therefore would weaken the strength of the evidence and make the causality questionable. Moreover, given the differences in the research questions and methodology of the individual studies, the external validity and generalizability of their findings might be compromised. Some studies have been published to explore how to increase adherence to the reporting guidelines. For example, Hopewell et al. (35) found that active implementation of the CONSORT guidelines for abstracts by journals could lead to improvements in the reporting quality, while passive endorsement through instruction to authors failed to improve the completeness of reporting. Likewise, a study by Barnes et al. (36) showed that a CONSORT-based writing aid tool comprising reminders of the CONSORT item(s), bullet points explaining all of the key components to be reported, and examples of good reporting could significantly improve the completeness of manuscripts reporting the results of RCTs, compared with no use of the writing-aid tool. These findings indicate opportunities to improve the guideline adherence and quality of RCT reporting. Other solutions to improving guideline adherence may include systematically promoting training and education, advancing journal endorsement of CONSORT with an active policy to implement reporting guidelines, and enhancing the use of guidelines in study design and the manuscript-writing process, among others. However, more evidence on how to cost-effectively and pragmatically implement the aforementioned strategies is clearly needed in the literature.

### Inconsistency Between Protocols or Registrations and Full Reports of RCTs

A prospective trial protocol or registration, if adequately reported, aids in improving reporting of RCTs, because of its prespecified information for assessment of and comparison with its published full reports (37). One of the most influential trial registries, ClinicalTrials.gov, was initially developed mainly because of lobbying by breast cancer survivors and their advocates (38). Subsequently, the Food and Drug Administration Modernization Act of 1997 was passed with the objective of regulating trial implementation and enhancing trial reporting (38). In 2004, the International Committee of Medical Journal Editors (ICMJE) announced that all clinical trials must be registered ahead of their enrollment before they can be considered for publication (39). Similarly, the World Health Organization International Clinical Trials Registry Platform was established in 2006, aiming to enhance global access to a registry with information about all clinical trials that is specifically available to patients, families, physicians, researchers, and others (40). Compared with the high acceptance of and public access to trial registrations, however, it is difficult to access the trial protocols. Some journals request authors to submit their protocols and give public access to the protocol when the article is published, while the trial protocols cannot be available in most of the other journals. Nevertheless, many journals have started establishing editorial policies for the publication of clinical trial protocols and have tried to make the protocol publications open access. In 2013, the Standard Protocol Items: Recommendations for Interventional Trials statement was therefore published to assist in transparent reporting and to improve the quality of the emerging protocol publications (41).

However, despite these multiple concerted efforts to improve the reporting of RCTs, inconsistent reporting was still found to be substantially high after comparing registrations or protocols with their full reports. Such inconsistent reporting poses a severe threat to the validity of trial findings, given that the reporting issues are generally subject to bias and may even lead to ethical impairment (3). One study comparing protocols and full reports in randomized trials found that 62% of full reports (51 of 82) had at least one predefined primary outcome changed, omitted, or introduced (42). Likewise, another study found a median inconsistency level
Box 1.
The current situations of the three reporting challenges and some suggestions for RCT-reporting improvement

Lack of adherence to guidelines for RCT reporting
First published in 1996, the CONSORT statement and subsequent extensions aim to improve transparent and complete reporting of trials. Endorsement of CONSORT and requirement to include the CONSORT checklist at the time of submission by journals are associated with enhanced quality of trial reporting and peer reviews.

Levels of adherence to CONSORT guidelines remain suboptimal, unsatisfactory, inadequate, or poor across various biomedical research fields. Suggestions for improvements: (i) Training of researchers and reviewers on CONSORT principles and the importance of using guidelines in the study design and writing process; (ii) endorsement of CONSORT by journals; (iii) journal requirements for authors to include the CONSORT checklist during manuscript submission; (iv) journal instructions for reviewers to review CONSORT checklist as part of the peer-review process; and (v) educators, healthcare providers, and other research consumers need to be vigilant about the problem of inadequate reporting.

Inconsistency between abstracts and full reports of RCTs
Abstracts usually prepared with the least care; reporting quality of abstracts remains unsatisfactory. Comparison of abstracts with full-text reports show: (i) high levels of abstract inaccuracy across various fields and (ii) a major “spin problem” (defined as an overinterpretation of trial findings in an attempt to show significant results or draw strong conclusions, despite findings for the primary outcome in full reports being clearly not significant). Suggestions for improvement: (i) Authors need to follow the CONSORT extension for reporting of trial abstracts; (ii) authors need to carefully check the abstract to ensure its accuracy and consistency with findings in the full report; (iii) journal requirements should include instruction for authors to confirm that information in the abstract has been verified against what is reported in the full report; (iv) editorial staff and reviewers need to be cautious about the “spin problem”; and (v) educators, healthcare providers, and other research consumers need to pay particular attention to the wording of conclusions or interpretations of findings.

Inconsistency between protocols or registrations and full reports of RCTs
Trial registration requirement and trial protocol publication aim to reduce publication bias.

Major inconsistencies exist between trial protocols or registrations and corresponding full trial reports across various fields, including outcome measures, subgroup analyses, statistical analyses, and other trial aspects.

Results not statistically significant for primary outcomes are significantly associated with more inconsistencies.

Suggestions for improvement: (i) Authors need to explain the protocol modifications that occur during trial conduct that may be seen as major discrepancies between full reports and the protocols or registrations; (ii) journal requirements for submission of trial reports should include a list of trial modifications; (iii) journal peer-review process of full trial reports needs to include a careful check of protocols or registrations for discrepancies; and (iv) educators, healthcare providers, and other research consumers need be cautious about the potential for discrepancies.

Inconsistency between protocols or registrations and full reports of RCTs
Trial registration requirement and trial protocol publication aim to reduce publication bias.

Major inconsistencies exist between trial protocols or registrations and corresponding full trial reports across various fields, including outcome measures, subgroup analyses, statistical analyses, and other trial aspects.

Results not statistically significant for primary outcomes are significantly associated with more inconsistencies.

Suggestions for improvement: (i) Authors need to explain the protocol modifications that occur during trial conduct that may be seen as major discrepancies between full reports and the protocols or registrations; (ii) journal requirements for submission of trial reports should include a list of trial modifications; (iii) journal peer-review process of full trial reports needs to include a careful check of protocols or registrations for discrepancies; and (iv) educators, healthcare providers, and other research consumers need be cautious about the potential for discrepancies.

Inconsistency between abstracts and full reports of RCTs
Abstracts usually prepared with the least care; reporting quality of abstracts remains unsatisfactory. Comparison of abstracts with full-text reports show: (i) high levels of abstract inaccuracy across various fields and (ii) a major “spin problem” (defined as an overinterpretation of trial findings in an attempt to show significant results or draw strong conclusions, despite findings for the primary outcome in full reports being clearly not significant). Suggestions for improvement: (i) Authors need to follow the CONSORT extension for reporting of trial abstracts; (ii) authors need to carefully check the abstract to ensure its accuracy and consistency with findings in the full report; (iii) journal requirements should include instruction for authors to confirm that information in the abstract has been verified against what is reported in the full report; (iv) editorial staff and reviewers need to be cautious about the “spin problem”; and (v) educators, healthcare providers, and other research consumers need to pay particular attention to the wording of conclusions or interpretations of findings.

CONSORT, Consolidated Standards of Reporting Trials; RCT, randomized controlled trial.

of 31% in primary outcomes in RCTs, based on 27 reviews that compared full reports with trial registries (43). One study in progress, entitled “COMPare Trials” (compare-trials.org/), compares each clinical trial report with its protocol or registry entry regarding study outcomes in the top five medical journals. The preliminary findings of this study show that, on average, each trial reported just 58.2% of its specified outcomes while, on average, each trial silently added 5.3 new outcomes (44). Several studies that identified inconsistencies in outcomes between protocols or registrations with full trial reports revealed that most of the discrepancies favored statistically significant findings in the full reports (45–48). Inconsistencies were also observed in subgroup and statistical analyses. For example, in the study by Kasenda et al. (49), 26% of trials (132 of 515) provided post hoc subgroup analyses that were not mentioned in their protocols, while 12% (64 of 515) did not include subgroup analyses that had been specifically planned in their protocols. Another study revealed that about half of the full trial reports (47%, 25 of 60) had at least one discrepancy in statistical analyses compared with the registrations, design, or protocol papers (50). Further discrepancies could be found in full reports with respect to funding, sample-size calculations, randomization, blinding, conclusions, or other aspects compared with their corresponding registrations or protocols (51). One study concluded that the inconsistent reporting problem of RCTs published in gastroenterology and hepatology journals may have improved from 2009 to 2012; however, as emphasized in the study’s Discussion section, this conclusion was probably influenced by sampling bias (52).

Two surveys observed that results that were nonsignificant for primary outcomes were related to higher odds of incomplete reporting in full reports, with an odds ratio of 2.7 [95% confidence interval (CI): 1.5–5.0] and 4.7 (95% CI: 1.8–12.0), respectively (42, 53). The positive relationship between an absence of statistical significance and incomplete reporting, together with the fact that most inconsistencies favored a statistically significant finding in the full reports, may be due mainly to a publication bias that favors a claim of statistically significant results (43, 54). These inconsistent reporting issues in full reports would therefore lead to research findings being questionable and irreproducible and result in evidence-biased syntheses for decision-making. Joint efforts from authors, sponsors, regulators, and journals are needed to mitigate and reverse this inconsistent reporting. For example, it is the authors’ obligation to explain the inevitable modifications in full reports compared with the protocols or registrations; sponsors and regulators should closely monitor the trial implementation, data analyses and interpretation, and manuscript writing; and journals (editorial staff and reviewers) should carefully check protocols or registrations to scrutinize inconsistencies in full reports for peer review and decision-making. Moreover, guidance and/or checklists are required to assist the multiple stakeholders in their prompt and easy identification and subsequent evaluation of inconsistencies between registration or protocols and full reports.

Inconsistency Between Abstracts and Full Reports of RCTs

Abstracts of RCTs usually provide key summaries of a trial including study design and writing process; (ii) endorsement of CONSORT by journals; (iii) journal requirements for authors to include the CONSORT checklist during manuscript submission; (iv) journal peer-review process of full trial reports needs to include a careful check of protocols or registrations for discrepancies; and (v) educators, healthcare providers, and other research consumers need be cautious about the potential for discrepancies.
meetings. Abstracts are used to guide readers to seek more information about the trials or even used to aid in readers’ decision-making, especially when the full reports cannot be accessed (55). However, despite being the most-read section in publications, abstracts are usually prepared with the least care (56). Even though guidelines such as the CONSORT extension to abstracts have been published, the reporting quality of abstracts in RCTs remains unsatisfactory (57, 58). Inaccurate information summarized in the abstracts will distract or mislead the audiences in the biomedical community. Nevertheless, some studies evaluating the inconsistencies between abstracts and full trial reports have found high levels of abstract inaccuracy (59–61).

Compared with the more detailed information found in the main texts of trial publications, one study of spinal journals revealed that 75% of abstracts (30 of 40) had at least one deficiency; the abstracts were considered to have a deficiency if their data were inconsistent with full trial reports, if the data could not be found at all in the full reports, or if the abstracts did not report pertinent negative findings (60). Among a sample of RCT conference abstracts in cardiology, Toma et al. (62) found that 24% (35 of 148) presented different sample sizes and 41% (60 of 148) had different treatment-effect estimates compared with the full reports. Similarly, another study observed that 16% of conference abstracts in oncology differed in defining a primary outcome, 54% differed in the number of participants randomized, and 78% differed in the number of participants analyzed, compared with the full reports (58). Boutron et al. (63) explored the so-called “spin problem” in abstract reporting, where “spin” was defined as an overinterpretation in an attempt to show significant results or draw strong conclusions despite clearly non-significant findings for the primary outcome in full reports. They found that 38% of the Results (27 of 72) and 58% of the Conclusions (42 of 72) sections in the abstracts demonstrated spin. To avoid such subtle spin in abstracts, it is therefore recommended that authors and journals should carefully check the abstract reporting to ensure its accuracy and consistency (64). After a submission is accepted, strict copyediting and proof-reading should also be conducted to improve the consistency of abstract reporting (61).

No study provided evidence on any factors related to improved abstract reporting in RCTs. However, for the conference abstracts, a longer time between abstracts and full reports being published was found to significantly increase the likelihood of an inconsistency, as reported in two studies that did not focus on trials (65, 66). These inconsistencies may be due to extended study duration when more data are collected, which would inevitably lead to the full reports being different from their previous conference abstracts. Nevertheless, to reduce or prevent inconsistencies between abstracts and full reports, it is expected that authors should provide details and explanations of the changes that have occurred and that journals should refer to the previous conference abstracts during their peer review.

Conclusions

In this article, we have presented an overview of the reporting of RCTs in the biomedical literature, with a focus on the three most common reporting problems (suboptimal guideline adherence, high inconsistencies between protocols or registrations and full reports, and high inconsistencies between abstracts and full reports). Box 1 highlights the three challenges and their current situations and provides some suggestions for RCT-reporting improvement. The incomplete or inconsistent reporting of full reports thwarts scientific progress, yields results that are irreproducible, wastes research resources, threatens the reliability and validity of evidence published, and impairs ethical integrity (3, 67). Although the biomedical community has made some efforts, including enforcing the ICMJE policy of trial registration and developing the CONSORT and corresponding extensions for protocols, abstracts, and full reports, the quality and transparency of RCT reporting still needs substantial improvement.

We have focused only on the three most common reporting issues of RCTs in this article. However, these reporting issues were not uncommon in other types of research or in other contexts. For example, inconsistencies between protocols or registrations and full reports can be found in observational studies (68) and systematic reviews (69), while inconsistencies between abstracts and full reports have also been identified in diagnostic studies (70) and veterinary science (66). Likewise, poor guideline adherence levels have been found in animal research regarding the guidelines of Animal Research: Reporting In Vivo Experiments (71) and in health economic studies regarding Consolidated Health Economic Evaluation Reporting Standards guidelines (72). Reporting improvement is clearly needed in the state-of-the-art biomedical literature.

Ioannidis (73) has commented that most claimed research findings in scientific fields are in fact false. The framework that he proposed could be helpful in evaluating the likelihood of research findings being true, in which relevant factors include the number of studies conducted in the field, the magnitude of effect sizes, the number of hypotheses, the flexibility in study designs and methodology, conflicts of interest, and the compulsion of multiple stakeholders to chase statistical significance. However, we cannot assess the likelihood of a research result being true if there is no transparent and consistent reporting in the published full reports.

Systematic training and intervention for all stakeholders will undoubtedly be beneficial to enhancing reporting. Several factors, based on cross-sectional associations, have been found to be associated with better reporting of RCTs, and they may provide some insight into the reporting problem. However, more evidence is largely required for how to cost-effectively and pragmatically improve RCT reporting in the biomedical literature.

Building public trust and credibility in science, saving resources, and enhancing the ethical integrity of research fall under the collective responsibility of the biomedical community including researchers, journal editors and reviewers, educators, healthcare providers, and other research consumers. Therefore, everyone needs to play a role in reducing and reversing the current suboptimal state of RCT reporting in the literature.

ACKNOWLEDGMENTS. We thank Dr. Stephen Walter for his professional editing of the manuscript.


