Peer reviewed evaluation of registered end-points of randomised trials (the PRE-REPORT study): protocol for a stepped-wedge, cluster-randomised trial

Christopher W Jones, Amanda Adams, Mark A Weaver, Sara Schroter, Benjamin S Misemer, David Schriger, Timothy F Platts-Mills

ABSTRACT

Introduction Clinical trials are critical to the advancement of medical knowledge. However, the reliability of trial conclusions depends in part on consistency between pre-planned and reported study outcomes. Unfortunately, selective outcome reporting, in which outcomes reported in published manuscripts differ from pre-specified study outcomes, is common. Trial registries such as ClinicalTrials.gov have the potential to help identify and stop selective outcome reporting during peer review by allowing peer reviewers to compare outcomes between registry entries and submitted manuscripts. However, the persistently high rate of selective outcome reporting among published clinical trials indicates that the current peer review process at most journals does not effectively address the problem of selective outcome reporting.

Methods and analysis PRE-REPORT is a stepped-wedge cluster-randomised trial that will test whether providing peer reviewers with a summary of registered, pre-specified primary trial outcomes decreases inconsistencies between prospectively registered and published primary outcomes. Peer reviewed manuscripts describing clinical trial results will be included. Eligible manuscripts submitted to each participating journal during the study period will comprise each cluster. After an initial control phase, journals will transition to the intervention phase in random order, after which peer reviewers will be emailed registry information consisting of the date of registration and any prospectively defined primary outcomes. Blinded outcome assessors will compare registered and published primary outcomes for all included trials. The primary PRE-REPORT outcome is the presence of a published primary outcome that is consistent with a prospectively defined primary outcome in the study's trial registry. The primary outcome will be analysed using a mixed effect logistical regression model to compare results between the intervention and control phases.

Ethics and dissemination The Cooper Health System Institutional Review Board determined that this study does not meet criteria for human subject research. Findings will be published in peer-reviewed journals.

Trial registration number ISRCTN41225307; Pre-results.

INTRODUCTION

Randomised trials can help to determine the impact of medical interventions on patient outcomes, and therefore form a critically important foundation on which much of the evidence-based medicine movement has been built. However, the reliability of clinical trial data depends on the consistent reporting of pre-specified trial outcomes. Changes between the pre-specified and reported outcome often reflect selective outcome reporting in which investigators or study sponsors report statistically significant treatment effects favouring the intervention which may result from multiple hypothesis testing, post hoc hypothesising and chance rather than actual efficacy of the intervention. Selective or incomplete outcome reporting is widespread throughout the published biomedical literature, occurring in an estimated 30% to 40% of published clinical trials.

Clinical trial registries were developed, in part, to solve the problem of selective outcome reporting. Registries are
publicly available databases that make trial information available to both the scientific community and the general public. This information includes descriptions of trial eligibility criteria and treatment arms, along with definitions of pre-specified primary and secondary outcomes. Since 2005, the International Committee of Medical Journal Editors (ICMJE) has mandated the prospective registration of clinical trials as a condition of publication in member journals, and in 2007 the Food and Drug Administration Amendments Act made prospective registration with ClinicalTrials.gov a requirement under federal law for many US clinical trials. Similar requirements have also been implemented by numerous other stakeholders and regulators, including the World Association of Medical Journal Editors, the WHO, the European Union and the National Institutes of Health.

Despite the widespread adoption of registration requirements, a substantial body of evidence shows that selective outcome reporting remains common, and is routinely observed among trials published in both general medical and specialty journals and across a wide range of medical specialties and funding sources. Because trial registry data are publicly available, selective outcome reporting can be detected during peer review. However, the persistence of this problem indicates that current peer review practices at most journals do not result in the consistent identification and correction of selective outcome reporting.

Several barriers likely impair the ability of standard peer review processes to detect and correct selective outcome reporting. First, some reviewers and journal editors are not fully aware of existing registry resources, or of best practices regarding trial registration and outcome reporting. Second, submitted manuscripts often fail to include the unique identifiers assigned to each trial at the time of registration, thereby necessitating an extensive search of multiple trial registries to identify a matching registry entry. Furthermore, many registries allow investigators to edit existing registry data at any time, meaning that the registered trial outcomes can be changed after trial completion to match the outcomes reported in a submitted manuscript. Such changes occur in more than 30% of registered trials. ClinicalTrials.gov and WHO-approved trial registries track these changes, but accessing the audit trail that captures changes to prospectively registered outcomes is more time-consuming than simply viewing the updated registry webpage. In addition, some reviewers may be hesitant to review registry sites because these sites typically list the study sponsor and participating enrolment sites and identify the principal study investigator. Thus, direct registry review is not compatible with blinded peer review. Finally, for many journals there are not clear guidelines delineating whether reviewers, editors or editorial staff members are responsible for comparing trial manuscripts to the relevant registry entries, and when these guidelines do exist, reviewers may not be aware of them.

The PRE-REPORT trial will test an intervention that is designed to address each of these barriers by directly providing information from the clinical trial registry to peer reviewers for use during the review of manuscripts that present the results of clinical trials. This intervention consists of a comprehensive third-party registry search, abstraction of information from the registry and provision of this information to peer reviewers. A cluster-randomised, stepped-wedge trial will be performed to test the effect of this intervention on selective outcome reporting. The goal of this study will be to determine whether providing reviewers with information about registered primary trial outcomes at the time of peer review improves clinical trial reporting by increasing the consistency between prospectively registered and reported trial outcomes. This submission reflects the study protocol as of 25 September, 2018 (version 2).

METHODS AND ANALYSIS

Study design

The PRE-REPORT trial is a stepped-wedge, cluster-randomised trial which will test the impact of providing peer reviewers with easily accessible registry information to facilitate the comparison between registered and published trial outcomes. Individual clusters within the study will consist of all eligible clinical trial manuscripts sent for peer review during the pre- or post-intervention phase for an individual journal. A cluster design, rather than manuscript-level randomisation, is necessary to minimise contamination of the intervention, as journals typically use a limited roster of decision editors and peer reviewers, and once an individual has participated in the intervention condition he or she may be more likely to seek out registry data when evaluating subsequent manuscripts. Manuscripts submitted to participating journals between 1 November, 2018, and 31 October, 2019, will be screened for inclusion.

Stepped-wedge randomisation

For the first 2 months of the trial, all participating journals will be in the control phase. Journals will then cross over to the experimental intervention phase in random order between months 3 and 10, and for the final 3 months all clusters will receive the experimental intervention (table 1). An important advantage of this study design is the ability to compare pre- and post-intervention outcomes within individual clusters, thereby controlling for potentially confounding characteristics unique to those clusters. For example, participating journals differ with respect to their existing peer review processes, as well as the volume, quality and type of individual manuscripts undergoing review. The stepped-wedge design also allows for partial (but not complete) control of secular trends in the peer review process over the yearlong study. Although temporal changes in registration, reporting and peer review practices are possible,
we think these changes are a smaller threat to the study than differences between clusters.

Journal selection:
Participating journals were identified by emailing the editors-in-chief of high impact journals across a broad range of medical specialities to describe the proposed study, gauge interest among the journal leadership and assess the feasibility of participation given each journal’s existing peer review processes. In order to be eligible for participation, journals could not have already implemented a robust process for ensuring the performance of a comprehensive registry analysis during peer review of clinical trial manuscripts. The determination of what constituted a robust existing process and whether there might be opportunity for each journal’s existing practices with regard to registry review to improve was made by the editor-in-chief of each journal approached about potential participation. Additionally, participating journals must routinely publish manuscripts describing clinical trials, which we defined as publishing a mean of at least 10 trials per year over the past 3 years. In an attempt to minimise any change in behaviour on behalf of submitting authors or peer reviewers due to participation in the study (ie, a Hawthorne effect), the identities of participating journals are not provided in the published version of the protocol or trial registration record.

Manuscript eligibility
Manuscripts that report the results of a clinical trial are eligible for inclusion if they are sent for peer review during the study period by any of the participating journals. We define clinical trials according to the criteria adopted by the ICMJE: any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes.17 We will exclude manuscripts that describe a protocol for a planned trial without reporting trial results. Manuscripts clearly stating that the manuscript is not intended to report on the trial’s primary outcome (ie, manuscripts describing only secondary analyses, secondary outcomes or re-analyses) will also be excluded. Finally, we will also exclude resubmitted manuscripts that were initially submitted prior to the 1 November 2018, start date. If any manuscripts are submitted to more than one participating journal during the study, they will be analysed in the first journal’s cluster, and will not be included a second time if resubmitted to a different participating journal.

Manuscript screening procedures at each participating journal have been individualised in order to accommodate existing editorial and peer review processes at the journals, while also allowing us to design processes that address the specific confidentiality concerns raised by editors of the participating journals. In some cases, journal staff members perform an initial screen for manuscript eligibility before alerting the PRE-REPORT investigators to a potentially eligible submission, and in some cases the PRE-REPORT investigators screen all submitted manuscripts for eligibility without involvement from the journal staff.

Control phase
Due to the stepped-wedge crossover design, each participating journal will initially be allocated to the control condition. During the control phase, the PRE-REPORT investigators will prospectively evaluate submitted manuscripts for inclusion and will perform data collection for those manuscripts meeting enrolment criteria. During this phase, there will be no change in the information that peer reviewers receive as part of the usual peer review practice at each participating journal.

Intervention phase
Between months 2 and 10, each journal will cross over into the intervention phase at monthly intervals according to a randomisation schedule created by the study statistician using computer-generated random numbers. For each manuscript submitted for peer review during the intervention phase, PRE-REPORT investigators will assess manuscript eligibility (figure 1). If eligible, the coordinator or investigator will perform a registry search to identify a registry entry matching the trial described in the submitted manuscript. After confirming a match between the submitted manuscript and a corresponding registry entry, the PRE-REPORT staff member will abstract information from the registry into a registry data form (online supplementary appendix 1), including the following information: whether the trial was registered, the date of initial registration and the registered primary outcome(s) at the time study enrolment began. At some journals this registry information sheet will be made available to reviewers at the same time that they accept an invitation to review an included trial and receive access.
Open access

Figure 1  Information flow between participating journals and PRE-REPORT team. RCT, randomised controlled trial.

to the manuscript, and at other journals reviewers will receive the registry information via email after they have already accepted the review assignment. In these cases the study team will coordinate with staff members at each journal to email the completed data form to the relevant peer reviewers; for some of the participating journals the PRE-REPORT team will directly send the registry information to reviewers using emails generated via the manuscript tracking system. In all cases our goal is to make the registry information available to all reviewers within 24 hours of accepting the review assignment. If the search fails to identify a registry entry for the study, the absence of registry data will be reported to the peer reviewers. The PRE-REPORT team will provide reviewers with information from the relevant trial registry entry, but will not provide a comparison between registered outcomes and the outcomes described in the manuscript under consideration and will not provide instructions regarding use of the registry information. Our goal is to make the relevant registry information available to peer reviewers, as we believe that one of the key responsibilities for reviewers of clinical trials is to rigorously evaluate the selection, definition and reporting of the primary trial outcomes. We anticipate that editors will also be exposed to the registry information included in our intervention, either through access to the emails providing reviewers with access to the registry information or by reading comments from reviewers who have taken this registry information into consideration when writing their reviews. Reviewers and editors will decide whether and how to use the information provided to them from the registry when evaluating the manuscript under review.

Registry data abstraction
For each manuscript, an investigator experienced in the use of trial registries will review the published manuscript for a trial registration number or other evidence of trial registration. If no registration information is provided within the manuscript, the investigator will then search ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform search portal and any national or regional registries corresponding to the principal investigators’ countries of origin (eg, Australian New Zealand Clinical Trials Registry) by keyword and title to identify a matching registry entry. Potential matches between registry entries and manuscripts will be assessed by comparing the study title, interventions, sample sizes, enrolment dates and trial locations between the registry and manuscript. Manuscripts will be classified as unregistered if they do not include a registry identification number and the registry search does not identify a matching registry entry. When the initial registry search fails to identify a matching registry entry, a second investigator will perform an additional registry search before the trial in question is labelled as unregistered. This registry search strategy has been previously used in prior registry-based studies.8 9 25 27 47

Data collection
Participating journals will supply the PRE-REPORT study team with a copy of the initial manuscript submitted for peer review. A single investigator will collect data from these initial manuscripts and from relevant registry entries for each trial, including the primary outcome listed in the initially submitted version of the manuscript as well as the registry used, registration date, study start date and registered study outcomes. We will follow manuscripts throughout the editorial process to determine the final publication decision. For accepted manuscripts, after publication of the finalised version of the manuscript we will abstract additional data from the final manuscript using a standardised data collection template. Data abstracted at this stage will include information about the sample size, description of the statistical plan and the
published primary and secondary outcome definitions. Any outcome(s) described by study authors within the abstract or manuscript as primary study outcomes will be considered primary outcomes. If the manuscript contains no outcome that is explicitly identified as the primary outcome but a sample size calculation was performed, the outcome used in this calculation will be considered the published primary outcome. If no outcome was explicitly identified as the primary outcome, and no sample size calculation was performed, the published primary outcome will be considered undefined. Data collection will continue until all manuscripts from participating trials accepted for publication have been published, after which blinded outcome assessors will determine the consistency of the prospectively registered and published outcomes.

Primary outcome

Our primary outcome is the presence of a clearly defined, prospectively registered primary trial outcome that is consistent with the primary outcome in the published manuscript, as determined by two independent outcome assessors. We define prospective registration as registration of a primary outcome with ClinicalTrials.gov or any of the Primary Registries in the WHO Registry Network (http://www.who.int/ictrp/network/primary/en/) prior to enrolment of the trial's first participant (or prior to 13 September, 2005, for trials beginning before 1 July, 2005, in keeping with ICMJE policy on trial registration). A clearly defined outcome provides sufficient information to reasonably allow its identification on review of the study results and to allow an independent investigator to design a study measuring the same parameter. In general, this requires that registration include both a specifically defined variable and a specifically defined period for assessment. In order for the outcome variable in question to meet this required level of specificity, in most cases it must specify a general domain, specific measurement and specific metric. For example, Zarin et al describe an example of an outcome measure at the following levels of specification: anxiety (domain), Hamilton Anxiety Rating Scale (specific measurement), change from baseline (specific metric).49 While registered outcomes will ideally also specify the method of aggregation for each outcome variable (eg, proportion of participants with decrease ≥50%), we do not require this level of specificity as this information is rarely included in prospectively defined registry entries, and it has been argued that the method of aggregation may be specified after data collection has been completed as part of the statistical analysis plan.48 A specifically defined period is not required if the nature of the study limits the outcome assessment to an obvious time frame.

We will characterise outcome inconsistencies according to the classification of outcome discrepancies developed by Chan et al50 and refined by Mathieu et al (box 1).51 Outcomes will be considered to be consistent if every primary outcome described in the registry is reported as a primary outcome in the manuscript, and every primary outcome reported in the manuscript is described as a primary outcome in the registry. When multiple primary outcomes are defined for a single trial, all primary outcomes in the registry must match the manuscript, and all primary outcomes in the manuscript must match the registry for the outcomes to be considered consistent. Two investigators will independently assess all registered and published outcomes for consistency. Both investigators will be blinded to whether the manuscript was in the control or intervention phase, to the content of the manuscript draft sent for initial peer review and the date on which the trial was submitted. Further, the ordering of pairs of registered and published outcomes will be randomised prior to sending to investigators to eliminate the potential for investigators to be influenced by the knowledge that manuscripts submitted later during the trial are more likely to have received the intervention. Inter-rater reliability will be assessed using a kappa value; our group has previously performed similar analyses of agreement between paired assessors evaluating outcome consistency between registry entries and published manuscripts with excellent inter-rater agreement (κ=0.87).27 Any discrepancies will be resolved by consensus after having both investigators review the full text of the manuscript and registry; persistent disagreements will be adjudicated by a third investigator. Trials not prospectively registered will be considered to have inconsistent outcomes, as these publications will introduce new outcomes by definition.

Secondary outcomes

We will record the final editorial decision for each included trial manuscript. Among trials with primary outcome inconsistencies present, we will assess whether the published manuscript included a disclosure of this change and an explanation of the reason for the change. Also, by comparing primary outcomes in the initial submitted manuscript to the primary outcome in the published version of the manuscript, we will be able to directly measure the impact of peer reviewer/editor feedback related to outcome consistency. We will also measure and report changes in acceptance rates for clinical trials over the course of our study period. Additionally, we will classify any observed primary outcome inconsistencies according to whether or not they impact the statistical significance (as defined in each included manuscript)
of the published outcome. In an additional secondary analysis we will assess included manuscripts to determine whether discrepancies are present between prospectively registered secondary outcomes and published secondary outcomes, and we will describe the nature of identified discrepancies. Finally, among trials with registered primary outcomes that were registered prospectively but unclarly, we will determine whether the registered outcomes are broadly consistent with the published outcomes.

Exploratory analyses
Results from two additional exploratory analyses will be presented in subsequent manuscripts following publication of the main study outcomes. The first exploratory analysis will assess the impact of the intervention on secondary outcomes within the cohort of included trials, by comparing registered and published secondary outcomes. These comparisons will be performed using the same methods described above for the comparison of primary outcomes. While the PRE-REPORT intervention involves providing peer reviewers with registered information related to primary trial outcomes only, we hypothesise that the intervention may increase reviewer attention to the issue of trial registration in general, and may result in closer scrutiny of registered secondary outcomes during peer review. Additionally, we plan to perform an exploratory analysis to assess for evidence of a Hawthorne effect causing a change in the baseline agreement between registered and published primary outcomes among manuscripts published in the participating journals. Because the editors-in-chief from the collaborating journals gave permission for participation of their journals and were not blinded to the study hypothesis, it is possible that journal behaviour may have changed due to participation in the study. Using the same methods described above to assess agreement between outcomes, we plan to compare primary registered and published outcomes for clinical trials published in the year before we initially contacted editors about study participation. Rates of outcome discrepancies observed during this pre-study period will be compared with rates observed during the course of the PRE-REPORT study.

Sample size and power
We used simulations to calculate power for comparing our primary outcome (outcome inconsistency) between intervention and control phases.\(^5\) We used Qaqish’s conditional linear family approach to generate 2000 simulated datasets with correlated binary outcomes corresponding to the stepped-wedge design described above.\(^5\) Based on data from a prior systematic review we assumed that 33% of published manuscripts would have inconsistent outcomes during the control phase,\(^6\) and based on 2017 data we assumed that the participating journals would accept for publication, on average, two trial manuscripts per month. We further assumed that responses from manuscripts from the same journal in the same phase would have an intra-cluster correlation of no more than 0.50 (ICC1), and that responses from manuscripts from the same journal but from different phases would have an intra-cluster correlation of at least 0.05 (ICC2). Generally, higher levels of ICC1 lead to decreased power whereas higher levels of ICC2 lead to increased power.\(^5\) Under these assumptions, enrolling eight journals would provide at least 80% power to detect an 80% reduction in outcome inconsistency using a one-sided test at the 0.05 significance level. Five additional participating journals were included, for a total of 13 participating journals, in order to accommodate the possibility of journals dropping out of the study, and to account for the possibility of lower rates of manuscript publication or a lower magnitude of impact of the intervention. We have elected to use a one-sided test because it is extremely unlikely that the intervention would increase the frequency of outcome inconsistencies.

Analytical plan
Consistent with an intention to treat analysis, all manuscripts will be analysed according to the study phase the relevant journal is in when the manuscript is submitted, regardless of whether the intervention is successfully distributed to reviewers. For our primary outcome of outcome inconsistency, we will use mixed effect logistic regression models to compare observations between intervention and control phases. Mixed models allow for different numbers of manuscripts per journal, and also account for correlated responses between manuscripts published within the same journal. The model will include fixed effects for study phase (control or intervention) and study month, and will include journal-specific random effects that allow for different levels of correlation depending on whether manuscripts are reviewed in the same month or in different months. A one-sided test at the 5% level will be conducted to compare the intervention and control phases. In addition, an OR will be estimated along with a 90% CI (to be consistent with the one-sided 5% level).

Patient and public involvement
Patients were not involved in the design of this study, and the study does not involve recruitment of individual patients. Following completion of the study we plan to present the findings and a draft version of the paper for publication to several BMJ patient and public reviewers to obtain their insight into the implications that the study might have for patient care.

Ethics and dissemination
Despite the decision of our local ethics board that this study does not meet criteria for human-subjects research, the investigators recognise that the study design necessitates the implementation of precautions to protect the confidentiality and intellectual property of authors whose work may be included in the study, along with peer reviewers and editors involved in the evaluation of
REFERENCES
34. Chauvin A, Flavaud P, Baro G, et al. The most important tasks for peer reviewers evaluating a randomized controlled trial are not congruent with the tasks most often requested by journal editors. *BMC Med* 2015;13:158.