BMJ Open

Peer Reviewed Evaluation of Registered End-Points of Randomised Trials (the PRE-REPORT study): a stepped wedge, cluster-randomised trial

Christopher W Jones ,[1] Amanda Adams,[2] Benjamin S Misemer,[3] Mark A Weaver,[4] Sara Schroter ,[5] Hayat Khan,[1] Benyamin Margolis,[6] David L Schriger,[7] Timothy F Platts-Mills[8]

To cite: Jones CW, Adams A, Misemer BS, *et al*. Peer Reviewed Evaluation of Registered End-Points of Randomised Trials (the PRE-REPORT study): a stepped wedge, cluster-randomised trial. *BMJ Open* 2022;**12**:e066624. doi:10.1136/bmjopen-2022-066624

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-066624).

Received 13 July 2022
Accepted 02 September 2022

For numbered affiliations see end of article.

Correspondence to
Dr Christopher W Jones; jones-christopher@cooperhealth.edu

ABSTRACT:

Objective To test whether providing relevant clinical trial registry information to peer reviewers evaluating trial manuscripts decreases discrepancies between registered and published trial outcomes.

Design Stepped wedge, cluster-randomised trial, with clusters comprised of eligible manuscripts submitted to each participating journal between 1 November 2018 and 31 October 2019.

Setting Thirteen medical journals.

Participants Manuscripts were eligible for inclusion if they were submitted to a participating journal during the study period, presented results from the primary analysis of a clinical trial, and were peer reviewed.

Interventions During the control phase, there were no changes to pre-existing peer review practices. After journals crossed over into the intervention phase, peer reviewers received a data sheet describing whether trials were registered, the initial registration and enrolment dates, and the registered primary outcome(s) when enrolment began.

Main outcome measure The presence of a clearly defined, prospectively registered primary outcome consistent with the primary outcome in the published trial manuscript, as determined by two independent outcome assessors.

Results We included 419 manuscripts (243 control and 176 intervention). Participating journals published 43% of control-phase manuscripts and 39% of intervention-phase manuscripts (model-estimated percentage difference between intervention and control trials = −10%, 95% CI −25% to 4%). Among the 173 accepted trials, published primary outcomes were consistent with clearly defined, prospectively registered primary outcomes in 40 of 105 (38%) control-phase trials and 27 of 68 (40%) intervention-phase trials. A linear mixed model did not show evidence of a statistically significant primary outcome effect from the intervention (estimated difference between intervention and control=−6% (90% CI −27% to 15%); one-sided p value=0.68).

Conclusions These results do not support use of the tested intervention as implemented here to increase agreement between prospectively registered and published trial outcomes. Other approaches are needed to improve the quality of outcome reporting of clinical trials.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Highly innovative study that uses a stepped wedge design to evaluate an intervention aimed at improving peer review.
⇒ The 13 participating high-impact journals are diverse with respect to medical specialty, journal size, location and peer review practices.
⇒ The intervention is simple and was deployed via pragmatic methods that could potentially be automated for scalability.
⇒ Pairing the tested intervention with an educational programme for editors and reviewers detailing specific recommendations for the assessment of registry information might have increased the intervention's effectiveness.

Trial registration number ISRCTN41225307.

INTRODUCTION

Clinical trial registries were developed more than 20 years ago, in large part to facilitate the unbiased reporting of results from clinical trials.[1 2] To further support this goal, in 2005, the International Committee of Medical Journal Editors (ICMJE) mandated the registration of all clinical trials before the start of enrolment as a condition of publication in ICMJE-member journals.[3] Shortly afterwards, the WHO also called for the prospective registration of all trials, and trial registration is now endorsed or required by many regulators, funding organisations and other stakeholders.[4–9] While trial registries have proven to be a valuable tool for identifying and quantifying some forms of reporting biases within the biomedical literature, evidence shows that biased reporting persists.[10–12] Specifically, discrepancies between prespecified trial outcomes and outcomes reported in published manuscripts are frequently

observed, indicating the presence of selective outcome reporting.[11–16]

In theory, publicly available trial registries should allow for the identification and correction of selective outcome reporting during peer review. In practice, however, the continued high prevalence of selective outcome reporting indicates that registries have not fulfilled this potential. Several possible barriers exist, which may limit the effective use of registries during peer review. These barriers include the absence of clear policies at some journals identifying specific individuals responsible for reviewing registry entries and lack of familiarity among some reviewers and editors with registration requirements and available registry resources, including audit trails showing changes made to registry entries over time.[17 18] Most reviewers are also volunteers with limited time available to devote to their reviews, and they may not feel that they have enough time to seek out registry information themselves. Additionally, because registry entries often include the names of trial investigators and sponsors, some reviewers may hesitate to access these sites in order to preserve blinded peer review.

This study was designed to test a solution to these barriers by pushing relevant trial registry information to peer reviewers assigned to evaluate clinical trial manuscripts. We hypothesised that providing manuscript reviewers with information about relevant registration requirements, the trial's registration status and the list of prospectively defined primary outcomes would improve clinical trial reporting by increasing the consistency between prospectively registered and published trial outcomes.

METHODS

Study design

This was a stepped wedge, cluster-randomised trial testing the impact of providing peer reviewers with registry information for the clinical trials they were reviewing on consistency between registered and published trial outcomes. The rationale and detailed description for the study methods have been published previously (online supplemental appendix).[19] Clusters consisted of all eligible clinical trial manuscripts sent for peer review during the study period by an individual participating journal. This study was prospectively registered at the ISRCTN Registry on 24 October 2018.

JOURNAL SELECTION

We emailed the editors-in-chief at journals across a broad range of medical specialties to assess eligibility, feasibility and interest in study participation. Journals were approached for possible participation if they published an average of at least 10 clinical trial manuscripts per year and had endorsed the ICMJE requirement for prospective trial registration as a condition for publication. In order to participate, each journal's editor-in-chief had to determine that there was an opportunity to improve on existing practices with respect to the use of registry information during peer review. To minimise the risk of a change in behaviour on behalf of submitting authors or peer reviewers due to participation in the study (ie, Hawthorne effect), we did not publicly disclose the identities of participating journals prior to study completion, and journals were asked not to provide reviewers with specific details about the purpose of the study. Additionally, in order to maintain the confidentiality of all relevant stakeholders and to encourage journal participation, we agreed not to publicly release outcome data identifying individual manuscripts or individual participating journals. Thirteen journals agreed to participate in the study (online supplemental table 1).

Manuscript eligibility

Manuscripts reporting results from a clinical trial were eligible for inclusion if they were submitted to a participating journal between 1 November 2018 and 31 October 2019 and were sent for external peer review. We defined clinical trials according to the definition used by the WHO and ICMJE: any research study that prospectively assigned human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes.[3 20] Manuscripts were not eligible if they described a study protocol without reporting trial results, or if they stated that the manuscript was not intended to report on the trial's primary outcome (ie, manuscripts describing only secondary analyses, secondary outcomes or reanalyses). Manuscripts were also ineligible if they had previously been peer reviewed by one of the other participating journals.

Screening procedures for eligible manuscripts were individualised for each participating journal in order to accommodate confidentiality requirements as well as existing editorial and peer review processes. At some participating journals, journal staff members screened submissions for manuscript eligibility before alerting Peer Reviewed Evaluation of Registered End-Points of Randomised Trials (PRE-REPORT) investigators to a potentially eligible submission; at others, the PRE-REPORT investigators directly screened all submitted manuscripts for eligibility via the journal's electronic manuscript management system.

Registry data abstraction

An investigator with expertise in the use of trial registries reviewed each included manuscript for a trial registration number or other evidence of trial registration. When manuscripts did not contain registration information, the investigator then performed a keyword and title search of ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform search portal and relevant national or regional registries to identify a matching registry entry. We assessed potential matches between registry entries and manuscripts by comparing the study title, interventions, sample sizes, enrolment dates, investigator names

and trial locations between the registry and manuscript. When this initial search failed to identify a match, a second investigator performed an additional registry search. If the second search also failed to identify a matching registry entry, the trial in question was labelled unregistered. After identifying the registry record for an included trial, we recorded the date of trial registration, enrolment dates and all registered outcomes as defined when enrolment began.

Randomisation

We used a stepped wedge, cluster-randomised design to control for potentially confounding factors at the journal level, including characteristics of submitted manuscripts and existing peer review processes.[21 22] All participating journals began the 12-month study in the control phase and then crossed over into the intervention phase between months 3 and 10 after study initiation according to a randomisation schedule created by the study statistician using computer-generated random numbers. Manuscript screening and data collection processes did not change between the control and intervention phases.

Interventions

For journals in the control phase, there was no change to the instructions or information provided to peer reviewers as part of each journal's usual peer review practices. For eligible manuscripts submitted for peer review during the intervention phase, a PRE-REPORT investigator completed a registry data form (Supplementary Methods) which described whether the submitted trial was registered, the timing of registration relative to study enrolment and a description of the registered primary outcome(s) at the time study enrolment began. In some cases, registries are amended after enrolment has commenced to change the listed primary outcome. When this occurred, the primary outcome listed in the registry when enrollment began was the outcome included in the data form. These registry data forms were then distributed to all peer reviewers who accepted an invitation to review the included manuscript. For each journal, the method of distribution of data forms was determined by the editor-in-chief based on existing peer review processes and consisted of either incorporating this form into the downloadable version of the manuscript available to reviewers (six journals) or directly emailing the registry data form to reviewers through the journal's online manuscript tracking system within 24 hours of accepting the invitation to review (seven journals). We were not able to determine whether individual reviewers actually accessed the registry data forms. The data forms also included a reminder of the ICMJE requirements for prospective clinical trial registration as a condition of publication, though they did not include recommended actions in order to promote case-by-case decision-making by reviewers and editors having expertise specific to the content of each manuscript.

Outcome assessment

We tracked each submitted manuscript throughout the editorial process until the journal editors reached a final publication decision. For accepted manuscripts, we recorded the published definitions of all primary and secondary trial outcomes. Outcomes were classified as primary if they were described as such by study authors within the published abstract or manuscript. If the manuscript did not contain an explicitly identified primary outcome but did include a sample size calculation, we classified the outcome used for this calculation as the published primary outcome. If no primary outcome was explicitly identified and no sample size calculation was performed, the published primary outcome was considered undefined.

Our primary outcome for each included trial was the presence of a clearly defined, prospectively registered primary outcome consistent with the primary outcome in that trial's published manuscript. Trials were considered prospectively registered if a primary outcome was recorded on 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). We classified an outcome as being clearly defined if it provided sufficient information to reasonably allow its identification on review of the study results and to allow an independent investigator to replicate the study. In most cases, this required the registry to include both a clearly defined time period for assessment and an outcome variable specifying a general domain (eg, 'pain'), specific measurement (eg, '11-point visual analogue scale') and specific metric (eg, 'change from baseline').[23]

Outcomes were considered consistent if every primary outcome described in the registry was reported as a primary outcome in the manuscript, and every primary outcome reported in the manuscript was described as a primary outcome in the registry. We characterised outcome inconsistencies using the following classification: registered primary outcome is reported as secondary in the publication; registered primary outcome not reported in the publication; published primary outcome was described as secondary in the registry; published primary outcome was not registered; timing of primary outcome assessment differs between the registry and publication.[15 24] By definition, trials which were not prospectively registered were considered to have introduced new primary outcomes in the publication.

Two investigators, who were blinded to whether manuscripts were in the control or intervention phase, independently evaluated all registered and published outcomes for clarity and consistency using a standardised assessment form. Before beginning the evaluation process, these outcome assessors participated in a series of training sessions focused on our standardised framework for performing outcome evaluations. Both outcome assessors could access relevant registry data

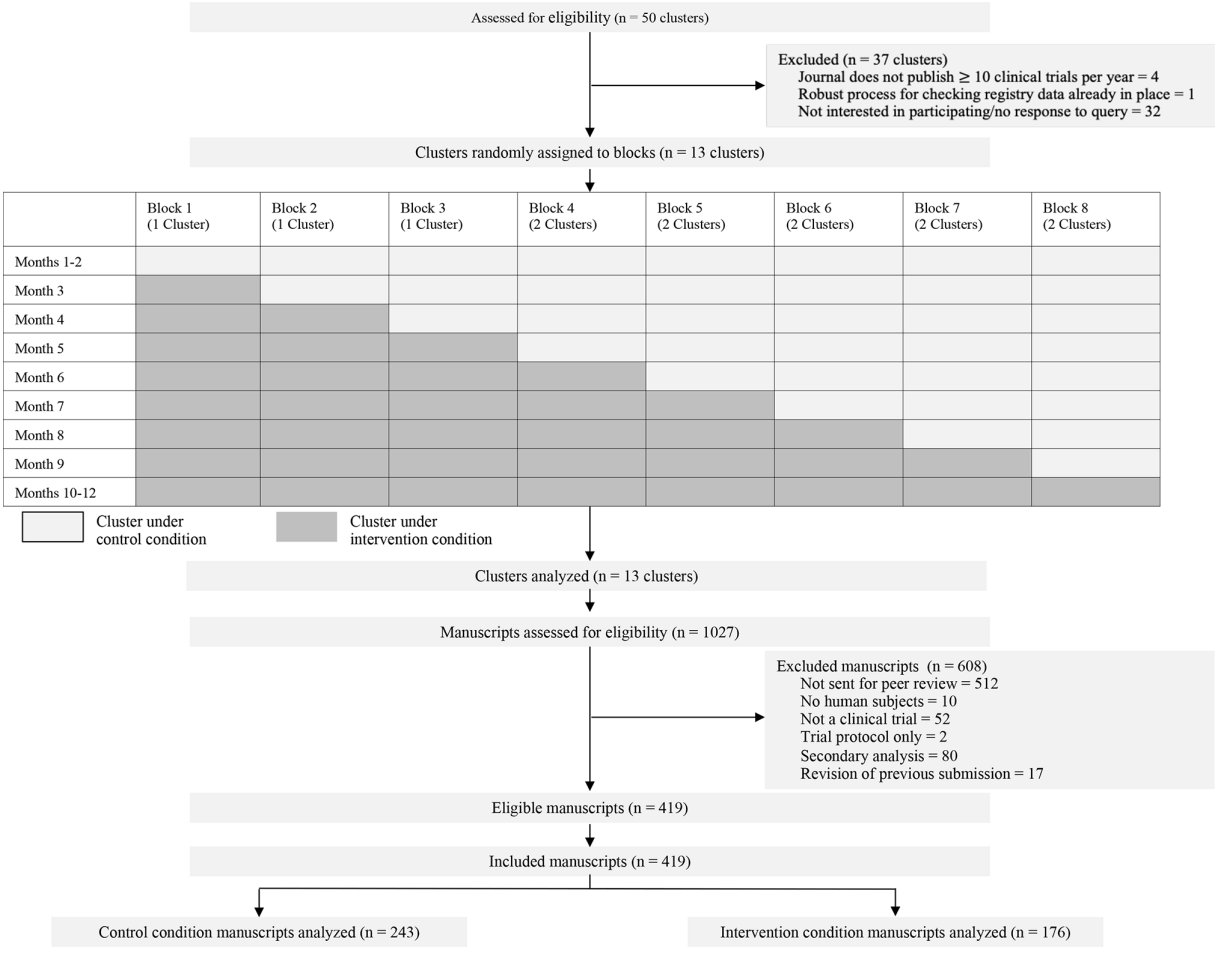

Figure 1 Flow diagram of eligible and included trials.

and a version of the published manuscript from which all submission and publication dates had been redacted. Dates were redacted in order to preserve blinding given the trial's stepped wedge format. After performing their initial independent evaluations, the assessors discussed all outcome discrepancies until a consensus was reached for each included trial. We measured inter-rater agreement with respect to both outcome clarity and consistency using Cohen's kappa.

Secondary outcomes of our study included the rates of acceptance for included manuscripts, the disclosure of a primary outcome change within the published version of the manuscript and time elapsed between initial manuscript submission and publication. As another secondary analysis involving those trials that were registered prospectively but without a clear primary outcome definition, we determined whether the vaguely defined registered outcome was broadly consistent with the published primary outcome. When possible, we also recorded the implications of primary outcome discrepancies for statistical significance of the primary outcome reported in the published manuscript. Discrepancies in secondary outcomes between prospectively completed registry entries and publications will be reported in a future manuscript.

Sample size

We used Qaqish's conditional linear family approach to generate 2000 simulated data sets with correlated binary outcomes corresponding to the trial's stepped wedge design in order to estimate power for the primary outcome.[25] Based on data from a prior systematic review, we expected that 33% of published manuscripts would have inconsistent outcomes during the control phase, and based on 2017 data, we assumed that the participating journals would accept for publication, on average, two trial manuscripts per month.[10] We further assumed that responses from manuscripts from the same journal in the same phase would have an intracluster correlation of no more than 0.50 (ICC1), and that responses from manuscripts from the same journal but from different phases would have an ICC of at least 0.05 (ICC2). Under these assumptions, enrolling eight journals would provide at least 80% power to detect a 26% absolute reduction (80% relative reduction) in outcome inconsistency using a one-sided test at the 0.05 significance level. We recruited an additional five participating journals, for a total of 13, in order to accommodate a lower magnitude impact of the intervention, lower rates of manuscript publication or the possibility of journals dropping out of the study.

ANALYSIS

All manuscripts were analysed based on the study phase of the relevant journal at the time of manuscript submission, in keeping with an intention-to-treat approach. For our primary outcome, we used a linear mixed model to compare observations between intervention and control phases. This approach was used rather than logistic regression to allow for the estimation of differences in proportions between intervention-condition and control-condition trials, as the cluster-randomised, stepped wedge design limits the utility of comparing raw differences in proportions. The model included fixed effects for study phase (control or intervention) and study month as well as 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. We prespecified the use of a one-sided test at the 5% level based on the assumption that the intervention would be very unlikely to increase the publication of manuscripts with primary outcome inconsistencies and estimated 90% CIs to be consistent with this one-sided 5% level. We used a similar linear mixed model for the secondary outcome of acceptance of submitted manuscripts; however, for this outcome, we report a two-sided test and 95% CI because the expected intervention effect was less clear. Because most of the remaining secondary outcomes were based on observations within a subset of the published trials, there were not sufficient data to support fitting a statistical model for these outcomes.

We also performed a post hoc exploratory analysis involving all included trials regardless of publication status in which we classified trials according to whether: (1) the manuscript was published without a primary outcome that matched a clearly defined, prospectively registered primary outcome or (2) the trial was either unpublished or was published with a primary outcome matching a clearly defined, prospectively registered primary outcome. The same linear mixed model as used for the primary outcome was also applied to this exploratory outcome.

Patient and public involvement

Patients or the public were not involved in the design, conduct, analysis or reporting of this research.

RESULTS

Thirteen journals initially agreed to participate in the trial, one of which dropped out after the enrolment of one control-condition manuscript, which was not accepted for publication, leaving 12 journals which contributed manuscripts that were eventually published. From the participating journals, we assessed 1027 submitted manuscripts for eligibility, and 419 manuscripts were included in the trial (figure 1). Of these included manuscripts, 243 were submitted to journals in the control phase of the study, and 176 were submitted to journals in the intervention phase.

Table 1 Characteristics of included trials

Trial characteristics	Control condition (n=243)	Intervention condition (n=176)
Number of participants; median (IQR)	78 (35–172)	75 (42–149)
Trial phase; n (%)		
Phase 0-II	44 (18%)	41 (23%)
Phase III	26 (11%)	19 (11%)
Phase IV	27 (11%)	25 (14%)
Other/undefined	146 (60%)	91 (52%)
Funding source; n (%)*		
Industry	66 (27%)	55 (31%)
Federal government	79 (33%)	46 (26%)
Foundation	54 (22%)	25 (14%)
Other/self-funded	74 (30%)	65 (37%)

*Total adds to more than 100% because trials could have multiple funding sources.

Analyses of all included trials

The control and intervention groups were similar with respect to study size and funding characteristics (table 1). Thirty-three trials (8%) were unregistered at the time of manuscript submission, 135 (32%) were registered retrospectively, and 251 (60%) were registered prospectively. Of the 419 included manuscripts, 173 (41%) were published by the participating journal and 246 (59%) were not.

Forty-three per cent (n=105) of manuscripts submitted to journals in the control phase were published, and 39% (n=68) of manuscripts submitted to journals during the intervention phase were published (linear mixed model-estimated percentage difference between intervention and control trials=−10%, 95% CI −25% to 4%) (online supplemental figure 1). Prospectively registered trials were more likely to be accepted for publication (117/251, 47%) than unregistered trials (7/33 trials, 21%; model-estimated difference=29% (95% CI 10% to 47%)). Among retrospectively registered trials, 49/135 (36%) were accepted for publication; the model-estimated difference between prospectively and retrospectively registered trials was 9% (95% CI −1% to 20%)) (table 2).

Analyses of published trials

We evaluated our primary outcome among each of the 173 published trials by assessing for the presence of a clearly defined, prospectively registered primary trial outcome that matched the primary outcome reported in the publication. Forty of 105 (38%) control-condition trials and 27 of 68 (40%) intervention-condition trials met this standard. The linear mixed model did not show evidence of a statistically significant effect from the intervention (model-estimated difference between the proportions of intervention-condition and control-condition

Table 2 Timing of registration relative to trial initiation

Registration timing	Total 419	Published* n=173	Not published* n=246
Registered before enrolment began; n (%)	251	117 (68%)	134 (54%)
Registered after enrolment began; n (%)	135	49 (28%)	86 (35%)
Registered up to 6 months after enrolment began; n (%)	66	31 (18%)	35 (14%)
Registered 6–12 months after enrolment began; n (%)	20	6 (3%)	14 (6%)
Registered over 12 months after enrolment began; n (%)	49	12 (7%)	37 (15%)
Unregistered; n (%)	33	7 (4%)	26 (11%)

*Refers to publication within the participating journal.

trials meeting the primary outcome =−6% (90% CI −27% to 15%); one-sided p value=0.68) (figure 2). Most trials in both the control and intervention arms which failed to meet criteria satisfying our primary outcome did so because they were not prospectively registered (table 3). Among the trials which were prospectively registered with a clearly defined primary outcome, 36 (21%) had discrepancies between the registered and published primary outcomes (online supplemental table 2). We did not observe a notable difference in the proportion of trials with outcome discrepancies between the control (n=21, 20%) and intervention (n=15, 22%) groups. Cohen's kappa revealed substantial interrater agreement with respect to both whether registered primary outcomes were clearly defined (k=0.88, p<0.001) and consistency between registered and published primary outcomes (k=0.74, p<0.001).

When primary outcome discrepancies were present, the discrepancy resulted in promotion of a new primary outcome in the publication that was statistically significant or demotion of a registered primary outcome that was not statistically significant in 11 cases (7/21 (33%) control; 4/15 (27%) intervention). Four outcome changes did not result in promotion of a statistically significant outcome or demotion of a non-significant outcome. In 21 cases, we were unable to determine the impact of the outcome discrepancy on the statistical significance of outcomes in the published manuscript. Among trials with discrepancies between a clearly registered primary outcome and a published primary outcome, the published manuscript disclosed or explained discrepancies between registered and published outcomes for one trial in the control group (5%) and three trials in the intervention group (20%).

The median time elapsed between submission and publication did not differ substantially between submissions made during the control phase (182 days, IQR 112–248) and the intervention phase (188 days, IQR 147–285).

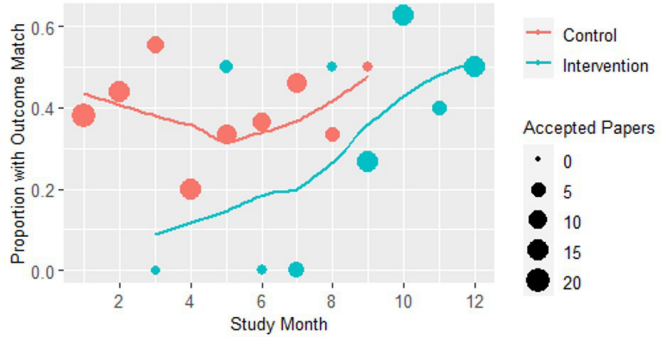

Figure 2 Proportion of accepted papers with clear, prospectively registered primary outcomes that match the published primary outcomes, by intervention phase and study month. Includes loess smoothing lines.

Sensitivity analysis

There were 14 published trials (10 control and 4 intervention), which were prospectively registered, but had registered primary outcomes that were not clearly defined. Three of these trials (two control and one intervention) had published outcomes that were consistent with the available outcome information in the registry, though in each of these cases, the registered trial record was ambiguous such that published outcomes could have been measured and reported in multiple ways while still broadly matching the primary outcome as defined in the registry. We performed a sensitivity analysis in which we considered these three trials to have met criteria for our primary outcome. In this analysis, 42 of 105 published trials in the control condition (40%) and 28 of 68 published trials in the intervention condition (41%) were prospectively registered with matches between all registered and published primary outcomes (model-estimated difference between proportions of intervention-condition and control-condition trials=−3% (90% CI −23% to 18%); one-sided p value=0.58).

Post hoc analysis

A post hoc exploratory analysis involving all of the included trials regardless of publication status did not show evidence that the tested intervention increased the proportion of trials that were either rejected by the journal or published with a primary outcome matching a clearly defined, prospectively registered primary outcome (control=178/243 (73%), intervention=135/176 (77%); model-estimated difference between intervention-condition and control-condition trials=−1% (95% CI −13% to 10%), one-sided p value=0.58).

Table 3 Registration and outcome data for trial manuscripts published in the index journal

Outcomes for included trials	Control trials (n=105)	Intervention trials (n=68)
Prospectively registered; n (%)	71 (68%)	46 (68%)
Prospectively registered with clear primary outcome; n (%)	61 (58%)	42 (62%)
Prospectively registered with clear primary outcome that matches published primary outcome; n (%)	40 (38%)	27 (40%)

DISCUSSION

This stepped wedge, cluster-randomised trial included manuscripts sent for peer review over a 1-year period at a diverse group of biomedical journals to test whether pushing information from trial registries to peer reviewers would decrease the incidence of inconsistencies between prospectively registered and published primary outcomes. Fewer than 40% of the manuscripts accepted for publication had published primary outcomes that matched clearly defined, prospectively registered primary outcomes. The tested intervention did not decrease the incidence of inconsistencies between registered and published outcomes.

In contrast to the critical role that peer review plays in the dissemination of findings from biomedical research, relatively few studies have been performed to test innovations aimed at improving current peer review processes.[26–34] Furthermore, the majority of trials which have tested interventions related to peer review are limited in their generalisability because they were performed at a single journal.[33 34] No prior randomised trials have examined interventions to ensure consistency between registered and published outcomes.[34] Our finding that fewer than half of the included trials had published primary outcomes matching prospectively registered primary outcomes emphasise the urgent need to identify interventions that are effective at increasing the quality of prospective trial registration, reducing post hoc outcome switching and improving trial reporting.[35 36] Although the current trial did not show evidence that the tested intervention was beneficial, it does demonstrate the feasibility of testing peer review interventions using rigorous methodology, and offers a potential framework for similar trials in the future. Specifically, applying a stepped wedge design with clusters defined by the participating journals allowed us to control for journal-specific confounders while balancing logistical considerations involved in integrating the intervention into existing editorial workflows at each journal.

Several lessons from the implementation of this study may increase the likelihood of ascertaining a benefit when testing future peer review-based interventions. First, in order to limit the risk that trial participation itself, rather than the intervention, might be responsible for changing reviewer behaviour (ie, risk of a Hawthorne effect), we deployed the intervention at participating journals without notifying peer reviewers about the study and without providing educational material to reviewers explaining the importance of trial registration and the scope or implications of outcome switching. In retrospect, these educational efforts could have taken place without disclosing the existence of the trial and would likely be incorporated into a roll-out of this type of intervention outside the context of a trial. Therefore, it would have been reasonable to include such efforts as part of the tested intervention. Additionally, among our participating journals, those that published the largest number of clinical trials were randomly assigned to cross into the intervention condition towards the end of the study period. Although it is unlikely that this impacted the study findings in this case, it did decrease the power of our trial to identify a beneficial treatment effect. Future trials randomising at the journal level should consider employing stratified randomisation based on the anticipated number of included manuscripts. Finally, we observed that 40% of the included trials which were published were either unregistered, registered retrospectively or registered with a vague primary outcome. In these cases, intervention at the time of peer review might prevent publication of such a trial in the participating journal, thereby aligning that specific journal with the guidance from the WHO and ICMJE requiring prospective registration as a condition of publication. Intervention during peer review, however, does not solve the more fundamental problem that unregistered or retrospectively registered trials, or trials with vaguely defined primary outcomes, are often conducted and regularly published. Instead, solving this problem will require interventions further upstream in the trial process, before trial enrolment is initiated in the first place. For example, we recommend that Institutional Review Boards confirm the timely and clear registration of clinical trials prior to granting study approval.[37]

This study has a number of strengths which should increase confidence in the validity of these findings. First, we included more than 400 clinical trials from a group of well-respected journals that represent a diverse range of clinical specialties. This group included journals published in both the USA (n=9) and Europe (n=4), and participating journals were also heterogeneous with respect to editorial and peer-review processes, number of trials published and impact factor, thereby increasing the external validity of these findings. The use of a cluster design, as opposed to randomisation of individual submitted manuscripts, helped minimise the risk that peer reviewers and decision editors assessing control-phase manuscripts would be contaminated by exposure to the intervention. Additionally, the stepped wedge

design allowed us to control for potentially confounding characteristics unique to each participating journal by facilitating within-cluster comparisons between outcomes in the preintervention and postintervention phases.[21 22]

This study also has several limitations, which should be considered when interpreting our findings. First, while the included journals were diverse in many important respects, it is our impression that the editors of each journal primarily or in part chose to participate because they had a particular interest in improving both the quality of trial reporting and current peer review processes. While it is likely that most journal editors share these characteristics, the intervention may perform differently among other journals. Second, the stepped wedge design allows for partial control of secular trends affecting peer review processes over the duration of the trial, but we are unable to entirely exclude the possibility that such trends influenced the study results. Third, the determination of whether registered outcomes were clearly defined and comparisons between registered and published outcomes both necessarily rely on subjective judgement. We addressed this concern by having two outcome assessors independently perform each comparison using standardised data forms and by blinding these assessors to the intervention status of the included trials. Both assessors were physicians with expertise in the use of registry data and with previous experience performing similar comparisons between registered and published outcomes. Fourth, the intervention was delivered to reviewers at some of the participating journals by email after they had already agreed to review; we were not able to confirm that these emails were received and considered when the reviews were written. As a result, we are unable to determine which step(s) within the peer review process allowed outcome discrepancies to persist despite our intervention. Future studies should consider incorporating the use of qualitative methods to better understand how reviewers and editors might use similar interventions most effectively.

In conclusion, this stepped wedge trial found that the tested method of distributing an information sheet containing registry data to peer reviewers was ineffective at increasing the proportion of published clinical trial manuscripts reporting primary outcomes that matched prospectively registered primary outcomes. Alternative methods of implementing the intervention may have increased its effectiveness, though this requires further study. The continued high prevalence of unregistered or retrospectively registered trials and of discrepancies between registered and published trial outcomes necessitate the identification of effective interventions for these problems.

Author affiliations
[1]Emergency Medicine, Cooper Medical School of Rowan University, Camden, New Jersey, USA
[2]Medical Library, Cooper Medical School of Rowan University, Camden, New Jersey, USA
[3]Emergency Medicine, University of Michigan, Ann Arbor, Michigan, USA
[4]Mathematics and Statistics, Elon University, Elon, North Carolina, USA
[5]BMJ, London, UK
[6]Rollins School of Public Health, Emory University, Atlanta, Georgia, USA
[7]Emergency Medicine, UCLA, Los Angeles, California, USA
[8]Ophirex, Inc, Corte Madera, California, USA

Contributors Concept and design: CJ, AA, SS, MAW, DLS, BSM, BM, TFP-M. Acquisition, analysis or interpretation of data: all authors. Drafting of the manuscript: CJ, MAW, TFP-M. Critical revision of the manuscript for important intellectual content: all authors. Statistical analysis: MAW. Obtained funding: CJ, TFP-M. Administrative, technical or material support: CJ, AA, SS, BSM, BM, TFP-M. Supervision: CJ, TFP-M. CJ is guarantor.

Funding This work was supported by the US Department of Health and Human Services Office of Research Integrity, grant number ORIIR180039. The sponsor has no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The contents of this manuscript are those of the authors and do not represent the official views of, nor an endorsement, by OASH, HHS or the U.S. Government. All authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Competing interests All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/disclosure-of-interest/ and declare: no support from any commercial organisation for the submitted work; CJ has received research grants from AstraZeneca, Vapotherm, Abbott, and Ophirex outside the submitted work. SS is a full time employee at BMJ but is not involved in editorial decision-making on manuscripts. BM is an employee of the Department of Health and Human Services in the Office of Research Integrity. DLS is an associate editor at JAMA and a deputy editor at Annals of Emergency Medicine. TFP-M is an employee of Ophirex; no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

ORCID iDs
Christopher W Jones http://orcid.org/0000-0001-9704-9094
Sara Schroter http://orcid.org/0000-0002-8791-8564

REFERENCES

1 Dickersin K, Rennie D. Registering clinical trials. *JAMA* 2003;290:516–23.
2 Tonks A. Registering clinical trials. *BMJ* 1999;319:1565–8.
3 DeAngelis CD, Drazen JM, Frizelle FA, *et al*. Clinical trial registration: a statement from the International Committee of medical Journal editors. *JAMA* 2004;292:1363–4.
4 Assembly WH, World Health Organization. Ministerial Summit on health research: report by the Secretariat, 2005. Available: https://apps.who.int/iris/handle/10665/20304

5 Regulation (EU) NO 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing directive 2001/20/ED. European Union, 2014. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0536
6 WHO Statement on Public Disclosure of Clinical Trial Results. World Health organization, 2015. Available: https://www.who.int/news/item/09-04-2015-japan-primary-registries-network
7 Food and Drug Administration Act of 2007. *USC*. United States, 2007.
8 The Registration of Clinical Trials. World association of medical Journal editors, 2005. Available: https://wame.org/registration-clinical-trials
9 Zarin DA, Tse T, Williams RJ, *et al*. Trial Reporting in ClinicalTrials.gov - The Final Rule. *N Engl J Med* 2016;375:1998–2004.
10 Jones CW, Keil LG, Holland WC, *et al*. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Med* 2015;13:282.
11 Gaudino M, Hameed I, Rahouma M, *et al*. Characteristics of contemporary randomized clinical trials and their association with the trial funding source in invasive cardiovascular interventions. *JAMA Intern Med* 2020;180:993–1001.
12 Robinson NB, Fremes S, Hameed I, *et al*. Characteristics of randomized clinical trials in surgery from 2008 to 2020: a systematic review. *JAMA Netw Open* 2021;4:e2114494.
13 Bourgeois FT, Murthy S, Mandl KD. Outcome reporting among drug trials registered in ClinicalTrials.gov. *Ann Intern Med* 2010;153:158–66.
14 Killeen S, Sourallous P, Hunter IA, *et al*. Registration rates, adequacy of registration, and a comparison of registered and published primary outcomes in randomized controlled trials published in surgery journals. *Ann Surg* 2014;259:193–6.
15 Mathieu S, Boutron I, Moher D, *et al*. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;302:977–84.
16 Hannink G, Gooszen HG, Rovers MM. Comparison of registered and published primary outcomes in randomized clinical trials of surgical interventions. *Ann Surg* 2013;257:818–23.
17 Chauvin A, Ravaud P, Baron 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.
18 Wager E, Williams P, Project Overcome failure to Publish nEgative fiNdings Consortium. "Hardly worth the effort"? Medical journals' policies and their editors' and publishers' views on trial registration and publication bias: quantitative and qualitative study. *BMJ* 2013;347:f5248.
19 Jones CW, Adams A, Weaver MA, *et al*. Peer reviewed evaluation of registered end-points of randomised trials (the PRE-REPORT study): protocol for a stepped-wedge, cluster-randomised trial. *BMJ Open* 2019;9:e028694.
20 Clinical Trials Q&A. World Health organization, 2020. Available: https://www.who.int/news-room/q-a-detail/clinical-trials
21 Ellenberg SS. The Stepped-Wedge clinical trial: evaluation by rolling deployment. *JAMA* 2018;319:607–8.
22 Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007;28:182–91.
23 Zarin DA, Tse T, Williams RJ, *et al*. The ClinicalTrials.gov results database--update and key issues. *N Engl J Med* 2011;364:852–60.
24 Chan A-W, Hróbjartsson A, Haahr MT, *et al*. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457–65.
25 Baio G, Copas A, Ambler G, *et al*. Sample size calculation for a stepped wedge trial. *Trials* 2015;16:354.
26 Rennie D. Let's make peer review scientific. *Nature* 2016;535:31–3.
27 Justice AC, Cho MK, Winker MA, *et al*. Does masking author identity improve peer review quality? A randomized controlled trial. peer Investigators. *JAMA* 1998;280:240–2.
28 van Rooyen S, Delamothe T, Evans SJW. Effect on peer review of telling reviewers that their signed reviews might be posted on the web: randomised controlled trial. *BMJ* 2010;341:c5729.
29 van Rooyen S, Godlee F, Evans S, *et al*. Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. *BMJ* 1999;318:23–7.
30 Godlee F, Gale CR, Martyn CN. Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *JAMA* 1998;280:237–40.
31 Schroter S, Black N, Evans S, *et al*. Effects of training on quality of peer review: randomised controlled trial. *BMJ* 2004;328:673.
32 van Rooyen S, Godlee F, Evans S, *et al*. Effect of blinding and unmasking on the quality of peer review: a randomized trial. *JAMA* 1998;280:234–7.
33 Jefferson T, Rudin M, Brodney Folse S, *et al*. Editorial peer review for improving the quality of reports of biomedical studies. *Cochrane Database Syst Rev* 2007;2:MR000016.
34 Bruce R, Chauvin A, Trinquart L, *et al*. Impact of interventions to improve the quality of peer review of biomedical journals: a systematic review and meta-analysis. *BMC Med* 2016;14:85.
35 Blanco D, Schroter S, Aldcroft A, *et al*. Effect of an editorial intervention to improve the completeness of reporting of randomised trials: a randomised controlled trial. *BMJ Open* 2020;10:e036799.
36 Speich B, Schroter S, Briel M, *et al*. Impact of a short version of the CONSORT checklist for peer reviewers to improve the reporting of randomised controlled trials published in biomedical journals: study protocol for a randomised controlled trial. *BMJ Open* 2020;10:e035114.
37 Levin LA, Palmer JG. Institutional review boards should require clinical trial registration. *Arch Intern Med* 2007;167:1576–80.