BMC Medicine

RESEARCH ARTICLE Open Access

Harm effects in non-registered versus registered randomized controlled trials of medications: a retrospective cohort study of clinical trials

Chang Xu[1*], Shiqi Fan[1], Luis Furuya-Kanamori[2], Sheyu Li[3], Lifeng Lin[4], Haitao Chu[5,9], Su Golder[6], Yoon Loke[7] and Sunita Vohra[8]

Abstract

Background Trial registration aims to address potential bias from selective or non-reporting of findings, and therefore has a vital role in promoting transparency and accountability of clinical research. In this study, we aim to investigate the influence of trial registration on estimated harm effects in randomized controlled trials of medication interventions.

Methods We searched PubMed for systematic reviews and meta-analyses of randomized trials on medication harms indexed between January 1, 2015, and January 1, 2020. To be included in the analyses, eligible meta-analyses should have at least five randomized trials with distinct registration statuses (i.e., prospectively registered, retrospectively registered, and non-registered) and 2 by 2 table data for adverse events for each trial. To control for potential confounding, trials in each meta-analysis were analyzed within confounder-harmonized groups (e.g., dosage) identified using the Directed Acyclic Graph method. The harm estimates arising from the trials with different registration statuses were compared within the confounder-harmonized groups using hierarchical linear regression. Results are shown as ratio of odds ratio (OR) and 95% confidence interval (CI).

Results The dataset consists of 629 meta-analyses of harms with 10,069 trials. Of these trials, 74.3% were registered, and 23.9% were not registered, and for those registered, 70.6% were prospectively registered, while 26.3% were retrospectively registered. In comparison to prospectively registered trials, both non-registered trials (ratio of OR = 0.82, 95%CI 0.68 to 0.98, $P = 0.03$) and retrospectively registered trials (ratio of OR = 0.75, 95%CI 0.66 to 0.86, $P < 0.01$) had lower OR for harms based on 69 and 126 confounders-harmonized groups. The OR of harms did not differ between retrospectively registered and non-registered trials (ratio of OR = 1.02, 95%CI 0.85 to 1.23, $P = 0.83$) based on 76 confounders-harmonized groups.

Conclusions Medication-related harms may be understated in non-registered trials, and there was no obvious evidence that retrospective registration had a demonstrable benefit in reducing such selective or absent reporting. Prospective registration is highly recommended for future trials.

Keywords Randomized controlled trials, Registration, Harm effects, Selective reporting

*Correspondence:
Chang Xu
xuchang2016@runbox.com
Full list of author information is available at the end of the article

Background

Randomized controlled trials are the main source of high-quality primary evidence for causal relationships between medical interventions and health outcomes [1]. Findings from randomized controlled trials have been widely used for intervention efficacy assessment and clinical guidelines development [2]. Randomized allocation provides a strong safeguard against selection bias and reduces the risk of confounding due to baseline differences between participants in the different arms [3]. However, randomization alone does not protect against bias from subsequent selective outcome reporting [4].

Trial registration promotes transparency and accountability of clinical research, primarily by making any selective or non-reporting potentially visible to readers [5]. Access to the summarized details of a trial (e.g., through a trial registry) allows public scrutiny of the trial design, analysis plan, and reporting of the results [6]. The first trial registry, ClinicalTrials.gov, was established and available to the public in 2000 by the US Food and Drug Administration Modernization Act of 1997 [7]. In 2004, The International Committee of Medical Journal Editors (ICMJE) announced a requirement for authors of its member journals to register their trials in a public registry [8]. The ongoing legal and regulatory process has led to trial registration gradually becoming a mandatory requirement for manuscript publication in many academic journals [9]. The registration of published randomized trials increased from 25% in 2005 to 52% in 2015 [10], and to 85–90% by 2020 [11].

Despite these advances, numerous trials remain unregistered, especially those that started in an earlier time period. Also, a large proportion of trials were registered retrospectively after the study had been started, or in some cases, after completion of the trial (which may therefore negate any of the enhanced transparency from prospective registration) [12]. There appears to be sizeable evidence suggesting that non-registered and retrospectively registered trials tend to report larger treatment effects than those registered prospectively due to selective reporting [13–15]. As a result, exaggerated treatment effects may occur when the findings of non-registered and retrospectively registered trials are incorporated into evidence synthesis.

In contrast to efficacy, harms are featured less prominently in both randomized controlled trials and meta-analyses of randomized trials [16]. However, selective or non-reporting of adverse outcomes could potentially be even more problematic because of the very wide range of spontaneously recorded harms in a trial. We hypothesized that non-registered and retrospectively registered trials may potentially be markedly affected by selective or non-reporting of adverse events, thus potentially further biasing estimates of harm. Since evidence of harms is important in guiding informed healthcare decisions [17], elucidating selective or non-reporting of harmful outcomes of registered and retrospectively registered trials is critical for evidence-based medicine and policy formulation.

In this study, based on a retrospective cohort design, we compared medication-related harms reported among trials that differed in their registration status. We aimed to answer two specific questions: (1) whether there are differences in reporting of harm effects for non-registered trials versus prospectively registered trials; (2) whether retrospectively registered trials were similar to or different from non-registered trials.

Methods

This study is part of a larger research program that aims to investigate potential confounders that affect estimates of harms reported in randomized controlled trials. The protocol for the program has been previously reported [18]. No major amendments were made; minor modifications are recorded in the Additional file 1. The reporting of the current study is in accordance with relevant reporting guidelines (e.g., PRIOR statement) [19, 20], whenever possible.

Data source

The current study was based on the recently established dataset [21]. In brief, systematic reviews and meta-analyses of adverse events that were indexed in PubMed between Jan 1, 2015, and Jan 1, 2020, were searched by an information specialist on Jul 28, 2020 (see Additional file 1). We included systematic reviews and meta-analyses of randomized controlled trials on medication interventions of any topic that focused exclusively on adverse events as their outcomes of interest. Systematic reviews in languages other than English or Chinese were excluded for pragmatic reasons. For inclusion in the final analysis, the meta-analyses must have involved at least five randomized controlled trials and compared effects between two intervention arms (i.e., pairwise meta-analysis), with 2 by 2 table adverse events data available for each trial.

We considered articles to be systematic reviews or meta-analyses on the basis of the article titles as stipulated by the original authors. We defined adverse events as 'any untoward medical occurrence in a patient or subject in clinical practice'. The representativeness of the search was verified through an additional search of PubMed of 292 systematic reviews or meta-analyses that were identified in multiple databases (e.g., PubMed, Embase, CENTRAL), with an estimated coverage ranging from 93.9 to 99.3% [18, 22].

The literature screen was conducted in duplicate by a program assistant (XQ) and the principal investigator (CX). Titles and abstracts were screened first and excluded based on full consensus; full texts were subsequently reviewed for the remaining articles. During this full-text screening process, any conflicts were resolved through discussion. The Rayyan web application (https://www.rayyan.ai/) was used for the literature screen in order to facilitate the sharing of papers for screening, with independent scoring and application of eligibility criteria being conducted by each screener.

Data collection

The extracted information fell into two separate workstreams: (1) the first category was items that could be directly extracted via the meta-analyses or trials without any further judgments, and (2) the second was items that needed additional assessment and judgment by the data extraction team. The first category included items such as author names, publication year, review outcomes, trial name of each trial, the 2 by 2 table data for each trial, analytic rules (e.g., intervention-to-treat), type of intervention and control, median treatment duration of each arm, dosage of the intervention, age of trial population (children, adult), source of funding, number of trial centers, and registration status of the trial. To minimize potential data extraction errors, all information was collected by y at least two reviewers from the team (CX, TQ, FZ, XY, RZ, YT, XX, YZ, XZ, LFK, YY, HD), with a further independent double- or triple-checking process. We recognized that our previous research has empirically demonstrated the presence of data extraction errors in up to 60% of the 2 by 2 table data of trials presented in the forest plot or table of the meta-analyses [23, 24]. Hence, we checked the accuracy of these 2 by 2 table of adverse events data by referring to the original data source (i.e., full texts, appendix, registration of the trial) and made necessary correction after the duplicate extraction [18]. The Additional file 1 has a record of these details.

Some published manuscripts may not have given any details regarding their trial registration status [12, 25]. In these situations, we searched (using the name of the trial's principal author, intervention, and control group) four registries to determine the registration status: ClinicalTrials.gov, WHO's International Clinical Trials Registry Platform, European Clinical Trials Register, and the International Standard Randomized Controlled Trial Number (ISRCTN) registry. The sample size, arms of interventions, and listed publications in the registries were used to match the trials. We did not contact the corresponding authors of the trials for the registration status since we expected very few replies. Thus, it is possible that we may have missed a few potentially registered trials, thus resulting in misclassification bias. This bias would lead to an underestimation of any difference in the harm estimates between non-registered and registered trials. Registered trials were coded as prospectively or retrospectively registered, based on the timing of registration versus the start date in the registries. A trial registered before or within 1 month of the trial start date was regarded as prospectively registered [13].

The second category of extracted information relates to the risk of bias for each trial and the type of outcomes evaluated by each meta-analysis. Since the risk of bias information for the included trials is not described by all systematic reviews [26], we re-assessed the risk of bias using an adapted RoB 2 with an emphasis on applicable components and domains based on the original reports. The assessment included (1) random sequence generation, (2) allocation concealment, (3) blinding of participants, (4) blinding of healthcare providers, and (5) blinding of outcome assessors [27]. We made a judgment on the type of adverse events by dichotomizing the outcomes as either subjective or objective. To minimize potential human errors when collecting the second type of information, a duplicate assessment strategy was used; the RoB was assessed by two groups of investigators independently (group 1: FY, TQ, YY; group 2: XY, ZR), whereas the type of outcomes was judged by two senior methodologists (CX and LFK) independently. Conflicting decisions were discussed online until a consensus was reached.

Outcomes

The primary outcomes were (1) the relative difference of the odds ratio (OR) of the harm effects for non-registered trials to prospectively registered trials; (2) the relative difference of the OR of the harm effects for retrospectively and prospectively registered trials. The secondary outcome was the relative difference of the OR of the harm effects between retrospectively registered trials and non-registered trials.

Confounders and adjusting

To identify potential variables that may bias our comparison of the relative differences between registered and unregistered trials, we applied a causal path analysis using directed acyclic graphs (http://dagitty.net/) [28]. This method connects potential variables with directed arrows to posit the relationship between variables, with the presence of an arrow denoting a causal relationship between two variables and vice versa. The components of the directed acyclic graph help researchers to recognize confounders that may bias the analysis from the exposure-endpoints pathway (see Additional file 1: Fig. S1 for further details). Based on previous studies [29, 30],

the following confounders were considered: intervention, control, treatment duration, dosage, allocation concealment, blinding status for participants, blinding status for care providers, blinding status for outcome assessors, population age, disease, number of centers, analytic rules, experience of trialists, country policies of registration, source of funding, and role of funding agencies.

After establishing the directed acyclic graph, one of the following four paths was used for the adjustment (Additional file 1: Fig. S1): (1) population age, blinding status for assessors, blinding status for participants, blinding status for care providers, allocation concealment, dosage, treatment duration, analytic rule, intervention, and control; (2) population age, blinding status for assessors, blinding status for participants, blinding status for care providers, allocation concealment, disease, treatment duration, analytic rule, intervention, and control; (3) blind assessors, blinding participants/care providers, concealment, funding; and (4) funding, and experience of trialists. As there was little information on the disease (e.g., severity), funding (e.g., role of funding), and experience of trialists of each trial, we used the first adjustment set, which contained ten confounders.

In order to control for confounding, we removed trials where the intervention and control arms had different treatment durations. The remaining trials in each meta-analysis were then grouped for harmonization based on the remaining nine confounders. Before the harmonization, we re-coded concealment and blinding status as two categories by combining 'Yes' and 'Probably Yes' into one and 'No', 'Probably No', 'No information' into one, as an effort to avoid sample loss. We further checked the registration status of each group, and only those groups of the trials with different statuses of registration would be further identified for eligibility. Each group thus had profiles matched on the key confounders and formed the basic unit for subsequent comparison between trials with distinct registration status (i.e., prospectively registered, retrospectively registered, and non-registered). For convenience, we refer to these matched groups as 'confounders-harmonized groups'. Further details are given in Additional file 1: Fig. S2.

We also considered the alternative method to adjust for these confounders by them as explanatory variables in a multivariable regression model. However, this proved not to be feasible in our case because some of the confounders involved hundreds of categories (i.e., intervention, control, and dosages) that would markedly influence the validity and robustness of the regression model [31].

Data analysis

Baseline characteristics were summarized descriptively, namely, by proportions, median value, and interquartile range (IQR). For the main analysis, we used the ratio of OR as effect estimates to measure the relative difference of the harm effects [32]. Since the current study focused on adverse outcomes, zero events may occur [33]. The continuity correction by adding 0.5 to each cell was used to approximate the OR if zero events occurred in single or both arms [34]. A hierarchical linear regression model was used to estimate the average ratio of OR across the confounders-harmonized groups, with trials as level 1 and groups as level 2. This is because some trials might be included in multiple meta-analyses, and some meta-analyses involved two or more groups. For the same reason, a robust error estimate was used by treating meta-analysis as a cluster [35].

Pre-planned subgroup analyses were conducted based on characteristics of the trials: type of outcomes (objective vs. subjective), source of funding (academic vs. industry), year of publication (dichotomized into 2001 to 2010 vs. 2011 to 2020), geographical region (Europe & North America, multiple-countries, others), and sample size (less than 500 vs. 500 and more). For the subgroup analysis on the year of publication, the trials published in 2000 or before were not included since ClinicalTrials.gov was not available to the public until 2000 [7]. The 2010 cutoff for stratification was based on the time of the release of the CONSORT 2010 statement [36]. The sample size stratification was made on the assumption that trials with a total sample size of 500 were unlikely to report zero events.

Nineteen variables in the dataset were affected by missing data, with proportions missing ranging from 3.1 to 27.5%. Missing data stemmed primarily from selective or incomplete reporting of the trials and our inability to access a small proportion of the full-text articles despite numerous efforts (Additional file 1: Table S3). For the variables used in the current study, the missing proportions ranged from 3.1 to 14.3%, with two exceeding 10%, namely, treatment duration in intervention and control arms. As there is currently no validated method for addressing this type of missing information, we removed these trials from the final analyses [37]. All data analyses were run via Stata/SE 16.0 (Stata Corp LCC, College Station, TX), with a two-sided alpha = 0.05 as the significance level.

Patient and public involvement

This study investigated the methodological issue of randomized trials. No patients or public were involved in setting the research question or the outcome measures, nor were they involved in design, implementation, interpretation or writing up of results.

Results

The initial search resulted in 18,636 records; 1967 were excluded for duplicate reporting, and 15,399 were excluded by screening the titles and abstracts. Full-text screening of the remaining 1330 records yielded 151 systematic reviews and 629 meta-analyses of 10,069 trials (Fig. 1). Here, the number of trials is reported on a repeated-counting (or non-exclusive) basis—a trial that is present in multiple meta-analyses or a trial that presents results of subgroups of different doses within a meta-analysis has been counted more than once. The list of included reviews and reasons for exclusion are listed in the Additional file 1. The characteristics of the included reviews and trials are shown in Table 1.

Among the 10,069 trials, 74.3% were registered, and 23.9% were not registered, and for those registered, 70.6% were prospectively registered, while 26.3% were retrospectively registered. The proportion of non-registration and retrospective registration has steadily decreased over time, see Fig. 2. Full-text reports in either English or Chinese were available for 9598 (97.5%) trials. Exclusion of trials with ambiguous data on outcome events [18] and those without treatment duration, 7541 (74.9%) trials remained. After eliminating trials with different treatment durations between intervention and control arms and grouping, the construction of confounder-harmonized groups enabled us to conduct the following analyses: 65 groups with 212 trials for the comparison of no registration vs. prospective registration, 126 groups with

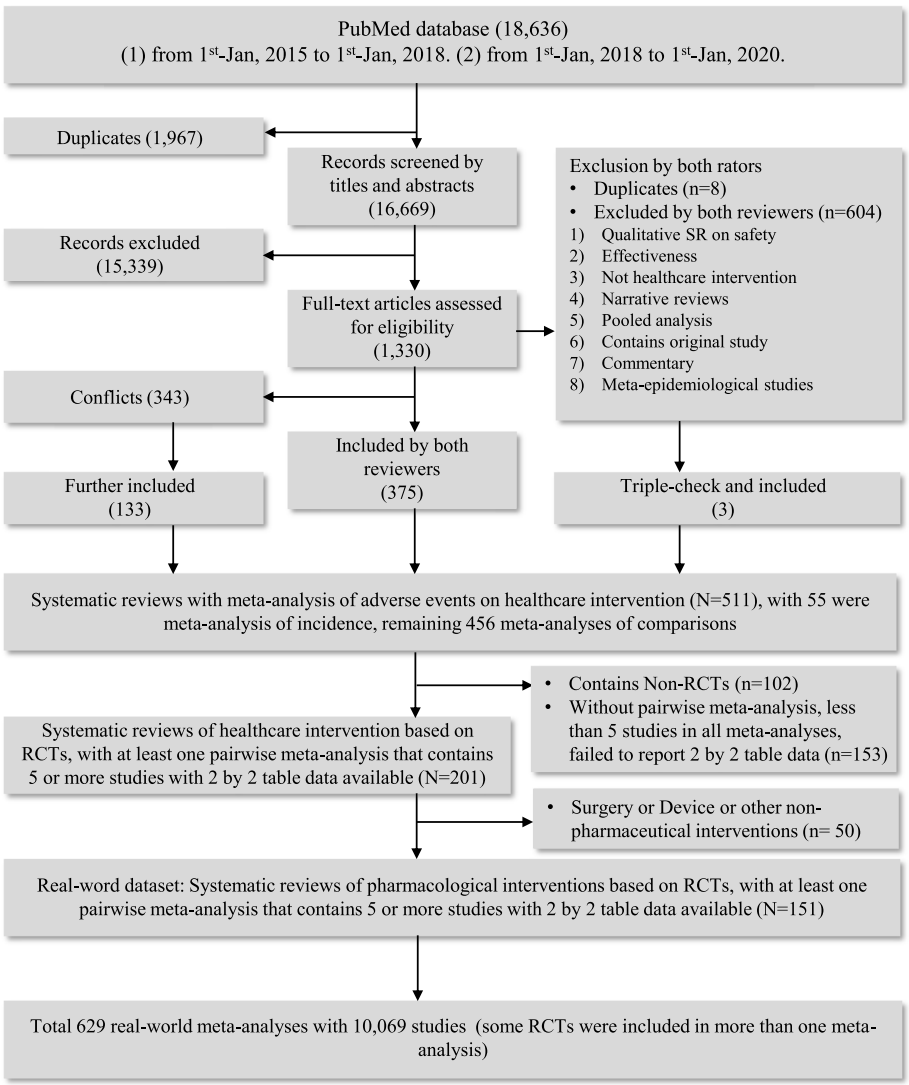

Fig. 1 The flow diagram of literature screening

Table 1 Characteristics of eligible systematic reviews and trials

Characteristics	Summary
Region of corresponding author (review level)	*N* = 151
Africa	9 (6.0%)
Americas (North and South)	32 (21.2%)
Asia	68 (45.0%)
Europe	40 (26.5%)
Oceania	2 (1.3%)
Topic of disease (review level)	
Cancer	68 (45.0%)
Diabetes and cardiovascular diseases	19 (12.6%)
Osteoarticular diseases	16 (10.6%)
Others	48 (31.8%)
Registration (study level)	*N* = 10,069
Yes	7483 (74.3%)
No	2408 (23.9%)
Not reported	178 (1.8%)
Registration forms (study level)	*N* = 7483
Prospective	5283 (70.6%)
Retrospective	1965 (26.3%)
Both (multiple trial registrations)	51 (0.7%)
Not applicable (unpublished trials)	184 (2.5%)
Center (study level)	*N* = 10,069
Multiple centers	7891 (78.4%)
Single center	350 (3.5%)
Not reported	1828 (18.2%)
Funding(study level)	
Industry	8440 (83.8%)
Industry & academic	108 (1.1%)
Academic	737 (7.3%)
No funding	3 (0.0%)
Not reported	781 (7.8%)
Region (study level)	
Africa	15 (0.2%)
Asia	850 (9.9%)
Australia	55 (0.6%)
Europe	759 (8.8%)
North America	1291 (15.0%)
South America	10 (0.1%)
Multi-regional	5617 (65.3%)
Not reported	1472 (14.6%)
Publication type of study (study level)	
Article	9848 (97.81%)
Abstract	18 (0.18%)
Registration only (unpublished)	198 (1.97%)
Non-randomized controlled trial (removed)	5 (0.05%)
Accessible of full texts (study level)	*N* = 9848
Yes	9598 (97.46%)
No	250 (2.54%)

342 trials for the comparison of retrospective registration vs. prospective registration, and 76 groups with 249 trials for the comparison of retrospective registration vs. no registration. The median number of trials within each group was 2 (IQR 2 to 4; range 2 to 17), 2 (IQR 2 to 3; range 2 to 17), and 3 (IQR 2 to 3; range 2 to 17) for each of the above comparisons.

Harm effects in non-registered trials versus prospectively registered trials

In the hierarchical linear regression analysis of the 65 confounder-harmonized groups of 212 trials that addressed all the ten confounders, the OR of harms was significantly lower in unregistered trials versus prospectively registered trials (ratio of OR = 0.82, 95%CI 0.68 to 0.98, $P = 0.03$).

In subgroup analyses, the underestimation of harms remains while not statistically significant when stratified by type of outcomes, year of publication, and source of funding, and interaction tests within the subgroups showed no substantial differences. However, when stratified by sample size and region, significant differences were observed among the subgroups: the ratio of OR was 0.70 (95%CI 0.50 to 0.98, $P = 0.04$) for trials with sample size less than 500, while was 1.09 (95%CI 0.82 to 1.46, $P = 0.54$) for trials with a sample size of 500 or more (P for interaction test = 0.05); similarly, the ratio of OR was 1.19 (95%CI 0.93 to 1.52, $P = 0.16$) for trials conducted in Europe or America, while was 0.85 (95%CI 0.69 to 1.05, $P = 0.13$) for trials conducted in multi-countries (P for interaction test = 0.04). See Fig. 3.

Harm effects in retrospectively registered versus prospectively registered trials

In the hierarchical linear regression analysis of the 126 confounder-harmonized groups of 342 trials that addressed all the ten confounders, the OR of harms in retrospectively registered trials was lower than that of prospectively registered trials (ratio of OR = 0.75, 95%CI 0.66 to 0.86, $P < 0.01$).

In subgroup analyses, the underestimation remains statistically significant when stratified by type of outcomes, year of publication, source of funding, and sample size. It was more prominent in objective outcomes (ratio of OR = 0.66, 95%CI 0.54 to 0.81, $P < 0.01$), trials published before 2011 (ratio of OR = 0.62, 95%CI 0.48 to 0.79, $P < 0.01$), trials received academic funding (ratio of OR = 0.59, 95%CI 0.44 to 0.79, $P < 0.01$), and trials with larger sample size (ratio of OR = 0.52, 95%CI 0.41 to 0.65, $P < 0.01$). However, when stratified by region, only trials conducted in multi-countries showed such an underestimation (Fig. 4).

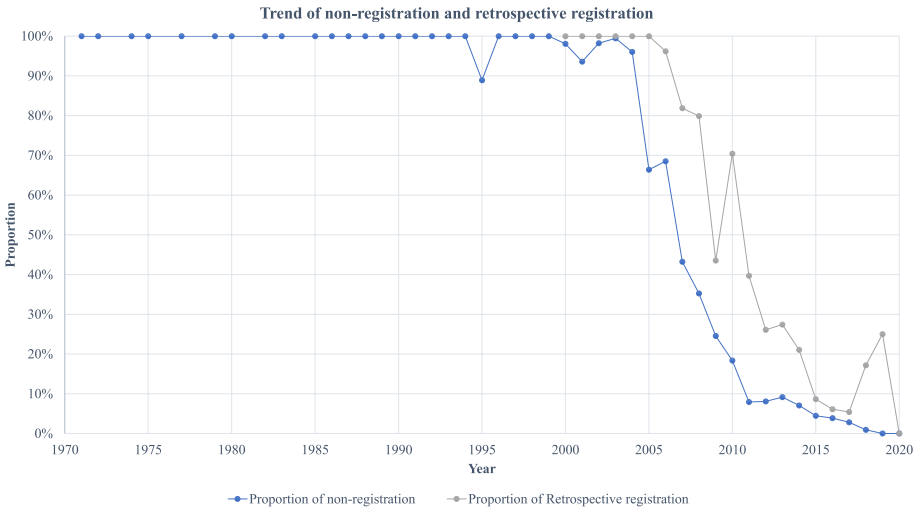

Fig. 2 The proportion of non-registration and retrospective registration over time

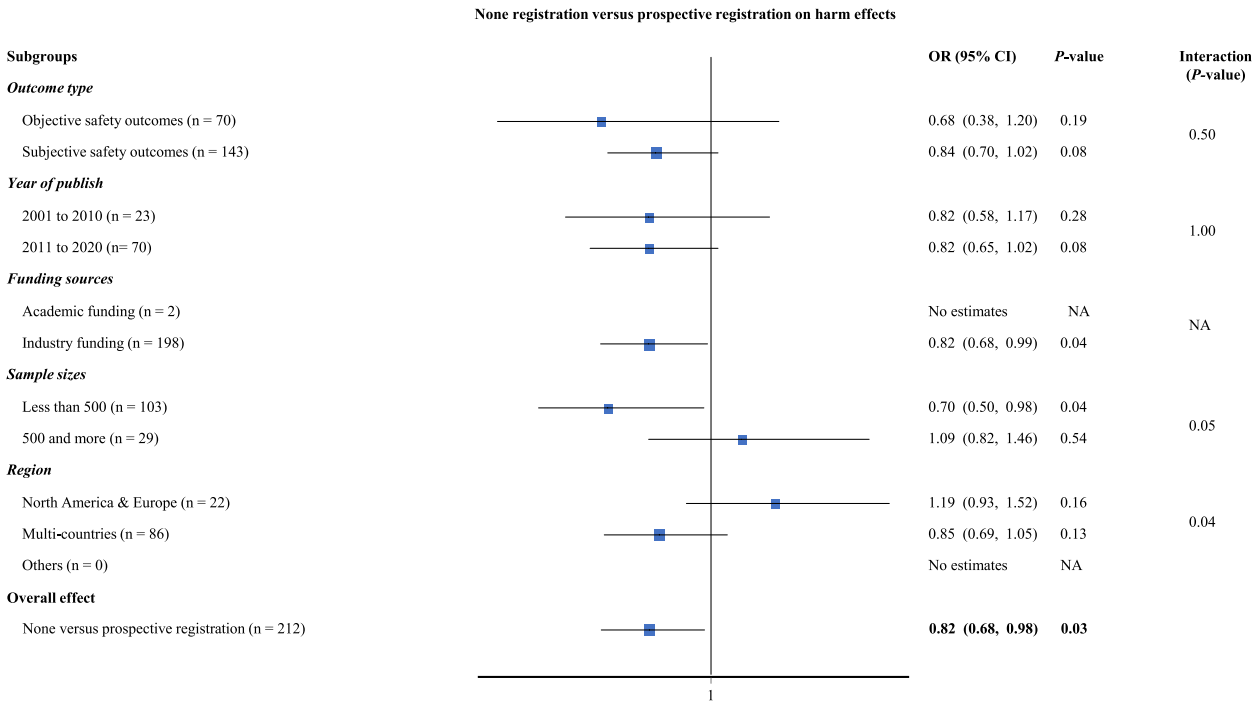

Fig. 3 Estimated harm effects in non-registered versus prospectively registered trials

Harm effects in retrospectively registered versus non-registered trials

In the hierarchical linear regression analysis of the 76 confounders-harmonized groups of 249 trials that addressed all the ten confounders, there was no significant difference in the OR for retrospectively registered and non-registered trials (ratio of OR = 1.02, 95%CI 0.85 to 1.23, $P = 0.83$).

In subgroup analyses, the null difference remained when stratified by type of outcomes and source of funding. However, when stratified by year of publication, in trials published before 2011, those retrospectively registered had larger OR than those not registered (ratio of OR = 1.12, 95%CI 0.75 to 1.68, $P = 0.58$). However, in trials published more recently, those retrospectively registered had smaller OR than those not registered (ratio of OR = 0.85, 95%CI 0.66 to 1.09, $P = 0.20$). Similar inconsistency was observed when stratified by region. See Fig. 5.

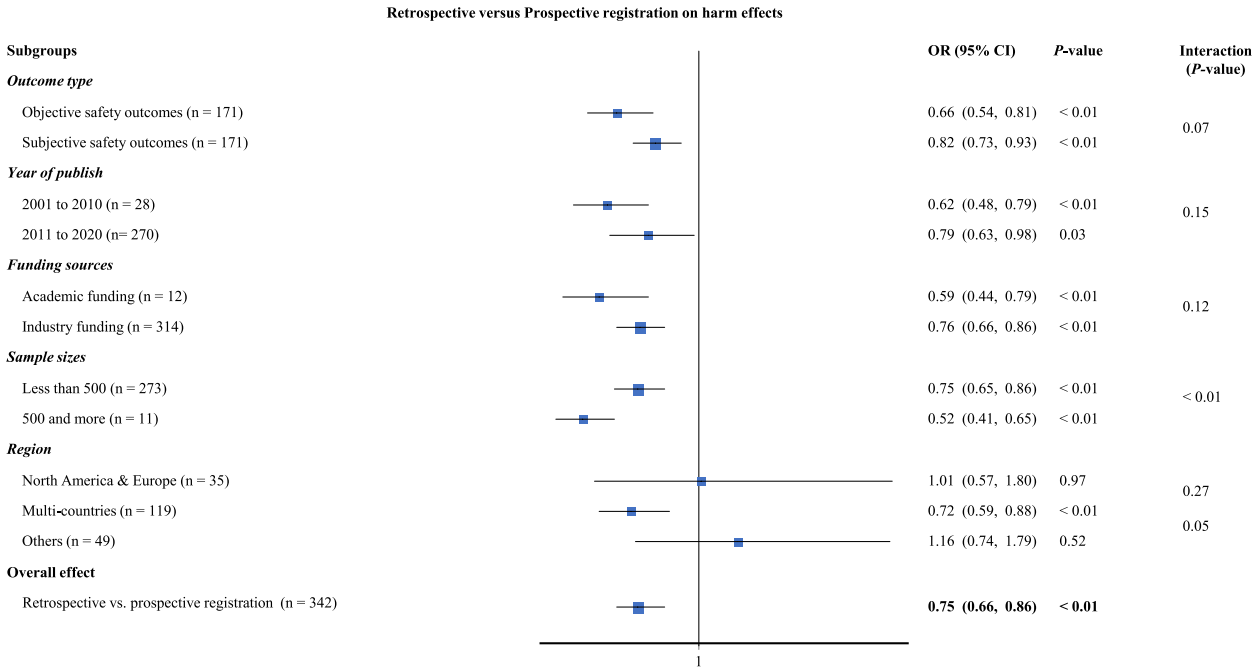

Fig. 4 Estimated harm effects in retrospectively registered versus prospectively registered trials

Discussion

Prospective trial registration has been widely highlighted and intensively promoted in the past decade, and the majority of current trials evaluated here have been prospectively registered. However, those trials that remained unregistered or registered retrospectively may still have an impact on the evidence, practice, and policy, which was less investigated. In this study, we evaluated differences in estimates of medication-related harms for non-registered trials as compared to registered trials after

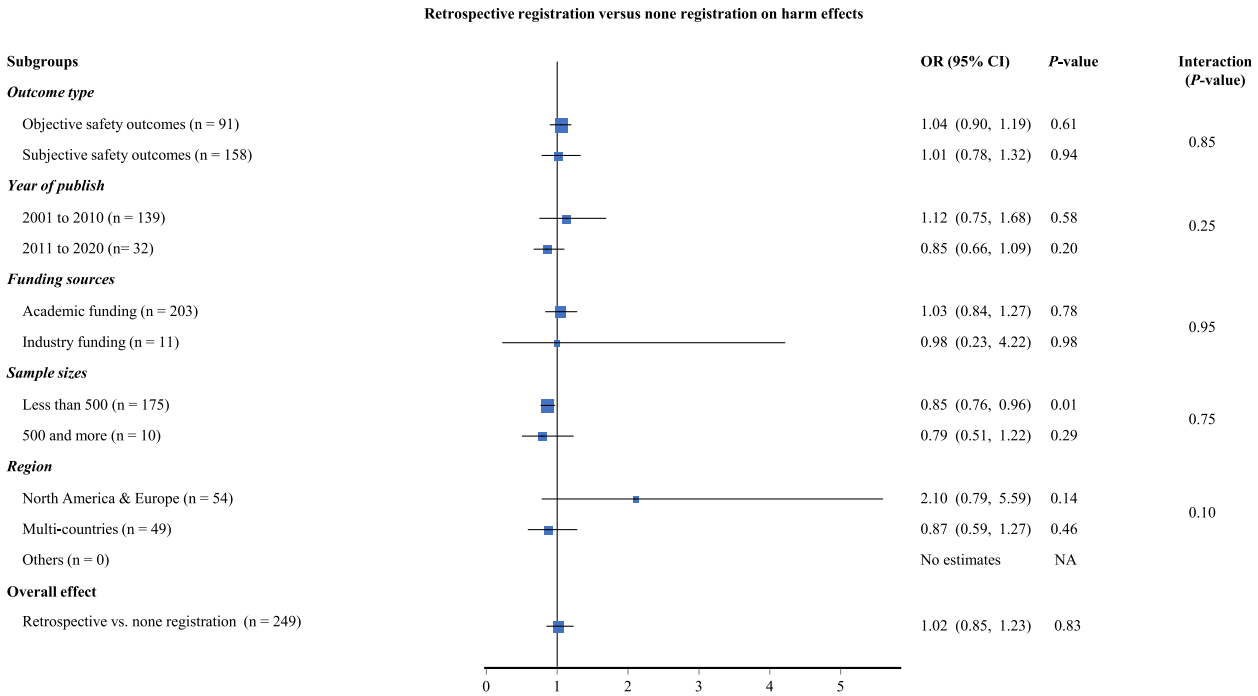

Fig. 5 Estimated harm effects in retrospectively registered versus non-registered trials

controlling for potential confounders. We found that for medication-related harms, non-registered trials and retrospectively registered trials had smaller effect estimates than prospectively registered ones. The findings are supported by a parallel analysis demonstrating that there was no significant difference in harm effects in trials registered retrospectively and trials that were non-registered. To this end, based on empirical evidence, on average, non-registered trials and retrospectively registered trials are associated with underestimates of the harm effects by 18 and 25%.

In our subgroup analyses, we observed some inconsistencies with regard to registration status on harm effects when stratified the comparisons by region. In all three comparisons, for trials conducted in Europe or America, there were no significant differences in the harm effects among non-registered, retrospectively registered, and prospectively registered trials. A potential explanation might be that trials conducted in Europe or North America had better reporting in terms of harm outcomes. This could be expected as currently most of the methodological and reporting guidelines (e.g., CONSORT harms [38]) of clinical trials were initiated in these regions where researchers' reporting behaviors were more likely to be affected. Further research to explore the impact of regulations and policies on clinical trial outcome reporting is worthwhile. Moreover, in the comparison of non-registered vs. prospectively registered trials, the sample size of the trials may have a subgroup effect as only trials with a sample size less than 500 underestimated harms. It is hard to distinguish whether this is by chance or a true difference, because the number of studies used in one of the subgroups is small (103 vs. 29), and similar divergence was not observed in the other two parallel analyses. Nevertheless, this may have little impact on generalizability as the majority of current trials had a sample size of less than 500 [39].

The effects of trial registration on reducing selective or absent reporting of efficacy outcomes may differ substantially from that of harm outcomes. Previous studies have demonstrated exaggerated efficacy estimates in non-registered and retrospectively registered trials versus prospectively registered trials [13–15]. The current study extended the topic to harms and showed harms are under-estimated in non-registered and retrospectively registered trials. If we take these findings together, it seems that non-registered trials and retrospectively registered trials are prone to exaggerate the beneficial effects while discounting the harmful effects. These findings are in line with the competing interests that drive the desire or optimistic expectation of delivering a successful treatment with a high level of benefits and only minimal harm. Our findings provide evidence of selective or absent reporting in non-registered trials and retrospectively registered trials, thus highlighting the critical importance of prospective trial registration.

Implications for future research

In evidence-based practice, registered and non-registered trials are routinely aggregated together. Inclusion of more trials possibly improves the precision of the pooled effects, but can also be associated with a risk of an additional flaw whereby the efficacy may become exaggerated while the harms are underestimated. We do not advocate discarding non-registered trials and retrospectively registered trials—they may still provide useful information, some of which could be of high quality, especially when the events of interest are rare. A valid and simple method would be to run subgroup analyses that separate data from non-registered and retrospectively registered trials from prospectively registered trials [40]. Alternatively, authors of the original publications could be contacted to verify the actual events of specific harm.

The current study showed a similar effect of harms between retrospectively registered and non-registered trials, suggesting retrospective registration did not prevent selective or absent reporting bias. This finding is consistent with a lack of improvement by retrospective registration in terms of efficacy reporting [14]. Regulatory agencies should consider mandatory requirements for prospective registration for clinical trials. Academic journals and peer reviewers are highly recommended to implement strict scrutiny of the results of non-registered trials to avoid potentially misleading healthcare practices. Review authors should be encouraged to use relevant tools (e.g., Outcome Reporting Bias in Trials II, ORBIT II [41]) to classify the risk of selective or absent reporting on harm outcomes of included trials.

Strengths and limitations

To the best of our knowledge, this is the first study that has focused on the association between reported harm effects and the trial registration status. To minimize human errors, we adopted a scrupulous data collection strategy, with all collected data being triple or even quadruple checked after double extraction. The application of directed acyclic graphs helped us to identify potential confounders which we subsequently were able to address within this large-scale dataset. In the data analysis, we considered both within-trial correlation and between-trial correlation to ensure robust parameter estimations. All these efforts strengthened the reliability of our findings.

There were three major limitations that may impact the robustness or representativeness of our findings. During the data collection process, we made every

effort to obtain the information of each trial, for example, by reviewing the supplementary file and registry of each trial. However, there remained a small proportion of trials afflicted by missing information of certain variables due to selective or incomplete reporting. Some very early trials (e.g., in the 1980s) provided limited access to full texts. The removal of these trials (those with missing information and without full texts) led to a 23.6% decrease in available trials for our evaluation. In addition, based on the study by Su et al., safety outcomes were only reported in 43% of the published trials, and the majority did not report any safety outcome [42]. This means the trials collected in our dataset refer to a smaller selection where safety outcomes were reported. While the representativeness of our findings may potentially be impacted, we recognize that there are currently no straightforward solutions for this issue. The release of the CONSORT Harms 2022 statement may help to improve the reporting issue [43]. The setting of minimal clinically important differences in our study may also help to partially address the impact. Moreover, the current study only included trials of pharmacological agents. It is unclear whether underestimation of harms is also a problem in trials of non-drug trials.

Conclusions

Current evidence suggests that non-registered trials and retrospectively registered trials may be linked to smaller estimated harm effects than prospectively registered ones. While trial registration had a beneficial effect on reducing selective or absent reporting for safety outcomes, there was no evidence that retrospective registration has a demonstrable benefit on better reporting. Prospective registration is highly recommended for future trials to reduce potential selective or absent reporting.

Supplementary information.

Abbreviations

CI	Confidence interval
CONSROT	Consolidated Standards of Reporting Trials
IQR	Interquartile ranges
ITT	Intention-to-treat
OR	Odds ratio
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RoB	Risk of Bias
ROR	Ratio of odds ratio

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-024-03621-7.

Additional file 1: Fig. S1. The DAG plot for identifying potential effect moderators. Fig. S2. Study design and analysis.

Acknowledgements

We thank Mr. Lu Cuncun from Lanzhou University for developing the search strategy. Xi Yang, Rui Zhang, Xing Xing, Yuan Tian, and Yi Zhu from Anhui Medical University, and Xiaoqin Zhou, Minghan Dai, Fengying Zhang, and Tianqi Yu from West China Hospital of Sichuan University provided assistance in data collection and verification.

Authors' contributions

Conception and design: CX; Manuscript drafting: CX; Data collection: CX, FY, TQ, XY, SF; Data analysis and result interpretation: CX, SF; Statistical guidance: LL, HC; Methodology guidance: LFK, LZ, SV, YL, SG, SL; Manuscript editing: LL, LFK, LZ, SV, YL, HC, SG, SL. All authors read and approved the final manuscript. All authors approved the final version to be published. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding

This study was supported by the National Natural Science Foundation of China (72204003) and the institutional funding by Shanghai Eastern Hepatobiliary Surgery Hospital of Navy Medical University ('TengFei Project', TF2024YZRH03). The funding bodies had no role in any process of the study (i.e., study design, analysis, interpretation of data, writing of the report, and the decision to submit the article for publication).

Availability of data and materials

Data can be found at https://osf.io/g3mdu/.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Proof of Concept Center, Eastern Hepatobiliary Surgery Hospital, Third Affiliated Hospital, Second Military Medical University, Naval Medical University, Shanghai, China. [2]UQ Center for Clinical Research, The University of Queensland, Herston, Australia. [3]Department of Endocrinology and Metabolism, Division of Guideline and Rapid Recommendation, Chinese Evidence-Based Medicine Centre, Cochrane China Centre, MAGIC China Centre, West China Hospital, Sichuan University, Chengdu, China. [4]Department of Epidemiology and Biostatistics, University of Arizona, Tucson, AZ, USA. [5]Statistical Research and Innovation, Global Biometrics and Data Management, Pfizer Inc, New York, NY, USA. [6]Department of Health Sciences, University of York, York, UK. [7]Norwich Medical School, University of East Anglia, Norwich, UK. [8]Department of Pediatrics, Faculty of Medicine and Dentistry, Department of Psychiatry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada. [9]Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN, USA.

Received: 31 January 2024 Accepted: 9 September 2024
Published online: 11 October 2024

References

1. Rothman KJ, Greenland S, Lash TL (editors). Modern epidemiology, 3rd Edition. Baltimore: Lippincott Williams & Wilkins; 2008.
2. Sibbald B, Roland M. Understanding controlled trials. Why are randomised controlled trials important? BMJ. 1998;316(7126):201.
3. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med. 2000;342(25):1887–92.
4. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for

Systematic Reviews of Interventions version 6.3. Cochrane. 2022. Available from www.training.cochrane.org/handbook. Updated Feb 2022.
5. Laine C, Horton R, DeAngelis CD, Drazen JM, Frizelle FA, Godlee F, et al. Clinical trial registration–looking back and moving ahead. N Engl J Med. 2007;356(26):2734–6.
6. Zarin DA, Tse T, Williams RJ, Rajakannan T. Update on trial registration 11 years after the ICMJE policy was established. N Engl J Med. 2017;376(4):383–91.
7. Rennie D. Trial registration: a great idea switches from ignored to irresistible. JAMA. 2004;292(11):1359–62.
8. De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, International Committee of Medical Journal Editors, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. N Engl J Med. 2004;351(12):1250–1.
9. Hunter KE, Seidler AL, Askie LM. Prospective registration trends, reasons for retrospective registration and mechanisms to increase prospective registration compliance: descriptive analysis and survey. BMJ Open. 2018;8(3):e019983.
10. Trinquart L, Dunn AG, Bourgeois FT. Registration of published randomized trials: a systematic review and meta-analysis. BMC Med. 2018;16(1):173.
11. Speich B, Gryaznov D, Busse JW, Gloy VL, Lohner S, Klatte K, et al. Nonregistration, discontinuation, and nonpublication of randomized trials: a repeated metaresearch analysis. PLoS Med. 2022;19(4):e1003980.
12. Al-Durra M, Nolan RP, Seto E, Cafazzo JA. Prospective registration and reporting of trial number in randomised clinical trials: global cross sectional study of the adoption of ICMJE and Declaration of Helsinki recommendations. BMJ. 2020;369:m982.
13. Odutayo A, Emdin CA, Hsiao AJ, Shakir M, Copsey B, Dutton S, et al. Association between trial registration and positive study findings: cross sectional study (Epidemiological Study of Randomized Trials-ESORT). BMJ. 2017;356:j917.
14. Dechartres A, Ravaud P, Atal I, Riveros C, Boutron I. Association between trial registration and treatment effect estimates: a meta-epidemiological study. BMC Med. 2016;14(1):100.
15. Chan AW, Pello A, Kitchen J, Axentiev A, Virtanen JI, Liu A, Hemminki E. Association of trial registration with reporting of primary outcomes in protocols and publications. JAMA. 2017;318(17):1709–11.
16. Ernst E, Pittler MH. Assessment of therapeutic safety in systematic reviews: literature review. BMJ. 2001;323(7312):546.
17. Helsingen LM, Zeng L, Siemieniuk RA, Lytvyn L, Vandvik PO, Agoritsas T, et al. Establishing thresholds for important benefits considering the harms of screening interventions. BMJ Open. 2020;10(12):e037854.
18. Xu C, Yu T, Furuya-Kanamori L, Lin L, Zorzela L, Zhou X, et al. Validity of data extraction in evidence synthesis practice of adverse events: reproducibility study. BMJ. 2022;377:e069155.
19. Gates M, Gates A, Pieper D, Fernandes RM, Tricco AC, Moher D, et al. Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. BMJ. 2022;378:e070849.
20. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. Evid Based Med. 2017;22(4):139–42.
21. Fan S, Yu T, Yang X, Zhang R, Furuya-Kanamori L, Xu C. The SMART safety: an empirical dataset for evidence synthesis of adverse events. Data Brief. 2023;51:109639.
22. Zorzela L, Golder S, Liu Y, Pilkington K, Hartling L, Joffe A, et al. Quality of reporting in systematic reviews of adverse events: systematic review. BMJ. 2014;348:f7668.
23. Gøtzsche PC, Hróbjartsson A, Maric K, Tendal B. Data extraction errors in meta-analyses that use standardized mean differences. JAMA. 2007;298(4):430–7.
24. Tendal B, Higgins JP, Jüni P, et al. Disagreements in meta-analyses using outcomes measured on continuous or rating scales: observer agreement study. BMJ. 2009;339:b3128.
25. van de Wetering FT, Scholten RJ, Haring T, Clarke M, Hooft L. Trial registration numbers are underreported in biomedical publications. PLoS One. 2012;7(11):e49599.
26. Vale CL, Tierney JF, Burdett S. Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews. BMJ. 2013;346:f1798.
27. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
28. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2021;50(2):620–32.
29. Page MJ, Higgins JP, Clayton G, Sterne JA, Hróbjartsson A, Savović J. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. PLoS One. 2016;11(7):e0159267.
30. Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ. 2008;336(7644):601–5.
31. Hanley JA. Simple and multiple linear regression: sample size considerations. J Clin Epidemiol. 2016;79:112–9.
32. Doi SA, Furuya-Kanamori L, Xu C, Lin L, Chivese T, Thalib L. Controversy and debate: questionable utility of the relative risk in clinical research: paper 1: a call for change to practice. J Clin Epidemiol. 2022;142:271–9.
33. Xu C, Zhou X, Zorzela L, Ju K, Furuya-Kanamori L, Lin L, et al. Utilization of the evidence from studies with no events in meta-analyses of adverse events: an empirical investigation. BMC Med. 2021;19(1):141.
34. Xu C, Furuya-Kanamori L, Zorzela L, Lin L, Vohra S. A proposed framework to guide evidence synthesis practice for meta-analysis with zero-events studies. J Clin Epidemiol. 2021;135:70–8.
35. Xu C, Doi SAR. The robust error meta-regression method for dose-response meta-analysis. Int J Evid Based Healthc. 2018;16(3):138–44.
36. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials0. BMJ. 2010;340:c332.
37. Mavridis D, White IR. Dealing with missing outcome data in meta-analysis. Res Synth Methods. 2020;11(1):2–13.
38. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 2010;2010(8):18.
39. Davey J, Turner RM, Clarke MJ, Higgins JP. Characteristics of meta-analyses and their component studies in the Cochrane Database of Systematic Reviews: a cross-sectional, descriptive analysis. BMC Med Res Methodol. 2011;11:160.
40. Stone J, Gurunathan U, Glass K, Munn Z, Tugwell P, Doi SA. Stratification by quality induced selection bias in a meta-analysis of clinical trials. J Clin Epidemiol. 2019;107:51–9.
41. The ORBIT Team. The ORBIT II classification system for missing or complete outcome reporting in harm outcomes. https://outcome-reporting-bias.org/HarmOutcomes. Accessed 12 June 2023.
42. Golder S, Loke YK, Wright K, Norman G. Reporting of adverse events in published and unpublished studies of health care interventions: a systematic review. PLoS Med. 2016;13(9):e1002127.
43. Junqueira DR, Zorzela L, Golder S, Loke Y, Gagnier JJ, Julious SA, et al. CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. BMJ. 2023;381:e073725.

Publisher's Note