

REVIEW

# Protocols for meta-analysis of intervention safety seldom specified methods to deal with rare events

You Zhou<sup>a</sup>, Bo Zhu<sup>b</sup>, Lifeng Lin<sup>c</sup>, Joey S.W. Kwong<sup>d</sup>, Chang Xu<sup>e,\*</sup>

<sup>a</sup>School of Materials Science and Engineering, University of Science and Technology Beijing, Beijing, China

<sup>b</sup>Department of Hepatobiliary Surgery, The Second Affiliated Hospital, Guangzhou Medical University, Guangzhou, China

<sup>c</sup>Department of Statistics, Florida State University, Tallahassee, FL, USA

<sup>d</sup>JC School of Public Health and Primary Care, Faculty of Medicine, Chinese University of Hong Kong, Hong Kong, China

<sup>e</sup>Department of Population Medicine, College of Medicine, Qatar University, Doha, Qatar

Accepted 21 September 2020; Published online 25 September 2020

## Abstract

**Objectives:** Meta-analyses of rare events often generate unstable results, and selective reporting of the results may mislead the health care decision. Developing a synthesis plan for rare events in protocol may help to formulate the reporting. We aim to investigate whether existing protocols specified methods to deal with rare events.

**Study Design and Setting:** Protocols (not including Cochrane protocols) for systematic reviews of health care interventions focused on the safety registered in PROSPERO were included. The proportion of protocols that specified methods to deal with rare events and the detailed methods were summarized. We compared the difference of proportions (PD) across different settings.

**Results:** We identified 1,004 eligible protocols, of which, 119 (11.85%, 95% confidence interval (CI): 9.92%, 14.01%) specified methods to deal with rare events. The three most commonly planned methods were the Mantel–Haenszel, Peto's odds ratio, and continuity correction. Protocols planned a quantitative analysis (PD = 0.07, 95% CI: 0.02, 0.12;  $P = 0.004$ ) and listed safety as a primary outcome (PD = 0.07, 95% CI: 0.01, 0.12;  $P = 0.018$ ) were more likely to specify methods to deal with rare events.

**Conclusion:** Protocols for systematic reviews of intervention safety seldom specified methods to deal with rare events. Future systematic reviewers should provide a detailed and rigorous synthesis plan for rare events in their protocols. © 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Protocols; Meta-analysis of rare events; Data synthesis plan; PROSPERO

## 1. Introduction

Meta-analysis is a valid method to quantitatively incorporate findings from multiple studies on the same topic, as an effort, to increase the robustness of the evidence [1–3]. This powerful method has been widely used to assess effectiveness and safety of health care intervention [4,5]. Meta-analysis of health care intervention for

effectiveness has been well recognized, as the results were mostly robust and credible. As for safety, in many cases, because of the low incidence, small sample size, and so forth, the adverse events observed in a single study are often rare and may even be zero; the estimation of the effect and the variance then suffer from small-sample bias or even the problem of separation, making the synthesis problematic [6–8]; the results of meta-analyses for rare events thus tend to be unstable that the direction on effect size or  $P$ -value could be easily altered under different dispositions [9].

Several methods have been proposed to deal with rare events in meta-analysis. These mainly include the continuity correction, Peto's odds ratio (Peto's OR), the Mantel–Haenszel (MH) method, the arcsine-based transformations, the generalized linear mixed models (GLMMs), and the Bayesian models [10–17]. These methods, although have unique advantages in dealing with rare

Conflicts of interests: None.

Authors' contributions: X.C. conceived and designed the study and drafted the manuscript; Z.Y. developed the code and acquired the data; X.C. and L.L. analyzed the data and interpreted the results; X.C. and Z.B. screened the literature; L.L. and J.K. contributed careful edits for the manuscript. All authors approved the final version.

\* Corresponding author. Department of Population Medicine, College of Medicine, Qatar University, Al Jamiaa Street, P. O. Box 2713, Doha, Qatar.

E-mail address: [xuchang2016@runbox.com](mailto:xuchang2016@runbox.com) (C. Xu).

### What is new?

#### Key findings

- The majority (88.15%) of the protocols for systematic reviews of intervention safety failed to specify the methods to deal with rare events, and this worrisome issue was not improved over times.
- Protocols that had a meta-analysis plan, specified safety as the primary outcome, and involved more authors were more likely to document methods for dealing with rare events.

#### What this study adds to what is known?

- Meta-analyses of rare events often generate unstable results, and selective reporting of the results may mislead the health care decision. Developing a data synthesis plan in protocol may help formulate the reporting. In this study, we investigated the data synthesis plans on dealing with rare events in protocols for systematic reviews of intervention safety registered in PROSPERO. Our findings may be helpful for future protocol formulation and guideline development for meta-analyses of rare events.

#### What are the implications and what should be changed?

- Systematic reviewers are recommended to provide a detailed and rigorous data synthesis plan in their protocols.
- We advocate future systematic reviewers to specify at least three methods for dealing with rare events, which may include some generic approaches (Peto, MH, or continuity correction), GLMMs, and Bayesian models with appropriate priors.

events, more or less have their own weakness (e.g., Bayesian can solve the separation problem of zero events, but the results are susceptible to the setting of prior parameters); as a result, a perfect solution for rare events in meta-analysis is still hard to be established. Coupled with the unstable nature of statistical inference on rare events, the use of different methods may lead to dramatically different conclusions. Even worse, the selection of the methods is often arbitrary in practice because of the absence of a formal guideline or consensus. On this base, a selective reporting of the results may easily alter the conclusions toward what potential stakeholders intended to obtain [18].

Some feasible measures can be used to avoid such selective reporting. Sweeting et al. advocated to use several data synthesis methods for meta-analysis of rare events as a sensitivity analysis [19]. This approach makes sense if the

results produced by different methods are generally consistent and credible. However, Sweeting et al. did not explicitly point out the minimal sets of methods that need to be considered [19]; it is often up to stakeholders to determine which and how many synthesis methods were used for analyses. Another valid approach is to develop a data synthesis plan in advance and share it with the public. The international prospective register of systematic reviews, PROSPERO (<https://www.crd.york.ac.uk/prospéro/>), serves as an important platform to make protocols publicly available [20,21].

Since its first launch in 2011, a large amount of protocols of meta-analyses on health care safety has been registered in PROSPERO, and the number is continuing to sharply increase. PROSPERO has a structured framework to formulate the details of these protocols, including review questions, main outcomes, strategy for data synthesis, additional analysis, risk of bias assessment, and so forth. Nevertheless, it is unclear whether the strategy for data synthesis in these protocols contained a clear plan for rare events. In this study, we conducted a comprehensive survey on systematic reviews of health care safety registered in PROSPERO and explored the data synthesis plans for rare events in their protocols. We aimed to investigate whether these protocols documented detailed methods to deal with rare events.

## 2. Methods

### 2.1. Eligibility criteria

This survey collected protocols for systematic reviews of health care intervention that focused on the safety, with or without a meta-analysis plan, and were registered in PROSPERO as of Feb-14, 2020. This required the protocols to specify safety as the primary or as one of the main outcomes. We choose protocols registered in PROSPERO because it is the most commonly (71.3%) used platform for protocol registration [22]. We defined the primary outcome as the one listed in front of all other outcomes [23]; for the main outcomes, we defined them as those outcomes listed in the domain of “main outcome(s)” within a protocol. We did not consider protocols without safety outcomes and those with safety outcomes but listed them in the domain of “additional outcome(s)”. Reviews of systematic reviews (i.e., overviews) as well as reviews of practical guidelines were excluded because they were obviously out of the scope of our survey. Cochrane protocols and protocols relating to animal experiments were also excluded.

### 2.2. Search for protocols

We searched for potential protocols in PROSPERO with two groups of keywords separately: “adverse” and “safety”. For the group of “adverse”, the following keywords were used: *adverse event OR adverse effect OR*

adverse outcome; and for the group of “safety”, we searched for the keyword *safety*. The following restrictions were used to make the search more accurate: type and method of the review—*intervention*; source of the review—*exclude animal, exclude Cochrane protocol*. No other restrictions were used (see [Appendix](#)). It should be highlighted that because of the difficulty on obtaining all protocols referred to safety, we only expected to find a proportion of the potentially eligible protocols. This would have little impact on the results as the sample is expected to be representative.

### 2.3. Screen for protocols

The search records were exported from PROSPERO in a TXT file, and a Python 3.8.2 program was used to download the full text (HTML file) of each protocol in accordance with the Centre for Reviews and Dissemination (CRD) number listed in the TXT file. Then, the content of each protocol was automatically transformed into an Excel sheet (Microsoft, USA) via the Python program, divided by the subheading of each domain.

The screen of these protocols was based on two steps. We first used the “IF” and “Find” functions to automatically search for any possible keywords of safety, adverse and complication from the title, main outcome(s), and additional outcome(s). Those contained these keywords in additional outcome(s) but not in main outcomes were excluded. Then, the remaining protocols that contained these keywords in title or main outcome domains were screened through full-text by two authors separately, one methodologist (X.C.) and one surgeon (Z.B.). Any disagreements were solved by online discussion.

### 2.4. Data extraction

The following information of each included protocol was extracted: CRD number, number of authors, data synthesis plan (quantitative synthesis or qualitative synthesis), planned methods for dealing with rare events, planned effect estimates, word count for strategy of data synthesis, funding source, region of the protocol (documented in the “country” domain), data of the registration, and stage of the review. The information of data synthesis plan, planned methods for dealing with rare event, and planned effect estimates was extracted by the lead author (X.C.), an experienced methodologist of evidence synthesis. Other information was extracted automatically by the Python program or Excel macro (Z.Y.).

### 2.5. Data analysis

Data analysis was conducted by the lead author (X.C.) and a statistician (L.L.). The baseline information (e.g., data of registration, number of authors, funding source, and region) was summarized in terms of the data type through either proportions or quantiles. The main outcome

of interest is the proportion and 95% confidence interval (CI) of protocols that specified the methods to deal with rare events. The methods planned to deal with rare events were also summarized.

We further compared the potential difference of the proportions (PD) in terms of data synthesis type (quantitative vs. qualitative), importance of safety outcome (primary vs. one of the main outcomes), funding source, year of registration (2016 to 2020 vs. 2011 to 2015), number of authors (quartiles 2, 3, and 4 vs. quartile 1), and review stage of the protocol (completed vs. ongoing). The cutoff point for year of registration was set by the year of the release of the PRISMA harms checklist [24]. We defined these variables in a priori, which were the baseline information that we could have extracted from the protocol and may associate with the proportion by our subjective judgment. Previous study had recorded that some of these variables may be associated with methodological validity of meta-analysis [25]. We used the absolute difference as the effect estimate because it is a valid method in cases of zero counts in a single arm [26]. Zero counts in both arms were not likely to occur, as they only happened when none of these protocols documented a plan to deal with rare events.

We used the *t*-test for the statistical inference and pre-specified the significance level at  $\alpha = 0.05$ . When there were rare events ( $n < 5$ , including 0) in any arm, Fisher’s exact test was used as a sensitivity analysis [27]. A *post hoc* sensitivity analysis was used by using multivariable generalized linear regression (binomial, with link function of identity) for aforementioned comparisons. The MetaXL 5.3 (EpiGear International, Australia) and Stata 14.0/SE (STATA, College Station, TX) were used to perform the analyses and visualize the results.

## 3. Results

We obtained 3,186 records from the PROSPERO database; 232 of them were identified as duplicates by the CRD number. We further removed 1,471 protocols with safety as “additional outcomes”. The remaining 1,483 protocols were screened by full-text, of which 1,004 protocols were included in our study. The kappa statistic was 0.87 between the two raters for the full-text screen, which suggested a high agreement ([Figure S1, Appendix 1](#)).

### 3.1. Baseline information

[Table 1](#) presents the baseline information of these 1,004 protocols. Among these, 178 (17.73%) were registered from Jan 7, 2011 to Dec 31, 2015, and the majority ( $n = 826$ , 82.27%) were registered from Jan 1, 2016 to Feb 14, 2020. The number of registrations dramatically increased in 2018 and 2019; [Figure 1](#) presents the number of registrations from Jan 7, 2011 to Dec 31, 2019. The principal authors of these protocols came from 54 countries,

**Table 1.** Baseline characteristics of included protocols for meta-analysis of intervention safety

Baseline characteristics	All publication (N = 1004)
<b>Year</b>	
2011 (Jan-7)–2015	178 (17.73%)
2016–2020 (Feb-14)	826 (82.27%)
<b>Region of the first author</b>	
Africa	19 (1.89%)
America (North and South)	254 (25.30%)
Asia	332 (33.07%)
European	348 (34.66%)
Oceania	51 (5.08%)
<b>Author number [median (Q1, Q3)]</b>	
≤ 3	373 (37.15%)
4–5	326 (32.47%)
≥ 6	305 (30.38%)
<b>Meta-analysis plan</b>	
Yes	885 (88.15%)
No (narrative synthesis)	107 (10.66%)
Not reported	12 (1.20%)
<b>Effect estimates for binary outcomes for those planned a meta-analysis</b>	
Relative risk (e.g., RR, OR, HR)	484 (54.69%)
Absolute risk (e.g., RD, incidence)	31 (3.50%)
Both relative and absolute scale	16 (1.81%)
Not reported	354 (40.00%)
<b>Word counts of strategy for data synthesis</b>	
≤ 50	194 (19.32%)
51–100	268 (26.69%)
101–200	331 (32.97%)
≥ 201	211 (21.02%)
<b>Safety as</b>	
Primary outcome	210 (20.92%)
One of main outcomes	794 (79.08%)
<b>Funding</b>	
Nonprofit funding	429 (42.73%)
Profit funding	22 (2.19%)
No funding	531 (52.89%)
Not reported	22 (2.19%)
<b>Stage of review</b>	
Review ongoing	823 (81.97%)
Review discontinued	7 (0.70%)
Review completed—published	96 (9.56%)
Review completed—not published	78 (7.77%)

and authors from China ( $n = 245$ ), the United Kingdom ( $n = 158$ ), Canada ( $n = 97$ ), and the United States ( $n = 86$ ) contributed the most. Figure 2 presents the map of the registrations by country.

For the 1,004 protocols, the median number of authors was 4 (interquartile range [IQR]: 3 to 6). There were 373 (37.15%) protocols with three or less authors, 326 (32.47%) with four or five authors, and 305 with six or

more authors. There were 885 (88.15%) protocols with meta-analysis plans, 107 (10.66%) planned qualitative descriptions (narrative syntheses) instead of meta-analyses, and 12 (1.20%) failed to provide such information. The relative measures (including risk ratios, rate ratios, odds ratios, and hazard ratios) were planned most commonly ( $n = 484$ , 54.69%) among protocols that planned a meta-analysis for rare events; the absolute measures (including risk differences, rate differences, incidences, rates, and proportions) were only planned in 31 (3.50%) protocols.

The median word count of strategy for data synthesis was 108 (IQR: 62 to 183). Almost a half ( $n = 462$ , 46.02%) had less than 100 words for the description of strategy for data synthesis. There were 210 (20.92%) protocols that documented safety as the primary outcome, and the remaining 794 (79.08%) documented safety as the one of the main outcomes. Funding information was available in 982 (97.81%) protocols: 429 (42.73%) received nonprofit funding, 22 (2.19%) received profit funding, and 531 (52.89%) did not receive any funding on the data of registration. For the stage of review, 823 (81.97%) protocols were ongoing, 174 (17.33%) completed, whereas seven (0.70%) discontinued.

### 3.2. Dealing with rare events

There were 119 (11.85%, 95% CI: 9.92% to 14.01%) protocols that specified methods to deal with rare events. Among them, 52 (43.70%) planned to use the MH method, 18 (15.13%) planned to use the Peto's OR, 13 (10.92%) planned to use the continuity correction, nine (7.56%) planned to use qualitative descriptions, eight planned to use trial sequential analyses (6.72%) to detect the power for meta-analysis of rare events, five (4.20%) planned to use GLMMs, five (3.36%) planned to use risk differences (without specify a continuity correction or MH method for synthesis), three (2.52%) planned to use the Peto and MH methods, two (1.68%) planned to use the arcsine transformation (for risk difference), one (0.84%) planned to use the double-arcsine transformation (for incidence), one (0.84%) planned to use Clopper–Pearson exact confidence limits, one (0.84%) planned to use both GLMMs and Fisher's exact method, one (0.84%) planned to use both the GLMM and MH, and one (0.84%) planned to use “specific methods for rare events” but without details. Figure 3 presents the rank of the proportions of these methods.

For the methods documented in the 119 protocols, three could be used to deal with double-arm-zero-event studies (as well as single-arm-zero-event studies). They were risk differences, GLMMs, and the arcsine transformation. Six methods could only be used to deal with single-arm-zero-event studies: the MH (for relative risks), Peto's OR, continuity correction (for relative risks), double-arcsine transformation (for incidence or prevalence), Fisher's exact method, and Clopper–Pearson exact confidence limits (for incidence or prevalence). There were only five (0.5%,

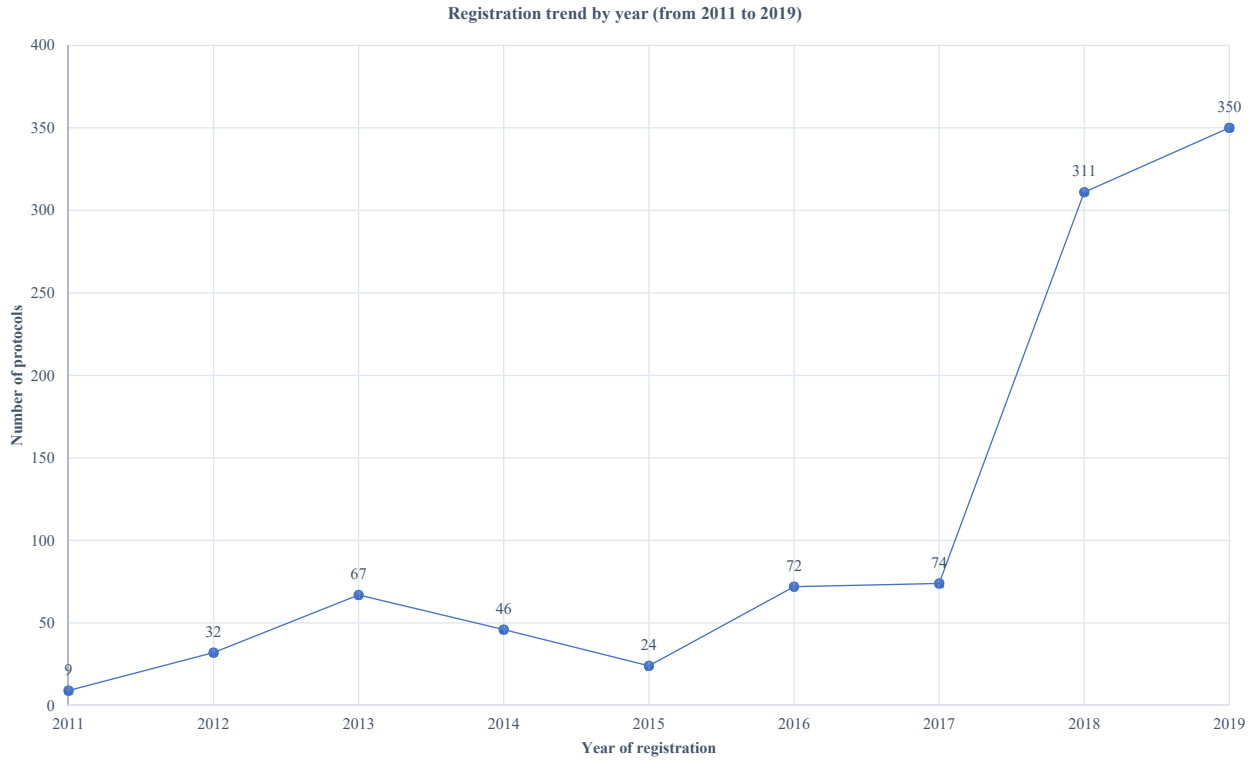


Fig. 1. The number of registration of protocols by years.

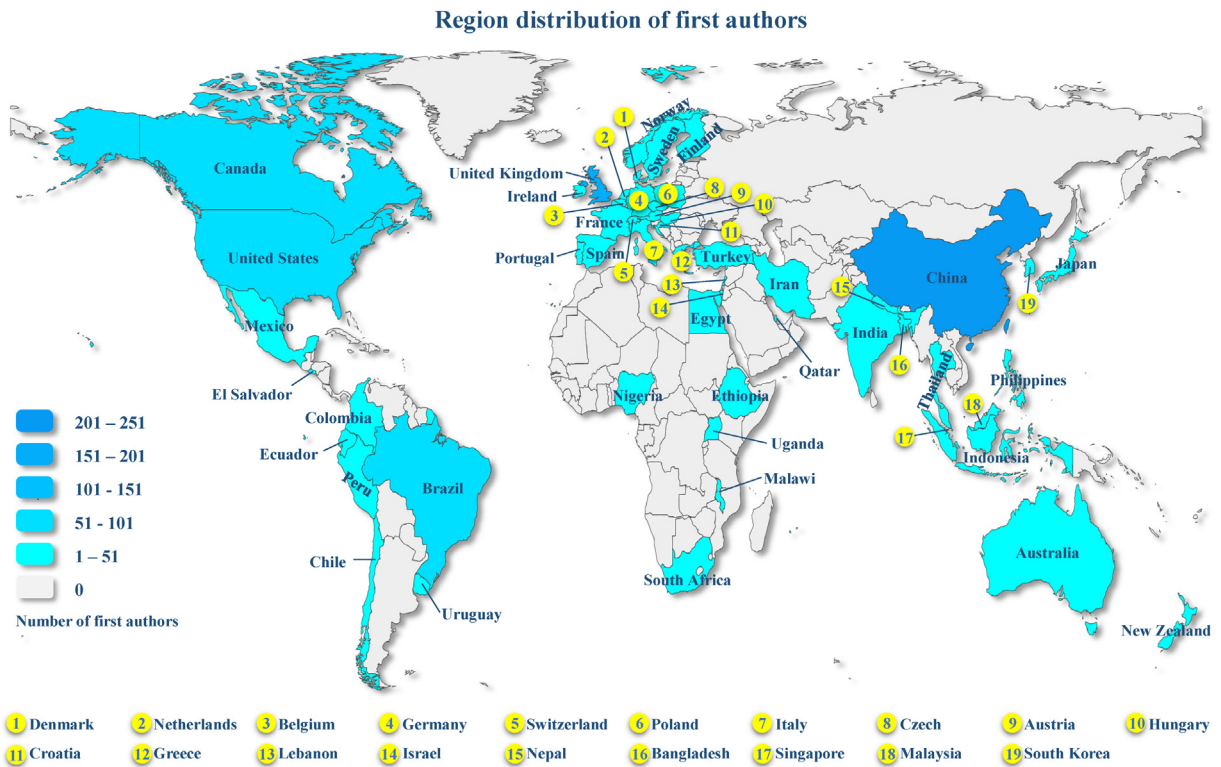


Fig. 2. The distribution of the registration by countries.

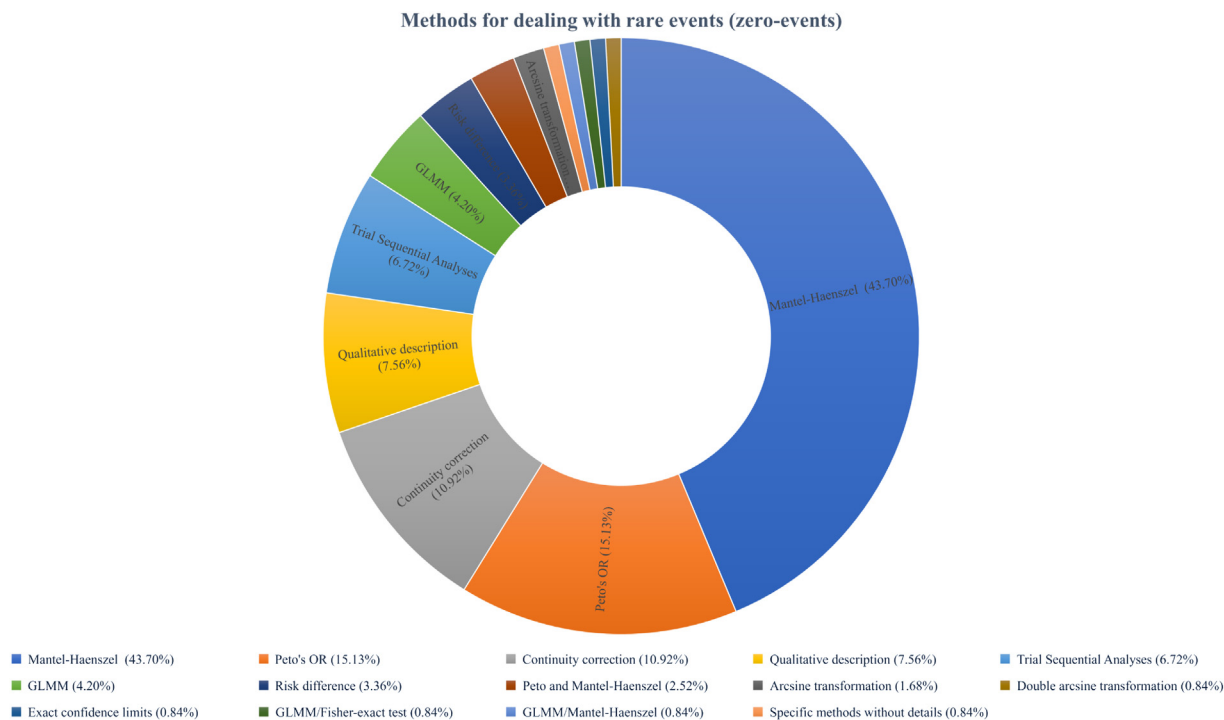


Fig. 3. A summary of the methods specified to deal with rare events.

95% CI: 0.16% to 1.16%) protocols that clearly specified methods to deal with double-arm-zero-event studies. Three of them planned to remove such studies directly, and two planned to use risk differences for estimation.

### 3.3. Comparisons in stratified groups

Figure 4 presents the comparisons among the proportions of protocols that specified methods to deal with rare events. Our results suggested that protocols that planned a quantitative analysis (i.e., meta-analysis, 12.77% vs. 5.61%; PD = 0.07, 95% CI: 0.02, 0.12; P = 0.004), specified safety as the primary outcome (17.14% vs. 10.45%; PD = 0.07, 95% CI: 0.01, 0.12; P = 0.018), had seven or more authors (15.27% vs. 8.85%; PD = 0.06, 95% CI: 0.01, 0.12; P = 0.028), and had been completed (17.24% vs. 10.69%; PD = 0.07, 95% CI: 0.01, 0.13; P = 0.032) were more likely to specify methods for dealing with rare events. There were no significant differences in terms of funding sources and year of registration. Two comparisons (funding sources, profit vs. nonprofit, and not reported vs. no funding) contained rare count cells and were further tested by Fisher's exact method as a sensitivity analysis. The sensitivity analysis suggested the results of these two comparisons were robust (P = 1.00 for profit vs. nonprofit; P = 0.495 for not reported vs. no funding). In our further multivariable regression, again, the results were consistent: those planned a quantitative analysis (PD = 0.07, 95% CI: 0.01, 0.14; P = 0.036), specified safety as the primary outcome (PD = 0.07, 95% CI: 0.02, 0.12; P = 0.004),

had seven or more authors (PD = 0.06, 95% CI: 0.003, 0.11; P = 0.037), and had been completed (PD = 0.06, 95% CI: 0.001, 0.11; P = 0.047) were more likely to specify methods for dealing with rare events.

## 4. Discussion

In this study, we investigated the data synthesis plans on dealing with rare events in protocols for systematic reviews of intervention safety registered in PROSPERO. We found that the majority (88.15%) of these protocols failed to specify the methods to deal with rare events, especially for dealing with double-arm-zero-event studies. This worrisome issue was not improved over times. Our further comparisons suggested that protocols that had a meta-analysis plan, specified safety as the primary outcome, and involved more authors were more likely to specify methods for dealing with rare events.

We summarized several methods for dealing with rare events. The three most commonly planned methods were the MH, the Peto's OR, and the continuity correction. These three were actually the most commonly used methods for rare events in practice. The Peto's OR and MH methods generally perform better than the continuity correction [19,28]. However, the Peto's OR could be only used to deal with studies with single-armed zero-events; for studies with both-arm zero-events, most software programs would remove such studies directly [26]. The continuity correction and MH can be used with risk difference to

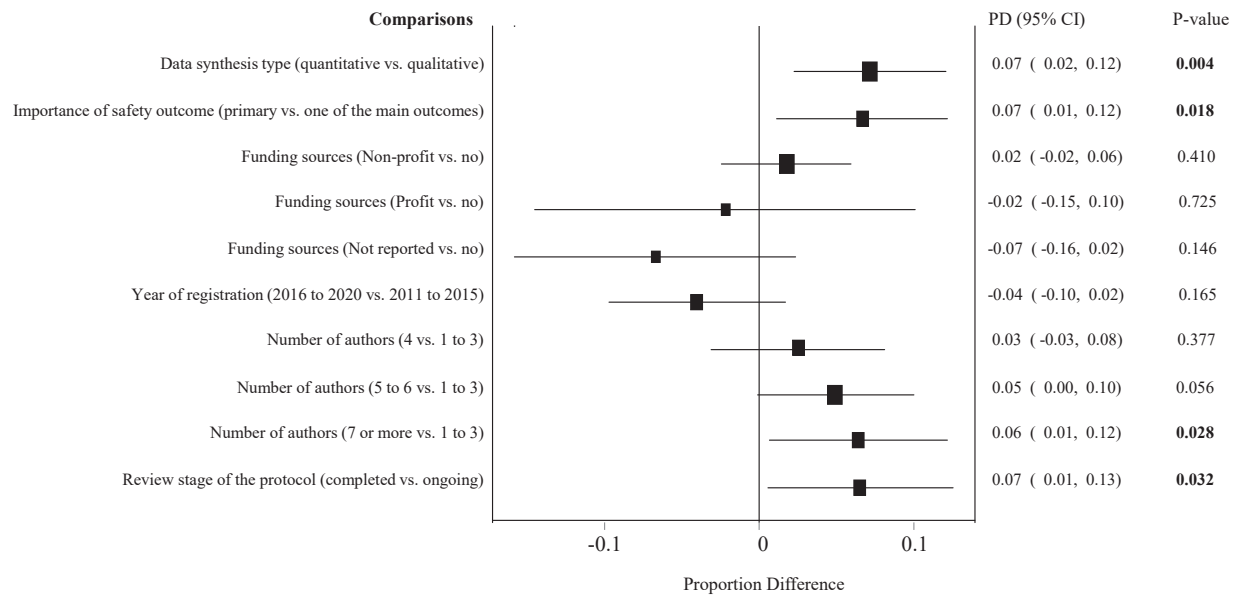


Fig. 4. The comparison among the differences of proportions in different settings.

deal with double-arm-zero-event studies, but the limitation is that these methods tend to generate large biases, especially when the sample size was unbalanced [19].

We found that GLMMs were seldom planned to deal with rare events. Indeed, GLMMs could be a good solution for synthesizing studies with zero events in either a single arm or both arms when there are sufficient total event counts (say,  $>10$  in both arms) [29]. Several simulation studies have verified the validity of GLMMs for meta-analyses of rare events [30–32]. The insufficient use of GLMMs in practice may be primarily attributable to the lack of introduction of this method in popular handbooks (e.g., the Cochrane handbook [26]) for meta-analysis. Another more practical reason is that, although this method could be used in many statistical software (including R, Stata, and SAS), it was not familiar among applied scientists to perform multilevel regression analyses by grouped data. In addition, this method involves more sophisticated theories than the classical synthesis methods. These factors hindered the promotion of GLMMs in meta-analyses of rare events.

In addition, in our survey, the arcsine-based transformations were planned in three protocols. It is also a good method to deal with rare events in terms of the statistical property, even for studies with single-arm zero-events or both-arm zero-events [13,33]. The main limitation of arcsine-based transformation is that the effect is hard to be explained as it was used to estimate the arcsine difference, say, the absolute difference on arcsine-transformed event rates of two comparative arms [33]. Therefore, it compares the “arcsine-transformed event rates”, rather than “event rates”, making the effect unintuitive. It should be highlighted that for meta-analysis of incidence or

prevalence on rare events, there was no such an issue as the arcsine incidence or prevalence could be back-transformed to incidence or prevalence itself [34].

The Bayesian method was planned in none of these protocols. Although Bayesian method is a promising solution to deal with rare events, the effect estimation is sensitive to the assumption of prior parameters as the sampling weight is small for rare events [35,36]. When the prior parameters are appropriately assumed, the Bayesian method is expected to produce satisfactory effect estimates; however, when not appropriately assumed, the results tend to be biased or misleading. The difficulty for the Bayesian method is the assumption of prior parameters. One possible way is to set these prior parameters based on the results of other methods (e.g., GLMM) and thus treat the Bayesian method as sensitivity analysis. Some sophisticated Bayesian method, such as beta-binomial model, has been proven to have good property for rare events, especially for those with zero counts [37,38].

We noticed that very few protocols specified two or more methods to deal with rare events. As we mentioned earlier, the selection of methods could affect the final results for meta-analyses of rare events. Great attention should be paid for this because an intentional or arbitrary use of the methods may seriously bias the conclusions and mislead health care decision. More effort is needed to improve the current situation. We advocate future systematic reviewers to specify at least three methods for dealing with rare events, which may include some generic approaches (Peto, MH, or continuity correction), GLMMs, and Bayesian models with appropriate priors.

To the best of our knowledge, this is the first study that investigated the data synthesis plans on dealing with rare

events in protocols for systematic reviews of intervention safety. We used a comprehensive search for protocols registered in PROSPERO; the retrieved samples are expected to be representative. Our findings may be helpful for future protocol formulation and guideline development for meta-analyses of rare events.

Some limitations of the present study should be highlighted. First, this study did not include Cochrane protocols, so our findings may not be applied to Cochrane protocols. Cochrane protocols often involve rigorous review processes by methodologists for both the rationale and the data synthesis methods planned and are expected to be more technologically sound. A further investigation on Cochrane protocols is worthwhile. Second, our conclusions were based on the protocols, but not on the final systematic reviews. Potential changes are likely to be made in the data synthesis methods in the final systematic reviews, which could not be recorded in this study. A further investigation on published systematic reviews is needed [39–42].

In conclusion, based on current evidence, the protocols for systematic review of intervention safety seldom specified methods to deal with rare events. This situation did not show improvements over times. Systematic reviewers are recommended to provide a detailed and rigorous data synthesis plan in their protocols. Considering the unstable nature of meta-analyses of rare events, we further suggest that at least three methods should be considered to deal with rare events.

### CRedit authorship contribution statement

**You Zhou:** Software, Data curation. **Bo Zhu:** Data curation. **Lifeng Lin:** Methodology, Data curation. **Joey S.W. Kwong:** Writing - review & editing. **Chang Xu:** Conceptualization, Methodology, Data curation, Writing - original draft.

### Acknowledgments

Funding: None.

Data sharing: The primary data are presented in [Appendix 2](#).

### Appendix A

#### Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jclinepi.2020.09.023>.

### References

- [1] Murad MH, Montori VM. Synthesizing evidence: shifting the focus from individual studies to the body of evidence. *JAMA* 2013;309:2217–8.
- [2] Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. *Nature* 2018;555:175–82.
- [3] Shadish WR, Lacy JD. The meta-analytic big bang. *Res Synth Methods* 2015;6(3):246–64.
- [4] Doi SA, Barendregt JJ, Khan S, Thalib L, Williams GM. Advances in the meta-analysis of heterogeneous clinical trials I: the inverse variance heterogeneity model. *Contemp Clin Trials* 2015;45:130–8.
- [5] Efthimiou O, Mavridis D, Nikolakopoulou A, Rücker G, Trelle S, Egger M, et al. A model for meta-analysis of correlated binary outcomes: the case of split-body interventions. *Stat Methods Med Res* 2019;28(7):1998–2014.
- [6] Xu C, Li L, Lin L, Chu H, Thabane L, Zou K, et al. Exclusion of studies with no events in both arms in meta-analysis impacted the conclusions. *J Clin Epidemiol* 2020.
- [7] Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A, et al. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med* 2007;26:53–77.
- [8] Kuss O. Statistical methods for meta-analyses including information from studies without any events-add nothing to nothing and succeed nevertheless. *Stat Med* 2015;34:1097–116.
- [9] Shuster JJ, Jones LS, Salmon DA. Fixed vs random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death. *Stat Med* 2007;26:4375–85.
- [10] Cox DR. The continuity correction. *Biometrika* 1970;57(1):217–9.
- [11] Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomised trials. *Prog Cardiovasc Dis* 1985;27(5):335–71.
- [12] Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959;22:719–48.
- [13] Rücker G, Schwarzer G, Carpenter J, Olkin I. Why add anything to nothing? The arcsine difference as a measure of treatment effect in meta-analysis with zero cells. *Stat Med* 2009;28:721–38.
- [14] Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. *J Epidemiol Community Health* 2013;67:974–8.
- [15] Jackson D, Law M, Stijnen T, Viechtbauer W, White IR. A comparison of seven random-effects models for meta-analyses that estimate the summary odds ratio. *Stat Med* 2018;37:1059–85.
- [16] Liu D, Liu RY, Xie M. Exact meta-analysis approach for discrete data and its application to  $2 \times 2$  tables with rare events. *J Am Stat Assoc* 2014;109:1450–65.
- [17] Kuss O, Hoyer A, Solms A. Meta-analysis for diagnostic accuracy studies: a new statistical model using beta-binomial distributions and bivariate copulas. *Stat Med* 2014;33:17–30.
- [18] Palpacuer C, Hammas K, Duprez R, Laviolle B, Ioannidis JPA, Naudet F. Vibration of effects from diverse inclusion/exclusion criteria and analytical choices: 9216 different ways to perform an indirect comparison meta-analysis. *BMC Med* 2019;17(1):174.
- [19] Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004;23:1351–75.
- [20] Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev* 2012;1:2.
- [21] Seidler AL, Hunter KE, Cheyne S, Ghersi D, Berlin JA, Askie L. A guide to prospective meta-analysis. *BMJ* 2019;367:15342.
- [22] Tawfik GM, Giang HTN, Ghozy S, Altibi AM, Kandil H, Le HH, et al. Protocol registration issues of systematic review and meta-analysis studies: a survey of global researchers. *BMC Med Res Methodol* 2020;20:213.
- [23] Li L, Xu C, Deng K, Zhou X, Liu Z, Busse JW, et al. The reporting of safety among drug systematic reviews was poor before the implementation of the PRISMA harms checklist. *J Clin Epidemiol* 2019;105:125–35.
- [24] Zorzela L, Loke YK, Ioannidis JP, Golder S, Santaguida P, Altman DG, et al. PRISMA harms checklist: improving harms reporting in systematic reviews. *BMJ* 2016;352:i157.



- [25] Xu C, Liu Y, Jia PL, et al. The methodological quality of dose-response meta-analyses needed substantial improvement: a cross-sectional survey and proposed recommendations. *J Clin Epidemiol* 2019;107:1–11.
- [26] Higgins JPT, Thomas J, Chandler J, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd Edition. Chichester (UK): John Wiley & Sons; 2019.
- [27] Shan G, Gerstenberger S. Fisher's exact approach for post hoc analysis of a chi-squared test. *PLoS One* 2017;12:e0188709.
- [28] Cheng LL, Ju K, Cai RL, et al. The use of one-stage meta-analytic method based on individual participant data for binary adverse events under the rule of three: a simulation study. *PeerJ* 2019;7:e6295.
- [29] Wynants L, Bouwmeester W, Moons KG, Moerbeek M, Timmerman D, Van Huffel S, et al. A simulation study of sample size demonstrated the importance of the number of events per variable to develop prediction models in clustered data. *J Clin Epidemiol* 2015; 68:1406–14.
- [30] Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med* 2010;29:3046–67.
- [31] Simmonds M, Higgins JPT. A general framework for the use of logistic regression models in meta-analysis. *Stat Methods Med Res* 2016; 35:2858–77.
- [32] Yu-Kang T. Use of generalized linear mixed models for network meta-analysis. *Med Decis Making* 2014;34:911–8.
- [33] Lin L, Xu C. Arcsine-based transformations for meta-analysis of proportions: pros, cons, and alternatives. *Health Sci Rep* 2020;3(3):e178.
- [34] Lin L, Chu H. Meta-analysis of proportions using generalized linear mixed models. *Epidemiology* 2020;31:713–7.
- [35] Pateras K, Nikolakopoulos S, Roes KCB. Prior distributions for variance parameters in a sparse-event meta-analysis of a few small trials. *Pharm Stat* 2020.
- [36] Günhan BK, Röver C, Friede T. Random-effects meta-analysis of few studies involving rare events. *Res Synth Methods* 2020;11(1):74–90.
- [37] Bakbergenuly I, Kulinskaya E. Beta-binomial model for meta-analysis of odds ratios. *Stat Med* 2017;36:1715–34.
- [38] Chu H, Nie L, Chen Y, Huang Y, Sun W. Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: methods for the absolute risk difference and relative risk. *Stat Methods Med Res* 2012;21(6):621–33.
- [39] Tricco AC, Cogo E, Page MJ, Poliseña J, Booth A, Dwan K, et al. A third of systematic reviews changed or did not specify the primary outcome: a PROSPERO register study. *J Clin Epidemiol* 2016;79: 46–54.
- [40] Rombey T, Doni K, Hoffmann F, Pieper D, Allers K. More systematic reviews were registered in PROSPERO each year, but few records' status was up-to-date. *J Clin Epidemiol* 2020;117:60–7.
- [41] Runjic E, Rombey T, Pieper D, Puljak L. Half of systematic reviews about pain registered in PROSPERO were not published and the majority had inaccurate status. *J Clin Epidemiol* 2019;116:114–21.
- [42] Sideri S, Papageorgiou SN, Eliades T. Registration in the international prospective register of systematic reviews (PROSPERO) of systematic review protocols was associated with increased review quality. *J Clin Epidemiol* 2018;100:103–10.