

Design, methods, and reporting of impact studies of cardiovascular clinical prediction rules are suboptimal: A systematic review

Jong-Wook Ban , Mei Sum Chan , Tonny Brian Muthee ,
Arsenio Paez , Richard Stevens , Rafael Perera

PII: S0895-4356(21)00027-5
DOI: <https://doi.org/10.1016/j.jclinepi.2021.01.016>
Reference: JCE 10404



To appear in: *Journal of Clinical Epidemiology*

Accepted date: 21 January 2021

Please cite this article as: Jong-Wook Ban , Mei Sum Chan , Tonny Brian Muthee , Arsenio Paez , Richard Stevens , Rafael Perera , Design, methods, and reporting of impact studies of cardiovascular clinical prediction rules are suboptimal: A systematic review, *Journal of Clinical Epidemiology* (2021), doi: <https://doi.org/10.1016/j.jclinepi.2021.01.016>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Design, methods, and reporting of impact studies of cardiovascular clinical prediction rules are suboptimal: A systematic review

Jong-Wook Ban^{a, b*}, Mei Sum Chan^c, Tonny Brian Muthee^d, Arsenio Paez^{a, b}, Richard Stevens^d, Rafael Perera^d

^a Centre for Evidence-Based Medicine, Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, United Kingdom

^b Department for Continuing Education, University of Oxford, Rewley House, 1 Wellington Square, Oxford, OX1 2JA, United Kingdom

^c Nuffield Department of Population Health, University of Oxford, Richard Doll Building, Old Road Campus, Oxford, OX3 7LF, United Kingdom

^d Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, United Kingdom

* Corresponding author

E-mail: jong-wook.ban@conted.ox.ac.uk (JWB)

Phone number: +44 (0)1865 600035 (UK), +1 360 970 3953 (USA),

Abstract

Objective

To evaluate design, methods, and reporting of impact studies of cardiovascular clinical prediction rules (CPRs).

Study Design and Setting

We conducted a systematic review. Impact studies of cardiovascular CPRs were identified by forward citation and electronic database searches. We categorized the design of impact studies as appropriate for randomized and non-randomized experiments, excluding uncontrolled before-after study. For impact studies with appropriate study design, we assessed the quality of methods and reporting. We compared the quality of methods and reporting between impact and matched control studies.

Results

We found 110 impact studies of cardiovascular CPRs. Of these, 65 (59.1%) used inappropriate designs. Of 45 impact studies with appropriate design, 31 (68.9%) had substantial risk of bias. Mean number of reporting domains that impact studies with appropriate study design adhered to was 10.2 of 21 domains (95% CI, 9.3 and 11.1). The quality of methods and reporting was not clearly different between impact and matched control studies.

Conclusion

We found most impact studies either used inappropriate study design, had substantial risk of bias, or poorly complied with reporting guidelines. This appears to be a common feature of complex interventions. Users of CPRs should critically evaluate evidence showing the effectiveness of CPRs.

Keywords

Clinical prediction rule, cardiovascular disease, study design, risk of bias, reporting guideline

Running title

A systematic review of impact studies

Word count

2995 words, excluding abstract, tables, figures, and references.

1. Introduction

Many clinical prediction rules (CPRs) for cardiovascular conditions exist [1, 2]. They are more commonly recommended by guidelines and used in practice than CPRs in other clinical areas [3, 4]. These CPRs may undergo derivation, external validation, and impact study where distinctive evidence about the CPR is generated [5, 6]. In a derivation study, a CPR is constructed using predictors for an outcome, and external validation studies assess its accuracy in different populations and settings. Because accurate and generalizable CPRs do not always improve care, conducting impact studies to evaluate CPRs' effectiveness is a crucial last step before they are implemented in practice.

Research waste in healthcare arises when research findings cannot contribute to decision making or future research [7] due to research questions irrelevant to users, inappropriate study design and methods, incomplete publication of results, or unclear reporting [8]. Recently, studies have demonstrated some of the research waste in the CPR development. For example, researchers often do not justify a new CPR by citing existing CPRs [9] leading to many similar redundant CPRs [10, 11]. For most of these CPRs, a timely external validation study by independent researchers is infrequently conducted [12] and furthermore, an impact study is rarely carried out [1, 13]. These redundant CPRs with unknown generalizability or effectiveness represent a key source of waste in CPR research.

Research waste can also arise from poor design, method or reporting. We searched Medline and found 15 systematic reviews that evaluated design, methods, or reporting of CPR studies. These systematic reviews showed that flaws in design, methods, and reporting were common among derivation studies [14-24] and validation studies [23-26]. However, our search did not identify any systematic review that assessed design, methods, or reporting of impact studies.

The primary aim of our systematic review was to assess whether impact studies of cardiovascular CPRs used appropriate study design. Additionally, we aimed to summarize the quality of methods and reporting for impact studies with appropriate study design. Lastly, we compared the quality of methods and reporting between impact studies with appropriate study design and matched control studies evaluating other types of nonpharmacologic intervention.

2. Methods

2. 1. Information source and search

We evaluated CPRs whose derivation, validation, or impact study was included in the cardiovascular domain of the International Register of Clinical Prediction Rules for Primary Care [1]. We identified impact studies of these cardiovascular CPRs by conducting forward citation searches of their derivation studies in Scopus [27]. Some CPRs were introduced by guidelines and never published in

a journal. For these CPRs, we searched for their impact studies in Medline (Ovid) and Embase (Ovid) using strategies provided in Appendix A. All searches were completed on 3 September 2018 without language restrictions. Searches were limited to publication years from 2013 because the extension of the Consolidated Standards of Reporting Trials (CONSORT) Statement to cluster randomized trials [28], which was one of the reporting guidelines used in this study, was published in 2012.

2. 2. Eligibility criteria and study selection

A study was eligible for inclusion if: (a) the aim stated in the title, abstract, or introduction was to evaluate the impact of an intervention that included using a cardiovascular CPR or promoting the use of it, (b) the CPR was one of the cardiovascular CPRs included in the International Register, and (c) it assessed the CPR's impact on patient outcome, process of care, efficiency, or healthcare provider outcome [6, 13, 29]. A study was excluded if: (a) the CPR was not the main component of intervention, (b) it was a validation study that only estimated a hypothetical impact by reporting a predictive performance, (c) it was a secondary analysis or duplicate publication of previously published primary results, or (d) it was a review article or a modelling study.

For each eligible impact study with appropriate study design, we identified a control study matched by intervention type, study design, journal, and publication date. We searched for a control study that evaluated the effectiveness of a nonpharmacologic intervention, used an identical design as the impact study, and was published in the same issue of the journal where the impact study was published. If a control study meeting these criteria was not found in the same issue, preceding and succeeding issues were searched. The study with the closest publication date to the publication date of the impact study was selected. We excluded studies that were published before 2013, evaluated an intervention in non-human subjects, or assessed the impact of an intervention that included using a CPR or promoting the use of it.

2. 3. Assessment of study design

We assessed whether the design of impact study was appropriate using definitions consistent with the Cochrane Effective Practice and Organisation of Care (EPOC) guideline [30] and the framework for the impact analysis of clinical prediction rule [6].

2. 3. 1. Appropriate study design

An ideal study design for evaluating the impact of a CPR is randomized experiment where researchers deliberately and randomly allocate participants to interventions [6, 30-34]. It may involve randomly assigning individual participants (randomized controlled trial) or groups of individuals (cluster randomized trial) to interventions. Conducting randomized experiments may not always be feasible. Certain non-randomized experiments are described in literature as acceptable alternatives to randomized experiments [6, 13, 30, 31, 33, 34]. These include non-randomized trial, controlled before-after study, interrupted time series study, and repeated measures study. In non-randomized trials, researchers assign participants to interventions using a

nonrandom method (e.g. alternating assignment) [30]. Interrupted time series studies assess the impact of an intervention while taking preexisting trends into account by measuring an outcome in multiple time points before and after implementing the intervention [30, 35, 36]. Repeated measures studies involve multiple assessments of an outcome within each participant after different interventions are applied [30, 37, 38]. In controlled before-after studies, an outcome is assessed in a group of participants before and after receiving an intervention and is compared with the outcome assessed in control group participants who did not receive the intervention [6, 30, 35].

2. 3. 2. Inappropriate study design

Other study designs such as uncontrolled before-after study, observational study, and non-comparative study are inappropriate for evaluating the impact of a CPR [30, 32, 34]. When a study assesses an outcome in a group of participants before and after receiving an intervention without a control group, it is referred as an uncontrolled before-after study. This design is discouraged because it is difficult to understand whether an improvement in the outcome is attributable to the preexisting trends, participants' awareness of taking part in the study, or intervention [30, 33, 35, 39-41]. Furthermore, nonexperimental designs, such as observational and non-comparative studies should be avoided because of inevitable bias and confounding [13, 30, 35, 42, 43]. Meta-epidemiological studies have shown that uncontrolled before-after studies and observation studies may overestimate the effects of interventions [44-46].

2. 4. Assessment of methods and reporting

To appraise the methods of impact studies with appropriate study design, we evaluated their risk of bias. For randomized experiments, we used the revised Cochrane risk-of-bias tool for randomized trials (RoB 2) [47] and additional considerations for cluster-randomized trials [48] to assess the risk of bias. For non-randomized experiments, the Risk Of Bias In Nonrandomized Studies of Interventions (ROBINS-I) tool [49] was used. We assessed the intention to treat effect of the first reported result of the primary outcome. Risk of bias is categorized by three levels (low, some concern, high) in the Cochrane risk of bias tools and four levels (low, moderate, serious, critical) in ROBINS-I. We considered the risk of bias was substantial and the results were untrustworthy when the assessment using the Cochrane risk of bias tools was high and when the assessment using ROBINS-I was serious or critical.

We also assessed whether impact studies with appropriate design complied with reporting guidelines that were available before 2013. For randomized experiments, these were the CONSORT 2010 Statement for reporting parallel group randomized trials [50], the extension of the CONSORT Statement for randomized trials of nonpharmacologic treatments [51], and the extension of the CONSORT Statement for cluster randomized trials [28]. For non-randomized experiments, the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) statement [52] was used. Although Berger et al. [53] published a guideline for reporting non-randomized studies of treatment effects more recently in 2009, it was primarily aimed at retrospective observational studies using secondary data source rather than non-randomized experiments. We applied criteria from all relevant reporting guidelines for each study design but excluded criteria that were not

always applicable, Appendix B. We also assessed the following additional items mentioned in reporting guidelines [28, 50, 51]: a justification for the new study with a citation to a systematic review of similar studies, an interpretation of results with a systematic examination of existing evidence, a flow diagram to present participant flow, how study protocol could be accessed, and the sources of funding and funders' role. We assessed the methods and reporting of matched control studies using the same approach. All studies meeting the selection criteria were randomly assigned to MSC, TBM, or AP for risk of bias and reporting assessment. JWB independently assessed risk of bias and reporting for all included studies. Any disagreements that could not be adjudicated through discussion were resolved by third reviewer.

2. 5. Analysis

We reported the number and proportion of impact studies using appropriate and inappropriate study design. For impact studies that were conducted using appropriate study design, we presented the proportion of studies with substantial risk of bias and the mean number of reporting domains that studies complied with. The proportion of impact and control studies with substantial overall risk of bias was compared using McNemar's test for paired proportions. We used paired *t* test to compare the mean number of reporting domains that impact and control studies complied with.

3. Results

The search and selection of impact studies are summarized in Figure 1. We found 110 impact studies of cardiovascular CPRs after screening 42,769 references and reviewing 788 full-text articles. Cardiovascular CPRs most commonly evaluated in impact studies were Framingham general cardiovascular risk score by D'Agostino [54] (20 studies), CHA₂DS₂-VASc and CHADS₂ score [55, 56] (13 and 10 studies), HAS-BLED score [57] (9 studies), and Wells rule for pulmonary embolism [58-60] (11 studies), Appendix C.

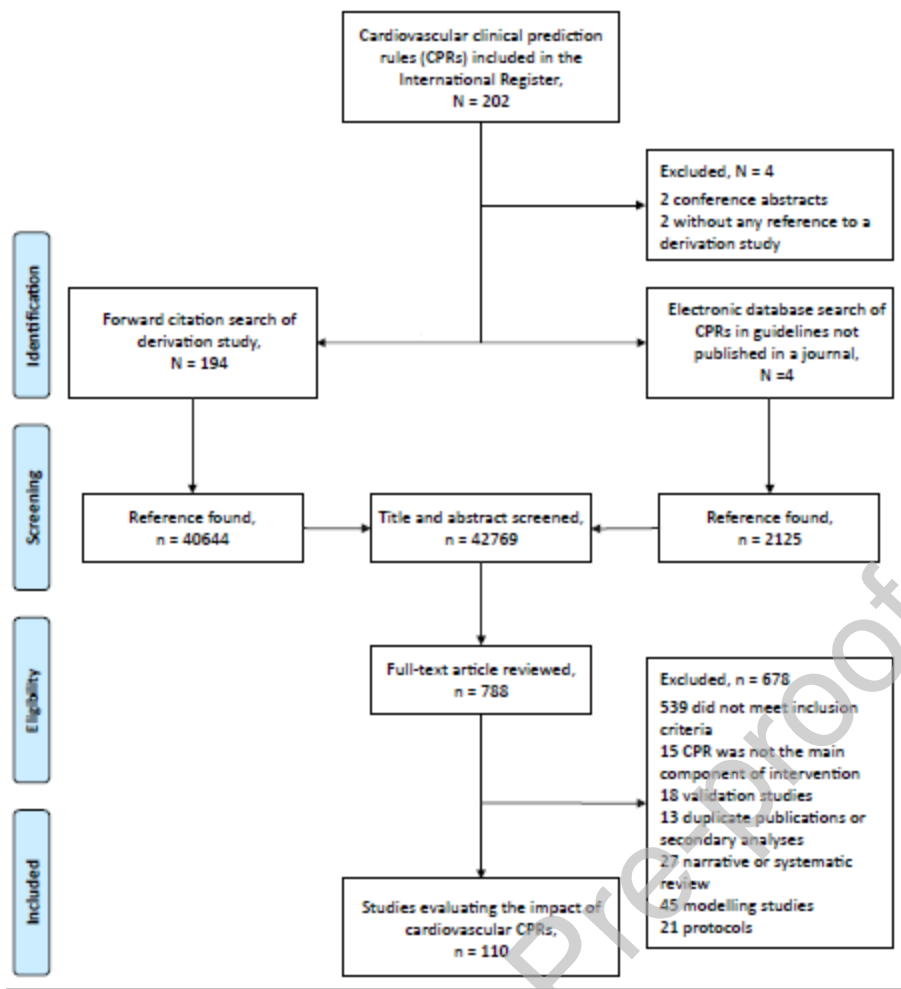


Figure 1. A PRISMA flow diagram summarizing the search and selection of impact studies of cardiovascular CPRs (N = number of cardiovascular CPR, n = number of reference).

3.1. Design of impact study

Of 110 impact studies included, 45 (40.9%) used appropriate study design and 65 (59.1%) used inappropriate study design. Characteristics of these impact studies are summarized in Table 1. Statistically significant results were reported in 71 of 90 (78.9%) impact studies that presented a primary outcome with a statistical hypothesis test using a p value or confidence interval (CI): 25 of 35 (71.4%) randomized experiments, 7 of 9 (77.8%) non-randomized experiments, and 39 of 46 (84.8%) studies with inappropriate design.

Table 1. Characteristics of impact studies of cardiovascular clinical prediction rules with appropriate and inappropriate study design, n (%) unless indicated otherwise.

3.2. Methods and reporting of impact study

The overall risk of bias was substantial in 31 of 45 (68.9%) impact studies with appropriate study design: 8 of 12 (66.7%) cluster randomized trials, 14 of 23 (60.9%) randomized controlled trials, and

10 of 10 (100.0%) non-randomized experiments, Table 2. The mean number of reporting domains that impact studies with appropriate study design complied with was 10.2 of 21 domains (95% CI, 9.3 and 11.1): 8.7 (95% CI, 7.0 and 10.4) for cluster randomized trials, 11.4 (95% CI, 10.1 and 12.7) for randomized controlled trials, and 9.3 (95% CI, 8.4 and 10.2) for non-randomized experiments.

Table 2. Impact studies with appropriate study design but substantial risk of bias, n (%).

There were 26 (57.8%) impact studies that complied with ten or less reporting domains, Figure 2. Only 24 (53.3%) impact studies provided a justification for the new study with a citation to a systematic review of similar studies (or noted that such study was absent) and 17 (37.8%) impact studies interpreted results with a systematic examination of existing evidence. Participant flow was presented using a flow diagram in 33 (73.3%) impact studies, how study protocol can be accessed was explained in 17 (37.8%) impact studies, and the sources of funding and funders' role were described in 29 (64.4%) impact studies.

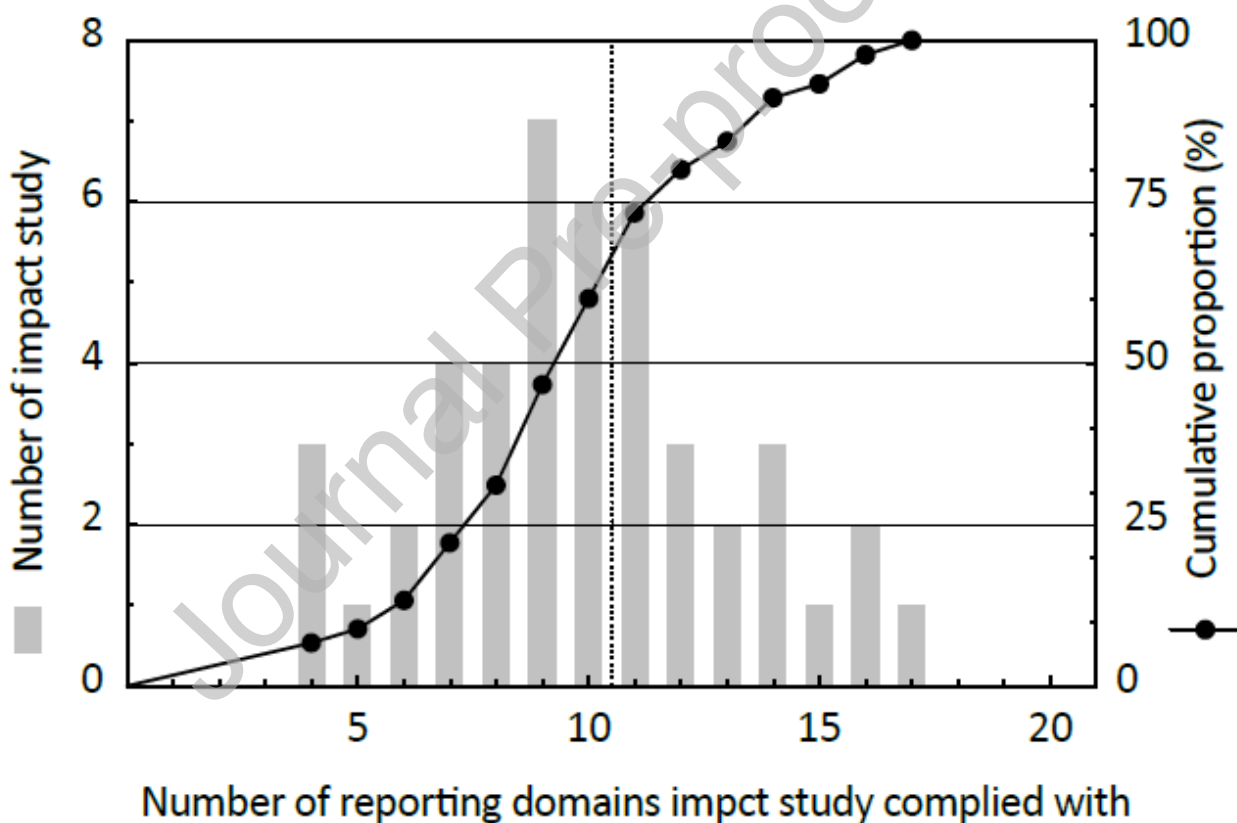


Figure 2. Reporting quality of impact studies of cardiovascular clinical prediction rules with appropriate study design.

Of 110 impact studies of cardiovascular CPRs, only six (5.5%) were conducted using appropriate study design, avoided substantial risk of bias, and complied with more than half of domains in reporting guidelines.

3.3. Comparison with matched control study

For 40 of 45 impact studies with appropriate study design, a matched control study was found. For one cluster randomized trials and four repeated measures studies, no matched control study was found. The median interval between the publication of an impact study and the matched control study was 0 month (Interquartile range, -3 and 1). In a matched analysis, there were 27 (67.5%) impact studies and 28 (70.0%) control studies with substantial risk of bias (difference = -2.5%, 95% CI, -27.4 and 22.4, $p = 0.83$). Similarly, the mean number of reporting domains that impact and control studies complied with was not clearly different (difference = -0.2, 95% CI, -1.3 and 1.0, $p = 0.79$), Figure 3. The full results of risk of bias and reporting assessment provided in Appendix D and E.

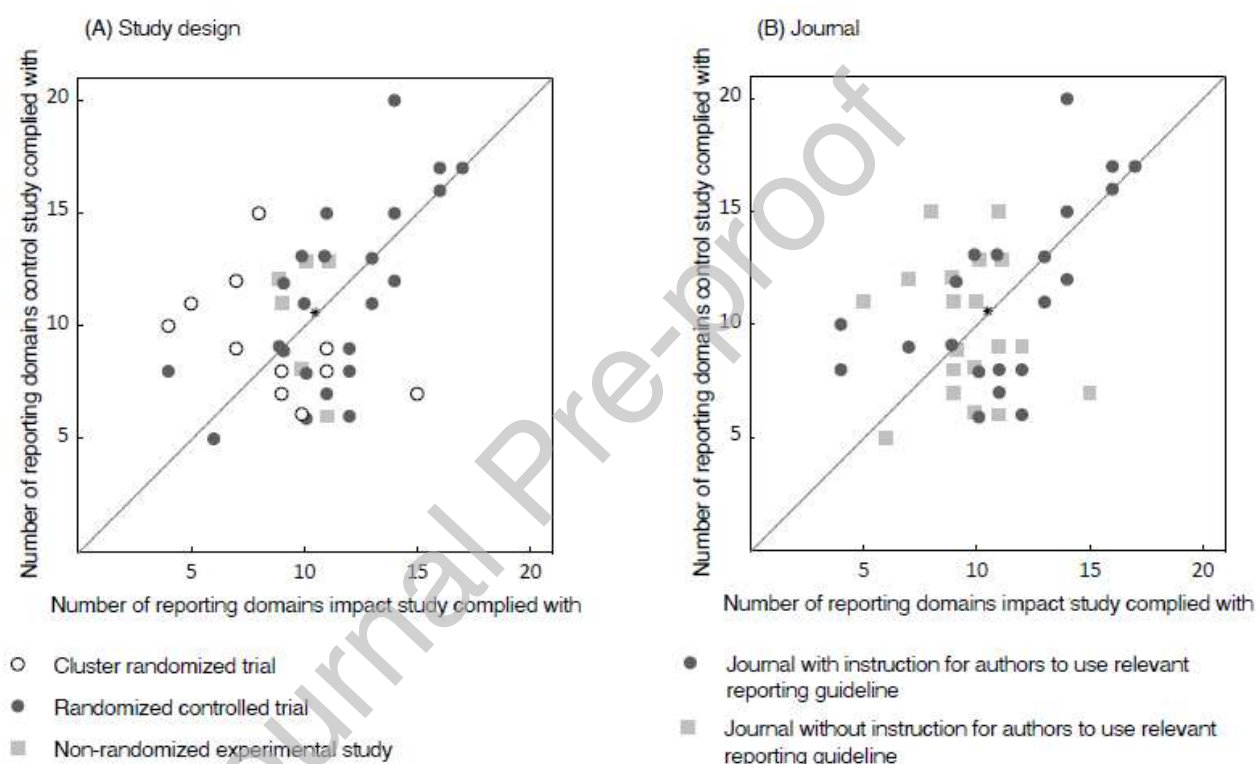


Figure 3. Number of reporting domains that impact and control studies complied with according to (A) study design and (B) Journal where studies were published. The star (★) represents the mean number of reporting domains impact and control studies complied with, 10.5 and 10.6 (difference = -0.2, 95% confidence interval, -1.3 and 1.0, $p = 0.79$).

4. Discussion

We presented the first study that systematically summarized design, methods, and reporting of impact studies of CPRs. We found the majority of impact studies of cardiovascular CPRs included in our systematic review were conducted using study designs inappropriate for evaluating the effectiveness of a CPR. Uncontrolled before-after study was the most frequently used study design. The majority of impact studies that conducted a hypothesis test for primary outcome reported

statistically significant results. When impact studies were conducted using appropriate study design, most used suboptimal methods leading to substantial risk of bias in results or poorly complied with reporting guidelines. The quality of methods and reporting was not clearly different between impact studies of cardiovascular CPRs and studies evaluating other types of nonpharmacologic interventions.

Along with existing systematic reviews that showed poor design, methods, and reporting were prevalent in derivation and validation studies of CPRs [14-26], our findings highlight persistent and worrisome trends throughout CPR development.

A recently published systematic review found cohort study was the most common study design for nonrandomized studies of interventions funded by UK National Institute for Health Research Biomedical Research Centres [61]. In contrast, we found uncontrolled before-after study was the most frequently used study design among impact studies of cardiovascular CPRs. Researchers conducting impact studies which are typically aimed at assessing the effectiveness of CPRs in a group of participants rather than individuals might have favored uncontrolled before- after study due to its simplicity. The same systematic review found the risk of bias assessed using ROBINS-I tool was either serious or critical in all nonrandomized trials [61] which is consistent with our findings.

The compliance to reporting guidelines of impact studies was worse than the ones described in recently conducted systematic reviews of other types of intervention [62, 63]. However, these reviews assessed only the presence of reporting using criteria from reporting guidelines. On the other hand, the compliance rate of impact studies in our study to reporting guidelines was comparable to a study that evaluated the completeness of reporting for domains from all relevant extensions in addition to the CONSORT statement [65].

QRISK scores were used by almost all participants who took part in a recent survey of general practitioners (GPs) in the UK [4]. However, we did not find any impact study with appropriate study design that demonstrated the effectiveness of QRISK scores. This might suggest GPs' use of QRISK scores was influenced by other factors such as recommendations from guideline or integration of CPRs in electronic health records [4], rather than evidence. We found the most impact studies included in the review reported the results favoring the use of cardiovascular CPRs. This is consistent with previous systematic reviews of impact studies that showed the effectiveness of using cardiovascular CPRs [64-67], although the presence of publication bias cannot be ruled out.

A linear relationship of reporting quality between impact and matched control studies in Figure 3 suggests that a publication in a journal might be a surrogate for reporting quality. Systematic reviews have shown that journal endorsement of reporting guidelines [62, 68-70] and impact factor [71, 72] are associated with reporting quality which might be the mechanisms for this relationship.

4. 1. limitations

We only evaluated impact studies of cardiovascular CPRs included in the International Register of CPRs for Primary Care that were published since 2013. Therefore, findings from our study may not be generalizable to impact studies of CPRs for other clinical areas or impact studies published before 2013. However, because the results of this study suggest the suboptimal methods and reporting might be a universal feature in impact studies of all nonpharmacologic complex interventions including CPRs, impact studies of CPRs in other clinical areas could also have similar weaknesses. Although we assessed impact studies of cardiovascular CPRs, we were not able to restrict control studies to ones that evaluated nonpharmacologic interventions for cardiovascular disease. In addition, constraints in time and resource precluded us from including more than one control study for each impact study.

We used criteria from reporting guidelines to assess the quality of reporting. These guidelines are intended to help authors prepare clear research reports but not necessarily meant to be used as tools for assessing the reporting quality. Although there are other studies that assessed reporting quality using reporting guidelines, it is often difficult to compare the findings due to heterogeneity in assessment methods between studies.

4. 2. Implication on practice and research

Clinicians should beware that high-quality impact studies of cardiovascular CPRs are uncommon. They should critically evaluate evidence showing a CPR could benefit patients, reduce cost, or improve efficiency and understand the uncertainties when making decisions using the CPR. As the development and use of CPRs accelerate, more high-quality impact studies that could inform users are needed. Researchers should prioritize evaluating CPRs that have been already implemented without clear evidence of impact. Journals and funders could play a crucial role in improving the value of CPR research by motivating researchers.

Authors statement

Jong-Wook Ban: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, writing - review and editing, Project administration. Mei Sum Chan: Methodology, Investigation, Writing - review and editing. Tonny Brian Muthee: Investigation, Writing - review and editing. Arsenio Paez: Investigation, Writing - review and editing. Richard Stevens: Conceptualization, Methodology, Writing - review and editing, Supervision. Rafael Perera: Conceptualization, Methodology, Writing - review and editing, Supervision.

All authors made substantial contributions to all of the following: (1) the conception and design the study, or acquisition of data, or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, and (3) final approval of the version to be submitted.

Declaration of interest

Rafael Perera receives funding from the NIHR Oxford Biomedical Research Council (BRC), the NIHR Oxford Medtech and In-Vitro Diagnostics Co-operative (MIC), and the Oxford Martin School. Richard Stevens and Rafael Perera receive funding from the NIHR Applied Research Collaboration (ARC) Oxford and Thames Valley.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Rafael Perera receives funding from the NIHR Oxford Biomedical Research Council (BRC), the NIHR Oxford Medtech and In-Vitro Diagnostics Co-operative (MIC), and the Oxford Martin School. Richard Stevens and Rafael Perera receive funding from the NIHR Applied Research Collaboration (ARC) Oxford and Thames Valley.

Acknowledgement

Authors are grateful to Dr Emma Wallace and Professor Tom Fahey of Royal College of Surgeons in Ireland for granting access to the current version of the International Register of Clinical Prediction Rules for Primary Care.

Supporting information

Appendix A: Electronic database search strategies for impact studies of cardiovascular clinical prediction rules included in guidelines.

Appendix B: Criteria for reporting assessment.

Table B-1. Criteria for assessing the report of cluster randomized trials.

Table B-2. Criteria for assessing the report of randomized controlled trials.

Table B-3. Criteria for assessing the report of non-randomized trials.

Appendix C: Cardiovascular clinical prediction rules evaluated in impact studies.

Appendix D: Risk of bias assessment for 40 impact studies and matched control studies.

Figure D-1. Risk of bias assessment for cluster randomized trials.

Figure D-2. Risk of bias assessment for randomized controlled trials.

Figure D-3. Risk of bias assessment for non-randomized trials.

Appendix E: Assessment of reporting for 40 impact studies and matched control studies.

Table E-1. Cluster randomized trials that complied with reporting criteria, n (%)

Table E-2. Randomized controlled trials that complied with reporting criteria, n (%)

Table E-3. Non-randomized experiments that complied with reporting criteria, n (%)

Journal Pre-proof

References

- [1] Keogh C, Wallace E, O'Brien KK, Galvin R, Smith SM, Lewis C, et al. Developing an international register of clinical prediction rules for use in primary care: a descriptive analysis. *Annals of family medicine*. 2014;12:359-66.
- [2] Wessler BS, Lai Yh L, Kramer W, Cangelosi M, Raman G, Lutz JS, et al. Clinical Prediction Models for Cardiovascular Disease: Tufts Predictive Analytics and Comparative Effectiveness Clinical Prediction Model Database. *Circ Cardiovasc Qual Outcomes*. 2015;8:368-75.
- [3] Plüddemann A, Wallace E, Bankhead C, Keogh C, Van der Windt D, Lasserson D, et al. Clinical prediction rules in practice: review of clinical guidelines and survey of GPs. *Br J Gen Pract*. 2014;64:e233-42.
- [4] Ban J-W, Perera R, Stevens R. Familiarity and use of cardiovascular clinical prediction rules: survey of general practitioners. University of Oxford; 2020.
- [5] Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Annals of internal medicine*. 2006;144:201-9.
- [6] Wallace E, Smith SM, Perera-Salazar R, Vaucher P, McCowan C, Collins G, et al. Framework for the impact analysis and implementation of Clinical Prediction Rules (CPRs). *BMC medical informatics and decision making*. 2011;11:62.
- [7] Ioannidis JP. Why Most Clinical Research Is Not Useful. *PLoS medicine*. 2016;13:e1002049.
- [8] Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet*. 2009;374:86-9.

- [9] Ban J-W, Wallace E, Stevens R, Perera R. Why do authors derive new cardiovascular clinical prediction rules in the presence of existing rules? A mixed methods study. *PLoS One*. 2017;12:e0179102.
- [10] Rahimi K, Bennett D, Conrad N, Williams TM, Basu J, Dwight J, et al. Risk prediction in patients with heart failure: a systematic review and analysis. *JACC Heart Fail*. 2014;2:440-6.
- [11] Damen JA, Hooft L, Schuit E, Debray TP, Collins GS, Tzoulaki I, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ*. 2016;353:i2416.
- [12] Ban J-W, Stevens R, Perera R. Predictors for independent external validation of cardiovascular risk clinical prediction rules: Cox proportional hazards regression analyses. *Diagnostic and Prognostic Research*. 2018;2:3.
- [13] Wallace E, Uijen MJ, Clyne B, Zarabzadeh A, Keogh C, Galvin R, et al. Impact analysis studies of clinical prediction rules relevant to primary care: a systematic review. *BMJ Open*. 2016;6:e009957.
- [14] Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Med*. 2011;9:103.
- [15] Collins GS, Omar O, Shanyinde M, Yu LM. A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods. *Journal of clinical epidemiology*. 2013;66:268-77.
- [16] Mallett S, Royston P, Dutton S, Waters R, Altman DG. Reporting methods in studies developing prognostic models in cancer: a review. *BMC Med*. 2010;8:20.
- [17] Mallett S, Royston P, Waters R, Dutton S, Altman DG. Reporting performance of prognostic models in cancer: a review. *BMC Med*. 2010;8:21.

- [18] Mushkudiani NA, Hukkelhoven CW, Hernandez AV, Murray GD, Choi SC, Maas AI, et al. A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. *Journal of clinical epidemiology*. 2008;61:331-43.
- [19] Sahle BW, Owen AJ, Chin KL, Reid CM. Risk Prediction Models for Incident Heart Failure: A Systematic Review of Methodology and Model Performance. *J Card Fail*. 2017;23:680-7.
- [20] Wen Z, Guo Y, Xu B, Xiao K, Peng T, Peng M. Developing Risk Prediction Models for Postoperative Pancreatic Fistula: a Systematic Review of Methodology and Reporting Quality. *Indian J Surg*. 2016;78:136-43.
- [21] Whittle R, Peat G, Belcher J, Collins GS, Riley RD. Measurement error and timing of predictor values for multivariable risk prediction models are poorly reported. *Journal of clinical epidemiology*. 2018;102:38-49.
- [22] Altman DG. Prognostic models: a methodological framework and review of models for breast cancer. *Cancer Invest*. 2009;27:235-43.
- [23] Heus P, Damen J, Pajouheshnia R, Scholten R, Reitsma JB, Collins GS, et al. Poor reporting of multivariable prediction model studies: towards a targeted implementation strategy of the TRIPOD statement. *BMC Med*. 2018;16:120.
- [24] Bouwmeester W, Zuithoff NP, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, et al. Reporting and methods in clinical prediction research: a systematic review. *PLoS medicine*. 2012;9:1-12.
- [25] Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol*. 2014;14:40.

- [26] Ban J-W, Ignacio Emparanza J, Urreta I, Burls A. Design Characteristics Influence Performance of Clinical Prediction Rules in Validation: A Meta-Epidemiological Study. *Plos One*. 2016;11.
- [27] Scopus. Elsevier. p. Web site.
- [28] Campbell MK, Piaggio G, Elbourne DR, Altman DG, Group C. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012;345:e5661.
- [29] Cochrane Effective Practice and Organisation of Care (EPOC). What outcomes should be reported in EPOC reviews? EPOC Resources for review authors. 2017.
- [30] Cochrane Effective Practice and Organisation of Care (EPOC). What study designs should be included in an EPOC review? EPOC Resources for review authors. 2017.
- [31] McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA : the journal of the American Medical Association*. 2000;284:79-84.
- [32] Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ*. 2009;338:b606.
- [33] Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart*. 2012;98:691-8.
- [34] Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. *Diagn Progn Res*. 2019;3:16.
- [35] Grimshaw J, Campbell M, Eccles M, Steen N. Experimental and quasi-experimental designs for evaluating guideline implementation strategies. *Fam Pract*. 2000;17 Suppl 1:S11-6.

- [36] Kontopantelis E, Doran T, Springate DA, Buchan I, Reeves D. Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. *BMJ*. 2015;350:h2750.
- [37] Sullivan LM. Repeated measures. *Circulation*. 2008;117:1238-43.
- [38] Kraska M. Repeated Measures Design. In: Salkind NJ, editor. *Encyclopedia of Research Design*. Thousand Oaks: SAGE Publications, Inc.; 2010. p. 1243-7.
- [39] Sedgwick P. Before and after study designs. *BMJ*. 2014;349:g5074.
- [40] Hendriksen JM, Geersing GJ, Moons KG, de Groot JA. Diagnostic and prognostic prediction models. *Journal of thrombosis and haemostasis : JTH*. 2013;11 Suppl 1:129-41.
- [41] Ho AMH, Phelan R, Mizubuti GB, Murdoch JAC, Wickett S, Ho AK, et al. Bias in Before-After Studies: Narrative Overview for Anesthesiologists. *Anesth Analg*. 2018;126:1755-62.
- [42] Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet*. 2002;359:248-52.
- [43] Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. *Lancet*. 2002;359:145-9.
- [44] Lipsey MW, Wilson DB. The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *Am Psychol*. 1993;48:1181-209.
- [45] Wilson DB, Lipsey MW. The role of method in treatment effectiveness research: evidence from meta-analysis. *Psychological methods*. 2001;6:413.
- [46] Shikata S, Nakayama T, Noguchi Y, Taji Y, Yamagishi H. Comparison of effects in randomized controlled trials with observational studies in digestive surgery. *Ann Surg*. 2006;244:668-76.

- [47] Higgins JPT, Savović J, Page MJ, Sterne JAC. Revised Cochrane risk-of-bias tool for randomized trials (RoB 2). In: Chandler J, Clarke M, McKenzie J, Boutron I, Welch V, editors. *Cochrane Methods: Cochrane Database of Systematic Reviews*; 2018.
- [48] Eldridge S, Campbell M, Campbell M, Dahota A, Giraudeau B, Higgins J, et al. Revised Cochrane risk of bias tool for randomized trials (RoB 2.0): Additional considerations for cluster-randomized trials. In: Chandler J, Clarke M, McKenzie J, Boutron I, Welch V, editors. *Cochrane Methods: Cochrane Database of Systematic Reviews*; 2016. p. 10 (Suppl 1).
- [49] Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919.
- [50] Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332.
- [51] Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, Group C. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of internal medicine*. 2008;148:295-309.
- [52] Des Jarlais DC, Lyles C, Crepaz N, Group T. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health*. 2004;94:361-6.
- [53] Berger ML, Mamdani M, Atkins D, Johnson ML. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report--Part I. *Value Health*. 2009;12:1044-52.

[54] D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation*. 2008;117:743-53.

[55] Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest*. 2010;137:263-72.

[56] Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. *JAMA : the journal of the American Medical Association*. 2001;285:2864-70.

[57] Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. *Chest*. 2010;138:1093-100.

[58] Wells PS, Ginsberg JS, Anderson DR, Kearon C, Gent M, Turpie AG, et al. Use of a clinical model for safe management of patients with suspected pulmonary embolism. *Annals of internal medicine*. 1998;129:997-1005.

[59] Wells PS, Anderson DR, Rodger M, Ginsberg JS, Kearon C, Gent M, et al. Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: increasing the models utility with the SimpliRED D-dimer. *Thromb Haemost*. 2000;83:416-20.

[60] Gibson NS, Sohne M, Kruip MJ, Tick LW, Gerdes VE, Bossuyt PM, et al. Further validation and simplification of the Wells clinical decision rule in pulmonary embolism. *Thromb Haemost*. 2008;99:229-34.

- [61] Dhiman P, Lee H, Kirtley S, Collins GS. A systematic review showed more consideration is needed when conducting nonrandomized studies of interventions. *Journal of clinical epidemiology*. 2020;117:99-108.
- [62] Wilson B, Burnett P, Moher D, Altman D, Salman RAS. Completeness of reporting of randomised controlled trials including people with transient ischaemic attack or stroke: A systematic review. *European Stroke Journal*. 2018;3.
- [63] Alamri HM, Alharbi F. Quality Assessment of Randomized Clinical Trials Reporting in Endodontic Journals: An Observational Study from 2012 to 2017. *J Endod*. 2018;44:1246-50.
- [64] Collins DR, Tompson AC, Onakpoya IJ, Roberts N, Ward AM, Heneghan CJ. Global cardiovascular risk assessment in the primary prevention of cardiovascular disease in adults: systematic review of systematic reviews. *BMJ Open*. 2017;7:e013650.
- [65] Sheridan SL, Crespo E. Does the routine use of global coronary heart disease risk scores translate into clinical benefits or harms? A systematic review of the literature. *BMC Health Serv Res*. 2008;8:60.
- [66] Sheridan SL, Viera AJ, Krantz MJ, Ice CL, Steinman LE, Peters KE, et al. The effect of giving global coronary risk information to adults: a systematic review. *Arch Intern Med*. 2010;170:230-9.
- [67] Usher-Smith JA, Silarova B, Schuit E, Moons KG, Griffin SJ. Impact of provision of cardiovascular disease risk estimates to healthcare professionals and patients: a systematic review. *BMJ Open*. 2015;5:e008717.
- [68] Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev*. 2012;11:MR000030.

- [69] Stevens A, Shamseer L, Weinstein E, Yazdi F, Turner L, Thielman J, et al. Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: systematic review. *BMJ*. 2014;348:g3804.
- [70] Sarkis-Onofre R, Poletto-Neto V, Cenci MS, Moher D, Pereira-Cenci T. CONSORT endorsement improves the quality of reports of randomized clinical trials in dentistry. *Journal of clinical epidemiology*.
- [71] Arra I, Velker V, Sexton T, Rotenberg BW, Boldt RG, Rodrigues G. A CONSORT Clinical Trial Reporting Compliance Audit of the Oncology Randomized Controlled Trial Literature. *Cureus*. 2013.
- [72] Dechartres A, Trinquart L, Atal I, Moher D, Dickersin K, Boutron I, et al. Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study. *BMJ*. 2017;357:j2490.

Table 1. Characteristics of impact studies of cardiovascular clinical prediction rules with appropriate and inappropriate study design, n (%) unless indicated otherwise.

Appropriate study design, n = 45				
Characteristic	Randomized experiment, n = 35	Non-randomized experiment, n = 10	Inappropriate study design, n = 65	
	<ul style="list-style-type: none"> • 12 cluster randomized trials • 23 randomized controlled trials 	<ul style="list-style-type: none"> • 2 non-randomized trials • 4 interrupted time series studies • 4 repeated measures studies 	<ul style="list-style-type: none"> • 40 uncontrolled before-after studies • 7 cohort studies • 18 non-comparative studies 	Total, n = 110
Publication date	Median, IQR September 2015, April 2014 – December 2016	November 2016, July 2015 – February 2018	August 2016, June 2014 – September 2017	March 2016, May 2014 – August 2017
Journal impact factor	Median, IQR 4.5, 2.7 – 11.7	2.1, 1.3 – 3.3	1.8, 0.8 – 4.0	2.6, 1.3 – 4.5
Location	North 13	4 (40.0)	26 (40.0)	43 (39.1)

America	(37.1)			
Europe	12 (34.3)	4 (40.0)	20 (30.8)	36 (32.7)
Oceania	7 (20.0)	1 (10.0)	7 (10.8)	15 (13.6)
Asia	2 (5.7)	1 (10.0)	7 (10.8)	10 (9.1)
Central and South America	0 (0.0)	0 (0.0)	3 (4.6)	3 (2.7)
Africa	1 (2.9)	0 (0.0)	1 (1.5)	2 (1.8)
International	0 (0.0)	0 (0.0)	1 (1.5)	1 (0.9)
Setting ^a				
Outpatient	17 (48.6)	7 (70.0)	21 (32.3)	45 (40.9)
ED	6 (17.1)	1 (10.0)	19 (29.2)	26 (23.6)
Inpatient	1 (2.9)	1 (10.0)	8 (12.3)	10 (9.1)
Community	6 (17.1)	0 (0.0)	4 (6.1)	10 (9.1)
Pharmacy	1 (2.9)	0 (0.0)	4 (6.1)	5 (4.5)
Multiple	3 (8.6)	0 (0.0)	4 (6.1)	7 (6.4)
Other	1 (2.9)	1 (10.0)	5 (7.7)	7 (6.4)
Sample size ^b				

Median, IQR	570, 240 – 1723	373, 144 – 832	353.5, 200 – 1556	499.5, 208.5 – 1611.5
Primary outcome				
Statistical test	35 (100.0)	9 (90.0)	46 (70.8)	90 (81.8)
No statistical test	0 (0.0)	1 (10.0)	19 (29.2)	20 (18.2)

^a Other settings were emergency medical technician service, health insurance, medical conference, military, and online.

^b One interrupted time series study and one uncontrolled before-and-after study did not report sample size.

Table 2. Impact studies with appropriate study design but substantial risk of bias, n (%).

Stage of study	Risk of bias domain	Randomized experiment		Non-ran experime
		Cluster randomized	Randomized controlled	
		trial ^b , n = 12	trial ^a , n = 23	
Pre-intervention or at-intervention	1. Randomization process	0 (0.0)	0 (0.0)	10
	2. Timing of identification and recruitment of individual participants	4 (33.3)	N/A	
	3. Confounding	N/A	N/A	
	4. Selection of participants	N/A	N/A	
	5. Classification of intervention	N/A	N/A	

<i>Stage</i>		4 (33.3)	0 (0.0)	10
Post-intervention	6. Deviation from the intended intervention	2 (16.7)	9 (39.1)	
	7. Missing data	1 (8.3)	4 (17.4)	
	8. Measurement of outcome	1 (8.3)	4 (17.4)	
	9. Selection of reported result	2 (16.7)	2 (8.7)	
	<i>Stage</i>	6 (50.0)	14 (60.9)	
Overall risk of bias		8 (66.7)	14 (60.9)	10

^a Studies with high risk of bias assessed by the revised Cochrane risk-of-bias tool for randomized trials (RoB 2.0).

^b Studies with high risk of bias assessed by additional considerations for cluster-randomized trials in the revised Cochrane risk-of-bias tool for randomized trials (RoB 2.0).

^c Studies with serious or critical risk of bias assessed by the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool.

^d No matching control study was found for one impact study with cluster randomized trial design and four impact studies.