BMJ Open

Conducting observational analyses with the target trial emulation approach: a methodological systematic review

Noémie Simon-Tillaux ,[1,2] Guillaume L Martin,[2] David Hajage,[2] Carole Scheifer,[2] Maxime Beydon,[2] Agnes Dechartres,[2] Florence Tubach[2]

To cite: Simon-Tillaux N, Martin GL, Hajage D, *et al*. Conducting observational analyses with the target trial emulation approach: a methodological systematic review. *BMJ Open* 2024;**14**:e086595. doi:10.1136/bmjopen-2024-086595

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2024-086595).

AD and FT contributed equally.

Received 19 March 2024
Accepted 08 October 2024

[1]Office of Biostatistics and Epidemiology, Gustave Roussy, Oncostat U1018, Inserm, University Paris-Saclay, labeled Ligue Contre le Cancer, Villejuif, France
[2]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Equipe PEPITES, AP-HP, Hôpital Pitié Salpêtrière, Département de Santé Publique, Paris, France

Correspondence to
Dr Noémie Simon-Tillaux; simon.noemie@ymail.com

ABSTRACT

Objectives Target trial emulation is an approach that is increasingly used to improve transparency in observational studies and help mitigate biases. For studies declaring that they emulated a target trial, we aimed to evaluate the specification of the target trial, examine its consistency with the observational emulation and assess the risk of bias in the observational analysis.

Design Methodological systematic review reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.

Data sources The database MEDLINE (Medical Literature Analysis and Retrieval System Online) was interrogated for all studies published from 1 January 2021 to 3 July 2022. We performed an additional manual search of 20 general medical and specialised journals that spanned the same period.

Eligibility criteria All studies that declared emulating a hypothetical or real randomised trial were eligible.

Data extraction and synthesis Two independent reviewers performed the whole systematic review process (screening and selection of studies, data extraction and risk of bias assessment). The main outcomes were the definition of the key protocol components of the target trial and its emulation, consistency between the target trial and its emulation and risk of bias according to the ROBINS-I (Risk Of Bias In Non-randomised Studies - of Interventions) tool.

Results Among the selected sample of 100 studies, 24 (24%) did not specify the target trial. Only 40 studies (40%) provided detailed information on all components of the target trial protocol. Eligibility criteria, intervention strategies and outcomes were consistent between the target trial and its emulation in 35 studies (46% of those specifying the target trial). Overall, 28 studies (28%) exhibited serious risk of bias and 41 (41%) had misalignments in the timing of eligibility assessment, treatment assignment and the start of follow-up (time-zero). As compared with studies that did not specify the target trial, those that did specify the trial less frequently seemed to have both time-zero issues (39% vs 52%) and serious risk of bias (26% vs 33%).

Conclusions One-quarter of studies declaring that they emulated a target trial did not specify the trial. Target trials and their emulations were particularly inconsistent for studies emulating a real randomised trial. Risk of methodological issues seemed lower in observational analyses that specified versus did not specify the target trial.

STRENGHS AND LIMITATIONS OF THIS STUDY

⇒ The reporting of this study complies with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.
⇒ The risk of bias in observational analyses was assessed with ROBINS-I, a dedicated tool for such studies.
⇒ The search strategy considered only studies published between 2021 and 2022, limiting the generalisability of the results and preventing from studying any time trend.
⇒ The evaluation of the consistency between the target trial and its emulation was limited to major differences.

INTRODUCTION

Observational studies often face criticisms in their ability to estimate the causal effect of an intervention as compared with randomised experiments.[1] Nevertheless, randomised trials may not always be feasible due to ethical considerations, the complexity of interventions that need to be assessed, and the substantial resources needed for measuring outcomes over a prolonged follow-up period or including a sufficient number of patients to detect a small yet significant intervention effect. They also take time to yield results, which may pose a challenge when the assessment of intervention is time-sensitive, such as during the urgency of the COVID-19 pandemic. Therefore, observational studies are the preferred option for certain research questions. Target trial emulation has been proposed as a structured methodological framework that would improve confidence in observational analyses by making more transparent the causal question at hand and helping to mitigate biases.[2–5] This method relies on defining key components of the protocol of the target randomised trial that would have been conducted to answer the causal question of interest and emulating each of the components using observational

data. Among the key components, the definition of eligibility criteria, intervention strategies to be compared, outcomes and the causal contrast may help accurately identify the research question and its associated causal effect.[4 6 7] These components also encompass the intervention assignment and start of follow-up, which are critical because their timing should align with when eligibility criteria are met to avoid time-zero-related biases.[6–10]

The number of published studies declaring to emulate a real or hypothetical randomised experiment as an objective or in their methods has been increasing over the past decade, especially after the publication of a paper that provided a comprehensive description of the approach.[5] While the approach may seem increasingly popular, there is a need to assess whether the method is being correctly applied and reported and to assess what is the methodological quality of these observational analyses. Therefore, in this methodological systematic review, for observational studies stating that they emulated a randomised trial, we aimed to evaluate the specification of the target trial, examine its consistency with the observational emulation and assess the risk of bias in the observational analysis.

METHODS

This is a methodological review complying with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for reporting[11] (checklists are provided in the online supplemental table S1 and S2). A protocol was written before starting the review and is available in the supplemental material (see online supplemental appendix 1).

Search strategy and study selection

We searched MEDLINE with the Embase search engine for studies declaring that they emulated a real randomised trial with observational data. We chose Embase over the PubMed search engine because only Embase search equations can include textual distance. The search algorithm is provided in online supplemental appendix 2 in the supplemental material. To obtain a comprehensive overview, we also interrogated search engines with the keywords “emulate”, “target trial” and “trial emulation” in the following general medical journals: *Journal of the American Medical Association* (*JAMA*), *New England Journal of Medicine* (*NEJM*), *the BMJ*, *The Lancet*, *Annals of Internal Medicine*, *JAMA Internal Medicine*, *Nature* and *Nature Medicine*. We also searched for additional references in specialised medical journals with the highest impact factor in their specialty in 2021: *The Lancet Infectious Diseases*, *The Lancet Neurology*, *The Lancet Diabetes and Endocrinology*, *The Lancet Respiratory Medicine*, *The Lancet Global Health*, *European Heart Journal*, *Journal of Hepatology*, *Gut*, *Journal of Clinical Oncology*, *World Psychiatry*, *Annals of the Rheumatic Diseases* and *Blood*. We added this manual search because abstracts may not always mention the use of the target trial emulation approach, and search engines of journal websites can perform full-text searches of keywords. This manual search was limited to high-impact factor journals, as we expected that target trial emulations would be perceived as high-quality studies and would predominantly be published in such journals. The search was performed on 3 July 2022 and was restricted to studies published after 1 January 2021.

Any study that declared emulating a hypothetical (ie, that has never been conducted for real) randomised trial or emulating a real one (ie, the randomised experiment does exist) using observational data and comparing two or more interventions was eligible. Studies interested in a prognostic factor or interventions delivered to clusters of patients were excluded, even if designed as an emulated target trial. We additionally excluded all interventional studies (or observational analyses that used data from it), before–after studies, narrative or systematic reviews, preclinical and veterinary medicine studies, genetic studies, pharmacokinetics and/or pharmacodynamics studies, studies that aim to validate a statistical model, health economic studies and case reports. The eligibility was assessed independently by two reviewers (NST and CS or GLM) by screening titles and abstracts, then full texts whenever necessary. Any disagreement was resolved by discussion to reach a consensus.

Data extraction and risk of bias assessment

Two reviewers (NST and GLM or MB) independently extracted data for all selected studies according to a standardised data extraction form. Disagreements were resolved by consensus. Information was retrieved from the article identified by the selection process. In addition, for observational analyses with a preregistered protocol, we retrieved the protocols on websites for study registration (ClinicalTrials.gov and the European Union electronic Register of Post-Authorisation Studies). In the case of a study attempting to emulate a real randomised trial, we also retrieved information on the randomised trial most of the time from the original article and for one analysis on study registration websites (ClinicalTrials.gov, International Clinical Trials Registry Platform) because the results of the trial were not published.

We collected the following general characteristics of the study: location of the corresponding author, source of funding, reason for conducting the study (eg, to emulate an unfeasible or unethical randomised trial in real life), source of observational data (eg, claim databases) and if the target trial was a hypothetical or real randomised trial. For both the target trial and its emulation, we collected information on the following seven key protocol components: eligibility criteria, intervention strategies, intervention assignment, outcomes, follow-up, causal contrasts of interest and statistical analysis plan.[5 12] The target trial was deemed specified if at least one of the aforementioned components was explicitly reported as a protocol component of the ‘target trial’ or reported as a characteristic of the randomised experiment that was emulated in the observational analysis. Special attention

was paid to the observational analysis features from which biases could arise, including the definition of time-zero (start of follow-up) and its alignment with the time when eligibility criteria were met and the intervention was assigned as well as the complexity of the interventions being studied. An intervention was deemed complex if it could change over time based on postbaseline parameters (dynamic strategy), involved multiple interventions in a single arm (joint strategy), or incorporated a grace period, allowing the intervention to be received within a specified timeframe after the start of follow-up. We also assessed the statistical models used to infer the effect of the intervention in the observational setting and the exploration of residual confounding, with the calculation of the E-value[13] or the use of other controls (eg, negative outcome controls with expected null associations with the intervention).[14] The risk of bias in studies was assessed with a slightly modified version of the ROBINS-I tool.[15] The ROBINS-I tool is designed to evaluate the risk of bias in non-randomised studies of interventions and is the preferred tool to be used in Cochrane reviews of such studies.[16] It takes into account the concept of a target trial for evaluation of the methodology of the observational analysis and involves answering specific questions related to potential biases before, during and after the intervention. The tool requires some expert knowledge from the reviewers relative to the potential confounders and cointerventions for the given research question. As we could not provide this knowledge for all research questions, our assessment was based on the confounders and cointerventions that were listed by authors in the articles and whether the statistical analysis accounted for them properly. We added a signalling question to this tool, 'Do time points of eligibility assessment, intervention assignment and start of follow-up align for most of participants?', that overlaps the signalling questions of the selection bias domain 'Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?' and 'Do start of follow-up and start of intervention coincide for most participants?', and evaluated the biases that might arise accordingly, as reported elsewhere.[8 9]

If several target trials were emulated in the article, we collected only information regarding the first reported target trial (and the first reported emulation in case of multiple emulations). For more information on data extraction and assessment of risk of bias, see online supplemental appendix 3.

Data synthesis

Characteristics of both the target trials and observational analyses were described according to the nature of the target trial: real target trial, hypothetical target trial or no target trial specified. We then assessed the consistency between the target trial specification and its emulation for the following protocol components: eligibility criteria, intervention strategies, outcome, follow-up, causal contrasts of interest and intervention effect measure chosen in statistical analysis. The full research question (PICO: population, intervention, control and outcome) was considered consistent if the eligibility criteria, intervention strategies and outcome were consistent. See online supplemental table S3 for more information on when the target trial and its emulation were deemed consistent for each component. We compared the risk of bias among studies that specified the target trial (at least one key protocol component was documented) and those that did not, for each domain of the ROBINS-I tool.[15] Finally, when target trials were real, we assessed the consistency in results by using the full statistical significance agreement (estimates and their 95% CIs on the same side of null in the target trial and its emulation) as proposed elsewhere.[17 18]

Patient and public involvement

No patient or public participant was involved in the design, conduct, analysis or the reporting of findings of this methodological review, which only reuses published data. There are no plans to disseminate the results of the study to patients and public participants.

RESULTS

Search results

The Embase search engine and manual searches in selected journals yielded 481 and 100 records, respectively. Of these, 100 records were included in the methodological review (see online supplemental figure S1 for the selection process).

Study characteristics

Infectious diseases were the predominant medical field (30 studies, 30%), with a high proportion of research questions related to COVID-19 (21 of 30 studies). Most studies (90 studies, 90%) had at least one author affiliated with a biostatistics or epidemiology department. In 64% of studies, the source of funding was from public or not-for-profit organisations. Reporting guidelines were followed in 28 studies (28%), and the protocol of the observational analysis was available for 12 studies (12%). The main motivations for conducting the research were to evaluate the real-world effect of the intervention (39 studies, 39%), the lack of reliable evidence in previous observational studies that were possibly biased (29 studies, 29%), the unfeasibility of a randomised trial (21 studies, 21%) or to compare the results of the observational analysis with those of randomised trials (19 studies, 19%). In 59 studies (59%), the authors mentioned in their article that the target trial emulation approach would enhance the reliability of the study (see table 1 and online supplemental table S4 for more information on general characteristics).

Characteristics of the target trials

In 24 studies (24%), none of the key components of the target trial protocol was specified. Eligibility criteria

Table 1 General characteristics of included articles (n=100)

Main medical field	
Infectious diseases, COVID-19	21 (21%)
Cardiology	12 (12%)
Oncology	11 (11%)
Other infectious diseases	9 (9%)
Endocrinology, metabolism and nutrition	9 (9%)
Nephrology	9 (9%)
Neurology	7 (7%)
Obstetrics, gynaecology and reproductive medicine	6 (6%)
Haematology	5 (5%)
Gastroenterology	4 (4%)
Others	7 (7%)
Location of the corresponding author	
North America	48 (48%)
Europe	36 (36%)
Asia	15 (15%)
Oceania	1 (1%)
At least one author affiliated with a biostatistics, epidemiology or data science department	90 (90%)
Funder	
Public or not-for-profit	64 (64%)
For profit	2 (2%)
Both	10 (10%)
No funding	10 (10%)
Not reported or unclearly reported	14 (14%)
Use of reporting guidelines	28 (28%)
STROBE guidelines	23 (82%)
RECORD or RECORD-PE guidelines	4 (14%)
ISPOR guidelines	3 (11%)
Reporting guidelines for propensity score analysis	1 (4%)
Registration of the study	7 (7%)
EU PAS register	2 (29%)
Clinicaltrials.gov	5 (71%)
Availability of the protocol of the observational analysis	12 (12%)

EU PAS register, European Union electronic Register of Post-Authorisation Studies; ISPOR, International Society for Pharmacoeconomics and Outcomes Research; RECORD, REporting of studies Conducted using Observational Routinely-collected Data; RECORD-PE, RECORD statement for pharmacoepidemiology; STROBE, STrengthening the Reporting of OBservational studies in Epidemiology.

(69%), intervention strategies (73%), intervention assignment (68%), outcome (64%), follow-up (56%), causal contrast (56%) and statistical analysis (47%) were individually specified in the 100 selected studies, respectively (figure 1). The full research question was documented in 62 studies (62%) and all seven key protocol components in 40 studies (40%). The reporting of protocol components did not vary according to the quartile of journal impact factors nor according to general and specialised medical journals (online supplemental figure S2). The target trial may have been fully specified more frequently if an epidemiologist or a biostatistician was present in the two first or last positions of authorship of the related article (44% vs 21% in studies where there were no epidemiologist or a biostatistician in the two first or last positions of authorship, (online supplemental figure S2). Of the 76 observational analyses that specified the target trial (ie, reported at least one of its protocol components), 18/76 (24%) aimed to emulate a real randomised trial (16 unique randomised trials because two were emulated twice by two separate observational analyses; see online supplemental table S5), and 58/76 (76%) targeted a hypothetical trial. Key components of the protocol of the target trial were less frequently specified for a hypothetical than a real target trial (online supplemental figure S3). Although all real target trials reported the objective of demonstrating the superiority or non-inferiority of the intervention, this feature was not reported for 56/58 (97%) hypothetical trials. Eligibility was based on postbaseline characteristics (measured after the start of follow-up) in 5/51 (10%) hypothetical target trials that specified eligibility criteria. Active comparators were used in 12/18 (67%) of real target trials but only 23/55 (41%) of hypothetical target trials that specified the comparator strategy. Among the 56/76 target trials (74%) that specified the causal contrast, per protocol (PP) effects were estimated less frequently for real target trials (in only 3/17 trials, 18%) than in hypothetical trials (24/39 trials, 62%). Further details can be found in online supplemental table S6.

Consistency of protocol components between the target trial and its emulation

In the 76 emulations of studies that specified at least one component of the target trial protocol, eligibility criteria (39/76, 51%), intervention strategies (58/76, 76%), outcome (55/76, 72%), follow-up (44/76, 58%), causal contrast (51/76, 67%) and intervention effect measure (36/76, 47%) were consistent between the target trial and its emulation, respectively (online supplemental table S7 and figure S4). Notably, the follow-up was longer in the target trial for only four studies (7% of 56 studies for which both follow-up of the target and its emulation were specified), shorter for 9/56 studies (16%) and equal for the remaining 44/56 studies (79%). The consistency was lower across most protocol components for 18 emulations of a real target trial than for the 58 emulations of a hypothetical one, particularly with respect to eligibility criteria (1/18, 6% vs 38/58, 66%) and follow-up (5/18, 28% vs 40/58, 69%). The full research question (eligibility criteria, intervention strategies and outcome) was consistent in 1/18 (6%) of studies emulating a real target trial and in 34/58 (59%) a hypothetical one. Overall,

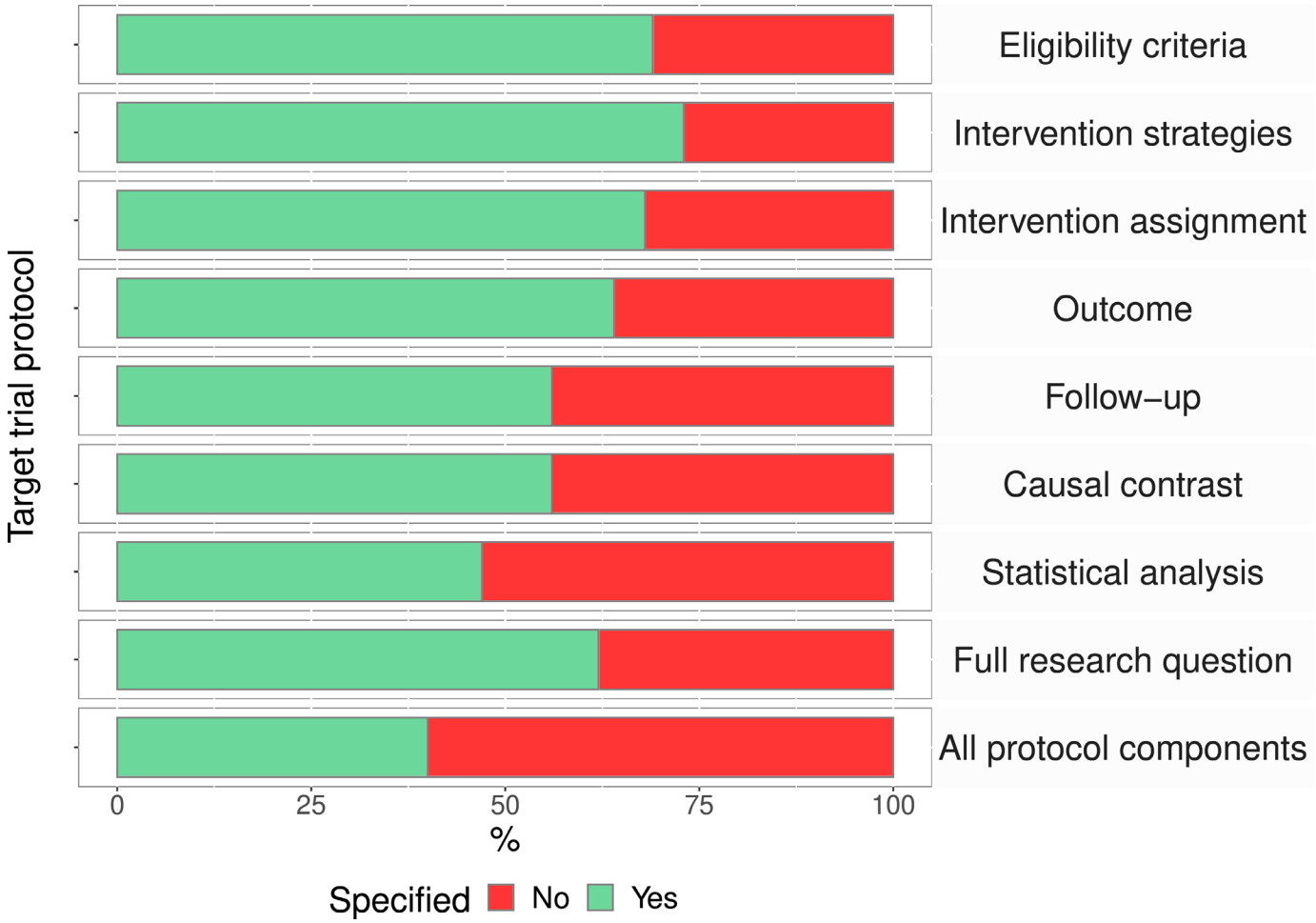


Figure 1 Summary plot of the specification of the target trial protocol components (n=100). Each bar represents the specification of the key protocol components of the target trial. The full research question was deemed specified if eligibility criteria, treatment strategies and outcome protocol components were specified for the target trial.

14/76 observational analyses (18%) achieved full consistency for all the evaluated protocol components, all emulating a hypothetical target trial (figure 2).

Characteristics of observational analyses

Overall, 79 articles (79%) reported the use of at least one database of routinely collected data to emulate the target trial (see online supplemental table S8 for more information). All observational analyses were performed with a cohort design. Eligibility of participants in the analysis was based on postbaseline characteristics in 22 studies (22%), whereas 14 studies (14%) did not exclude prevalent users of the intervention. Complex intervention strategies were studied in 33 studies (33%). The authors reported the causal contrasts of interest in 70 studies (70%), with 43/70 (61%) analyses estimating the observational analogue of the intention-to-treat (ITT) effect, 32/70 (46%) the per protocol (PP) effect and 12/70 (17%) both (table 2).

In terms of statistical analysis, 13 studies (13%) used directed acyclic graphs to identify potential biases (online supplemental table S9). Regarding the design, 18 studies (18%) used a sequential trials analysis and 10 studies (10%) the cloning, censoring and weighting method. 58 studies (58%) used a sole propensity score estimated with baseline confounders, and 18 (18%) used g-methods (g-formula, inverse probability weighting with time-varying weights, or doubly robust methods). Of the 76 studies performing survival analyses, 51/76 (68%) reported the intervention effect as an adjusted HR, and 35/76 (47%) presented weighted or standardised survival curves. Controls for residual confounding were reported in 21 studies (21%), and 13 studies (13%) reported an E-value (table 2). The median number of participants included in the analysis was 9286 (Q1;Q3 1264;70 570), with fewer participants in studies that emulated a real target trial (median 2346, 1015;14 076), see online supplemental table S10 for further details.

Risk of bias of observational analyses and time-zero issues

Overall, the risk of bias was serious in 28 studies (28%) and moderate for all others. No study was classified at low risk of bias for confounding (online supplemental table S11 and figure S5). The risk of bias due to the selection of participants was serious for 16 studies (16%); no study exhibited a serious risk of bias due to deviations from intended interventions. The information bias for intervention and outcome was serious in 10 (10%) and four (4%) studies, respectively, and the risk of bias due to missing data was

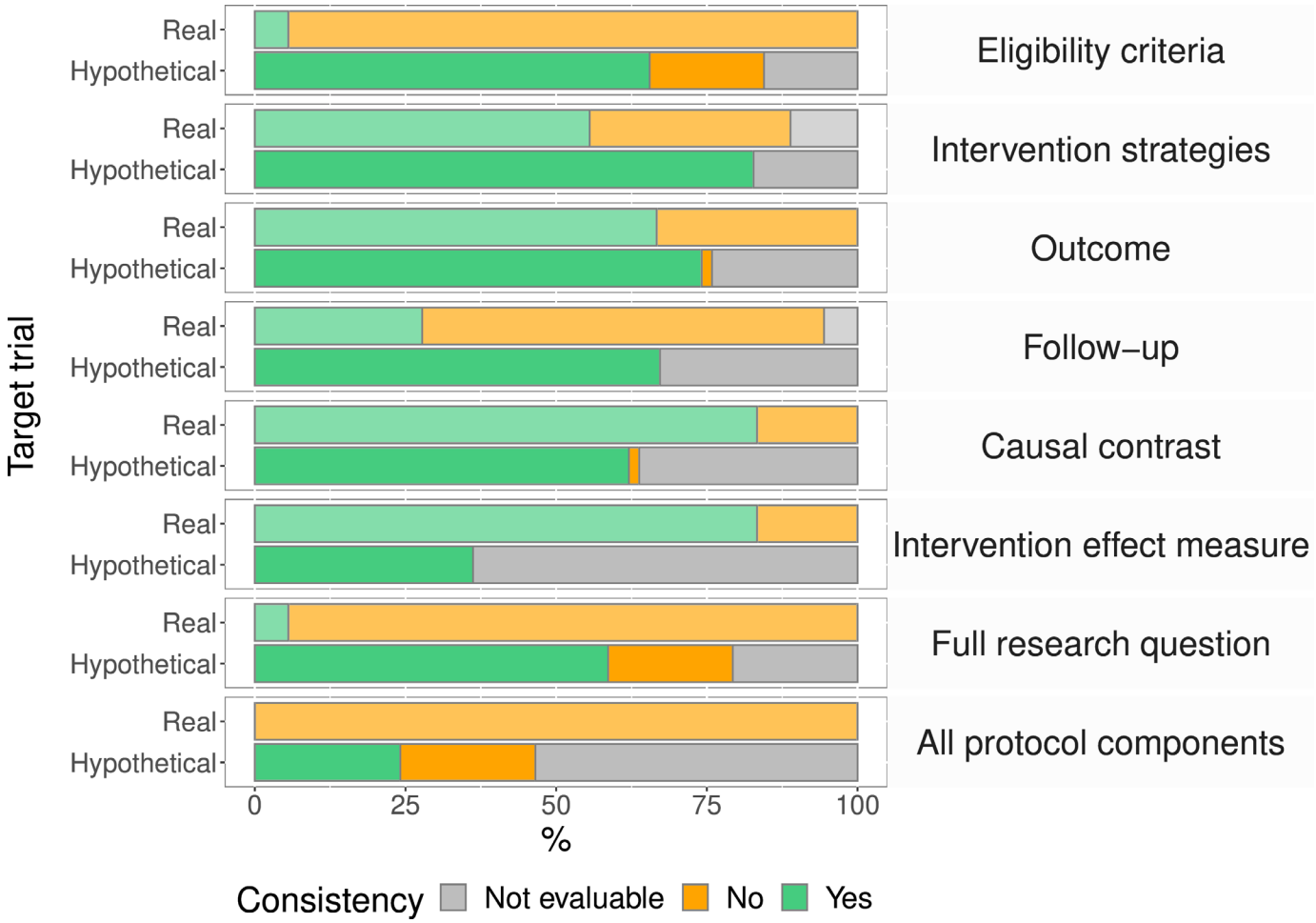


Figure 2 Summary plot of the consistency of the target trial and its emulation for studies that emulated a real or a hypothetical target trial (n=76). Bars in light colours represent the consistency of studies that emulated a real target trial and bars in dark colours represent the consistency of studies that emulated a hypothetical target trial. Consistency was assessed only if both the target trial and its emulation specified the considered protocol component, and deemed not evaluable otherwise.

serious in six studies (6%). The risk of serious bias seemed lower in studies that specified the target trial compared with those that did not overall (20/76, 26% vs 8/24, 33%) and in most domains, especially in the classification of interventions (4/76, 5% vs 6/24, 25%) and measurement of outcomes (1/76, 1% vs 3/24, 12%; figure 3).

When focusing on time-zero issues, 41 observational analyses did not align time-zero (start of follow-up) with the time when eligibility criteria were met or with intervention assignment (online supplemental table S12). The most common problem was immortal-time bias and selection bias due to postintervention eligibility (16/41 studies, 39%; see online supplemental table S12). The risk of bias of these 41 observational analyses with time-zero misalignments was deemed serious for 20/41 of them (49%). Time-zero issues seemed less frequent in studies that specified the target trial compared with those that did not (29/76, 39% vs 12/24, 52%).

Comparison of intervention effect in real randomised trials and their observational emulation

Among the 18 studies that emulated a real randomised trial, 15 observational analyses (83%) had a similar intervention effect measure and could be compared (online supplemental figure S6). All randomised trial analyses were performed on the ITT or modified ITT population, whereas 3/15 emulations (20%) estimated an observational analogue of a PP effect. We found full statistical significance agreement in 7/15 studies (47%). In the remaining studies, contrary to what was estimated in the randomised trial, the intervention effect in the observational emulation was not statistically significant in 6/15 studies (40%) and was statistically significant in 2/15 studies, in favour of the intervention of interest (13%). In 5/15 emulations (33%), the 95% CI of the estimate was wider than the one of the estimates of the targeted randomised trial.

DISCUSSION

Main findings

We conducted a methodological review of a recent sample of 100 studies declaring that they emulated a target trial. Among these studies, 24% did not, strictly speaking, emulate target trials because the target trial was not

Table 2 Characteristics of observational analyses

	Overall (n=100)	Studies emulating a real target trial (n=18)	Studies emulating a hypothetical target trial (n=58)	No target trial specified (n=24)
Eligibility criteria				
Users of the intervention				
New users	86 (86%)	14 (78%)	52 (90%)	20 (83%)
Prevalent users	14 (14%)	4 (22%)	6 (10%)	4 (17%)
Eligibility defined on characteristics observed during follow-up	22 (22%)	5 (28%)	10 (17%)	7 (29%)
Look-back period to assess eligibility (in years)	54 (54%)	10 (56%)	36 (62%)	8 (33%)
Median(Q1;Q3)	1([1 1])	1(0.49;1)	1([1 1])	1(0.44;1)
Minimum/maximum	0.08/9.5	0.49/9.5	0.25/3	0.08/1
Intervention strategy*				
Complex strategy	33 (33%)	9 (50%)	19 (33%)	5 (21%)
Dynamic	4 (4%)	0 (0%)	4 (7%)	0 (0%)
Joint strategies	17 (17%)	8 (44%)	6 (10%)	3 (12%)
Grace period	16 (16%)	2 (11%)	12 (21%)	2 (8%)
Pharmacological intervention	67 (67%)	15 (83%)	36 (62%)	16 (67%)
Comparator strategy*				
Nature of comparator				
Active comparator	46 (46%)	14 (78%)	23 (40%)	9 (38%)
No intervention or usual care	40 (40%)	2 (11%)	24 (41%)	14 (58%)
Other comparator strategies	14 (14%)	2 (11%)	11 (19%)	1 (4%)
Pharmacological intervention	46 (46%)	15 (83%)	25 (43%)	6 (25%)
Outcome†				
Efficacy and/or safety assessment				
Efficacy	50 (50%)	9 (50%)	22 (38%)	19 (79%)
Safety	13 (13%)	1 (6%)	11 (19%)	1 (4%)
Both efficacy and safety	37 (37%)	8 (44%)	25 (43%)	4 (17%)
Nature of outcome				
Binary	19 (19%)	4 (22%)	6 (10%)	9 (38%)
Continuous	5 (5%)	0 (0%)	2 (3%)	3 (12%)
Time-to-event	76 (76%)	14 (78%)	50 (86%)	12 (50%)
Follow-up				
Length of follow-up (in years, if applicable)				
Median(Q1;Q3)	2(0.5;6)	2.2[1 5]	2(0.43;5.75)	1(0.28;5.75)
Minimum/maximum	0/20	0.3/11.6	0/20	0/18
Causal contrast(s) reported by authors	70 (70%)	14 (78%)	46 (79%)	10 (42%)
Observational analogue of ITT effect	43 (61%)	9 (64%)	25 (54%)	9 (90%)
Observational analogue of PP effect	32 (46%)	4 (29%)	26 (57%)	2 (20%)
Both observational analogue of ITT and PP effects	12 (17%)	1 (7%)	9 (20%)	2 (20%)

Continued

Table 2 Continued

	Overall (n=100)	Studies emulating a real target trial (n=18)	Studies emulating a hypothetical target trial (n=58)	No target trial specified (n=24)
Other causal contrasts	10 (14%)	4 (29%)	5 (11%)	1 (10%)
Design for emulation				
Sequential trials analysis with emulation several times of a trial	18 (18%)	0 (0%)	12 (21%)	6 (25%)
Cloning, censoring and weighting	10 (10%)	1 (6%)	8 (14%)	1 (4%)
Statistical methodology				
A priori sample size or computation of anticipated power	5 (5%)	3 (17%)	2 (3%)	0 (0%)
Propensity score analysis (with baseline confounders)	58 (58%)	16 (89%)	26 (45%)	16 (67%)
Use of g-methods‡	20 (20%)	3 (17%)	13 (22%)	4 (15%)
Exploration of residual confounding				
Controls for confounding	21 (21%)	5 (28%)	14 (24%)	2 (8%)
Reporting of E-value	13 (13%)	0 (0%)	8 (14%)	5 (21%)

*First reported ones.
†Primary outcome or first reported outcome.
‡g-formula, inverse probability weighting with time-varying weights and/or doubly robust methods; studies that used inverse probability of treatment weighting only with a single weight at baseline were not counted in g-methods. Statistical techniques were not mutually exclusive (eg, a sequential trials analysis with the use of a propensity score weighting in each sequential trial).
ITT, intention-to-treat; PP, per protocol; Q1;Q3, first and third quartiles.

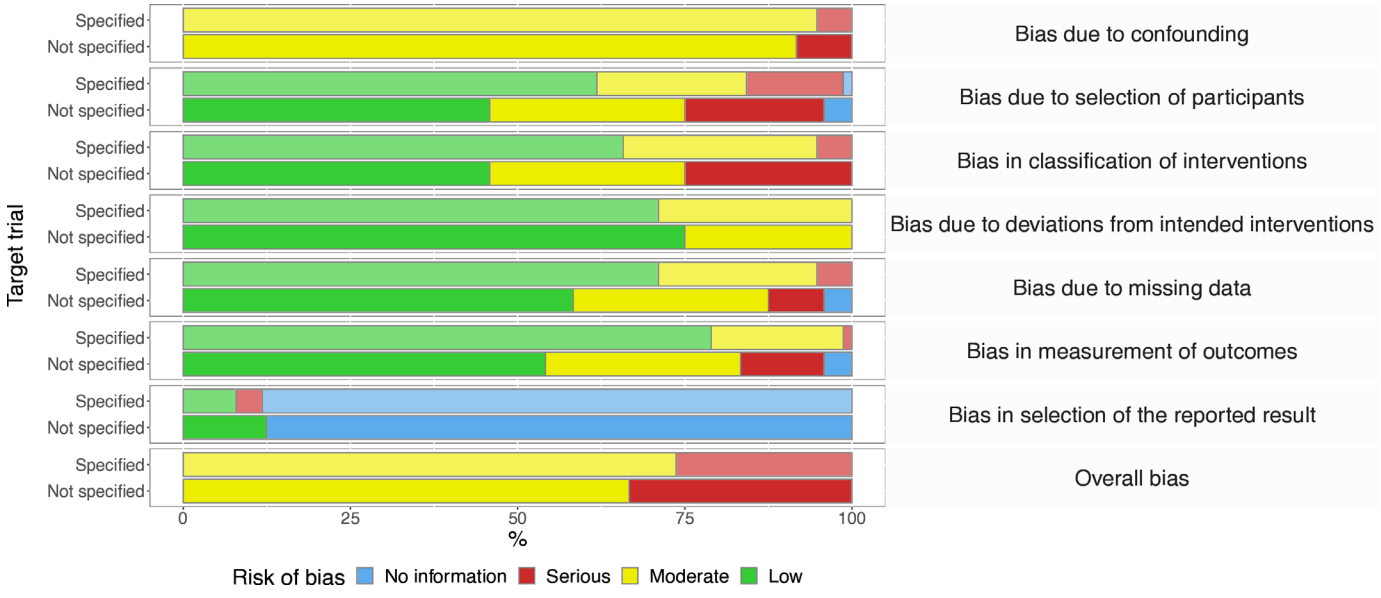


Figure 3 Summary plot of ROBINS-I assessment of risk of bias according to specification of the target trial. Bars in light colours represent the risk of bias of studies that specified the target trial and bars in dark colours represent the risk of bias of studies that did not specify the target trial. The target trial was deemed specified if at least one protocol component was documented. The risk of bias in the selection of the reported result was assessed only when a protocol was available for the observational study (n=12). For a given study, the overall risk of bias was the highest risk of bias observed in one of the individual risk of bias domains.

specified at all. The full research question was specified in 62% of studies and all seven key protocol components in 40%. Regarding the consistency of the target trial and its emulation, the full research question was consistent in 46% of studies specifying the target trial, but this was rarer for studies emulating real target trials versus hypothetical ones (6% vs 59%). The overall risk of bias was considered serious in 28% of studies and moderate in the remaining studies. Time-zero issues were present in 41% of observational analyses, with approximately half of these resulting in a serious risk of bias. Risk of serious biases and time-zero misalignments seemed less frequent in studies that specified the target trial than those that did not. When comparing the results of real target trials with their observational emulations, less than half reached the full statistical significance agreement.

Strengths and limitations of the study

This methodological review has several strengths. First, we conducted a comprehensive review of the use of the target trial emulation approach. Second, we reported this study in accordance with the guidance specific to this type of work.[11] Third, we thoroughly assessed various aspects of the selected studies, including the specification of the target trial for each key protocol component, its consistency with its emulation, the causal inference techniques used by authors and the risk of bias of the observational analysis, with a focus on time-zero issues. Fourth, the use of the ROBINS-I tool was particularly relevant in this review because it is based on the definition of the hypothetical target trial that is being emulated in the evaluated observational analysis.[15]

Our study also has some limitations. Regarding the selection of studies, we assessed only a recent sample of studies due to limited resources and could not evaluate if the understanding of the target trial emulation approach has evolved over time. Consequently, our study sample is also biased towards research questions related to COVID-19, thereby limiting the generalisability of our findings. Regarding the risk of bias assessment, the ROBINS-I tool relies on expert knowledge to define confounders and cointerventions for a given research question, which we were not able to provide in our evaluation. Instead, we considered only the confounders and cointerventions reported by the authors of the studies, thereby limiting our assessment of the confounding domain, which might have the most significant influence on the overall risk of bias evaluated by the ROBINS-I tool, according to a previous study.[19] Another limitation was to consider the target trial as specified if at least one protocol component (not all) was documented. This definition is practical and was used for this review to classify studies that were clearly not an emulated target trial. However, this definition should not be considered the minimum requirement to qualify an observational analysis as a target trial emulation. Additionally, our evaluation of consistency could only identify strong deviations from what was targeted, as this assessment was related to the reporting of intervention strategies, which was overall poor. Similarly to what Hansford *et al* reported, we found a very low reporting of dose, duration and frequency of treatment strategies (data not shown)[20] and decided to assess only major deviations from intervention strategies that were actually targeted. For example, the control arm was an active comparator in the observational analysis whereas it was a placebo arm in the target randomised trial. In addition, we used an empirical threshold of three discrepant eligibility criteria to deem an emulation inconsistent with its target trial about this protocol component, which might have been too strict, in particular for observational analyses that emulate a real target trial. Finally, our evaluation of the risk of bias was not blinded to target trial specification, thereby leaving open the possibility that the observed differences between studies that specified or not the target trial could be attributed to this absence of blinding.

Difference in results with existing studies

Recently, three methodological reviews were published almost simultaneously, focusing on the target trial emulation approach.[20–22] The initial review identified 38 studies attempting to emulate a trial up to the beginning of 2021, while the subsequent review identified 96 studies up to the beginning of 2022.[21 22] Across these two overlapping samples of studies, the results were quite consistent with ours, with studies pertaining to the same medical fields (cardiology, infectious diseases and oncology studies were the most represented ones). The statistical analysis features were also similar, notably regarding the use of time-to-event outcomes and propensity scores to adjust for confounding. The authors also focused on methods to account for immortal time-biases and not surprisingly showed a frequent utilisation of the cloning, censoring and weighting and sequential trials approaches (12% and 29% of studies respectively in[21]). Even more recently, an international initiative has been proposed to establish reporting recommendations for target trial emulations.[23] This initiative began with a systematic review of studies using the target trial framework.[20] Dates of publication of the 200 included studies spanned from 2012 to the end of 2022, again with the inclusion of only studies that explicitly applied the framework. The results of this review echoed our findings regarding general characteristics and statistical analysis features. Interestingly, this last review additionally focused on the reporting of the approach and showed that only 87 studies (44%) described both the target trial protocol and its emulation protocol separately, which roughly corresponds to the proportion of studies (40%) in our study sample that defined all the target trial components.

A significant difference in our review compared with previous ones lies in our study selection criteria. We opted to evaluate studies that claimed to emulate either a hypothetical or a real randomised trial, regardless of whether such emulation was indeed conducted. This broader eligibility allowed us to compare the risk of

bias of studies that specified the target trial to those that actually did not, which was not performed in the other reviews, and provided some insights about the potential effect on bias mitigation of the target trial emulation approach. Another key difference is our assessment of the consistency between the target trial and its emulation. Indeed, one of the goals of the target trial emulation is to ensure transparency regarding the intervention strategies targeted and those assessed in observational analysis. This evaluation is intricately linked to the counterfactual theory of causality, and in particular to the consistency assumption.[3]

Specification of the target trial

In our review, one-quarter of studies did not define the target trial. This situation is a significant concern because specifying the components of the target trial serves the dual purpose of enhancing transparency and mitigating biases. This may be attributed, in part, to a misunderstanding of the method. Indeed, confusion may have emerged between the target trial emulation framework and sequential (or nested) trials analysis, a statistical design based on the specification of the target trial that helps mitigate immortal-time bias because the first studies entitled 'emulated target trials' used this particular technique.[2 24] In our sample, 25% of studies did not explicitly specify the target trial despite declaring that they performed a sequential trials analysis. In addition, the target trial could have been actually conceptualised by authors but not reported in the article. Of note, the likelihood of complete reporting of the target trial was greater with target trials that were real randomised trials than hypothetical ones because we extracted their characteristics in the original publication of the randomised trial and not necessarily in the article of their observational emulation. Finally, thinking of epidemiological research questions within a causal inference framework is still uncommon, and the term 'emulated target trial' may have been only strategically included in certain studies to enhance their chances of publication.

Use of causal inference techniques

To mitigate biases in the observational analysis, the studies often used advanced statistical techniques of causal inference, such as g-methods for time-varying confounding[3] or inverse probability of censoring weighting for selection bias.[25] Regarding the immortal-time bias, a sequential trials design[24] or the cloning, censoring and weighting strategy[26 27] were also applied quite frequently. Methods to explore residual confounding (eg, negative control outcomes[14] or calculation of the E-value[13]) were reported in almost one-third of studies. The causal contrast of interest was defined in 70% of studies versus 19% in a previous methodological review of observational studies that did not specify the target trial,[28] thereby suggesting improvement of the transparency about the causal effect at hand with the framework.

Risk of bias of observational analyses

The risk of bias in our study sample can seem high in that no study was considered a low risk of bias, but these results must be mitigated by the stringency of the ROBINS-I tool, particularly for the confounding domain: the risk of bias can be considered low only if no confounding is expected for the relation between the intervention and the given outcome.[15] In a methodological review[29] of a sample of systematic reviews of observational studies that used the ROBINS-I tool, the average risk of serious/critical risk of bias across studies was 54% as compared with 28% of observational analyses in our review. Therefore, the risk of bias in our study sample may be actually lower than in previous assessments. In addition, the risk of serious bias was less prevalent in studies that defined the target trial: even if we could attribute this effect to the use of the framework itself, there may be other reasons for differences (eg, expertise of researchers, involvement of biostatistics department affiliated authors, etc). Finally, we found a higher proportion of time-zero issues than in a previous methodological review of observational studies:[9] 41% versus 19%. However, in our assessment, we collected all time-zero misalignments, even the ones that did not have a serious risk of bias, for example, in the case of excluding patients on postbaseline characteristics, but in a very limited proportion. Consequently, this apparent discrepancy may be explained by considering only time-zero issues that induced a serious risk of bias in the previous work.[9]

Consistency of the target trial and its emulation

For studies that emulated a hypothetical target trial, the consistency was quite good not surprisingly, as it was easy to adapt the protocol of the target trial to fit in the available observational data. This should not be considered a real issue, as the target trial is typically a pragmatic trial in case of a hypothetical target trial and therefore close to a real-life setting.[5] For the emulation of a real randomised trial, consistency between the target and emulated trial protocols becomes more of a focal point. Indeed, benchmarking the observational analysis with the results of the real randomised experiment is often an objective. In our study, we observed poor consistency in key protocol components between real target trials and their emulations, particularly for eligibility criteria. While our assessment of consistency may have been stringent for this specific aspect, it is important to note that this result can also be attributed to the inherent limitations of emulating randomised trials within observational settings. Randomised trials typically involve restricted populations with specific characteristics that are challenging to replicate in observational data.[30 31] The difficulty in emulating real randomised trials was explored in a recent comprehensive work, called the RCT-DUPLICATE initiative (Randomised Controlled Trials Duplicated Using Prospective Longitudinal Insurance Claims: Applying Techniques of Epidemiology).[18] In a sample of 32 studies that emulate a real target trial and were performed by the same team

of researchers, 16 (50%) were qualified as 'close' emulations of the real randomised trial. This qualification overlaps with our definition of consistency between the target trial and its emulation but is not completely identical (eg, the authors assessed the existence of a run-in period in the real randomised trial that may have selected adherent participants to interventions), which may explain the observed differences. The full statistical significance agreement also seemed better in their sample (75% vs 47% in ours). These results may be explained by the particular expertise in observational analyses of this team, whereas our study sample comprised trial emulations from various teams. Moreover, the sample size of observational emulations they performed was larger than in our review, thus improving the precision of the estimations and the odds of concluding to the full statistical significance agreement. Finally, a recent meta-analysis of the RCT-DUPLICATE initiative data suggested that the heterogeneity in results between the randomised trials and observational analyses could be explained by three characteristics of the target trial: treatment started in the hospital, discontinuation of some baseline treatments at randomisation and delayed onset of drug effects.[32] We unfortunately did not assess these characteristics in the target trial, but a different distribution of them in our sample of studies may also explain the apparent discrepancy with the results of the RCT-DUPLICATE initiative regarding consistency.

CONCLUSIONS

In this methodological systematic review investigating the target trial emulation approach, one-quarter of observational analyses did not specify the target trial. For the remaining studies, target trials and their emulations were particularly inconsistent in those emulating a real randomised trial. Observational analyses that defined the target trial seemed to have a lower risk of methodological issues than those that did not.

Furthers directions

While the target trial emulation appears to be a very promising approach to improve transparency and mitigate bias in observational analyses, our data underscore an urgent need for the upcoming recommendations for its reporting.[23] These guidelines will serve as a crucial step towards enhancing both the utilisation and comprehension of this approach. Still, if our aim is to enhance the level of evidence derived from observational studies, conducting an observational analysis as a target trial emulation might not be enough, and a recent proposal suggests that a more comprehensive process of inferential analysis, integrating the target trial approach, is necessary to obtain causal effects from observational data.[33] Therefore, such a framework will need to be evaluated as well.

Contributors AD and FT are the co-last authors of the manuscript. Contributors: NST, DH, AD and FT designed the study. NST, GLM and CS did the study selection. NST, GLM and MB extracted the data of selected studies. NST performed the statistical analysis. NST, DH, AD and FT interpreted the data. NST wrote the draft of the manuscript. All authors contributed to its critical revision. NST, AD and FT are guarantors of the manuscript and attest that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare that they have no competing interest related to the study. GLM declares consulting fees from Synapse Medicine, unrelated to the study.

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

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The database that contains the product of study extraction is available upon reasonable request.

ORCID iD

Noémie Simon-Tillaux http://orcid.org/0000-0001-7138-5364

REFERENCES

1 Vickers AJ, Assel M, Dunn RL, *et al*. Guidelines for Reporting Observational Research in Urology: The Importance of Clear Reference to Causality. *Eur Urol* 2023;84:147–51.
2 Hernán MA, Alonso A, Logan R, *et al*. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiol Camb Mass* 2008;19:766–79.
3 Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton: Chapman&Hall/CRC, 2020.
4 Hernán MA. Methods of Public Health Research - Strengthening Causal Inference from Observational Data. *N Engl J Med* 2021;385:1345–8.
5 Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* 2016;183:758–64.
6 Fu EL. Target Trial Emulation to Improve Causal Inference from Observational Data: What, Why, and How? *J Am Soc Nephrol* 2023;34:1305–14.
7 García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol* 2017;32:495–500.
8 Hernán MA, Sauer BC, Hernández-Díaz S, *et al*. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol* 2016;79:70–5.
9 Nguyen VT, Engleton M, Davison M, *et al*. Risk of bias in observational studies using routinely collected data of comparative effectiveness research: a meta-research study. *BMC Med* 2021;19:279.
10 Farjat AE, Virdone S, Thomas LE, *et al*. The importance of the design of observational studies in comparative effectiveness research:

Lessons from the GARFIELD-AF and ORBIT-AF registries. *Am Heart J* 2022;243:110–21.
11 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
12 Dickerman BA, García-Albéniz X, Logan RW, *et al*. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med* 2019;25:1601–6.
13 VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med* 2017;167:268–74.
14 Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiol Camb Mass* 2010;21:383–8.
15 Sterne JA, Hernán MA, Reeves BC, *et al*. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.
16 Higgins JPT, Thomas J, Chandler J, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. 1st edn. Wiley, 2019.
17 Franklin JM, Pawar A, Martin D, *et al*. Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. *Clin Pharmacol Ther* 2020;107:817–26.
18 Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative, *et al*. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA* 2023;329:1376–85.
19 Dhiman P, Lee H, Kirtley S, *et al*. A systematic review showed more consideration is needed when conducting nonrandomized studies of interventions. *J Clin Epidemiol* 2020;117:99–108.
20 Hansford HJ, Cashin AG, Jones MD, *et al*. Reporting of Observational Studies Explicitly Aiming to Emulate Randomized Trials: A Systematic Review. *JAMA Netw Open* 2023;6:e2336023.
21 Scola G, Chis Ster A, Bean D, *et al*. Implementation of the trial emulation approach in medical research: a scoping review. *BMC Med Res Methodol* 2023;23:186.
22 Zuo H, Yu L, Campbell SM, *et al*. The implementation of target trial emulation for causal inference: a scoping review. *J Clin Epidemiol* 2023;162:29–37.
23 Hansford HJ, Cashin AG, Jones MD, *et al*. Development of the TrAnsparent ReportinG of observational studies Emulating a Target trial (TARGET) guideline. *BMJ Open* 2023;13:e074626.
24 Danaei G, Rodríguez LAG, Cantero OF, *et al*. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res* 2013;22:70–96.
25 Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* 2000;56:779–88.
26 Hernán MA. How to estimate the effect of treatment duration on survival outcomes using observational data. *BMJ* 2018;360:k182.
27 Maringe C, Benitez Majano S, Exarchakou A, *et al*. Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. *Int J Epidemiol* 2020;49:1719–29.
28 Zhao SS, Lyu H, Solomon DH, *et al*. Improving rheumatoid arthritis comparative effectiveness research through causal inference principles: systematic review using a target trial emulation framework. *Ann Rheum Dis* 2020;79:883–90.
29 Igelström E, Campbell M, Craig P, *et al*. Cochrane's risk of bias tool for non-randomized studies (ROBINS-I) is frequently misapplied: A methodological systematic review. *J Clin Epidemiol* 2021;140:22–32: .
30 Matthews AA, Dahabreh IJ, Fröbert O, *et al*. Benchmarking Observational Analyses Before Using Them to Address Questions Trials Do Not Answer: An Application to Coronary Thrombus Aspiration. *Am J Epidemiol* 2022;191:1652–65.
31 Kirchgesner J, Desai RJ, Beaugerie L, *et al*. Calibrating Real-World Evidence Studies Against Randomized Trials: Treatment Effectiveness of Infliximab in Crohn's Disease. *Clin Pharmacol Ther* 2022;111:179–86.
32 Heyard R, Held L, Schneeweiss S, *et al*. Design differences and variation in results between randomised trials and non-randomised emulations: meta-analysis of RCT-DUPLICATE data. *BMJ Med* 2024;3:e000709.
33 Desai RJ, Wang SV, Sreedhara SK, *et al*. Process guide for inferential studies using healthcare data from routine clinical practice to evaluate causal effects of drugs (PRINCIPLED): considerations from the FDA Sentinel Innovation Center. *BMJ* 2024;384:e076460.

Supplemental material

Simon-Tillaux N, *et al. BMJ Open* 2024; 14:e086595. doi: 10.1136/bmjopen-2024-086595

Table S1. PRISMA 2020 checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	NA
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	p3-4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	p5-6
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	P6
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	p8 and Appendix 3
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	p7 and Appendix 3
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Appendix 2 and 3
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	p8
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	p8
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	p8-9
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	p8-9 and Appendix 3
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	p9-10 and Appendix 3
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	NA
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	NA
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	NA
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	NA
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	NA
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	NA
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	NA
Reporting bias	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	NA

assessment			
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	NA
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	p12 and Figure S1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Figure S1
Study characteristics	17	Cite each included study and present its characteristics.	Appendix 4, Tables 1 and 2, Figure S2, Tables S4-S11
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Table 3, S11, Figure 3 and S5
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Figures 1-2 and 4, Figure S3-4
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	NA
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	NA
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	NA
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	NA
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	NA
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	NA
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	p25-26
	23b	Discuss any limitations of the evidence included in the review.	p24-25
	23c	Discuss any limitations of the review processes used.	p24
	23d	Discuss implications of the results for practice, policy, and future research.	p30
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	p7
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	p7
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	NA
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	p31
Competing interests	26	Declare any competing interests of review authors.	p31
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	p31

Table S2. PRISMA 2020 for abstracts checklist

Section and Topic	Item #	Checklist item	Reported (Yes/No)
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	Yes
Information sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.	Yes
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	Yes
Synthesis of results	6	Specify the methods used to present and synthesise results.	Yes
RESULTS			
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	Yes
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favoured).	Yes
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).	Yes
Interpretation	10	Provide a general interpretation of the results and important implications.	Yes
OTHER			
Funding	11	Specify the primary source of funding for the review.	Yes
Registration	12	Provide the register name and registration number.	NA

Table S3. Reasons for inconsistencies in the target trial and its emulation

	Reasons
Eligibility criteria	• Post-baseline eligibility criteria in emulation • Major deviations in reproduction of eligibility criteria (3 or more discrepant eligibility criteria, not related to the observational setting)
Intervention	• Not the same interventions assessed
Users of intervention	• New-users versus prevalent users
Comparator	• Not the same comparators assessed
Intervention strategies	• At least one inconsistency for intervention, users of intervention or comparator
Primary (or first reported) outcome	• Primary outcomes were different in target trial and its emulation • If co-primary outcome in the target trial, and emulation considered only one of these outcomes as main outcome, the target trial and its emulation were deemed consistent
Follow-up	• Exclusion of beginning of follow-up (for example, the first 3 months of follow-up) • Length of follow-up discrepant of more than 10%
Causal contrast	• ITT in the target trial versus only observational analogue of a PP effect in its emulation • If the target trial reported both ITT and PP effects, and its emulation only one, the two of them were deemed consistent
Intervention effect measure	• Hazard ratio in the target trial versus risk difference in the emulation

List of abbreviations: ITT, intention-to-treat; PP, per protocol. The target trial protocol seldom provided clear details on how outcomes were measured. Similarly, information on the precise route and/or dosage of pharmacological interventions was rarely available for both the target trial and its emulation. Consequently, these aspects were not taken into consideration during our consistency evaluation.

Table S4. Additional general characteristics

	All studies (N=100)
Reasons reported by authors for conducting the study	
To assess the real-world effect of the intervention	39 (39%)
Previous observational studies were possibly biased	29 (29%)
The randomised trial is unfeasible	21 (21%)
To benchmark observational analysis against randomised trial results	19 (19%)
To explore discrepancies in the previously published literature	16 (16%)
Results of randomised trials are pending	11 (11%)
Randomised trials were inconclusive on this particular research question	9 (9%)
To compare two interventions that have never been compared to each other in randomised trials	6 (6%)
The effect of the intervention needs to be assessed rapidly	5 (5%)
Other reasons	33 (33%)
Reference for using the emulated target trial framework provided	81 (81%)
Hernán et al. "Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available"[1]	65 (81%)
Hernán and Robins "Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses"[2]	13 (16%)
Dickerman et al. "Avoidable flaws in observational analyses: an application to statins and cancer"[3]	7 (9%)
Hernán et al. "Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease"[4]	7 (9%)
Other references	7 (9%)
Number of target trials in the article	
Median [Q1;Q3]	1 [1;1]
Minimum / maximum	1 / 262
Number of emulated trials in the article (median [Q1;Q3])	
Median [Q1;Q3]	1 [1;2]
Minimum / maximum	1 / 262

Abbreviation: Q1;Q3, first and third quartiles.

Table S5. Characteristics of real target trials

	Real target trials (N=16)
Main medical field	
Oncology	7 (44%)
Cardiology	3 (19%)
Gastroenterology	2 (12%)
Rheumatology	2 (12%)
Dermatology	1 (6%)
Endocrinology, Metabolism and Nutrition	1 (6%)
Results of randomised trial published	15 (94%)
Availability of the protocol of the study (including registration on a clinical trial website)	16 (100%)
Multicentre study	16 (100%)
Location of study sites	
Europe	14 (88%)
North America	12 (75%)
South America	4 (25%)
Africa	3 (19%)
Asia	8 (50%)
Oceania	5 (31%)
Duration of the inclusion period (in years, median [Q1;Q3])	
Median [Q1;Q3]	3.5 [3;5.25]
Minimum / maximum	1 / 8
Blinding	
Open-label trial	11 (69%)
Blinding of participants	5 (31%)
Blinding of investigators and/or caregivers	5 (31%)
Blinding of outcome assessors	7 (44%)
Number of participants included in the analysis	
Median [Q1;Q3]	616 [324.5;3681]
Minimum / maximum	36 / 12834
Statistical significance for the intervention effect	8 (50%)

There are only 16 unique real target trials to describe, as two of them were targeted twice by two separated observational analyses. Abbreviation: Q1;Q3, first and third quartiles.

Table S6. Characteristics of the target trial for the studies specifying it

	Overall (N=76)	Studies emulating a real target trial (N=18*)	Studies emulating a hypothetical target trial (N=58)
Design			
Experimental design			
Parallel	75 (99%)	18 (100%)	57 (98%)
Not evaluable	1 (1%)	0 (0%)	1 (2%)
Statistical hypothesis tested			
Superiority	15 (20%)	13 (72%)	2 (3%)
Non-inferiority	5 (7%)	5 (28%)	0 (0%)
Not reported	56 (74%)	0 (0%)	56 (97%)
Eligibility criteria specified	69 (91%)	18 (100%)	51 (88%)
Users of the intervention			
New users	61 (88%)	18 (100%)	43 (84%)
Prevalent users	8 (12%)	0 (0%)	8 (16%)
Eligibility is defined on characteristics observed after the start of follow-up	5 (7%)	0 (0%)	5 (10%)
Intervention strategy specified	75 (99%)	18 (100%)	57 (98%)
Complex strategy	28 (37%)	9 (50%)	19 (33%)
Dynamic	4 (5%)	0 (0%)	4 (7%)
Joint strategies	13 (17%)	7 (39%)	6 (11%)
Grace period	12 (16%)	2 (11%)	10 (18%)
Pharmacological intervention	51 (68%)	15 (83%)	36 (63%)
Comparator strategy specified	73 (96%)	18 (100%)	55 (95%)
Active comparator	35 (47%)	12 (67%)	23 (41%)
Placebo	3 (4%)	2 (11%)	1 (2%)
No intervention or usual care	25 (33%)	2 (11%)	23 (41%)
Other comparator strategies	11 (15%)	2 (11%)	9 (16%)
Outcome specified	64 (84%)	18 (100%)	46 (79%)
Number of primary or co-primary outcomes			
Median [Q1;Q3]	1 [1;1]	1 [1;1]	1 [1;2]
Minimum / maximum	1 / 11	1 / 2	1 / 11
Time point of outcome is defined	55 (72%)	18 (100%)	37 (64%)
Time point of outcome (if applicable, in years)			
Median [Q1;Q3]	2 [0.5;5.25]	2.9 [0.6;7.38]	2 [0.46;5]

Minimum / maximum	0 / 20	0.08 / 10	0 / 20
Efficacy and/or safety assessment			
Efficacy	24 (38%)	9 (50%)	15 (33%)
Safety	10 (16%)	1 (6%)	9 (20%)
Both efficacy and safety	29 (46%)	8 (44%)	21 (47%)
Nature of outcome			
Binary	10 (16%)	4 (22%)	6 (13%)
Continuous	1 (2%)	0 (0%)	1 (2%)
Time-to-event	43 (68%)	14 (78%)	29 (64%)
Not evaluable	9 (14%)	0 (0%)	9 (20%)
Follow-up specified	56 (74%)	18 (100%)	38 (66%)
Length of follow-up (in years, if applicable)			
Median [Q1;Q3]	2 [0.5;5.88]	4.15 [1.85; 8]	1.88 [0.42;5]
Minimum / maximum	0 / 20	0.49 / 10	0 / 20
Causal contrast(s) specified	56 (74%)	17 (94%)	39 (67%)
ITT	41 (73%)	13 (76%)	28 (72%)
Modified ITT	3 (5%)	2 (12%)	1 (3%)
PP	27 (48%)	3 (18%)	24 (62%)
Other causal contrast	7 (12%)	4 (24%)	3 (8%)
Statistical analysis specified	47 (62%)	18 (100%)	29 (50%)
First reported intervention effect measure			
Hazard ratio	27 (73%)	14 (78%)	13 (68%)
Risk difference	4 (11%)	0 (0%)	4 (21%)
Relative risk	2 (5%)	1 (6%)	1 (5%)
Odds ratio	1 (3%)	1 (6%)	0 (0%)
Other	3 (8%)	2 (11%)	1 (5%)
A priori sample size calculation	19 (25%)	18 (100%)	1 (2%)

A hypothetical target trial was considered specified if at least one of the seven key protocol component was documented. * Two real target trials were emulated twice. As a result, the description of real target trials corresponds to a total of 16 unique randomised trials. List of abbreviations: ITT, intention-to-treat; PP, per protocol; Q1;Q3, first and third quartiles.

Table S7. Consistency between the target trial and its emulation

	Overall (N=76)	Studies emulating a real target trial (N=18)	Studies emulating a hypothetical target trial (N=58)
Eligibility criteria consistent			
Yes	39 (51%)	1 (6%)	38 (66%)
No	28 (37%)	17 (94%)	11 (19%)
Not evaluable	9 (12%)	0 (0%)	9 (16%)
Users of intervention consistent			
Yes	65 (86%)	14 (78%)	51 (88%)
No	4 (5%)	4 (22%)	0 (0%)
Not evaluable	7 (9%)	0 (0%)	7 (12%)
Intervention consistent			
Yes	72 (95%)	15 (83%)	57 (98%)
No	1 (1%)	1 (6%)	0 (0%)
Not evaluable	3 (4%)	2 (11%)	1 (2%)
Comparator consistent			
Yes	69 (91%)	14 (78%)	55 (95%)
No	3 (4%)	3 (17%)	0 (0%)
Not evaluable	4 (5%)	1 (6%)	3 (5%)
Intervention strategies consistent			
Yes	58 (76%)	10 (56%)	48 (83%)
No	6 (8%)	6 (33%)	0 (0%)
Not evaluable	12 (16%)	2 (11%)	10 (17%)
Outcome consistent			
Yes	55 (72%)	12 (67%)	43 (74%)
No	7 (9%)	6 (33%)	1 (2%)
Not evaluable	14 (18%)	0 (0%)	14 (24%)
Follow-up consistent			
Yes	44 (58%)	5 (28%)	39 (67%)
No	12 (16%)	12 (67%)	0 (0%)
Not evaluable	20 (26%)	1 (6%)	19 (33%)
Causal contrast consistent			
Yes	51 (67%)	15 (83%)	36 (62%)
No	4 (5%)	3 (17%)	1 (2%)
Not evaluable	21 (28%)	0 (0%)	21 (36%)
Intervention effect measure consistent			
Yes	36 (47%)	15 (83%)	21 (36%)

No	3 (4%)	3 (17%)	0 (0%)
Not evaluable	37 (49%)	0 (0%)	37 (64%)
All previous protocol components consistent			
Yes	14 (18%)	0 (0%)	14 (24%)
No	18 (24%)	15 (83%)	3 (5%)
Not evaluable	44 (58%)	3 (17%)	41 (71%)
Full research question consistent*			
Yes	35 (46%)	1 (6%)	34 (59%)
No	23 (30%)	15 (83%)	8 (14%)
Not evaluable	18 (24%)	2 (11%)	16 (28%)

The consistency between the target trial and its emulation is reported in this table, for studies specifying at least one protocol component of the target trial. If the protocol component was not specified either in the target trial or in its emulation, the consistency was not evaluable. For more information on the definitions of inconsistent protocol components, see Table S3. *The full research question was consistent if eligibility criteria, intervention strategies and outcome were consistent.

Table S8. Data used in observational analyses

	Overall (N=99*)	Studies emulating a real target trial (N=18)	Studies emulating a hypothetical target trial (N=58)	No target trial specified in the studies (N=23*)
Number of databases				
Median [Q1;Q3]	1 [1;2]	2 [1;3]	1 [1;2]	1 [1;2]
Minimum / maximum	1 / 6	1 / 5	1 / 6	1 / 4
Routinely collected data**	79 (80%)	18 (100%)	43 (74%)	18 (78%)
Data source**				
Claims and/or EHR databases of public or not-for-profit healthcare providers	40 (40%)	7 (39%)	25 (43%)	8 (35%)
Claims and/or EHR databases of for-profit healthcare providers	17 (17%)	8 (44%)	7 (12%)	2 (9%)
Other patient's record databases	7 (7%)	0 (0%)	2 (3%)	5 (22%)
Registries	27 (27%)	7 (39%)	15 (26%)	5 (22%)
Cohorts	21 (21%)	0 (0%)	16 (28%)	5 (22%)
Other sources	12 (12%)	4 (22%)	3 (5%)	5 (22%)
Specificities of the databases**				
National databases	46 (46%)	11 (61%)	24 (41%)	11 (48%)
Profession-specific databases	14 (14%)	1 (6%)	8 (14%)	5 (22%)
Disease-specific databases	39 (39%)	7 (39%)	23 (40%)	9 (39%)
Participant's location**				
Europe	38 (38%)	7 (39%)	24 (41%)	7 (30%)
North America	48 (48%)	12 (67%)	24 (41%)	12 (52%)
Asia	14 (14%)	1 (6%)	9 (16%)	4 (17%)
Oceania	14 (14%)	0 (0%)	0 (0%)	1 (4%)
Transcontinental database	3 (3%)	0 (0%)	3 (5%)	0 (0%)

*One study did not report the data used for the analysis. **At least one database with these characteristics for the study. List of abbreviations: EHR, electronic health records; Q1;Q3, first and third quartiles.

Table S9. Characteristics of statistical analysis in observational analyses

	Overall (N=100)	Studies emulating a real target trial (N=18)	Studies emulating a hypothetical target trial (N=58)	No target trial specified in the studies (N=24)
Statistical hypothesis tested				
Superiority	4 (4%)	1 (6%)	2 (3%)	1 (4%)
Superiority and non-inferiority	1 (1%)	1 (6%)	0 (0%)	0 (0%)
Not defined	95 (95%)	16 (89%)	56 (97%)	23 (96%)
Participant's eligibility over time				
One time during follow-up	63 (63%)	18 (100%)	34 (59%)	11 (46%)
Several times during follow-up	34 (34%)	0 (0%)	22 (38%)	12 (50%)
Unclear definition of start of follow-up	3 (3%)	0 (0%)	2 (3%)	1 (4%)
*If eligibility criteria are met several times during follow-up, choice of time-zero**				
First eligible time	5 (15%)	0 (0%)	4 (18%)	1 (8%)
Random choice	1 (3%)	0 (0%)	1 (5%)	0 (0%)
Sequential trials analysis	18 (53%)	0 (0%)	12 (55%)	6 (50%)
Time-zero of the matched exposed participants for the unexposed ones	8 (24%)	0 (0%)	3 (14%)	5 (42%)
Other time points	2 (6%)	0 (0%)	2 (9%)	0 (0%)
Statistical modelling				
Use of a Directed Acyclic Graph	13 (13%)	1 (6%)	8 (14%)	4 (17%)
Selection bias				
IPCW	27 (27%)	2 (11%)	20 (34%)	5 (21%)
Adjustment for confounding				
Adjustment for baseline confounding	99 (99%)	18 (100%)	58 (100%)	23 (96%)
Adjustment for time-varying confounding	12 (12%)	0 (0%)	11 (19%)	1 (4%)
Propensity score analysis with baseline confounders	58 (58%)	16 (89%)	26 (45%)	16 (67%)
PS weighting	42 (72%)	12 (75%)	22 (85%)	8 (50%)
PS matching	25 (43%)	10 (62%)	8 (30%)	7 (44%)
PS stratification	1 (2%)	1 (6%)	0 (0%)	0 (0%)

PS adjustment	2 (3%)	1 (6%)	0 (0%)	1 (6%)
Use of high-dimension PS	2 (3%)	1 (6%)	1 (4%)	0 (0%)
Use of G-methods	20 (20%)	3 (17%)	13 (22%)	4 (15%)
G-formula	8 (8%)	2 (11%)	6 (10%)	0 (0%)
Doubly-robust methods	6 (6%)	1 (6%)	2 (3%)	3 (12%)
IPTW with time-varying weights	8 (8%)	0 (0%)	7 (12%)	1 (4%)
Controls for confounding	21 (21%)	5 (28%)	14 (24%)	2 (8%)
Negative outcomes	14 (67%)	3 (60%)	10 (71%)	1 (50%)
Positive outcomes	5 (24%)	2 (40%)	3 (21%)	0 (0%)
Negative exposures	1 (5%)	0 (0%)	1 (7%)	0 (0%)
Survival analysis	76 (76%)	14 (78%)	50 (86%)	12 (50%)
Adjusted hazard ratios	51 (68%)	11 (79%)	33 (66%)	7 (64%)
Weighted or standardised survival curves	35 (47%)	7 (50%)	25 (50%)	3 (27%)
Competing risks analysis	27 (27%)	5 (28%)	17 (29%)	5 (21%)
Subdistribution HR	9 (33%)	3 (60%)	6 (35%)	0 (0%)
Cause-specific HR	5 (19%)	0 (0%)	5 (29%)	0 (0%)
Use of composite outcome	13 (48%)	4 (80%)	5 (29%)	4 (80%)
Other methods to account for competing event or unclear method	5 (19%)	0 (0%)	4 (24%)	1 (20%)

*Time-zero is the start of follow-up. List of abbreviations: IPCW, inverse probability of censoring weighting; IPTW, inverse probability of treatment weighting; PS, propensity score; HR, hazard ratios. Studies that used IPTW only with a single weight at baseline were not deemed to apply g-methods but a regular propensity score analysis. Statistical techniques were not mutually exclusive (e.g. a sequential trial analysis with the use of a propensity score weighting in each sequential trial).

Table S10. Results of observational analyses

	Overall (N=100)	Studies emulating a real target trial (N=18)	Studies emulating a hypothetical target trial (N=58)	No target trial specified in the studies (N=24)
Reporting of flow diagram	81 (81%)	17 (94%)	48 (83%)	16 (67%)
Number of eligible participants to the analysis*				
Median [Q1;Q3]	9533 [1302;66739]	6137 [1015;17747]	11049 [2934;73373]	15318 [1101;178562]
Minimum / maximum	141 / 7619682	236 / 822640	178 / 3327088	141 / 7619682
Number of included participants in the analysis				
Median [Q1;Q3]	9286 [1264;70570]	2346 [1015;14076]	12693 [2872;87094]	8388 [481;105806]
Minimum / maximum	141 / 2698176	142 / 168692	178 / 2698176	141 / 2066296
First reported intervention effect				
Hazard ratio	45 (45%)	11 (61%)	28 (48%)	6 (25%)
Risk difference	27 (27%)	3 (17%)	18 (31%)	6 (25%)
Relative risk	8 (8%)	4 (22%)	3 (5%)	1 (4%)
Odds ratio	7 (7%)	0 (0%)	3 (5%)	4 (17%)
Mean difference	4 (4%)	0 (0%)	2 (3%)	2 (8%)
Other	9 (9%)	0 (0%)	4 (7%)	5 (21%)
Result statistically significant for the primary (or first reported) outcome	57 (57%)	6 (33%)	36 (62%)	15 (62%)

*In six studies, the number of eligible participants to the analysis was not evaluable. Abbreviation: Q1;Q3, first and third quartiles.

Table S11. Risk of bias of observational analyses

	Overall (N=100)	Target trial specified in studies (N=76)	Target trial not specified in studies (N=24)
Causal contrast of interest			
Assignment to intervention	52 (52%)	37 (49%)	15 (62%)
Starting and adhering intervention	48 (48%)	39 (51%)	9 (38%)
Bias due to confounding			
Low	0 (0%)	0 (0%)	0 (0%)
Moderate	94 (94%)	72 (95%)	22 (92%)
Serious	6 (6%)	4 (5%)	2 (8%)
No information	0 (0%)	0 (0%)	0 (0%)
Bias due to selection of participants			
Low	58 (58%)	47 (62%)	11 (46%)
Moderate	24 (24%)	17 (22%)	7 (29%)
Serious	16 (16%)	11 (14%)	5 (21%)
No information	2 (2%)	1 (1%)	1 (4%)
Bias in classification of interventions			
Low	61 (61%)	50 (66%)	11 (46%)
Moderate	29 (29%)	22 (29%)	7 (29%)
Serious	10 (10%)	4 (5%)	6 (25%)
No information	0 (0%)	0 (0%)	0 (0%)
Bias due to deviations from intended interventions			
Low	72 (72%)	54 (71%)	18 (75%)
Moderate	28 (28%)	22 (29%)	6 (25%)
Serious	0 (0%)	0 (0%)	0 (0%)
No information	0 (0%)	0 (0%)	0 (0%)
Bias due to missing data			
Low	68 (68%)	54 (71%)	14 (58%)
Moderate	25 (25%)	18 (24%)	7 (29%)
Serious	6 (6%)	4 (5%)	2 (8%)
No information	1 (1%)	0 (0%)	1 (4%)
Bias in measurement of outcomes			
Low	73 (73%)	60 (79%)	13 (54%)
Moderate	22 (22%)	15 (20%)	7 (29%)
Serious	4 (4%)	1 (1%)	3 (12%)
No information	1 (1%)	0 (0%)	1 (4%)

Bias in selection of the reported result			
Low	9 (9%)	6 (8%)	3 (12%)
Moderate	0 (0%)	0 (0%)	0 (0%)
Serious	3 (3%)	3 (4%)	0 (0%)
No information	88 (88%)	67 (88%)	21 (88%)
Overall bias			
Low	0 (0%)	0 (0%)	0 (0%)
Moderate	72 (72%)	56 (74%)	16 (67%)
Serious	28 (28%)	20 (26%)	8 (33%)

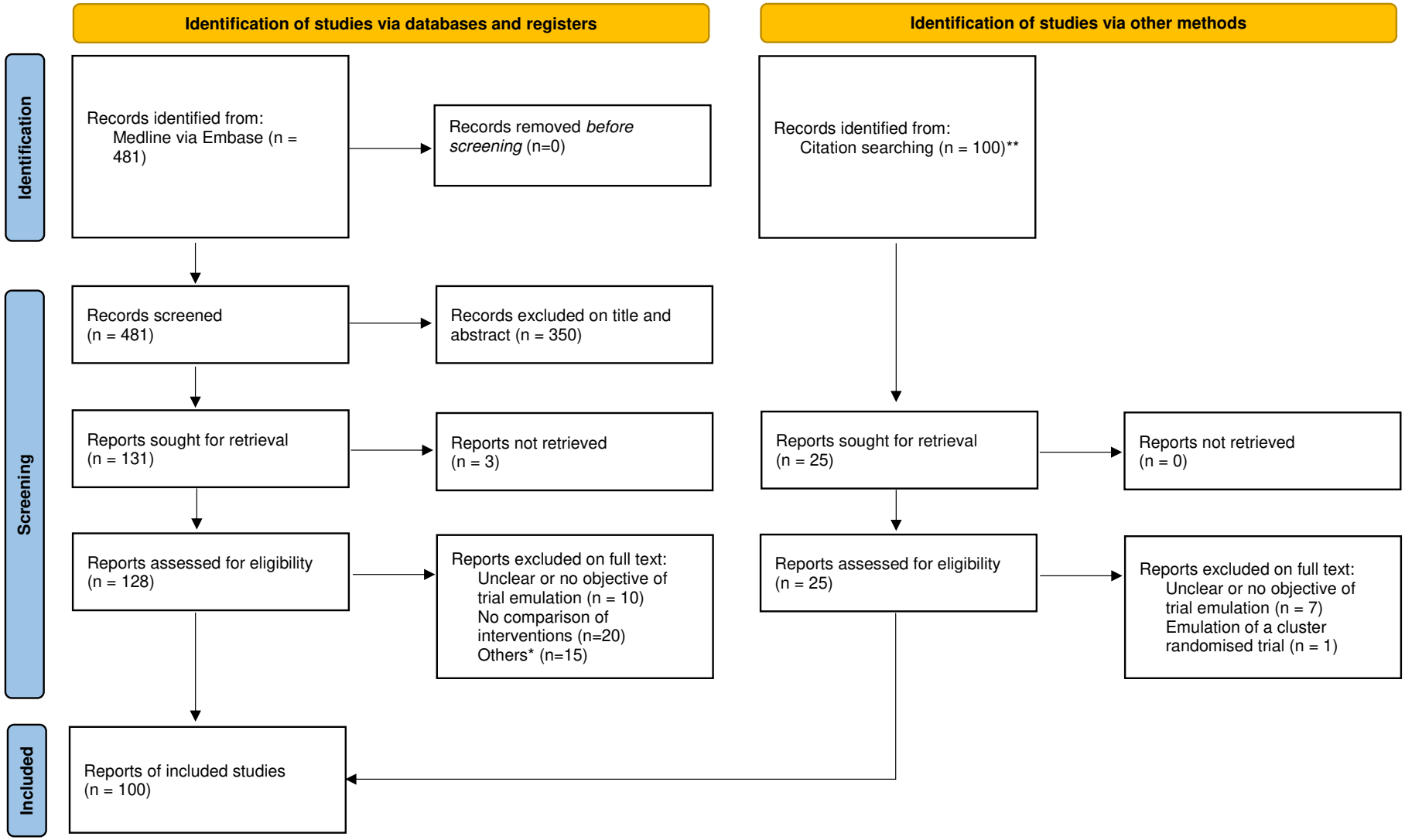
The target trial was deemed specified if at least one protocol component was documented. The risk of bias in the selection of the reported result was evaluated only in cases where an observational analysis protocol was available (N=12). For a given study, the overall risk of bias was the highest risk of bias observed in one of the individual risk of bias domains.

Table S12. Time-zero issues and biases that might arise, according to specification of the target trial

	Overall (n=100)	Target trial specified in studies (n=76)	Target trial not specified in studies (n=24)
Time-zero (start of follow-up) is defined	97 (97%)	74 (97%)	23 (96%)
Time-zero aligns with the moment of eligibility assessment	70 (72%)	56 (76%)	14 (61%)
Time-zero aligns with the moment of intervention assignment	66 (68%)	52 (70%)	14 (61%)
Alignment of moments of eligibility assessment, intervention assignment, and time-zero	56 (58%)	45 (61%)	11 (48%)
Biases that might arise from time-zero issues	41 (42%)	29 (39%)	12 (52%)
Prevalent user bias	6 (15%)	3 (10%)	3 (25%)
Prevalent user bias and selection bias due to post-intervention eligibility	12 (29%)	10 (34%)	2 (17%)
Immortal time bias and selection bias due to post-intervention eligibility	16 (39%)	10 (34%)	6 (50%)
Immortal time bias and misclassification of intervention	7 (17%)	6 (21%)	1 (8%)

The target trial was deemed specified if at least one protocol component was documented. The risk of bias that might arise was classified according to the four failures emulation categories proposed elsewhere[8,9]

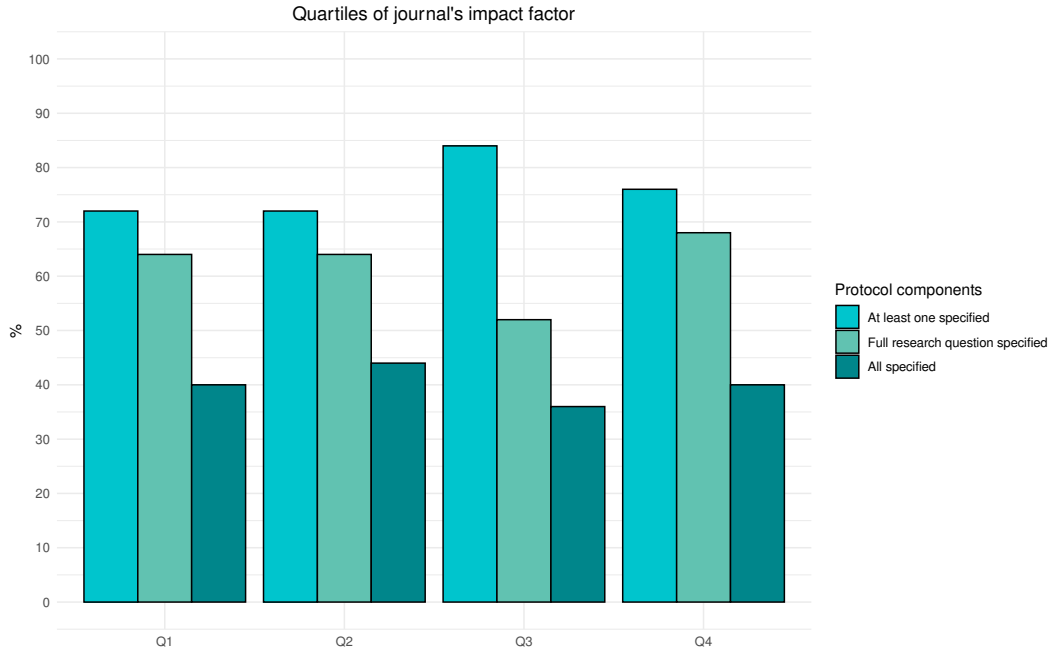
Figure S1. PRISMA 2020 flow diagram

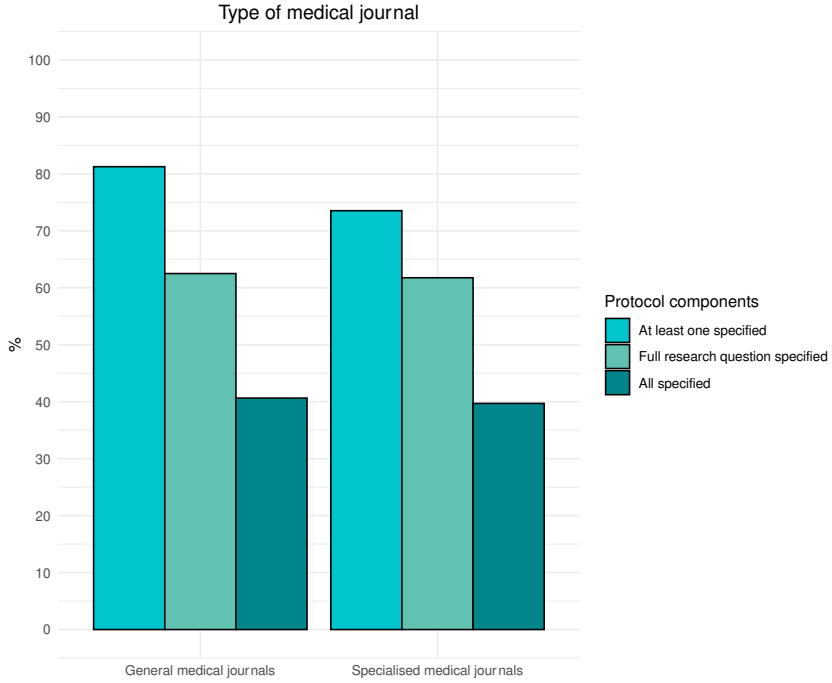


*Others: evaluation of statistical models by simulation studies (n=6), randomised trial, before-after study, preclinical study (n=3), emulation of a cluster randomised trial (n=2), use of randomised trial data (n=4).
**manual search using the search engine of the following journals: the Journal of the American Medical Association (JAMA), the New England Journal of Medicine (NEJM), the BMJ, The Lancet, Annals of Internal Medicine, JAMA Internal Medicine, Nature, Nature Medicine, the Lancet infectious Diseases, The Lancet Neurology, The Lancet diabetes and endocrinology, The Lancet Respiratory Medicine, The Lancet Global Health, the European Heart Journal, The Journal of Hepatology, Gut, The Journal of Clinical Oncology, World Psychiatry, Annals of the Rheumatic Diseases, and Blood.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/

Figure S2. Reporting of the target trial according to quartiles of journal's impact factor, type of journal and position of epidemiologist or biostatistician in authorship





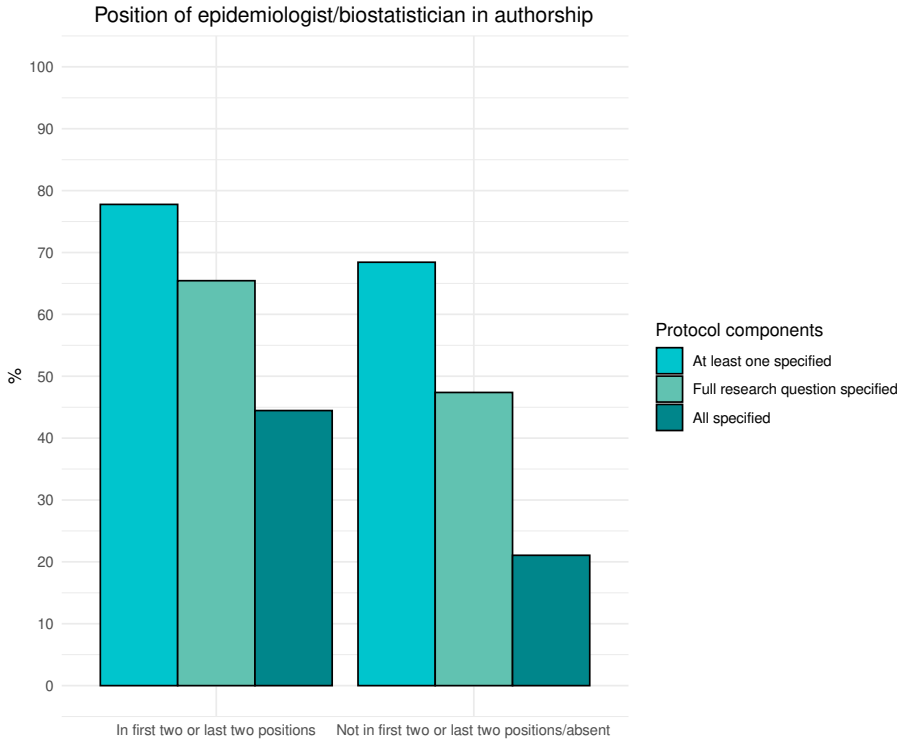
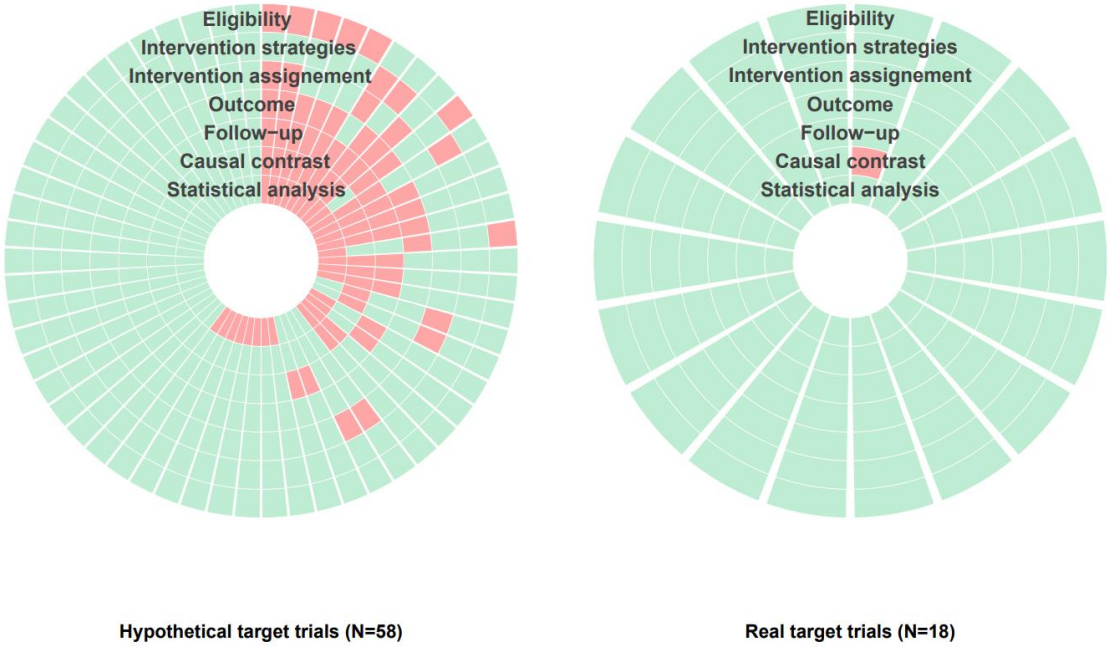

Position of epidemiologist/biostatistician in authorship
100
90
80
70
60
50
40
30
20
10
0
%
In first two or last two positions
Not in first two or last two positions/absent
Protocol components
At least one specified
Full research question specified
All specified

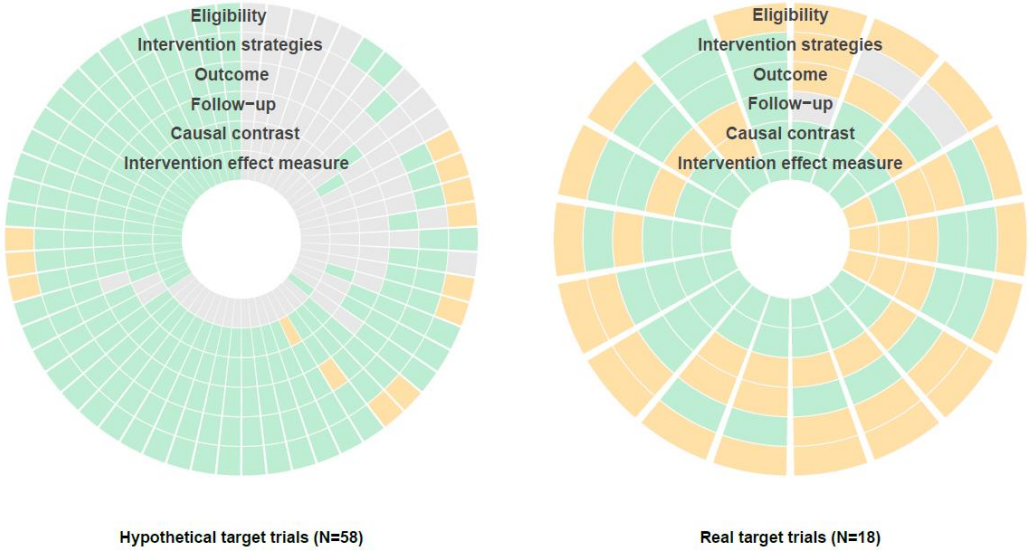
Figure S3. Specification of the target trial characteristics, for hypothetical and real target trials



Hypothetical target trials (N=58) **Real target trials (N=18)**

Legend: Each radius of the disk represents an observational analysis and each concentric layer one of the key protocol components of the target trial to specify. Green indicates that the protocol component is specified in the description of the target trial and red indicates that no information is available. Two real target trials were emulated twice in two separate observational analyses. As a result, the description of real target trials corresponds to a total of 16 unique randomised trials.

Figure S4. Consistency of the target trial and its emulation for hypothetical and real target trials.



Legend: Each radius of the disk represents an observational analysis and each concentric layer one of the key protocol components of the target trial that should be consistent for the target trial and its emulation. Green indicates consistency between the target trial and its emulation and orange inconsistency between the target trial and its emulation. Grey indicates that the protocol component was not specified in the target trial or in its emulation, so the comparison was not possible. See supplementary Table S3 for the reasons for deeming the target trial and its emulation inconsistent.

Figure S5. Traffic-light plot of ROBINS-I assessment

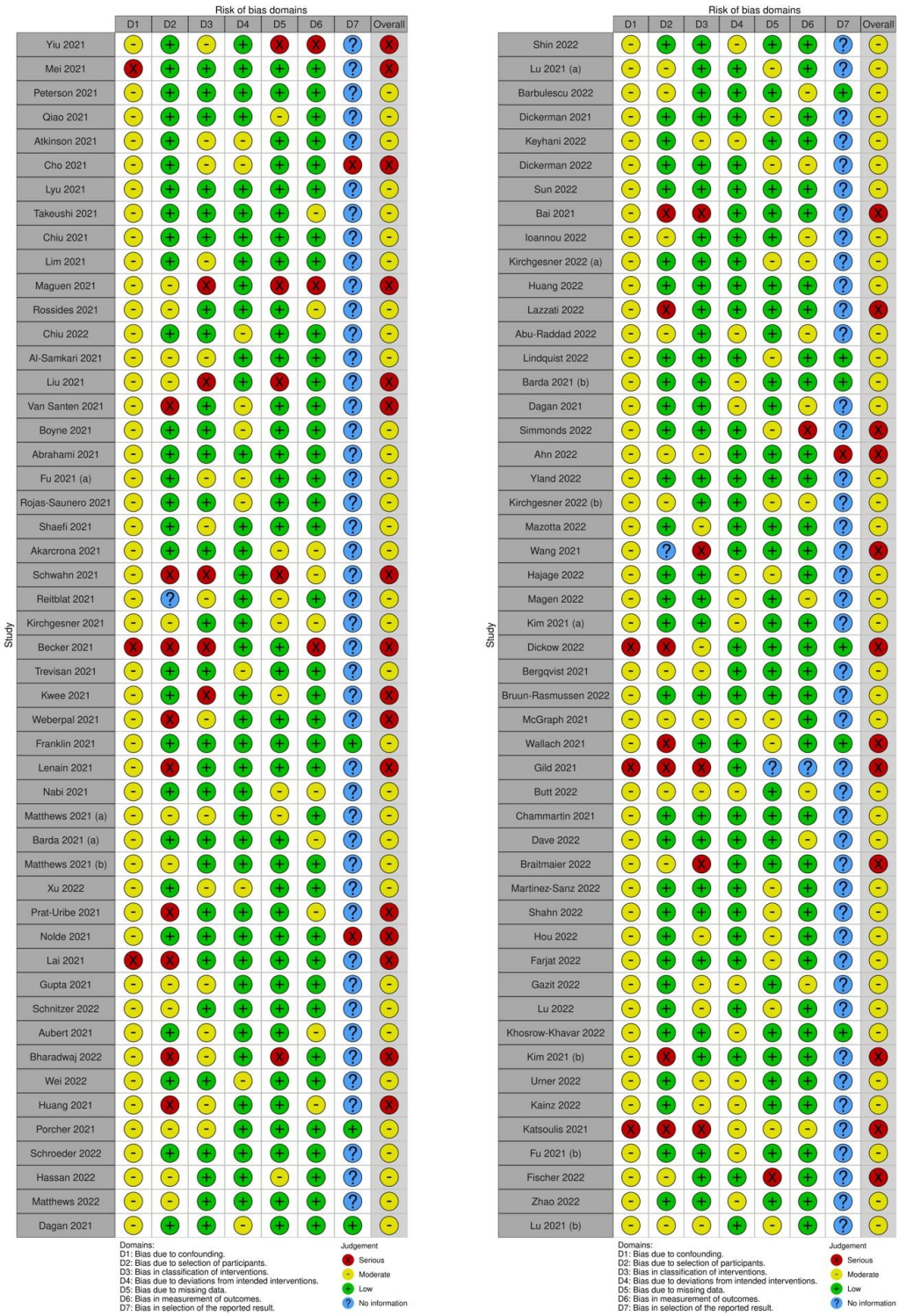
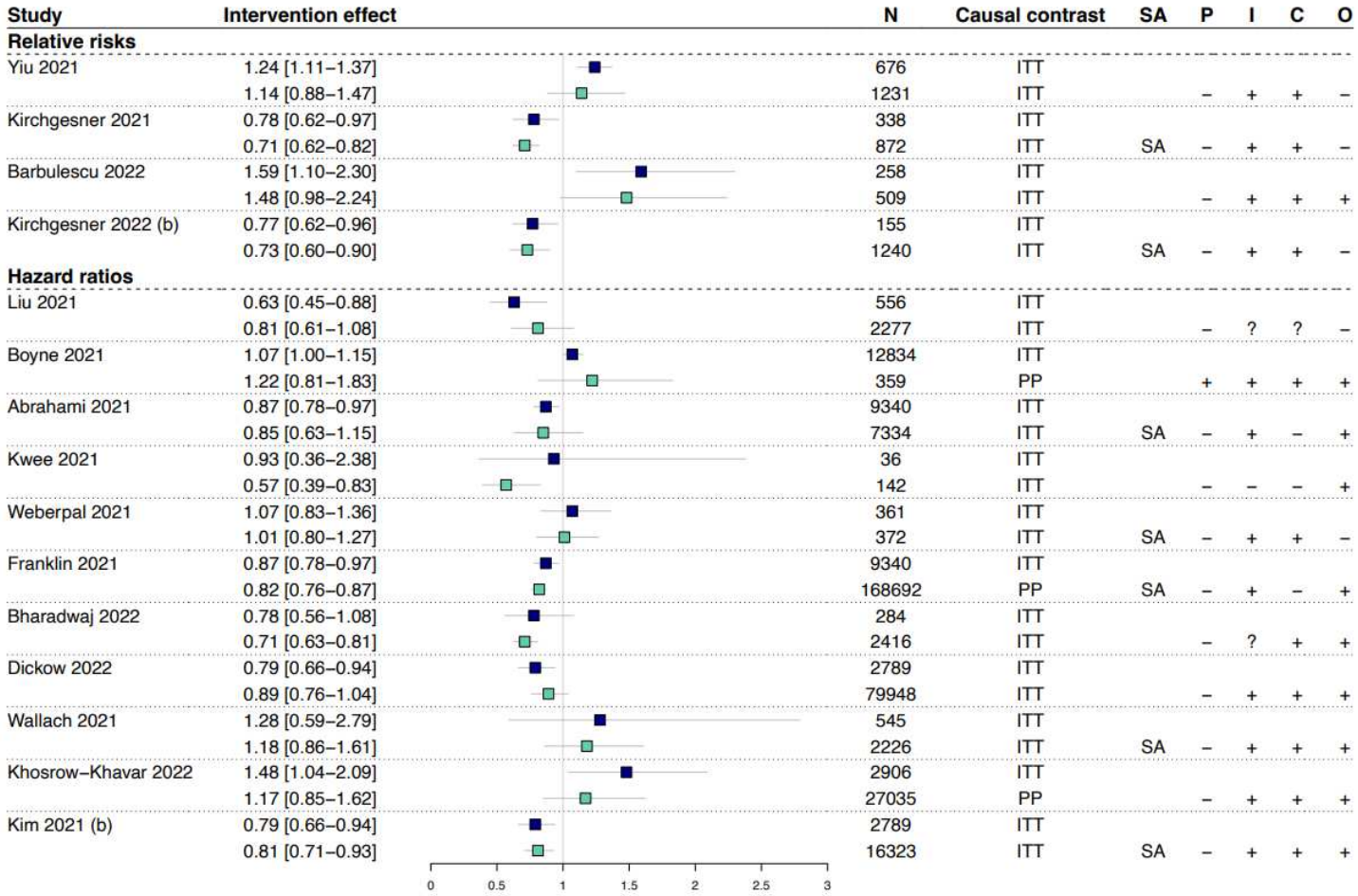


Figure S6. Comparison of intervention effects in real target trials and their observational emulations.



Legend: the intervention effects in navy blue are the ones produced in the randomised trials, and the ones in aquamarine blue are those from their emulations. The column “intervention effect” contains the point estimate for each analysis and its 95% confidence interval, except for the randomised trial point estimates of studies 76 and 93, for which a 96% confidence interval was reported. The column SA reports the full statistical significance agreement (SA if present). The column “N” contains the number of participants included in the analysis to estimate the intervention effect. The column “Causal contrast” reports whether the analysis was an intention-to-treat (ITT) or modified ITT analysis or a per protocol (PP) analysis. Of note, the causal contrasts are those from our own assessment (because authors did not always report it). For the rules to classify the causal contrast, see Appendix 3. The columns P, I, C, O represent the consistency (consistent: +, inconsistent: -, not evaluable: ?) of the target trial and its emulation for the definition of eligibility criteria, intervention, control and outcome, respectively. In Boyne 2021, Abrahami 2021, Franklin 2021 and Khosrow-Khavar 2022, a non-inferiority trial was emulated; all other observational analyses emulated a superiority trial. In Abrahami 2021 and Franklin 2021, the same trial was emulated (ClinicalTrials.gov identifier NCT01179048) as well as in Dickow 2022 and Kim 2021 (b) (NCT01288352). For references of studies, see Appendix 4.

Appendix 1. Protocol and deviations

The protocol was written before starting the study. We initially planned a meta-epidemiological analysis to compare the magnitude of treatment effect between real target trials and their observational counterpart, which was not conducted due to lack of studies that emulated a real target trial in our study sample. The search algorithm mentioned as “in appendix 1” is identical to the one provided in the appendix 2 of this supplemental material.

Protocol

Introduction

Pharmacoepidemiology focuses on the effectiveness and safety of healthcare products (drugs or medical devices) in real-world settings, after marketing authorization. This assessment is central for decision makers[1], but also for patients and caregivers, to get the most accurate information about the expected efficacy and safety according to the patient's medical background[2].

Although randomized controlled trials are considered the gold standard for evaluating a healthcare product, this design is often synonym of restrictive eligibility criteria, small sample size, and short-term outcomes, and may not be adapted for the aims of pharmacoepidemiologic studies[3]. On the other side, even if they are inherently far much closer from the real-world conditions, observational studies are considered to provide a lower level of evidence[4], because of the biases they can suffer from. To address the issue of biases and warrant causal inference in observational studies, Hernán and colleagues[5] formalized a framework to emulate randomized trials in the context of large datasets, including claim databases, cohorts or registers. Basically, this approach consists in fitting as much as possible to a hypothetical or real clinical trial protocol (the "target" trial). All protocol components i.e. eligibility criteria, treatment strategies, treatment assignment, outcomes, follow-up, causal contrast and statistical analysis are defined for the target trial and then emulated within observational data. In addition, the adequate statistical strategy can be applied only if the emulated trial is defined precisely[6]. Some well-conducted emulated trials were able to replicate the treatment effect of their randomized controlled trial counterparts[7]. However, like any observational study, these emulated target trials can be jeopardized by methodological flaws, and more specifically, like any pharmacoepidemiology study, by immortal-time bias, that can occurs if time-zero, the specification of eligibility criteria and treatment assignment do not match[8,9].

The number of new publications self-reported emulated target trials based on observational data is rapidly growing, and may represent a subset of heterogeneous studies. Here we aim to methodologically review for the first time all the self-reported emulated trials, and studies that claims to replicate real clinical trials with observational data, and to evaluate their methodology.

Methods

This is a methodological review of emulated target trials complying with the PRISMA statement[10] (the checklist will be provided with the final report). This protocol was written before starting the review.

Search strategy

We will search Medline and Pubmed (not Medline) with the Embase search engine, for emulated target trials or studies that aim to replicate a clinical trial with observational data. We chose Embase over Pubmed search engine because of broader options for the Embase search equations. The search algorithm is provided in appendix 1. To get a comprehensive overview, we will also interrogate search engines with the key words “emulate”, “target trial”, “trial emulation” of the following general journals: the Journal of the American Medical Association (JAMA), the New England Journal of Medicine (NEJM), the BMJ, The Lancet, Annals of Internal Medicine, JAMA Internal Medicine, Nature, Nature Medicine. We will also search for additional references in specialized journals: the Lancet infectious Diseases, The Lancet Neurology, The Lancet diabetes and endocrinology, The Lancet Respiratory Medicine, The Lancet Global Health, the European Heart Journal, The Journal of Hepatology, Gut, The Journal of Clinical Oncology, World Psychiatry, Annals of the Rheumatic Diseases, and Blood. These journals have been selected because they have the highest impact factor among the journals that publish original research of each medical specialty in 2021. We added this manual search because abstracts may not always mention the use of the target trial emulation framework, and search engines of journal websites can perform full-text search for key words. The search was performed on July 3rd 2022.

Eligibility criteria

Any study that states to aim or is designed to emulate a hypothetical clinical trial or replicate an existing one from observational data and comparing two interventions (or exposures) or more will be eligible, published between 1st January 2021 and 3rd July 2022. Studies interested in a prognostic factor or interventions delivered in cluster of patients will be excluded, even if designed as an emulated target trial. We will

additionally exclude all interventional studies (or observational analyses that used data from it) , before-after studies, narrative or systematic reviews, preclinical and veterinary medicine studies, genetic studies (including genome-wide association studies and mendelian randomization designs), pharmacokinetics and / or pharmacodynamics studies, studies that aims to validate a statistical model (e.g. with Monte-Carlo simulations), health economic studies, and case reports.

Selection process

The screening of titles and abstracts, as well as the assessment of full texts will be independently performed by two reviewers (NST and CS or GM) and any disagreement will be solved by discussion to reach consensus. The selection process will be represented with a flow diagram[11].

Search of corresponding sources

We will interrogate the databases of registered studies (clinicaltrials.gov, International Clinical Trials Registry Platform: ICTRP) and the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) website to retrieve the protocols of the selected studies.
In case of emulated trial replicating a real clinical trial, we will search the corresponding target trial from the reference list of the emulated trial and by interrogating study registries.

Data extraction

Two reviewers (NST and GM) will independently extract data for all selected studies according to a standardized data extraction form. Disagreements will be solved by consensus with the help of a third reviewer if needed.
We will collect the following items:

- General characteristics: name of authors, date and journal of publication, source of funding
- Author(s) affiliated to a department of Biostatistics, Epidemiology, Public Health, or Data sciences, or a Clinical Research Unit

- Aim of emulation: e.g., emulate an unfeasible or unethical clinical trial in real life, find a potential new indication of an already approved intervention, explore discrepancies between previously published observational and interventional studies, initial assessment of efficacy before completion of clinical trial, as reported in the introduction or in methods
- Reporting by authors of increased reliability of the study by the use of the target trial emulation framework
- Cited references for methodology of target trial emulation (articles, others)
- Source of data:
    - Routinely collected data[12], claim and/or electronic health records databases of for profit or public and not-for-profit healthcare providers, registries, cohort studies
    - Patients covered by the database: specific population (e.g. veterans), countries of residence of patients.
- Pre-registered or published protocol for the observational analysis
- Target trial:
    - Specification in the article (at least one protocol component of the target trial defined)
    - Real or fictional, reference of the potential published target(s) trial(s)
    - Only for real target trials:
        - Blinding (open-label trial, blinding of participants, caregivers, investigators, outcome assessors)
        - Number of centers
        - Countries in which the study was conducted
- Methodology of both the target trial and its emulation:
    - Definition of the objective:
        - Population: new- users, prevalent users
        - Intervention: pharmacological, non-pharmacological, static or dynamic treatment strategies
        - Control: active comparator, no treatment
        - Outcomes
    - Design of study
        - Parallel or cross-over design
        - Superiority, non-inferiority or equivalence trial

- Eligibility criteria
- Modalities of treatment assignment i.e. randomization or statistical methods to simulate randomization (see below)
- Definition of the primary outcome:
  - Unique, multiple or composite
  - Type of variable (binary, continuous, time-to-event)
  - Aim of analysis: efficacy, safety or both
- Definition of time-zero (start of follow-up) and duration of follow-up
- Statistical analysis strategy applied to the primary (or targeted) outcome; for co-primary outcomes, only the first reported one will be studied:
  - Adjustment variables (baseline or post-baseline confounding, time-varying confounding) and methods to adjust
  - Causal contrast(s) employed in the statistical analysis and their definitions
    - Intention-to-treat analysis
    - Per protocol analysis
  - Treatment effect measure (RR, HR, OR, difference in means, difference in probability, etc.)
  - Computed sample size, actual sample size for real clinical trials

- For the emulated trial, we will also assess the following points:
  - Eligibility criteria: timing for eligibility, and time spent in the database before being potentially eligible (look-back period), use of post-baseline characteristics (collected after time-zero of the study, see below for its definition)
  - Definition of time-zero (start of follow-up):
    - Unique or multiple eligible time for patients
    - Modalities to define time-zero: first eligible time, random choice of eligible time, nested trials including all eligible times
  - Grace period, which is the delay between meeting all eligibility criteria and starting treatment and maximal duration of this grace period
  - Statistical analysis (only for the first reported targeted outcome):
    - Modalities for adjusting for confounders: e.g. use of a propensity score, inverse probability weighting, g-methods, doubly robust methods

        - Use of cloning, censoring and weighting[13]
        - Methods to take into account informative censoring (e.g. by inverse probability of censoring weighting)
        - Internal controls in the analysis for residual confounding
- Results for the targeted outcome in the emulated target trial, and in real clinical trials, if applicable
- Consistency between the target trial specification and target trial emulation:
    - In methodology, for key components (eligibility criteria, treatment strategies, outcomes, follow-up, chosen causal contrast and intervention effect measure for statistical analysis) and reason for.
    - In results in case of real target trial with evaluation of the regulatory agreement[14]

Evaluation of consistency will be performed according to our own evaluation.

Risk of bias

Two reviewers (NST and GM) will assess the risk of bias of included studies according to an adapted version of the ROBINS-tool[15]. Disagreements will be solved by consensus.

Briefly, the ROBINS-tool proposes to assess the risk of bias of non-randomized studies of interventions, according to consideration of a target trial for the observational analysis, and answering signaling questions relative to pre-intervention, at intervention and post intervention potential biases.

Due to the specific risk of bias of emulated trials according to the timing of assessment of eligibility, the timing of assignment to treatment strategy and time-zero in the study[8], we will add one signaling question that will assess if these three times coincide, and evaluate the risk of selection and/or immortal-time bias associated with[8].

Statistical analysis

Descriptive statistics

The methodological features (including the risk of bias) of both the target and the emulated trials will be described with numbers and percentages for qualitative variables and median with first and third quartiles for quantitative variables. We will also present qualitatively the causal inference statistical techniques employed in the observational analysis and methods to deal with immortal-time bias. Characteristics of emulated target trials will be described according to the following subgroups: replicate a real clinical trial, emulate a fictional target trial or no target trial specified.

Meta-epidemiological analysis

If the number of real target trials is sufficient, with available results and with the same outcome between emulated trials and real target trials, the difference in treatment effects between the target trial and the emulated trial will be evaluated by a two-step meta-epidemiological approach[16]. First, for each real target trial and its emulation we will compute the difference between treatment effects (ratio of odds ratio for binary outcomes, or difference in standardized mean difference for continuous outcomes).

The difference of treatment effects will be then pooled by a random-effects meta-analysis model[17].

Statistical analysis will be performed with R software, version 4 (R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/). The package meta will be used for the meta-epidemiological analysis.

Funding

No dedicated funding will be used to conduct this review.

References

1 Skydel JJ, Luxkaranayagam AT, Dhruva SS, *et al.* Analysis of Postapproval Clinical Trials of Therapeutics Approved by the US Food and Drug Administration Without Clinical Postmarketing Requirements or Commitments. *JAMA Netw Open*. 2019;2:e193410.

2 Dreyer NA. Making Observational Studies Count: Shaping the Future of Comparative Effectiveness Research. *Epidemiology*. 2011;22:295–7.

3 Strom BL, Kimmel SE, Hennessy S, editors. *Pharmacoepidemiology*. 1st ed. Wiley 2019. https://doi.org/10.1002/9781119413431

4 Guyatt GH, Oxman AD, Vist GE, *et al.* GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924–6.

5 Hernán MA, Alonso A, Logan R, *et al.* Observational Studies Analyzed Like Randomized Experiments: An Application to Postmenopausal Hormone Therapy and Coronary Heart Disease. *Epidemiology*. 2008;19:766–79.

6 Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available: Table 1. *Am J Epidemiol*. 2016;183:758–64.

7 Dickerman BA, García-Albéniz X, Logan RW, *et al.* Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med*. 2019;25:1601–6.

8 Hernán MA, Sauer BC, Hernández-Díaz S, *et al.* Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70–5.

9 Suissa S. Immortal Time Bias in Pharmacoepidemiology. *Am J Epidemiol*. 2008;167:492–9.

10 Page MJ, McKenzie JE, Bossuyt PM, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;n71.

11 Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6:e1000097.

12 Nicholls SG, Langan SM, Benchimol EI. Routinely collected data: the importance of high-quality diagnostic coding to research. *Can Med Assoc J*. 2017;189:E1054–5.

13 Maringe C, Benitez Majano S, Exarchakou A, *et al.* Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit

of major surgery for elderly lung cancer patients using observational data. *Int J Epidemiol*. 2020;49:1719–29.

14 Franklin JM, Pawar A, Martin D, *et al.* Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. *Clin Pharmacol Ther*. 2020;107:817–26.

15 Sterne JA, Hernán MA, Reeves BC, *et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;i4919.

16 Sterne JAC, Jüni P, Schulz KF, *et al.* Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. *Stat Med*. 2002;21:1513–24.

17 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7:177–88.

Appendix 2. Search equation used in Embase

#1 ([medline]/lim OR [pubmed-not-medline]/lim)
#2 (‘simulat*’ OR ‘emulat*’ OR ‘replicat*’) NEAR/12 (‘trial’ OR ‘trials’ OR 'intervention study' OR ‘randomiz*’ OR ‘randomis*’)
#3 (‘target trial’:ti,ab,kw OR ‘target trials’:ti,ab,kw)
#4: #2 OR #3
#5 (‘real-world’:ti,ab,kw OR ‘real-world’:ti,ab,kw OR 'comparative effectiveness':ti,ab,kw OR 'comparative effectiveness'/exp OR 'observational':ti,ab,kw OR 'observational study'/exp OR ‘non-interventional’:ti,ab,kw OR ‘noninterventional’:ti,ab,kw OR ‘non-interventional’:ti,ab,kw OR ‘non-experimental’:ti,ab,kw OR 'quasi experimental study'/exp OR 'quasi experimental':ti,ab,kw OR 'quasi-experimental':ti,ab,kw OR 'quasiexperimental':ti,ab,kw OR ‘nonexperimental’:ti,ab,kw OR ‘non-experimental’:ti,ab,kw OR 'natural experiment'/exp OR 'natural experiment':ti,ab,kw OR ‘non-randomized’:ti,ab,kw OR ‘non-randomised’:ti,ab,kw OR ‘nonrandomised’:ti,ab,kw OR ‘non-randomized’:ti,ab,kw OR ‘non-randomised’:ti,ab,kw OR ‘health-care database’:ti,ab,kw OR ‘healthcare database’:ti,ab,kw OR ‘health-care database’:ti,ab,kw OR ‘health-care databases’:ti,ab,kw OR ‘healthcare databases’:ti,ab,kw OR ‘health-care databases’:ti,ab,kw OR ‘claim-data’:ti,ab,kw OR ‘claim-data’:ti,ab,kw OR ‘claims-data’:ti,ab,kw OR ‘claims data’:ti,ab,kw OR 'administrative health data':ti,ab,kw OR 'administrative health data'/exp OR 'patient coding':ti,ab,kw OR 'patient coding'/exp OR 'administrative claims (health care)'/exp OR ‘administrative claim’:ti,ab,kw OR ‘administrative-claim’:ti,ab,kw OR ‘administrative-claims’:ti,ab,kw OR ‘administrative-claims’:ti,ab,kw OR ‘claim database’:ti,ab,kw OR ‘claim-database’:ti,ab,kw OR ‘claims-database’:ti,ab,kw OR ‘claims database’:ti,ab,kw OR 'electronic medical record'/exp OR ‘electronic medical record’:ti,ab,kw OR ‘electronic medical file’:ti,ab,kw OR ‘electronic medical records’:ti,ab,kw OR 'electronic health records':ti,ab,kw OR ‘electronic medical files’:ti,ab,kw OR 'electronic health record'/exp OR 'electronic health record':ti,ab,kw OR ‘electronic health records’:ti,ab,kw OR ‘electronic database’:ti,ab,kw OR ‘electronic-database’:ti,ab,kw OR ‘electronic-databases’:ti,ab,kw OR ‘electronic-databases’:ti,ab,kw OR ‘routinely collected data’:ti,ab,kw OR 'routinely collected health data'/exp OR ('patient registry'/exp OR ‘registries’:ti,ab,kw OR ‘registry’:ti,ab,kw) OR ('cohort analysis'/exp OR ‘cohort’:ti,ab,kw OR ‘cohorte’:ti,ab,kw OR ‘cohorts’:ti,ab,kw) OR ('longitudinal study'/exp OR (‘longitudinal’:ti,ab,kw AND ‘studies’:ti,ab,kw) OR ‘longitudinal studies’:ti,ab,kw OR

('longitudinal':ti,ab,kw AND 'study':ti,ab,kw) OR 'longitudinal study':ti,ab,kw) OR ('prospective study'/exp OR ('prospective':ti,ab,kw AND 'studies':ti,ab,kw) OR 'prospective studies':ti,ab,kw OR ('prospective':ti,ab,kw AND 'study':ti,ab,kw) OR 'prospective study':ti,ab,kw) OR ('retrospective study'/exp OR ('retrospective':ti,ab,kw AND 'studies':ti,ab,kw) OR 'retrospective studies':ti,ab,kw OR ('retrospective':ti,ab,kw AND 'study':ti,ab,kw) OR 'retrospective study':ti,ab,kw))

#6 ('phase 1 clinical trial'/exp OR 'phase 2 clinical trial'/exp OR 'phase 3 clinical trial'/exp OR 'systematic review':ti OR 'systematic review'/exp OR 'Mendelian randomization analysis'/exp OR 'meta-analysis':ti OR 'economic evaluation'/exp)

#7: #1 AND #4 AND #5 NOT #6

#8: #7 AND (2021:py OR 2022:py)

Appendix 3. Supplemental methods

Search of corresponding sources

For observational analyses with a preregistered protocol, we retrieved the protocols of the selected studies on websites for study registration, as for example Clinicaltrial.gov or the European Union electronic Register of Post-Authorisation Studies (EU PAS register).
In case of a study attempting to replicate a real randomised trial, we also retrieved information on this clinical trial most of the time from the original article; and for one analysis on study registration websites (clinicaltrials.gov, International Clinical Trials Registry Platform: ICTRP), as the results of the trial were not reported in a publication.

Data extraction and risk of bias assessment

Two reviewers (NST, GLM or MB) independently extracted data for all selected studies according to a standardised data extraction form. Disagreements were solved by consensus. We collected the following items:

- General characteristics: name of authors, date and journal of publication, location of the corresponding author, source of funding
- Author(s) affiliated to a department of Biostatistics, Epidemiology, Public Health, or Data sciences, or a Clinical Research Unit
- Reason to conduct the study, e.g., emulate an unfeasible or unethical clinical trial in real life, as reported in the introduction or in methods
- Reporting by authors of increased reliability of the study by the use of the target trial emulation framework
- Cited references for methodology of target trial emulation (articles, others)
- Source of data:
    - Routinely collected data[5], claim and/or electronic health records databases of for profit or public and not-for-profit healthcare providers, registries, cohort studies
    - Patients covered by the database: specific population (e.g., veterans), continents of residence of patients.
- Pre-registered or published protocol for the observational analysis

- Definition of the target trial, that is, at least one characteristic defined among the following ones: eligibility criteria, intervention strategies, intervention assignment, outcomes, follow-up, causal contrasts and statistical analysis.
- Target trial (if they were several target trials in the article, only information on the first reported one was collected):
    - Real or hypothetical, reference of the potential published target trial
    - Only for real target trials:
        - Period of recruitment and duration
        - Blinding (open-label trial, blinding of participants, caregivers, investigators, outcome assessors)
        - Number of centres
        - Countries in which the study was conducted
- Methodology of both the target trial and its emulation:
    - Study design:
        - Parallel or cross-over design
        - Superiority, non-inferiority or equivalence trial
    - Eligibility criteria:
        - New- users, prevalent users
        - Definition of eligibility criteria on characteristics observed after the start of the follow-up (post-baseline characteristics)
    - Intervention strategies to be compared (regarding the first reported ones):
        - Intervention: pharmacological, non-pharmacological, static or dynamic (i.e. the intervention can be modified according to post-baseline characteristics), joint interventions (several interventions at the same time are started at baseline)[1,6], grace period to receive the intervention, that is, a delay between meeting all eligibility criteria and starting of intervention. The intervention was deemed complex if it met any of the following criteria: being dynamic, joint, or implemented with a grace period.
        - Control: active comparator, no intervention or usual care, other comparator strategies (e.g., delayed intervention)
        - Main outcome and its time point
    - Modalities of intervention assignment (randomisation, index date for the intervention strategies)

- Definition of the primary outcome:
  - Type of variable (binary, continuous, time-to-event)
  - Time point of the outcome
  - Aim of analysis: efficacy, safety or both
- Definition of time-zero (start of follow-up) and duration of follow-up
- Causal contrast(s) (as reported by authors and according to our own evaluation):
  - Intention-to-treat analysis
  - Per protocol analysis
  - Other causal contrast
- Statistical analysis strategy applied to the primary (or targeted) outcome; for co-primary outcomes, only the first reported one was studied:
  - Adjustment variables (baseline or post-baseline confounding, time-varying confounding) and methods to adjust
  - First reported intervention effect measure (relative risk, hazard ratio, odds ratio, difference in means, difference in probability, etc.)
  - Computed sample size, actual sample size for real randomised trials

- For the observational analyses, we also assessed the following features:
  - Eligibility criteria: moment(s) when eligibility criteria are met, and time spent in the database before being potentially eligible (look-back period)
  - Definition of time-zero (start of follow-up): first eligible time, random choice of an eligible time, emulation of sequential trials[4,7]
  - Statistical analysis (only for the first reported targeted outcome):
    - Methods for adjusting for confounders, e.g., use of a propensity score[8,9] or g-methods[6]
    - Emulation of sequential trials[4,7]
    - Use of the cloning, censoring and weighting method[10,11]
    - Use of inverse probability of censoring weighting to account for selection bias[12]
    - Methods to account for competing risks[13]
    - Internal controls (positive or negative outcomes, positive or negative exposures)[14] and reporting of E-value, that is, the minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the intervention and outcome, conditional

on the measured confounders, to fully explain away the intervention–outcome association[15],

- Results for the targeted outcome in the emulated target trial, and statistical significance (if the tested statistical hypothesis was not specified, the absence of a difference between groups was assumed to be the null hypothesis), and in the real targeted randomised trial, if applicable
- Consistency between the target trial specification and target trial emulation:
    - In the methods, for key components (eligibility criteria, intervention strategies, outcomes, follow-up, causal contrasts, intervention effect measure chosen for statistical analysis) and reason for. The rules used to deem the target trial and its emulation consistent or inconsistent are reported in the Table S3.
    - In the results in case of emulation of a real target trial, by assessing the full statistical significance agreement, as reported elsewhere[16]. The full statistical significance agreement was met if the estimates of the intervention effect in the target trial and its emulation were on the same side of the null (zero in case of superiority target trial or the non-inferiority margin in case of a non-inferiority target trial)

Risk of bias

Two reviewers (NST, GLM or MB) assessed the risk of bias of included studies according to an slightly modified version of the ROBINS-tool[17]. Disagreements were solved by consensus.

The ROBINS-I tool is designed to evaluate the risk of bias in non-randomized studies of interventions. It takes into account the concept of a target trial for the observational analysis and involves answering specific questions related to potential biases before, during, and after the intervention. The tool also requires also some expert knowledge from the reviewers relative to the potential confounders and cointerventions for the given research question. As we could not provide this knowledge for all research questions, our assessment was based on the confounders and cointerventions that were listed by authors in the articles and if the statistical analysis account for them properly.

We added a signalling question to this tool, “Do timepoints of eligibility assessment, intervention assignment and start of follow-up align for most of participants?”, that overlaps the signalling questions of the selection bias domain “Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?” and “Do start of follow-up and start of intervention coincide for most participants?”, and evaluated the biases that might arise accordingly, as reported elsewhere[2,18].

In the ROBINS-I tool, signalling questions are adapted according to the causal contrast of interest in the analysis (“to assess the effect of assignment to intervention”, which is an observational analogue of intention-to-treat effect, or “to assess the effect of starting and adhering to intervention”, which is an observational analogue of the per protocol effect). In the review, we chose to assess the risk of bias of the first reported causal contrast for the main outcome (or first reported outcome) in the analysis. If the causal contrast was not defined by authors, we classified the studies according the following rules: for as-treated at baseline analyses of sustained interventions, or assignment to interventions based on initial prescription, the effect was classified as an observational analogue of an intention-to-treat effect; for on-treatment analyses of sustained or not-sustained strategies, the effect was classified as an observational analogue of a per protocol effect. The risk of bias in selection of the reported result was assessed only when a protocol was available for the observational analysis. Finally, the risk of bias was assessed among studies that specified at least one key protocol component of the target trial and those that did not. The ROBINS-I assessment was reported with summary plots and traffic-light plots, using the robvis tool[19].

Appendix 4. List of the 100 studies included in the review

	Reference
Yiu 2021	Yiu, Zenas Z. N., Kayleigh J. Mason, Philip J. Hampton, Nick J. Reynolds, Catherine H. Smith, Mark Lunt, Christopher E. M. Griffiths, Richard B. Warren, et BADBIR Study Group. « Randomized Trial Replication Using Observational Data for Comparative Effectiveness of Secukinumab and Ustekinumab in Psoriasis: A Study From the British Association of Dermatologists Biologics and Immunomodulators Register ». JAMA Dermatology 157, no 1 (January 1, 2021): 66-73. https://doi.org/10.1001/jamadermatol.2020.4202.
Mei 2021	Mei, Hao, Jiping Wang, et Shuangge Ma. « An Emulated Target Trial Analysis Based on Medicare Data Suggested Non-Inferiority of Dabigatran versus Rivaroxaban ». Journal of Clinical Epidemiology 139 (November 2021): 28-37. https://doi.org/10.1016/j.jclinepi.2021.07.001.
Peterson 2021	Peterson, Rosemary G., Rui Xiao, Hannah Katcoff, Brian T. Fisher, et Pamela F. Weiss. « Effect of First-Line Biologic Initiation on Glucocorticoid Exposure in Children Hospitalized with New-Onset Systemic Juvenile Idiopathic Arthritis: Emulation of a Pragmatic Trial Using Observational Data ». Pediatric Rheumatology Online Journal 19, no 1 (July 5, 2021): 109. https://doi.org/10.1186/s12969-021-00597-z.
Qiao 2021	Qiao, Nidan, Ming Shen, Wenqiang He, Min He, Zhaoyun Zhang, Hongying Ye, Xuefei Shou, Yongfei Wang, et Yao Zhao. « Comparative Effectiveness of Endoscopic versus Microscopic Transsphenoidal Surgery for Patients with Growth Hormone Secreting Pituitary Adenoma: An Emulated Trial ». Clinical Neurology and Neurosurgery 207 (August 2021): 106781. https://doi.org/10.1016/j.clineuro.2021.106781.
Atkinson 2021	Atkinson, Andrew, Marcel Zwahlen, Diana Barger, Antonella d'Arminio Monforte, Stephane De Wit, Jade Ghosn, Enrico Girardi, et al. « Withholding Primary Pneumocystis Pneumonia Prophylaxis in Virologically Suppressed Patients With Human Immunodeficiency Virus: An Emulation of a Pragmatic Trial in COHERE ». Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America 73, nº 2 (July 15, 2021): 195-202. https://doi.org/10.1093/cid/ciaa615.
Cho 2021	Cho, Kelly, Sarah C. Keithly, Katherine E. Kurgansky, Arin L. Madenci, Hanna Gerlovin, Helen Marucci-Wellman, Annie Doubleday, et al. « Early Convalescent Plasma Therapy and Mortality Among US Veterans Hospitalized With Nonsevere COVID-19: An Observational Analysis Emulating a Target Trial ». The Journal of Infectious Diseases 224, nº 6 (September 17, 2021): 967-75. https://doi.org/10.1093/infdis/jiab330.
Lyu 2021	Lyu, Beini, Micah R. Chan, Alexander S. Yevzlin, Ali Gardezi, et Brad C. Astor. « Arteriovenous Access Type and Risk of Mortality, Hospitalization, and Sepsis Among Elderly Hemodialysis Patients: A Target Trial Emulation Approach ». American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation 79, nº 1 (January 2022): 69-78. https://doi.org/10.1053/j.ajkd.2021.03.030.
Takeushi 2021	Takeuchi, Yoshinori, Hiraku Kumamaru, Yasuhiro Hagiwara, Hiroki Matsui, Hideo Yasunaga, Hiroaki Miyata, et Yutaka Matsuyama. « Sodium-Glucose Cotransporter-2 Inhibitors and the Risk of Urinary Tract Infection among Diabetic Patients in Japan: Target Trial Emulation Using a Nationwide Administrative Claims Database ». Diabetes, Obesity & Metabolism 23, nº 6 (June 2021): 1379-88. https://doi.org/10.1111/dom.14353.
Chiu 2021	Chiu, Yu-Han, Jorge E. Chavarro, Barbra A. Dickerman, JoAnn E. Manson, Kenneth J. Mukamal, Kathryn M. Rexrode, Eric B. Rimm, et Miguel A. Hernán. « Estimating the Effect of Nutritional Interventions Using Observational Data: The American Heart Association's 2020 Dietary Goals and Mortality ». The American Journal of Clinical Nutrition 114, nº 2 (August 2, 2021): 690-703. https://doi.org/10.1093/ajcn/nqab100.
Lim 2021	Lim, Cherry, Yin Mo, Prapit Teparrukkul, Maliwan Hongsuwan, Nicholas P. J. Day, Direk Limmathurotsakul, et Ben S. Cooper. « Effect of Delays in Concordant Antibiotic Treatment on Mortality in Patients With Hospital-Acquired Acinetobacter Species Bacteremia: Emulating a Target Randomized Trial With a 13-Year Retrospective

	Cohort ». American Journal of Epidemiology 190, nº 11 (November 2, 2021): 2395-2404. https://doi.org/10.1093/aje/kwab158.
Maguen 2021	Maguen, Shira, Erin Madden, Nicholas Holder, Yongmei Li, Karen H. Seal, Thomas C. Neylan, Callan Lujan, Olga V. Patterson, Scott L. DuVall, et Brian Shiner. « Effectiveness and Comparative Effectiveness of Evidence-Based Psychotherapies for Posttraumatic Stress Disorder in Clinical Practice ». Psychological Medicine 53, nº 2 (January 2023): 419-28. https://doi.org/10.1017/S0033291721001628.
Rossides 2021	Rossides, Marios, Susanna Kullberg, Daniela Di Giuseppe, Anders Eklund, Johan Grunewald, Johan Askling, et Elizabeth V. Arkema. « Infection Risk in Sarcoidosis Patients Treated with Methotrexate Compared to Azathioprine: A Retrospective "target Trial" Emulated with Swedish Real-World Data ». Respirology (Carlton, Vic.) 26, nº 5 (May 2021): 452-60. https://doi.org/10.1111/resp.14001.
Chiu 2022	Chiu, Yu-Han, Jennifer J. Yland, Paolo Rinaudo, John Hsu, Sean McGrath, Sonia Hernández-Díaz, et Miguel A. Hernán. « Effectiveness and Safety of Intrauterine Insemination vs. Assisted Reproductive Technology: Emulating a Target Trial Using an Observational Database of Administrative Claims ». Fertility and Sterility 117, nº 5 (May 2022): 981-91. https://doi.org/10.1016/j.fertnstert.2022.02.003.
Al-Samkari 2021	Al-Samkari, Hanny, Shruti Gupta, Rebecca Karp Leaf, Wei Wang, Rachel P. Rosovsky, Samantha K. Brenner, Salim S. Hayek, et al. « Thrombosis, Bleeding, and the Observational Effect of Early Therapeutic Anticoagulation on Survival in Critically Ill Patients With COVID-19 ». Annals of Internal Medicine 174, nº 5 (May 2021): 622-32. https://doi.org/10.7326/M20-6739.
Liu 2021	Liu, Ruishan, Shemra Rizzo, Samuel Whipple, Navdeep Pal, Arturo Lopez Pineda, Michael Lu, Brandon Arnieri, et al. « Evaluating Eligibility Criteria of Oncology Trials Using Real-World Data and AI ». Nature 592, nº 7855 (April 2021): 629-33. https://doi.org/10.1038/s41586-021-03430-5.
Van Santen 2021	Santen, Daniela K. van, Anders Boyd, Amy Matser, Lisa Maher, Matthew Hickman, Sara Lodi, et Maria Prins. « The Effect of Needle and Syringe Program and Opioid Agonist Therapy on the Risk of HIV, Hepatitis B and C Virus Infection for People Who Inject Drugs in Amsterdam, the Netherlands: Findings from an Emulated Target Trial ». Addiction (Abingdon, England) 116, nº 11 (November 2021): 3115-26. https://doi.org/10.1111/add.15503.
Boyne 2021	Boyne, Devon J., Winson Y. Cheung, Robert J. Hilsden, Tolulope T. Sajobi, Atul Batra, Christine M. Friedenreich, et Darren R. Brenner. « Association of a Shortened Duration of Adjuvant Chemotherapy With Overall Survival Among Individuals With Stage III Colon Cancer ». JAMA Network Open 4, nº 3 (March 1, 2021): e213587. https://doi.org/10.1001/jamanetworkopen.2021.3587.
Abrahami 2021	Abrahami, Devin, Richeek Pradhan, Hui Yin, Peter Honig, Elodie Baumfeld Andre, et Laurent Azoulay. « Use of Real-World Data to Emulate a Clinical Trial and Support Regulatory Decision Making: Assessing the Impact of Temporality, Comparator Choice, and Method of Adjustment ». Clinical Pharmacology and Therapeutics 109, nº 2 (February 2021): 452-61. https://doi.org/10.1002/cpt.2012.
Fu 2021 (a)	Fu, Edouard L., Marie Evans, Catherine M. Clase, Laurie A. Tomlinson, Merel van Diepen, Friedo W. Dekker, et Juan J. Carrero. « Stopping Renin-Angiotensin System Inhibitors in Patients with Advanced CKD and Risk of Adverse Outcomes: A Nationwide Study ». Journal of the American Society of Nephrology: JASN 32, nº 2 (February 2021): 424-35. https://doi.org/10.1681/ASN.2020050682.
Rojas-Saunero 2021	Rojas-Saunero, Liliana Paloma, Saima Hilal, Eleanor J. Murray, Roger W. Logan, Mohammad Arfan Ikram, et Sonja A. Swanson. « Hypothetical Blood-Pressure-Lowering Interventions and Risk of Stroke and Dementia ». European Journal of Epidemiology 36, nº 1 (January 2021): 69-79. https://doi.org/10.1007/s10654-020-00694-5.
Shaefi 2021	Shaefi, Shahzad, Samantha K. Brenner, Shruti Gupta, Brian P. O'Gara, Megan L. Krajewski, David M. Charytan, Sobaata Chaudhry, et al. « Extracorporeal Membrane Oxygenation in Patients with Severe Respiratory Failure from COVID-19 ». Intensive Care Medicine 47, nº 2 (February 2021): 208-21. https://doi.org/10.1007/s00134-020-06331-9.
Akarcrona 2021	Ankarcrona, V., H. Zhao, B. Jacobsson, et S. Brismar Wendel. « Obstetric Anal Sphincter Injury after Episiotomy in Vacuum Extraction: An Epidemiological Study

	Using an Emulated Randomised Trial Approach ». BJOG: An International Journal of Obstetrics and Gynaecology 128, nº 10 (September 2021): 1663-71. https://doi.org/10.1111/1471-0528.16663.
Schwahn 2021	Schwahn, Christian, Stefan Frenzel, Birte Holtfreter, Sandra Van der Auwera, Christiane Pink, Robin Bülow, Nele Friedrich, et al. « Effect of Periodontal Treatment on Preclinical Alzheimer's Disease-Results of a Trial Emulation Approach ». Alzheimer's & Dementia: The Journal of the Alzheimer's Association 18, nº 1 (January 2022): 127-41. https://doi.org/10.1002/alz.12378.
Reitblat 2021	Reitblat, Chanan, Aaron Fleishman, Irving A. Kaplan, Kristian D. Stensland, Anthony V. D'Amico, Aria F. Olumi, Andrew A. Wagner, et al. « Radical Prostatectomy versus External Beam Radiation Therapy for High-Grade, Clinically Localized Prostate Cancer: Emulation of a Target Clinical Trial ». Urologic Oncology 39, nº 11 (November 2021): 785.e1-785.e10. https://doi.org/10.1016/j.urolonc.2021.03.017.
Kirchgesner 2021	Kirchgesner, Julien, Rishi J. Desai, Laurent Beaugerie, Seoyoung C. Kim, et Sebastian Schneeweiss. « Calibrating Real-World Evidence Studies Against Randomized Trials: Treatment Effectiveness of Infliximab in Crohn's Disease ». Clinical Pharmacology and Therapeutics 111, nº 1 (January 2022): 179-86. https://doi.org/10.1002/cpt.2304.
Becker 2021	Becker, William C., Yu Li, Ellen C. Caniglia, Rachel Vickers-Smith, Termeh Feinberg, Brandon D. L. Marshall, et E. Jennifer Edelman. « Cannabis Use, Pain Interference, and Prescription Opioid Receipt among Persons with HIV: A Target Trial Emulation Study ». AIDS Care 34, nº 4 (April 2022): 469-77. https://doi.org/10.1080/09540121.2021.1944597.
Trevisan 2021	Trevisan, Marco, Edouard L. Fu, Yang Xu, Gianluigi Savarese, Friedo W. Dekker, Lars H. Lund, Catherine M. Clase, Arvid Sjölander, et Juan J. Carrero. « Stopping Mineralocorticoid Receptor Antagonists after Hyperkalaemia: Trial Emulation in Data from Routine Care ». European Journal of Heart Failure 23, nº 10 (October 2021): 1698-1707. https://doi.org/10.1002/ejhf.2287.
Kwee 2021	Kwee, Sandi A., Linda L. Wong, Miles M. Sato, Jared D. Acoba, Young Soo Rho, Avantika Srivastava, et Douglas P. Landsittel. « Transarterial Radioembolization for Hepatocellular Carcinoma with Major Vascular Invasion: A Nationwide Propensity Score-Matched Analysis with Target Trial Emulation ». Journal of Vascular and Interventional Radiology: JVIR 32, nº 9 (September 2021): 1258-1266.e6. https://doi.org/10.1016/j.jvir.2021.07.001.
Weberpal 2021	Weberpals, Janick, Tim Becker, Jessica Davies, Fabian Schmich, Dominik Rüttinger, Fabian J. Theis, et Anna Bauer-Mehren. « Deep Learning-Based Propensity Scores for Confounding Control in Comparative Effectiveness Research: A Large-Scale, Real-World Data Study ». Epidemiology (Cambridge, Mass.) 32, nº 3 (May 1, 2021): 378-88. https://doi.org/10.1097/EDE.0000000000001338.
Franklin 2021	Franklin, Jessica M., Elisabetta Patorno, Rishi J. Desai, Robert J. Glynn, David Martin, Kenneth Quinto, Ajinkya Pawar, et al. « Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies: First Results From the RCT DUPLICATE Initiative ». Circulation 143, nº 10 (March 9, 2021): 1002-13. https://doi.org/10.1161/CIRCULATIONAHA.120.051718.
Lenain 2021	Lenain, Rémi, Julie Boucquemont, Karen Leffondré, Cécile Couchoud, Mathilde Lassalle, Marc Hazzan, et Yohann Foucher. « Clinical Trial Emulation by Matching Time-Dependent Propensity Scores: The Example of Estimating Impact of Kidney Transplantation ». Epidemiology (Cambridge, Mass.) 32, nº 2 (March 1, 2021): 220-29. https://doi.org/10.1097/EDE.0000000000001308.
Nabi 2021	Nabi, Hafsah, Stylianos Georgiadis, Anne Gitte Loft, Oliver Hendricks, Dorte Vendelbo Jensen, Marlene Andersen, Stavros Chrysidis, et al. « Comparative Effectiveness of Two Adalimumab Biosimilars in 1318 Real-World Patients with Inflammatory Rheumatic Disease Mandated to Switch from Originator Adalimumab: Nationwide Observational Study Emulating a Randomised Clinical Trial ». Annals of the Rheumatic Diseases 80, nº 11 (November 2021): 1400-1409. https://doi.org/10.1136/annrheumdis-2021-219951.
Matthews 2021 (a)	Mathews, Kusum S., Howard Soh, Shahzad Shaefi, Wei Wang, Sonali Bose, Steven Coca, Shruti Gupta, et al. « Prone Positioning and Survival in Mechanically Ventilated Patients With Coronavirus Disease 2019-Related Respiratory Failure ». Critical Care

	Medicine 49, nº 7 (July 1, 2021): 1026-37. https://doi.org/10.1097/CCM.0000000000004938.
Barda 2021 (a)	Barda, Noam, Noa Dagan, Yatir Ben-Shlomo, Eldad Kepten, Jacob Waxman, Reut Ohana, Miguel A. Hernán, et al. « Safety of the BNT162b2 MRNA Covid-19 Vaccine in a Nationwide Setting ». The New England Journal of Medicine 385, nº 12 (September 16, 2021): 1078-90. https://doi.org/10.1056/NEJMoa2110475.
Matthews 2021 (b)	Matthews, Anthony A., Karolina Szummer, Issa J. Dahabreh, Bertil Lindahl, David Erlinge, Maria Feychting, Tomas Jernberg, Anita Berglund, et Miguel A. Hernán. « Comparing Effect Estimates in Randomized Trials and Observational Studies From the Same Population: An Application to Percutaneous Coronary Intervention ». Journal of the American Heart Association 10, nº 11 (June 2021): e020357. https://doi.org/10.1161/JAHA.120.020357.
Xu 2022	Xu, Yang, Edouard L. Fu, Marco Trevisan, Tomas Jernberg, Arvid Sjölander, Catherine M. Clase, et Juan-Jesus Carrero. « Stopping Renin-Angiotensin System Inhibitors after Hyperkalemia and Risk of Adverse Outcomes ». American Heart Journal 243 (January 2022): 177-86. https://doi.org/10.1016/j.ahj.2021.09.014.
Prat-Uribe 2021	Prats-Uribe, Albert, Marc Tobed, José Miguel Villacampa, Adriana Agüero, Clara García-Bastida, José Ignacio Tato, Laura Rodrigáñez, et al. « Timing of Elective Tracheotomy and Duration of Mechanical Ventilation among Patients Admitted to Intensive Care with Severe COVID-19: A Multicenter Prospective Cohort Study ». Head & Neck 43, nº 12 (December 2021): 3743-56. https://doi.org/10.1002/hed.26863.
Nolde 2021	Nolde, Michael, Nayeon Ahn, Tobias Dreischulte, Ina-Maria Rückert-Eheberg, Florian Güntner, Alexander Günter, Roman Gerlach, et al. « The Long-Term Risk for Myocardial Infarction or Stroke after Proton Pump Inhibitor Therapy (2008-2018) ». Alimentary Pharmacology & Therapeutics 54, nº 8 (October 2021): 1033-40. https://doi.org/10.1111/apt.16565.
Lai 2021	Lai, Alvina G., Wai Hoong Chang, Constantinos A. Parisinos, Michail Katsoulis, Ruth M. Blackburn, Anoop D. Shah, Vincent Nguyen, et al. « An Informatics Consult Approach for Generating Clinical Evidence for Treatment Decisions ». BMC Medical Informatics and Decision Making 21, nº 1 (October 12, 2021): 281. https://doi.org/10.1186/s12911-021-01638-z.
Gupta 2021	Gupta, Shruti, Wei Wang, Salim S. Hayek, Lili Chan, Kusum S. Mathews, Michal L. Melamed, Samantha K. Brenner, et al. « Association Between Early Treatment With Tocilizumab and Mortality Among Critically Ill Patients With COVID-19 ». JAMA Internal Medicine 181, nº 1 (January 1, 2021): 41-51. https://doi.org/10.1001/jamainternmed.2020.6252.
Schnitzer 2022	Schnitzer, Mireille E., Steve Ferreira Guerra, Cristina Longo, Lucie Blais, et Robert W. Platt. « A Potential Outcomes Approach to Defining and Estimating Gestational Age-Specific Exposure Effects during Pregnancy ». Statistical Methods in Medical Research 31, nº 2 (February 2022): 300-314. https://doi.org/10.1177/09622802211065158.
Aubert 2021	Aubert, Carole E., Jeremy B. Sussman, Timothy P. Hofer, William C. Cushman, Jin-Kyung Ha, et Lillian Min. « Adding a New Medication Versus Maximizing Dose to Intensify Hypertension Treatment in Older Adults : A Retrospective Observational Study ». Annals of Internal Medicine 174, nº 12 (December 2021): 1666-73. https://doi.org/10.7326/M21-1456.
Bharadwaj 2022	Bharadwaj, Maheetha, Sumedh Kaul, Aaron Fleishman, Ruslan Korets, Peter Chang, Andrew Wagner, Simon Kim, et al. « Adjuvant Chemotherapy versus Observation Following Radical Cystectomy for Locally Advanced Urothelial Carcinoma of the Bladder ». Urologic Oncology 40, nº 6 (June 2022): 274.e15-274.e23. https://doi.org/10.1016/j.urolonc.2022.02.002.
Wei 2022	Wei, Jie, Hyon K. Choi, Tuhina Neogi, Nicola Dalbeth, Robert Terkeltaub, Lisa K. Stamp, Houchen Lyu, et al. « Allopurinol Initiation and All-Cause Mortality Among Patients With Gout and Concurrent Chronic Kidney Disease : A Population-Based Cohort Study ». Annals of Internal Medicine 175, nº 4 (April 2022): 461-70. https://doi.org/10.7326/M21-2347.
Huang 2021	Huang, Jonathan Yinhao, Shirong Cai, Zhongwei Huang, Mya Thway Tint, Wen Lun Yuan, Izzuddin M. Aris, Keith M. Godfrey, et al. « Analyses of Child Cardiometabolic

	Phenotype Following Assisted Reproductive Technologies Using a Pragmatic Trial Emulation Approach ». Nature Communications 12, nº 1 (September 23, 2021): 5613. https://doi.org/10.1038/s41467-021-25899-4.
Porcher 2021	Porcher, Raphaël, Isabelle Desguerre, Helge Amthor, Brigitte Chabrol, Frédérique Audic, François Rivier, Arnaud Isapof, et al. « Association between Prophylactic Angiotensin-Converting Enzyme Inhibitors and Overall Survival in Duchenne Muscular Dystrophy-Analysis of Registry Data ». European Heart Journal 42, nº 20 (May 21, 2021): 1976-84. https://doi.org/10.1093/eurheartj/ehab054.
Schroeder 2022	Schroeder, Emily B., Romain Neugebauer, Kristi Reynolds, Julie A. Schmittdiel, Linda Loes, Wendy Dyer, Noel Pimentel, et al. « Association of Cardiovascular Outcomes and Mortality With Sustained Long-Acting Insulin Only vs Long-Acting Plus Short-Acting Insulin Treatment ». JAMA Network Open 4, nº 9 (September 1, 2021): e2126605. https://doi.org/10.1001/jamanetworkopen.2021.26605.
Hassan 2022	Hassan, Waleed, Prabin Shrestha, Keiichi Sumida, Fridtjof Thomas, Patrick L. Sweeney, Praveen K. Potukuchi, Connie M. Rhee, Elani Streja, Kamyar Kalantar-Zadeh, et Csaba P. Kovesdy. « Association of Uric Acid-Lowering Therapy With Incident Chronic Kidney Disease ». JAMA Network Open 5, nº 6 (June 1, 2022): e2215878. https://doi.org/10.1001/jamanetworkopen.2022.15878.
Matthews 2022	Matthews, Anthony A., Issa J. Dahabreh, Ole Fröbert, Bertil Lindahl, Stefan James, Maria Feychting, Tomas Jernberg, Anita Berglund, et Miguel A. Hernán. « Benchmarking Observational Analyses Before Using Them to Address Questions Trials Do Not Answer: An Application to Coronary Thrombus Aspiration ». American Journal of Epidemiology 191, nº 9 (August 22, 2022): 1652-65. https://doi.org/10.1093/aje/kwac098.
Dagan 2021	Dagan, Noa, Noam Barda, Eldad Kepten, Oren Miron, Shay Perchik, Mark A. Katz, Miguel A. Hernán, Marc Lipsitch, Ben Reis, et Ran D. Balicer. « BNT162b2 MRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting ». The New England Journal of Medicine 384, nº 15 (April 15, 2021): 1412-23. https://doi.org/10.1056/NEJMoa2101765.
Shin 2022	Shin, HoJin, Sebastian Schneeweiss, Robert J. Glynn, et Elisabetta Patorno. « Cardiovascular Outcomes in Patients Initiating First-Line Treatment of Type 2 Diabetes With Sodium-Glucose Cotransporter-2 Inhibitors Versus Metformin : A Cohort Study ». Annals of Internal Medicine 175, nº 7 (July 2022): 927-37. https://doi.org/10.7326/M21-4012.
Lu 2021 (a)	Lu, Haidong, Stephen R. Cole, Daniel Westreich, Michael G. Hudgens, Adaora A. Adimora, Keri N. Althoff, Michael J. Silverberg, et al. « Clinical Effectiveness of Integrase Strand Transfer Inhibitor-Based Antiretroviral Regimens Among Adults With Human Immunodeficiency Virus: A Collaboration of Cohort Studies in the United States and Canada ». Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America 73, nº 7 (October 5, 2021): e1408-14. https://doi.org/10.1093/cid/ciaa1037.
Barbulescu 2022	Barbulescu, Andrei, Johan Askling, Saedis Saevarsdottir, Seoyoung C. Kim, et Thomas Frisell. « Combined Conventional Synthetic Disease Modifying Therapy vs. Infliximab for Rheumatoid Arthritis: Emulating a Randomized Trial in Observational Data ». Clinical Pharmacology and Therapeutics 112, nº 4 (October 2022): 836-45. https://doi.org/10.1002/cpt.2673.
Dickerman 2021	Dickerman, Barbra A., Hanna Gerlovin, Arin L. Madenci, Michael J. Figueroa Muñiz, Jessica K. Wise, Nimish Adhikari, Brian R. Ferolito, et al. « Comparative Effectiveness of Third Doses of MRNA-Based COVID-19 Vaccines in US Veterans ». Nature Microbiology 8, nº 1 (January 2023): 55-63. https://doi.org/10.1038/s41564-022-01272-z.
Keyhani 2022	Keyhani, Salomeh, Eric M. Cheng, Katherine Hoggatt, Peter C. Austin, Erin Madden, Paul L. Hebert, Ethan A. Halm, et al. « Comparative Effectiveness of Carotid Stenting to Medical Therapy Among Patients With Asymptomatic Carotid Stenosis ». Stroke 53, nº 4 (April 2022): 1157-66. https://doi.org/10.1161/STROKEAHA.121.036178.
Dickerman 2022	Dickerman, Barbra A., Arin L. Madenci, Hanna Gerlovin, Katherine E. Kurgansky, Jessica K. Wise, Michael J. Figueroa Muñiz, Brian R. Ferolito, et al. « Comparative Safety of BNT162b2 and MRNA-1273 Vaccines in a Nationwide Cohort of US

	Veterans ». JAMA Internal Medicine 182, nº 7 (July 1, 2022): 739-46. https://doi.org/10.1001/jamainternmed.2022.2109.
Sun 2022	Jenny W., Jessica G. Young, Aaron L. Sarvet, L. Charles Bailey, William J. Heerman, David M. Janicke, Pi-I. Debby Lin, Sengwee Toh, et Jason P. Block. « Comparison of Rates of Type 2 Diabetes in Adults and Children Treated With Anticonvulsant Mood Stabilizers ». JAMA Network Open 5, no 4 (April 1, 2022): e226484. https://doi.org/10.1001/jamanetworkopen.2022.6484.
Bai 2021	Bai, Lawrence, Madeleine K. D. Scott, Ethan Steinberg, Laurynas Kalesinskas, Aida Habtezion, Nigam H. Shah, et Purvesh Khatri. « Computational Drug Repositioning of Atorvastatin for Ulcerative Colitis ». Journal of the American Medical Informatics Association: JAMIA 28, nº 11 (October 12, 2021): 2325-35. https://doi.org/10.1093/jamia/ocab165.
Ioannou 2022	Ioannou, George N., Emily R. Locke, Ann M. O'Hare, Amy S. B. Bohnert, Edward J. Boyko, Denise M. Hynes, et Kristin Berry. « COVID-19 Vaccination Effectiveness Against Infection or Death in a National U.S. Health Care System : A Target Trial Emulation Study ». Annals of Internal Medicine 175, nº 3 (March 2022): 352-61. https://doi.org/10.7326/M21-3256.
Kirchgesner 2022 (a)	Kirchgesner, Julien, Rishi J. Desai, Maria C. Schneeweiss, Laurent Beaugerie, Sebastian Schneeweiss, et Seoyoung C. Kim. « Decreased Risk of Treatment Failure with Vedolizumab and Thiopurines Combined Compared with Vedolizumab Monotherapy in Crohn's Disease ». Gut 71, nº 9 (September 2022): 1781-89. https://doi.org/10.1136/gutjnl-2022-327002.
Huang 2022	Huang, Huei-Kai, Peter Pin-Sung Liu, Shu-Man Lin, Jin-Yi Hsu, Jih-I. Yeh, Edward Chia-Cheng Lai, Carol Chiung-Hui Peng, Kashif M. Munir, Ching-Hui Loh, and Yu-Kang Tu. "Diabetes-Related Complications and Mortality in Patients With Atrial Fibrillation Receiving Different Oral Anticoagulants : A Nationwide Analysis." Annals of Internal Medicine 175, no. 4 (April 2022): 490–98. https://doi.org/10.7326/M21-3498.
Lazzati 2022	Lazzati, Andrea, Salomé Epaud, Matthieu Ortala, Sandrine Katsahian, and Emilie Lanoy. "Effect of Bariatric Surgery on Cancer Risk: Results from an Emulated Target Trial Using Population-Based Data." The British Journal of Surgery 109, no. 5 (April 19, 2022): 433–38. https://doi.org/10.1093/bjs/znac003.
Abu-Raddad 2022	Abu-Raddad, Laith J., Hiam Chemaitelly, Houssein H. Ayoub, Sawsan AlMukdad, Hadi M. Yassine, Hebah A. Al-Khatib, Maria K. Smatti, et al. "Effect of MRNA Vaccine Boosters against SARS-CoV-2 Omicron Infection in Qatar." The New England Journal of Medicine 386, no. 19 (May 12, 2022): 1804–16. https://doi.org/10.1056/NEJMoa2200797.
Lindquist 2022	Lindquist, Anthea, Roxanne Hastie, Amber Kennedy, Lyle Gurrin, Anna Middleton, Jon Quach, Jeanie Cheong, Susan P. Walker, Richard Hiscock, and Stephen Tong. "Developmental Outcomes for Children After Elective Birth at 39 Weeks' Gestation." JAMA Pediatrics 176, no. 7 (July 1, 2022): 654–63. https://doi.org/10.1001/jamapediatrics.2022.1165.
Barda 2021 (b)	Barda, Noam, Noa Dagan, Cyrille Cohen, Miguel A. Hernán, Marc Lipsitch, Isaac S. Kohane, Ben Y. Reis, and Ran D. Balicer. "Effectiveness of a Third Dose of the BNT162b2 MRNA COVID-19 Vaccine for Preventing Severe Outcomes in Israel: An Observational Study." Lancet (London, England) 398, no. 10316 (December 4, 2021): 2093–2100. https://doi.org/10.1016/S0140-6736(21)02249-2.
Dagan 2021	Dagan, Noa, Noam Barda, Tal Biron-Shental, Maya Makov-Assif, Calanit Key, Isaac S. Kohane, Miguel A. Hernán, et al. "Effectiveness of the BNT162b2 MRNA COVID-19 Vaccine in Pregnancy." Nature Medicine 27, no. 10 (October 2021): 1693–95. https://doi.org/10.1038/s41591-021-01490-8.
Simmonds 2022	Simmonds, Kent P., James Burke, Allan J. Kozlowski, Michael Andary, Zhehui Luo, and Mathew J. Reeves. "Emulating 3 Clinical Trials That Compare Stroke Rehabilitation at Inpatient Rehabilitation Facilities With Skilled Nursing Facilities." Archives of Physical Medicine and Rehabilitation 103, no. 7 (July 2022): 1311–19. https://doi.org/10.1016/j.apmr.2021.12.029.
Ahn 2022	Ahn, Nayeon, Michael Nolde, Alexander Günter, Florian Güntner, Roman Gerlach, Martin Tauscher, Ute Amann, et al. "Emulating a Target Trial of Proton Pump Inhibitors and Dementia Risk Using Claims Data." European Journal of Neurology 29, no. 5 (May 2022): 1335–43. https://doi.org/10.1111/ene.15284.

Yland 2022	Yland, Jennifer J., Yu-Han Chiu, Paolo Rinaudo, John Hsu, Miguel A. Hernán, and Sonia Hernández-Díaz. "Emulating a Target Trial of the Comparative Effectiveness of Clomiphene Citrate and Letrozole for Ovulation Induction." Human Reproduction (Oxford, England) 37, no. 4 (April 1, 2022): 793–805. https://doi.org/10.1093/humrep/deac005.
Kirchgesner 2022 (b)	Kirchgesner, Julien, Rishi J. Desai, Maria C. Schneeweiss, Laurent Beaugerie, Seoyoung C. Kim, and Sebastian Schneeweiss. "Emulation of a Randomized Controlled Trial in Ulcerative Colitis with US and French Claims Data: Infliximab with Thiopurines Compared to Infliximab Monotherapy." Pharmacoepidemiology and Drug Safety 31, no. 2 (February 2022): 167–75. https://doi.org/10.1002/pds.5356.
Mazotta 2022	Mazzotta, Valentina, Alessandro Cozzi-Lepri, Francesca Colavita, Simone Lanini, Silvia Rosati, Eleonora Lalle, Ilaria Mastrorosa, et al. "Emulation of a Target Trial From Observational Data to Compare Effectiveness of Casirivimab/Imdevimab and Bamlanivimab/Etesevimab for Early Treatment of Non-Hospitalized Patients With COVID-19." Frontiers in Immunology 13 (2022): 868020. https://doi.org/10.3389/fimmu.2022.868020.
Wang 2021	Wang, Lei, Zixing Wang, Yuguang Huang, Yipeng Wang, Zhong Liu, Shijie Xin, Guanghua Lei, et al. "Expanding Restrictive Transfusion Evidence in Surgical Practice: A Multicentre, Prospective Cohort Study." Blood Transfusion = Trasfusione Del Sangue 20, no. 5 (September 2022): 382–94. https://doi.org/10.2450/2021.0172-21.
Hajage 2022	Hajage, David, Alain Combes, Christophe Guervilly, Guillaume Lebreton, Alain Mercat, Arthur Pavot, Saad Nseir, et al. "Extracorporeal Membrane Oxygenation for Severe Acute Respiratory Distress Syndrome Associated with COVID-19: An Emulated Target Trial Analysis." American Journal of Respiratory and Critical Care Medicine 206, no. 3 (August 1, 2022): 281–94. https://doi.org/10.1164/rccm.202111-2495OC.
Magen 2022	Magen, Ori, Jacob G. Waxman, Maya Makov-Assif, Roni Vered, Dror Dicker, Miguel A. Hernán, Marc Lipsitch, Ben Y. Reis, Ran D. Balicer, and Noa Dagan. "Fourth Dose of BNT162b2 MRNA Covid-19 Vaccine in a Nationwide Setting." The New England Journal of Medicine 386, no. 17 (April 28, 2022): 1603–14. https://doi.org/10.1056/NEJMoa2201688.
Kim 2021 (a)	Kim, Dae Hyun, Ajinkya Pawar, Joshua J. Gagne, Lily G. Bessette, Hemin Lee, Robert J. Glynn, and Sebastian Schneeweiss. "Frailty and Clinical Outcomes of Direct Oral Anticoagulants Versus Warfarin in Older Adults With Atrial Fibrillation : A Cohort Study." Annals of Internal Medicine 174, no. 9 (September 2021): 1214–23. https://doi.org/10.7326/M20-7141.
Dickow 2022	Dickow, Jannis, Paulus Kirchhof, Holly K. Van Houten, Lindsey R. Sangaralingham, Leon H. W. Dinshaw, Paul A. Friedman, Douglas L. Packer, Peter A. Noseworthy, and Xiaoxi Yao. "Generalizability of the EAST-AFNET 4 Trial: Assessing Outcomes of Early Rhythm-Control Therapy in Patients With Atrial Fibrillation." Journal of the American Heart Association 11, no. 11 (June 7, 2022): e024214. https://doi.org/10.1161/JAHA.121.024214.
Bergqvist 2021	Bergqvist, Rita, Viktor H. Ahlqvist, Michael Lundberg, Maria-Pia Hergens, Johan Sundström, Max Bell, and Cecilia Magnusson. "HMG-CoA Reductase Inhibitors and COVID-19 Mortality in Stockholm, Sweden: A Registry-Based Cohort Study." PLoS Medicine 18, no. 10 (October 2021): e1003820. https://doi.org/10.1371/journal.pmed.1003820.
Bruun-Rasmussen 2022	Bruun-Rasmussen, Peter, Per Kragh Andersen, Karina Banasik, Søren Brunak, and Pär Ingemar Johansson. "Intervening on the Storage Time of RBC Units and Its Effects on Adverse Recipient Outcomes Using Real-World Data." Blood 139, no. 25 (June 23, 2022): 3647–54. https://doi.org/10.1182/blood.2022015892.
McGraph 2021	McGrath, Leah J., Carrie Nielson, Bradley Saul, Alexander Breskin, Ying Yu, Sia K. Nicolaisen, Karynsa Kilpatrick, et al. "Lessons Learned Using Real-World Data to Emulate Randomized Trials: A Case Study of Treatment Effectiveness for Newly Diagnosed Immune Thrombocytopenia." Clinical Pharmacology and Therapeutics 110, no. 6 (December 2021): 1570–78. https://doi.org/10.1002/cpt.2399.
Wallach 2021	Wallach, Joshua D., Yihong Deng, Rozalina G. McCoy, Sanket S. Dhruva, Jeph Herrin, Alyssa Berkowitz, Eric C. Polley, et al. "Real-World Cardiovascular Outcomes Associated With Degarelix vs Leuprolide for Prostate Cancer Treatment." JAMA

	Network Open 4, no. 10 (October 1, 2021): e2130587. https://doi.org/10.1001/jamanetworkopen.2021.30587.
Gild 2021	Gild, Philipp, Raisa S. Pompe, Thomas Seisen, Jacob Keeley, Hoang J. Tang, Alberto Bossi, Derya Tilki, Mani Menon, and Firas Abdollah. "Regression Discontinuity Analysis of Salvage Radiotherapy in Prostate Cancer." European Urology Oncology 4, no. 5 (October 2021): 817–20. https://doi.org/10.1016/j.euo.2019.08.005.
Butt 2022	Adeel A. Butt, Victor B. Talisa, Obaid S. Shaikh, Saad B. Omer, and Florian B. Mayr. "Relative Vaccine Effectiveness of a Severe Acute Respiratory Syndrome Coronavirus 2 Messenger RNA Vaccine Booster Dose Against the Omicron Variant." Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America 75, no. 12 (December 19, 2022): 2161–68. https://doi.org/10.1093/cid/ciac328.
Chammartin 2021	Chammartin, Frédérique, Sara Lodi, Roger Logan, Lene Ryom, Amanda Mocroft, Ole Kirk, Antonella d'Arminio Monforte, et al. "Risk for Non-AIDS-Defining and AIDS-Defining Cancer of Early Versus Delayed Initiation of Antiretroviral Therapy : A Multinational Prospective Cohort Study." Annals of Internal Medicine 174, no. 6 (June 2021): 768–76. https://doi.org/10.7326/M20-5226.
Dave 2022	Dave, Chintan V., Gary M. Brittenham, Jeffrey L. Carson, and Soko Setoguchi. "Risks for Anaphylaxis With Intravenous Iron Formulations : A Retrospective Cohort Study." Annals of Internal Medicine 175, no. 5 (May 2022): 656–64. https://doi.org/10.7326/M21-4009.
Braitmaier 2022	Braitmaier, Malte, Sarina Schwarz, Bianca Kollhorst, Carlo Senore, Vanessa Didelez, and Ulrike Haug. "Screening Colonoscopy Similarly Prevented Distal and Proximal Colorectal Cancer: A Prospective Study among 55-69-Year-Olds." Journal of Clinical Epidemiology 149 (September 2022): 118–26. https://doi.org/10.1016/j.jclinepi.2022.05.024.
Martinez-Sanz 2022	Martínez-Sanz, Javier, Raquel Ron, Elena Moreno, Matilde Sánchez-Conde, Alfonso Muriel, Luis Fernando López Cortés, José Ramón Blanco, et al. "Similar CD4/CD8 Ratio Recovery After Initiation of Dolutegravir Plus Lamivudine Versus Dolutegravir or Bictegravir-Based Three-Drug Regimens in Naive Adults With HIV." Frontiers in Immunology 13 (2022): 873408. https://doi.org/10.3389/fimmu.2022.873408.
Shahn 2022	Shahn, Zach, Phoebe Spear, Helen Lu, Sharon Jiang, Suki Zhang, Neil Deshmukh, Shenbo Xu, Kenney Ng, Roy Welsch, and Stan Finkelstein. "Systematically Exploring Repurposing Effects of Antihypertensives." Pharmacoepidemiology and Drug Safety 31, no. 9 (September 2022): 944–52. https://doi.org/10.1002/pds.5491.
Hou 2022	Hou, Jue, Rachel Zhao, Tianrun Cai, Brett Beaulieu-Jones, Thany Seyok, Kumar Dahal, Qianyu Yuan, et al. "Temporal Trends in Clinical Evidence of 5-Year Survival Within Electronic Health Records Among Patients With Early-Stage Colon Cancer Managed With Laparoscopy-Assisted Colectomy vs Open Colectomy." JAMA Network Open 5, no. 6 (June 1, 2022): e2218371. https://doi.org/10.1001/jamanetworkopen.2022.18371.
Farjat 2022	Farjat, Alfredo E., Saverio Virdone, Laine E. Thomas, Ajay K. Kakkar, Karen S. Pieper, and Jonathan P. Piccini. "The Importance of the Design of Observational Studies in Comparative Effectiveness Research: Lessons from the GARFIELD-AF and ORBIT-AF Registries." American Heart Journal 243 (January 2022): 110–21. https://doi.org/10.1016/j.ahj.2021.09.003.
Gazit 2022	Gazit, Sivan, Roei Shlezinger, Galit Perez, Roni Lotan, Asaf Peretz, Amir Ben-Tov, Esma Herzel, et al. "The Incidence of SARS-CoV-2 Reinfection in Persons With Naturally Acquired Immunity With and Without Subsequent Receipt of a Single Dose of BNT162b2 Vaccine : A Retrospective Cohort Study." Annals of Internal Medicine 175, no. 5 (May 2022): 674–81. https://doi.org/10.7326/M21-4130.
Lu 2022	Lu, Yan, Aaron W. Gehr, Rachel J. Meadows, Bassam Ghabach, Latha Neerukonda, Kalyani Narra, and Rohit P. Ojha. "Timing of Adjuvant Chemotherapy Initiation and Mortality among Colon Cancer Patients at a Safety-Net Health System." BMC Cancer 22, no. 1 (May 31, 2022): 593. https://doi.org/10.1186/s12885-022-09688-w.
Khosrow-Khavar 2022	Khosrow-Khavar, Farzin, Rishi J. Desai, Hemin Lee, Su Been Lee, and Seoyoung C. Kim. "Tofacitinib and Risk of Malignancy: Results From the Safety of Tofacitinib in Routine Care Patients With Rheumatoid Arthritis (STAR-RA) Study." Arthritis &

	Rheumatology (Hoboken, N.J.) 74, no. 10 (October 2022): 1648–59. https://doi.org/10.1002/art.42250.
Kim 2021 (b)	Kim, Daehoon, Pil-Sung Yang, Seng Chan You, Jung-Hoon Sung, Eunsun Jang, Hee Tae Yu, Tae-Hoon Kim, et al. "Treatment Timing and the Effects of Rhythm Control Strategy in Patients with Atrial Fibrillation: Nationwide Cohort Study." BMJ (Clinical Research Ed.) 373 (May 11, 2021): n991. https://doi.org/10.1136/bmj.n991.
Urner 2022	Urner, Martin, Adrian G. Barnett, Gianluigi Li Bassi, Daniel Brodie, Heidi J. Dalton, Niall D. Ferguson, Silver Heinsar, et al. "Venovenous Extracorporeal Membrane Oxygenation in Patients with Acute Covid-19 Associated Respiratory Failure: Comparative Effectiveness Study." BMJ (Clinical Research Ed.) 377 (May 4, 2022): e068723. https://doi.org/10.1136/bmj-2021-068723.
Kainz 2022	Kainz, Alexander, Michael Kammer, Roman Reindl-Schwaighofer, Susanne Strohmaier, Vojtěch Petr, Ondrej Viklicky, Daniel Abramowicz, Marcel Naik, Gert Mayer, and Rainer Oberbauer. "Waiting Time for Second Kidney Transplantation and Mortality." Clinical Journal of the American Society of Nephrology: CJASN 17, no. 1 (January 2022): 90–97. https://doi.org/10.2215/CJN.07620621.
Katsoulis 2021	Katsoulis, Michail, Bianca D. Stavola, Karla Diaz-Ordaz, Manuel Gomes, Alvina Lai, Pagona Lagiou, Goya Wannamethee, et al. "Weight Change and the Onset of Cardiovascular Diseases: Emulating Trials Using Electronic Health Records." Epidemiology (Cambridge, Mass.) 32, no. 5 (September 1, 2021): 744–55. https://doi.org/10.1097/EDE.0000000000001393.
Fu 2021 (b)	Fu, Edouard L., Marie Evans, Juan-Jesus Carrero, Hein Putter, Catherine M. Clase, Fergus J. Caskey, Maciej Szymczak, et al. "Timing of Dialysis Initiation to Reduce Mortality and Cardiovascular Events in Advanced Chronic Kidney Disease: Nationwide Cohort Study." BMJ (Clinical Research Ed.) 375 (November 29, 2021): e066306. https://doi.org/10.1136/bmj-2021-066306.
Fischer 2022	Fischer, Urs, Mattia Branca, Leo H. Bonati, Emmanuel Carrera, Maria I. Vargas, Alexandra Platon, Zsolt Kulcsar, et al. "Magnetic Resonance Imaging or Computed Tomography for Suspected Acute Stroke: Association of Admission Image Modality with Acute Recanalization Therapies, Workflow Metrics, and Outcomes." Annals of Neurology 92, no. 2 (August 2022): 184–94. https://doi.org/10.1002/ana.26413.
Zhao 2022	Zhao, Jingcheng, Arvid Sjölander, and Gustaf Edgren. "Mortality Among Patients Undergoing Blood Transfusion in Relation to Donor Sex and Parity: A Natural Experiment." JAMA Internal Medicine 182, no. 7 (July 1, 2022): 747–56. https://doi.org/10.1001/jamainternmed.2022.2115.
Lu 2021 (b)	Lu, Haidong, Stephen R. Cole, Daniel Westreich, Michael G. Hudgens, Adaora A. Adimora, Keri N. Althoff, Michael J. Silverberg, et al. "Virologic Outcomes among Adults with HIV Using Integrase Inhibitor-Based Antiretroviral Therapy." AIDS (London, England) 36, no. 2 (February 1, 2022): 277–86. https://doi.org/10.1097/QAD.0000000000003069.

Appendix 5. Supplemental references

1 Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available: Table 1. *Am J Epidemiol*. 2016;183:758–64.

2 Hernán MA, Sauer BC, Hernández-Díaz S, *et al.* Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70–5.

3 Dickerman BA, García-Albéniz X, Logan RW, *et al.* Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med*. 2019;25:1601–6.

4 Hernán MA, Alonso A, Logan R, *et al.* Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiol Camb Mass*. 2008;19:766–79.

5 Nicholls SG, Langan SM, Benchimol EI. Routinely collected data: the importance of high-quality diagnostic coding to research. *Can Med Assoc J*. 2017;189:E1054–5.

6 Hernán MA, Robins JM. *Causal Inference: What If.* Boca Raton: Chapman&Hall/CRC. 2020.

7 Danaei G, Rodríguez LAG, Cantero OF, *et al.* Observational data for comparative effectiveness research: An emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res*. 2013;22:70–96.

8 Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivar Behav Res*. 2011;46:399–424.

9 Schneeweiss S, Rassen JA, Glynn RJ, *et al.* High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiol Camb Mass*. 2009;20:512–22.

10 Maringe C, Benitez Majano S, Exarchakou A, *et al.* Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. *Int J Epidemiol*. 2020;49:1719–29.

11 Hernán MA. How to estimate the effect of treatment duration on survival outcomes using observational data. *BMJ*. 2018;k182.

12 Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics*. 2000;56:779–88.

13 Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, *et al.* A causal framework for classical statistical estimands in failure-time settings with competing events. *Stat Med*. 2020;39:1199–236.

14 Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiol Camb Mass*. 2010;21:383–8.

15 VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med*. 2017;167:268.

16 Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative, *et al.* Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA*. 2023;329:1376.

17 Sterne JA, Hernán MA, Reeves BC, *et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;i4919.

18 Nguyen VT, Engleton M, Davison M, *et al.* Risk of bias in observational studies using routinely collected data of comparative effectiveness research: a meta-research study. *BMC Med*. 2021;19:279.

19 McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Res Synth Methods*. 2020;n/a. doi: 10.1002/jrsm.1411