

## Review

# The State of Use and Utility of Negative Controls in Pharmacoepidemiologic Studies

Zafar Zafari\*, Jeong-eun Park, Chintal H. Shah, Susan dosReis, Emily F. Gorman, Wei Hua, Yong Ma, and Fang Tian

\* Correspondence to Dr. Zafar Zafari, 220 N. Arch Street, Baltimore, Maryland, 21201 (e-mail: zzafari@rx.umaryland.edu).

Initially submitted December 16, 2022; accepted for publication October 6, 2023.

Uses of real-world data in drug safety and effectiveness studies are often challenged by various sources of bias. We undertook a systematic search of the published literature through September 2020 to evaluate the state of use and utility of negative controls to address bias in pharmacoepidemiologic studies. Two reviewers independently evaluated study eligibility and abstracted data. Our search identified 184 eligible studies for inclusion. Cohort studies (115, 63%) and administrative data (114, 62%) were, respectively, the most common study design and data type used. Most studies used negative control outcomes (91, 50%), and for most studies the target source of bias was unmeasured confounding (93, 51%). We identified 4 utility domains of negative controls: 1) bias detection (149, 81%), 2) bias correction (16, 9%), 3) *P*-value calibration (8, 4%), and 4) performance assessment of different methods used in drug safety studies (31, 17%). The most popular methodologies used were the 95% confidence interval and *P*-value calibration. In addition, we identified 2 reference sets with structured steps to check the causality assumption of the negative control. While negative controls are powerful tools in bias detection, we found many studies lacked checking the underlying assumptions.

**This article is part of a Special Collection on Pharmacoepidemiology.**

hidden bias; negative control exposure; negative control outcome; unmeasured confounding

Abbreviations: ADR, adverse drug reaction; AUC, area under the receiver-operating-characteristic curve; CI, confidence interval; DiD, difference-in-differences; EU-ADR, Exploring and Understanding Adverse Drug Reactions; FDA, Food and Drug Administration; NCE, negative control exposure; NCO, negative control outcome; OMOP, Observational Medical Outcome Partnership; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

A major concern in conducting observational studies is the threat to internal validity. While adequate adjustments for unbalanced covariates such as propensity score matching allow an observational study to potentially mimic an experimental study, the presence of unmeasured confounding, selection bias, and measurement error nonetheless remain an ongoing challenge in observational studies.

The use of negative controls for identifying and addressing potential residual bias has gained attraction in drug effectiveness and safety research in recent years. In such studies, negative controls can be used either in place of an exposure (negative control exposure or NCE) or an outcome (negative control outcome or NCO) in an alternative hypothesis for examining potential bias. The premise underlying the use of negative controls to detect bias is that while negative controls should not have a causal association with

the primary exposure or outcome of interest, they should share a same bias with the primary association of interest (1).

In recent years, there have been advancements in epidemiologic methods for applying negative controls in observational studies (1–9). Nevertheless, it is not well known to what extent these methodologies have been applied in pharmacoepidemiologic studies. In this work, we conducted a scoping review of the state of use, applications, and utility of negative controls in pharmacoepidemiologic studies. Findings from this work may inform the development of a guidance framework for future use of negative controls.

## METHODS

We conducted a scoping review of published literature on pharmacoepidemiologic studies that deployed negative

controls to detect or correct bias. Our review was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) (10).

### Identifying articles

The search strategy was initially developed by our librarian (E.F.G.) and modified by all team members (Z.Z., C.H.S., J.P., S.D.R., F.T., W.H., Y.M.). The search terms were chosen based on key concepts of focus in this review, including terms related to negative control, pharmacoepidemiology, and bias. We initiated the search in multiple electronic databases, including PubMed, EMBASE, CINAHL (Cumulative Index to Nursing and Allied Health Literature), Cochrane Library, Scopus, and Dissertations and Theses Global, from inception to September 4, 2020. In addition, we manually searched the articles that cited the key method papers that are often cited for the negative control methods using Scopus. Any relevant articles that were identified from the included studies and not retrieved from the database searches were screened and reviewed for inclusion additionally. Based on our review of the literature, we used multiple key terms that were used for negative controls, including negative control, proxy outcome, bias indicator, and probe variable. We utilized the review management software Covidence (Covidence systematic review software; Veritas Health Innovation, Melbourne, Australia; <https://www.covidence.org/>) throughout the review process to manage the data and facilitate discussion among the investigators (see Web Appendix 1 for search strategies, available at <https://doi.org/10.1093/aje/kwad201>).

### Selection of studies

We included studies that were published in English, had a focus on pharmacoepidemiology, and clearly expressed the use of negative controls to address bias in their analysis or added novel utilities of negative control methods. We excluded experimental studies either of a preclinical or clinical nature, or the ones that used negative controls as active comparators. Duplicate records were removed after comparing titles, published dates, authors, and journal fields. The selection criteria are provided in Web Appendix 1. Two reviewers (J.P. and C.H.S.) independently screened the titles and abstracts of all records retrieved from the database searches to determine studies to be further assessed for full-text screening. The full texts of the included articles from the abstract screening were reviewed by the same reviewers (J.P. and C.H.S.) independently for eligibility according to the inclusion criteria. Any discrepancies between the reviewers were resolved through further discussion, and if agreement was not reached, a third reviewer (Z.Z.) made the final determination throughout the selection process.

### Data extraction and analysis

Three reviewers (J.P., C.H.S., and Z.Z.) independently extracted data from the texts of the included articles. The team members crosschecked each other's data extraction.

Disagreements among the reviewers on extracted information were resolved through discussion at team meetings.

We obtained descriptive statistics on the negative control use in the included studies and abstracted data on key methodological items including study characteristics such as publication year, study region, study design/pharmacovigilance method, type of data source, type and number of negative controls used, causal question under investigation (e.g., a drug-outcome association of interest), measure of association used, negative control assumptions, results of the negative control analysis, and negative control method used for bias correction or *P*-value calibration. Based on our quantitative and qualitative review of the extracted data, we identified distinct domains of negative control use in pharmacoepidemiologic research and summarized the studies based on our domain classification. We also summarized the approaches undertaken in the included studies for validating the assumptions of negative controls.

## RESULTS

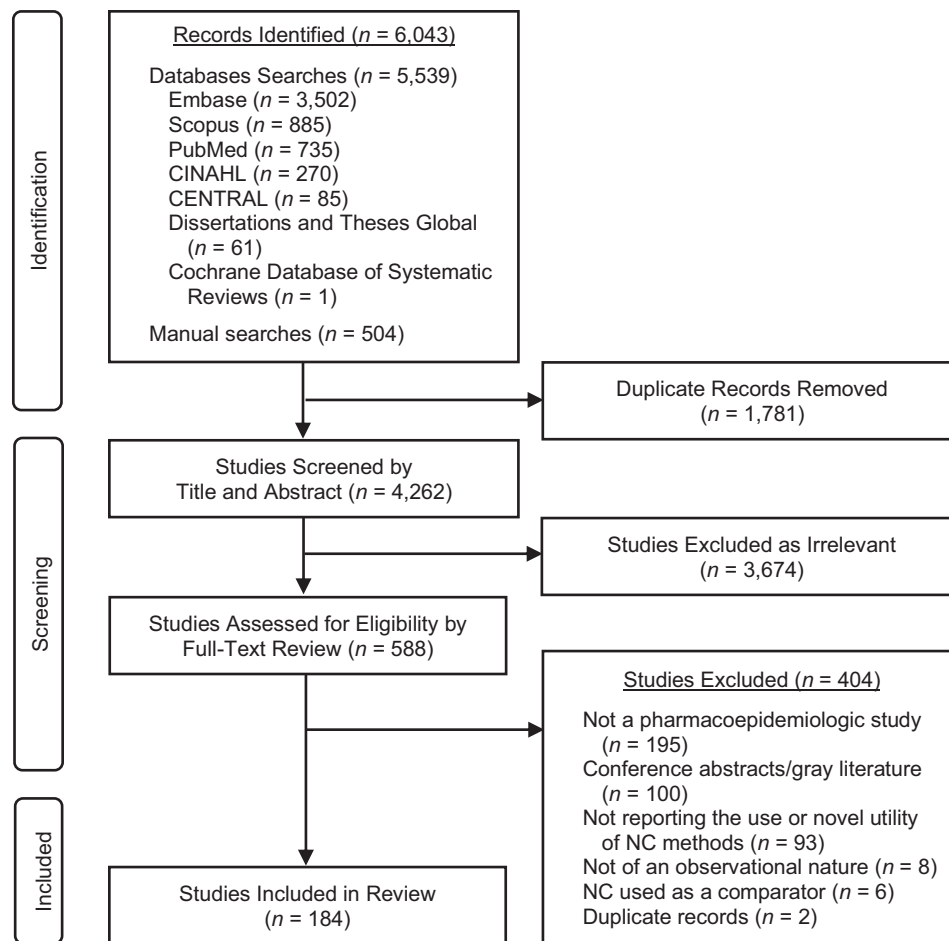
We initially identified 6,043 records from the database and the manual searches and selected 588 studies for further full-text assessment after the title and abstract screening. After the full-text review, we selected 184 papers for data extraction (Figure 1, showing a PRISMA flow diagram).

### Reviewing key definitions and assumptions

Under the counterfactual or potential outcome framework, which models and contrasts potential outcomes under alternative exposure levels, possibly contrary to fact (6), there are 2 major assumptions necessary for negative controls and their use for bias detection (1, 6, 9). The first assumption (lack of causality or exclusion restriction assumption) is necessary to define negative controls and ensures lack of causal associations between the NCE and the primary outcome and between the primary exposure and the NCO. The second assumption necessitates that the negative control association has a similar underlying bias structure to that of the primary exposure-outcome association (U-comparability assumption) required for bias detection. In addition to these main assumptions specific to negative controls, there are other assumptions required for causal inference while acknowledging the presence of unmeasured confounding, mainly the consistency assumption and the latent ignorability assumption, as described elsewhere (9). While the exclusion restriction and the U-comparability assumptions are sufficient for bias detection, there may be additional assumptions required for bias correction. For example, for identification of the causal effect using the double negative control methods, there are identification conditions described as positivity and completeness assumptions (9). Directed acyclic graphs of negative controls and formal representations of these assumptions are provided in Web Appendix 2, including Web Figure 1.

### Basic characteristics of the included studies

Table 1 shows the study characteristics. Out of 184 included studies, the majority were published in recent years



**Figure 1.** Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram for selection for inclusion in a review pharmacoepidemiologic studies involving negative controls (NCs). CENTRAL, Cochrane Central Register of Controlled Trials; CINAHL, Cumulative Index to Nursing and Allied Health Literature.

(126, 68.5%; between 2017 and 2020). The most common study design was cohort study ( $n = 115$ , 62.5%), followed by case-control study ( $n = 12$ , 6.5%) and self-controlled case series ( $n = 9$ , 4.9%). Most of the studies used health-care administrative data (e.g., insurance claims) ( $n = 114$ , 62.0%) or electronic health record data ( $n = 62$ , 33.2%). Of the 184 studies, 54 (29.3%) used NCE, 91 (49.5%) used NCO, and 34 (18.5%) used both NCE and NCO in their analyses. Unmeasured confounding was the most frequently stated reason ( $n = 93$ , 50.5%) for use of negative controls in the included studies, followed by information bias (including misclassification, measurement error) ( $n = 10$ , 5.4%) or selection bias ( $n = 7$ , 3.8%).

We identified different nomenclature used for negative controls. While most of the included studies used the term negative control or similar terms (e.g., negative outcome/exposure, control outcome/exposure), some studies used other terms, such as falsification endpoint/analysis, borrowed from the econometrics literature (10 studies). We identified 1 study that used the term indicator conditions to refer to negative control outcomes.

### The utility and application of negative controls in pharmacoepidemiologic studies

Based on our review of the 184 studies, we identified 4 main domains for the application of negative controls in pharmacoepidemiologic studies. The first 3 domains directly address bias in the effect estimate of the exposure-outcome association and include: 1) detection of bias, 2) correction of bias (by correcting the point estimate or calibrating its confidence interval (CI)), and 3) calibration of  $P$  value. The fourth domain is not directly related to a particular exposure-outcome association; instead, it deploys negative controls to assess the performance of different pharmacoepidemiologic methods.

*Detection of bias in the effect estimate of the exposure-outcome association.* Most of the studies (149 (81%)) used negative controls to detect potential bias in the effect estimate of the exposure-outcome association (4, 5, 7, 11–156). This can be done by testing the negative control association against the null hypothesis. Most studies indicated the presence of a bias by rejecting a null hypothesis

**Table 1.** Characteristics of the Included Studies in the Scoping Review for the State of Use and Utility of Negative Controls in Pharmacoepidemiologic Studies

Characteristic	Study (n = 184)	
	No.	%
Publication year		
Before 2012	12	6.5
2013–2016	46	25.0
2017–2020	126	68.5
Study region		
North America	94	51.1
Europe	61	33.2
Asia	13	7.1
Others	4	2.2
Multiple regions (including more than 1 continent)	12	6.4
Study design		
Cohort study	115	62.5
Case-control study	12	6.5
Self-controlled case series	9	4.9
Nested case-control study	7	3.8
Ecological study	5	2.7
Self-controlled cohort study	2	1.1
Case-crossover study	1	0.5
Cross-sectional study	1	0.5
Self-controlled risk interval study	1	0.5
Pharmacovigilance study		
Disproportionality analysis	4	2.2
Sequence symmetry analysis	3	1.6
Sequential statistical testing	3	1.6
Multiple studies with different designs/methods	14	7.6
Reference set identification	5	2.7
Others	2	1.1
Type of NC		
NCE	54	29.3
NCO	91	49.5
Both NCE and NCO	34	18.5
Not applicable <sup>a</sup>	5	2.7
No. of NCs used		
Single (either NCE or NCO)	84	45.7
Multiple	94	51.0
Not applicable <sup>a</sup> /not reported	6	3.3
Type of bias <sup>b</sup>		
Unmeasured confounding	93	50.5
Information bias (e.g., misclassification, measurement error)	10	5.4
Selection bias	7	3.8
Protopathic bias	2	1.1
Bias in general/not specified	80	43.5

Table continues

Table 1. Continued

Characteristic	Study (n = 184)	
	No.	%
Type of data source used <sup>b</sup>		
Health-care administrative data (e.g., insurance claims, national health system data)	114	62.0
Electronic health records	61	33.2
Research data (e.g., primary data)	26	14.1
Disease registry	11	6.0
Passive surveillance data for drug safety	7	3.8
Simulated data set	6	3.3
Public health survey data	4	2.2
Not applicable <sup>a</sup>	5	2.7

Abbreviations: NC, negative control; NCE, negative control exposure; NCO, negative control outcome.

<sup>a</sup> Reference set identification studies with no empirical application of NC.

<sup>b</sup> Categories are not mutually exclusive.

if the effect estimate of the negative control association was statistically significant given the type I error of 5% ( $P < 0.05$ ). However, we found some studies that concluded presence of bias if the direction of the point estimate of their negative control association was considered indicative of bias even when its  $P$  value was not statistically significant (14, 71, 116). Of the 149 studies included in this category, 63 (42%) reported detecting a presence of bias. For types of bias detected see Table 1 and Web Table 1.

*Negative controls to correct the effect estimate and CI for bias.* We found 16 studies that used negative controls to obtain bias-adjusted estimates (point estimates or CIs) for the exposure-outcome association of interest (Table 2) (5, 7, 47, 48, 50, 53, 63, 70, 73, 74, 89, 107, 111, 126, 139, 154). Fourteen out of 16 studies were published between 2017 and 2020, which may suggest expansion of the use of negative controls for bias correction in pharmacoepidemiology in recent years. The most common practice for bias correction was a difference-in-differences (DiD) approach by subtracting (or dividing) the effect estimate of the negative control association from that of the primary association of interest. This approach requires a restrictive assumption of additive equi-confounding, which necessitates that the magnitude and direction of the bias for the negative control association be equal to that of the primary exposure-outcome association on the additive scale (157). For example, a group of studies examining drug-drug interactions (DDIs) obtained the bias-corrected effect estimates by taking a ratio of rate ratio (RR) (or ratio of reporting odds ratio (ROR)) from an association with suspected DDI and the RR from negative control associations with no suspected DDI (50, 73, 74, 154). The delta method was commonly used to construct the 95% CIs for the corrected estimate in these studies (50, 73, 154).

Gruber et al. (48) used the DiD approach in examining the effect of different vaccine dose timing on rotavirus

vaccine effectiveness. The authors used the placebo group in a randomized controlled trial as a negative control for measuring and correcting for bias (48). Rodgers et al. (107) demonstrated the utility of prior event rate ratio (PERR), a bias correction method based on the DiD approach, to correct for bias using the same outcome measured in a negative control period. The study used PERR-adjusted hazard ratio ( $HR_{PERR}$ ), defined as the ratio between the hazard ratio for the study period and the hazard ratio for period before exposure, in adjusting for time-invariant unmeasured confounding for estimating the association between second-line antidiabetics and known side effects (edema and weight gain) (107). Using the outcome measured in the prior period as an NCO, the authors ensured the similarity in the bias measured in the negative control association with regard to time-invariant factors (107). Other studies described the use of the DiD approach in the context of survival analysis and its conditions (158–160). Bootstrapping was used to estimate 95% CIs for the corrected effect estimates for these studies (48, 107).

In another group of studies, an approach of utilizing more than one negative control was taken in correcting the effect estimates to account for variations in bias estimates from different negative control associations (5, 53, 70, 89, 111, 126, 139). Schuemie et al. (5) formalized the idea of using multiple negative controls in bias estimation into the empirical calibration method for 95% CI, which constructs a parametric empirical null distribution of bias using multiple negative controls and synthetic positive controls to calibrate the effect estimate and 95% CI. Table 2 shows the methodological details of the studies that corrected effect estimates and CIs for bias.

In addition, Shi et al. (7) leveraged the use of double negative controls, one NCE and one NCO, for constructing a semiparametric model for a multiply robust estimator of the average treatment effect in the presence of a categorical

**Table 2.** Pharmacoepidemiologic Studies That Used Negative Controls to Correct for (or Attenuate) Bias or Calibrate P Values

First Author, Year (Reference No.)	Study Design	Type of Data	Hypothesized Causal Association	Primary Study Effect Measure	NC(s) Used	Type of Bias as Reported by Study	Methods Used to Correct for Effect of Bias (in Point Estimates, CIs, or P Values)
<i>Correction of Point Estimate and 95% CI</i>							
Johanson, 2012 (63)	Ecological study	Public health survey data; administrative claims; other publicly available data	Exposure: buprenorphine/naloxone; outcome: report of diversion (the percentage of applicants who reported knowing that buprenorphine/naloxone was sold on the street) and abuse (the percentage who reported knowing that it was used to get high)	Proportion	NCE: amitriptyline; PCE: methadone, oxycodone, and heroin	Measurement error	To correct for the point estimate, the following formulae was used: Relative abuse of buprenorphine/naloxone = (abuse of buprenorphine or naloxone – abuse of NCE) / (abuse of PCE – abuse of NCE)).
Greene, 2013 (47)	Cohort study	EHR	Exposure: use of oseltamivir; outcome: neuropsychiatric adverse events (ataxia, psychiatric, encephalitis, disturbance of consciousness) and non-neuropsychiatric events (arrhythmia, syncope, convulsions, movement disorder, stroke)	OR, risk difference	NCO: cellulitis, anemia, injury/trauma	Not specified	To reduce bias (bias attenuation), the history of NCOs was included in the propensity score matching model.
Han, 2017 (50)	SCCS	Administrative claims	Exposure: concomitant use of a precipitant of interest (vs. not receiving a precipitant) with secretagogues; outcome: serious hypoglycemia	Rate ratio	NCE: concomitant use of a precipitant of interest with metformin	Confounding bias by inherent hypoglycemic effects of the precipitants	The ratio of the semi-Bayes-adjusted rate ratio associated with the exposure to the semi-Bayes-adjusted rate ratio associated with the corresponding NCE was estimated. The delta method was used for 95% CI calibration (204).
Leonard, 2017 (74)	SCCS	Administrative claims	Exposure: discontinuation of the antihyperlipidemic drugs in the presence of ongoing warfarin therapy; outcome: composite of hospitalization for venous thromboembolism or ischemic stroke	IRR	NCE: discontinuation of pravastatin in the presence of ongoing warfarin therapy	Confounding bias by inherent effects on the outcome of the precipitants	To correct for the point estimate and 95% CI, ratio of IRR for the exposure vs. IRR for the NCE was calculated.

Table continues

Table 2. Continued

First Author, Year (Reference No.)	Study Design	Type of Data	Hypothesized Causal Association	Primary Study Effect Measure	NC(s) Used	Type of Bias as Reported by Study	Methods Used to Correct for Effect of Bias (in Point Estimates, CIs, or P Values)
Gruber, 2018 (48)	Cohort study	Clinical trial data	Exposure: vaccine dose timing; outcome: Severe rotavirus gastroenteritis incidence	Risk difference, risk ratio	NCE: placebo	Confounding bias; administrative censoring	To correct for bias in the estimated risk difference, the effect estimate for the placebo arm was subtracted from the effect estimate for the vaccination arm (for the ratio outcomes, these estimates were divided). A nonparametric bootstrap with 2,000 sample draws with replacement was used to obtain the point estimates and empirical 95% CIs.
Schuemie, 2018 (5)	Southworth, Graham replications: cohort study; Tata case-control replication; case-control study; Tata SCCS replication: SCCS	Southworth, Graham replications: Administrative claims; Tata case-control, SCCS replications: EHR	Southworth replication: Exposure: use of dabigatran (vs. warfarin) in the Southworth and Graham replication and use of selective serotonin reuptake inhibitors in the Tata case-control and SCCS replications; outcome: gastrointestinal bleeding, in the Southworth and Graham replications and upper gastrointestinal bleeding in the in the Tata case-control and SCCS replications	Southworth IRR; Graham replication: HR Tata case-control replication: OR; Tata SCCS replication: IRR	50 NCOs and 150 synthetic PCOs for each replication	Any residual bias (including confounding bias, misclassification, selection bias)	This study developed a methodology for empirical calibration of 95% CI of the effect estimate when a set of negative controls were used in observational studies. The calibration procedure first estimated the distribution of systematic error using the observed estimates for negative and positive controls, then generated calibrated 95% CIs considering both random and systematic error and assuming Gaussian distribution with a mean and log standard deviation linearly related to the true effect size. We refer to this method as "Schuemie's empirical CI calibration method."

Table continues

Table 2. Continued

First Author, Year (Reference No.)	Study Design	Type of Data	Hypothesized Causal Association	Primary Study Effect Measure	NC(s) Used	Type of Bias as Reported by Study	Methods Used to Correct for Effect of Bias (in Point Estimates, CIs, or P Values)
Schuemie, 2018 (111)	Cohort study	Administrative claims	Exposure: use of duloxetine (vs. sertraline); outcome: stroke	HR	52 NCOs and 156 PCOs	Any residual bias (including confounding bias)	The study used Schuemie's empirical CI calibration method (5).
Thorrington, 2018 (126)	Ecological study	Hospital admission data	Exposure: 24-month post-PCV period (vs. pre-PCV period); outcome: pneumonia, empyema, sepsis, otitis media	IRR	NCO: urinary tract infections, the skin and subcutaneous tissue, disorders of the thyroid gland, diseases of the blood, and fractures	Bias arising from potential secular trend	For each outcome of interest, the age-specific ratio of the IRR for the outcome of interest vs. the geometric mean of the IRR for the NCOs was calculated. The minimum and maximum incidence rate ratio across all NCOs were used to represent uncertainty.
Leonard, 2019 (73)	SCCS	Administrative claims	Exposure: concomitant use of clopidogrel and a precipitant drug of interest (clopidogrel-precipitant pairs); outcome: gastrointestinal bleeding or intracranial hemorrhage	Rate ratio	NCE: concomitant use of pravastatin and a precipitant drug of interest (pravastatin-precipitant pairs)	Confounding bias by inherent bleeding effects of the precipitant drug	The ratio of the semi-Bayes-adjusted rate ratio associated with the exposure to the semi-Bayes-adjusted rate ratio associated with the corresponding NCE was estimated. The delta method was used for 95% CI calibration (204).
Hripcsak, 2020 (53)	Cohort study	Administrative claims; EHR	Exposure: use of chlorthalidone (vs hydrochlorothiazide); outcome: multiple cardiovascular outcomes (e.g., acute myocardial infarction, hospitalization for heart failure, etc.)	HR	76 NCOs and 228 PCOs (3 PCOs are generated for each NCO, with true relative risk 1.5, 2, and 4).	Any residual bias	The study used Schuemie's empirical CI calibration method (5).

Table continues



Table 2. Continued

First Author, Year (Reference No.)	Study Design	Type of Data	Hypothesized Causal Association	Primary Study Effect Measure	NC(s) Used	Type of Bias as Reported by Study	Methods Used to Correct for Effect of Bias (in Point Estimates, CIs, or P Values)
Lane, 2020 (70)	Cohort study	Administrative claims; EHR	Exposure: initiation of hydroxychloroquine (vs. initiation of sulfasalazine); outcome: 16 severe adverse events (e.g., gastrointestinal bleeding, acute renal failure, acute pancreatitis, myocardial infarction, stroke, etc.)	HR	NCO: 67 conditions	Confounding bias	The study used Schuemie's empirical CI calibration method (5).
Rodgers, 2020 (107)	Cohort study	EHR	Exposure: use of thiazolidinediones (vs. sulfonylureas); outcome: edema, weight gain	HR	NCO: gastroin- testinal side effects	Confounding bias	To correct for bias, the ratio of the HR for the post period vs. the HR of the prior period was calculated. They verified the removal of bias by replicating the same bias correction method for the NCO (ratio of the HRs for the post vs. prior periods for the NCO). The standard error for the estimates was obtained by bootstrapping.
Shi, 2020 (7)	Cohort study	Integrated health system data	Exposure: receipt of DTap-IPV-Hib vaccine (vs. DTap containing comparator vaccine); outcome: fever; NCO: injury or trauma; NCE: ringworm	Relative risk	1 NCO and 1 NCE	Confounding bias	The study developed a methodology for correcting for a categorical unmeasured confounding. The methodology uses an NCE and an NCO to build a semiparametric model and propose multiply robust estimator for the average treatment effect. The study demonstrated the application of their method in a pharmacoepidemiologic vaccine safety study and showed that the multiply robust estimator provided a smaller bias and protected against the model misspecification.

Table continues

Table 2. Continued

First Author, Year (Reference No.)	Study Design	Type of Data	Hypothesized Causal Association	Primary Study Effect Measure	NC(s) Used	Type of Bias as Reported by Study	Methods Used to Correct for Effect of Bias (in Point Estimates, CIs, or P Values)
Zhou, 2020 (154)	SCCS	Administrative claims	Exposure: concomitant use of a precipitant of interest (vs. not receiving a precipitant) with anticoagulants (anticoagulant-precipitant pairs); outcome: thromboembolism (a composite outcome of stroke and venous thromboembolism)	Rate ratio	NCE: concomitant use of a precipitant of interest with pravastatin (pravastatin-precipitant pairs)	Confounding bias by inherent effect of the precipitant on the primary outcome	The ratio of the semi-Bayes-adjusted rate ratio associated with the exposure to the semi-Bayes-adjusted rate ratio associated with the corresponding NCE was estimated. The delta method was used for 95% CI calibration (204).
Schuemie, 2014 (4)	Cohort study (example 1); case-control study (example 2); SCCS (example 3)	Administrative claims (example 1); EHR (examples 2 and 3)	Exposure: use of isoniazid (example 1); use of selective serotonin reuptake inhibitors (examples 2 and 3); outcome: acute liver injury (example 1); upper gastrointestinal bleeding (examples 2 and 3)	OR	37 NCEs (example 1); 67 NCEs (examples 2 and 3)	Most forms of bias (including confounding bias, misclassification, selection bias)	This study developed a methodology for calibrating the P value of the effect estimate when a set of negative controls were used to account for bias. The method first derived an empirical null distribution from the observed effect estimates for the negative controls, then generated calibrated P values assuming Gaussian distribution to the estimates and taking into account the sampling error of each estimate. We refer to this method "Schuemie's empirical P-value calibration method."
Duke, 2017 (38)	Cohort study	Administrative claims; EHR	Exposure: use of levetiracetam; outcome: risk of angioedema	HR	100 NCOs	Any residual bias	The study used Schuemie's empirical P-value calibration method (4).

Table continues

Table 2. Continued

First Author, Year (Reference No.)	Study Design	Type of Data	Hypothesized Causal Association	Primary Study Effect Measure	NC(s) Used	Type of Bias as Reported by Study	Methods Used to Correct for Effect of Bias (in Point Estimates, CIs, or P Values)
Yuan, 2018 (151)	Cohort study	Administrative claims	Exposure: use of sodium glucose co-transporter 2 inhibitors (overall), and canagliflozin (specifically) (vs. non-sodium glucose co-transporter 2 inhibitor antihyperglycemic agents); outcome: below-knee lower extremity amputation	HR	NCO (number not reported)	Confounding bias	The study used Schuemie's empirical <i>P</i> -value calibration method (4).
Schuemie, 2019 (112)	Case-control studies	Administrative claims	Exposure: use of isotretinoin (Crockett study replication); use of dipeptidyl peptidase-4 inhibitors (Chou study replication); outcome: ulcerative colitis (Crockett study replication); acute pancreatitis (Chou study replication)	OR	35 NCEs	Any residual bias	The study used Schuemie's empirical <i>P</i> -value calibration method (4).
Kim, 2020 (65)	Cohort study	Administrative claims; EHR	Exposure: use of alendronate (vs. use of raloxifene); outcome: hip fracture, vertebral fracture, esophageal cancer, osteonecrosis of the jaw	HR	147 NCOs	Confounding bias	The study used Schuemie's empirical <i>P</i> -value calibration method (4) to construct the empirical null distribution and compare it with the theoretical null distribution to check the presence of bias.
You, 2020 (148)	Cohort study	Administrative claims	Exposure: ACEIs/ARB + CCB vs. ACEI/ARB + TZD vs. CCB+TZD; outcome: all-cause mortality, myocardial infarction, heart failure, stroke, a composite outcome of major adverse cardiac and cerebrovascular events	HR	39 NCOs	Any residual bias	The study used Schuemie's empirical <i>P</i> -value calibration method (4).

Table continues

Table 2. Continued

First Author, Year (Reference No.)	Study Design	Type of Data	Hypothesized Causal Association	Primary Study Effect Measure	NC(s) Used	Type of Bias as Reported by Study	Methods Used to Correct for Effect of Bias (in Point Estimates, CIs, or P Values)
Weinstein, 2020 (139)	Cohort study	EHR	Exposure: use of any paracetamol (vs. use of ibuprofen); outcome: diagnoses of gastrointestinal bleeding, myocardial infarction, ischemic or hemorrhagic stroke, or acute or chronic renal disease	HR	39 NCOs and 147 synthetic PCOs (3 for each NCO)	Channeling bias	The study used Schuemie's empirical CI and P-value calibration method (4, 5).
Morales, 2021 (89)	Cohort study	Administrative claims; EHR	Exposure: use of ACEIs and ARBs (vs. CCBs and thiazide or THDs); outcome: COVID-19 diagnosis; hospital admission with COVID-19; hospital admission with pneumonia; hospital admission with pneumonia, acute respiratory distress syndrome, acute kidney injury, or sepsis	HR	123 NCOs	Confounding bias	The study used Schuemie's empirical CI and P-value calibration method (4, 5).

Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin-receptor blocker; CCB, calcium channel blocker; CI, confidence interval; COVID-19, coronavirus disease 2019; D TaP-IPV-Hib, diphtheria and tetanus toxoids and acellular pertussis adsorbed, inactivated poliovirus, and haemophilus B conjugate vaccine; EHR, electronic health records; HR, hazard ratio; IRR, incidence rate ratio; NC, negative control; NCE, negative control exposure; NCO, negative control outcome; OR, odds ratio; PCE, positive control exposure; PCO, positive control outcome; PCV, pneumococcal conjugate vaccine; SCCS, self-controlled case-series; TZZ, thiazide-type diuretic.

unmeasured confounding. Their study showed the multiply robust estimator reduced bias in a vaccine safety study (7).

*Negative controls to calibrate the P value.* We found 8 pharmacoepidemiologic studies that used the empirical calibration method to calibrate *P* value to correct for the type I error (erroneously rejecting the null hypothesis) (4, 38, 65, 89, 112, 139, 148, 151). Based on the same theoretical background as the CI calibration method, this method proposed by Schuemie et al. (4) uses a large number of negative controls to estimate the empirical null (bias) distribution using the maximum likelihood approach to calibrate *P* values and accounts for both random and systematic error in drawing a statistical inference (Table 2).

*Negative controls to assess the performance of different methods used in drug safety and effectiveness studies.* We found 31 studies that deployed negative controls as a reference standard to compare the performance of different methods for drug safety and effectiveness studies (31, 60, 64, 86, 141, 161–186). The details of these studies are presented in Table 3. The first group of studies in this domain used negative controls specifically to evaluate and validate methods used to identify potential adverse drug reactions (ADRs) for pharmacovigilance purposes (161–186). The studies commonly used a large set of negative controls and known drug reactions (positive controls) to quantify the variations in discriminatory power of different pharmacovigilance methods in detecting safety signals. For example, one of the earliest studies of this kind by Schuemie et al. compared the performance of different pharmacovigilance methods and study designs to identify ADRs, including cohort methods, case-based methods, and other data-mining methods commonly used in spontaneous reporting systems for ADRs (182). The authors used a set of known drug-ADR pairs and negative control pairs to compare the discriminatory power of the aforementioned methods by measuring the area under the receiver-operating-characteristic curve (AUC) across 7 European electronic health record databases (182). Some studies in this group used this approach to examine the transferability of the methods and assess the performance of established ADR identification methods in new settings. For example, Pratt et al. (174) evaluated the consistency of the prescription sequence symmetry analysis method to detect ADRs across settings in 5 countries using positive and negative controls.

The other group of 5 studies in this domain (lower section of Table 3) used negative controls to compare the performance of different analytical methods to mitigate bias in pharmacoepidemiologic studies (31, 60, 64, 86, 141). For example, Davies et al. (31) used an NCO and an NCE to measure and compare the relative bias in the effect estimates from 2 different analytical methods (instrumental variable vs. conventional regression) in assessing the comparative safety of smoking cessation therapies. Weinstein et al. (141) used 31 NCOs to compare residual bias resulting from the 6 different confounding adjustment methods in evaluating comparative safety of frequently used over-the-counter analgesics. The authors reported the proportion of statistically significant NCOs as an indicator of bias (141).

### Validating the assumptions of negative controls in pharmacoepidemiologic studies

We identified 3 common strategies for selecting negative control associations satisfying the first assumption (lack of causality assumption) in the included studies: 1) logically impossible association (e.g., causally implausible timing of the event), 2) clinically implausible association, and 3) association with no previous evidence of potential causal association. The use of a prespecified set of drug-outcome pairs classified as either positive or negative controls based on prior scientific evidence has become common practice in validation of pharmacovigilance methods. The Exploring and Understanding Adverse Drug Reactions (EU-ADR) reference set was developed as part of the EU-ADR project and composed of 44 positive and 50 negative control drug-outcome associations across 10 outcomes that were considered important from a pharmacovigilance perspective (187). The EU-ADR researchers started the selection by identifying drugs with enough exposure in the EU-ADR database network and checked the lack of causality for each drug-outcome pair against 3 resources: 1) information from published literature and drug product labels, 2) review of a spontaneous reporting system, and 3) manual verification by expert physicians. The Observational Medical Outcome Partnership (OMOP) reference set was another commonly used set as a reference standard in the OMOP studies (188, 189). This set included 165 positive control and 234 negative control associations across 4 common adverse outcomes (acute kidney injury, acute liver injury, acute myocardial infarction, and upper gastrointestinal bleeding) and similarly checked the lack of causality using evidence from: 1) Food and Drug Administration (FDA) structured product labeling; 2) systematic literature reviews by Tisdale et al. (190) and 3) manual search of the literature. More recently, a semiautomated process for identifying negative controls, in which the algorithm automatically combines up-to-date evidence from various sources (spontaneous reports, scientific literature, and both American and European product labeling), has been introduced (191). We also identified reference sets that were built for specific population groups (e.g., pediatric population) (172, 192, 193) or therapeutic areas (e.g., oncology, vaccine) (164, 193). More information on these reference sets is presented in Web Table 2.

For the second assumption of negative controls (shared bias structure assumption), while there is no empirical test to check this assumption, we found 8 studies that explicitly provided, and compared, the distributions of major measured covariates/confounders between the primary exposure or outcome of interest and their corresponding NCE or NCO (Table 4) (34, 50, 57, 72, 83, 104, 128, 154). While the true similarity in the bias structure of the exposure-outcome association and the negative control association remains unobserved, the premise behind this approach is that showing similar distributions and effects of observed confounders can provide some evidence for potential similarities in unmeasured confounders. Out of the 8 studies, one study provided additional steps to improve the comparability of the bias structure. Dillon et al. (34), in examining the

**Table 3.** Studies That Used Negative Controls to Assess the Performance of Methods for Drug Safety and Effectiveness Studies

First Author, Year (Reference No.)	Description of the Use of NCs in the Study
Brown, 2007 (165)	<p><i>Studies That Used NCs as Reference Standards in Evaluating the Performance of Pharmacovigilance Methods</i></p> <p>The study assessed the performance of sequential analysis for active surveillance of ADR in large observational databases. The performance was measured by applying sequential analysis to 5 PC and 2 NC drug-outcome pairs in administrative data from 9 health plans participating in the HMO Research Network's Center for Education and Research on Therapeutics. The signal was determined by testing the maximized sequential probability ratio against the null hypothesis, and the number of significant signals among PC and NC drug-outcome pairs was reported.</p>
Brown, 2009 (166)	<p>The study assessed the variation in the performance of sequential analysis with alternative analytical choices with regard to exclusion criteria and surveillance period, and the results were compared with the performance of sequential analysis with the original set of analytical choices, which was assessed in an earlier study (Brown et al. (165)). The same data source and methods from the previous study were used for the assessment including the choice of PC and NC drug-outcome pairs.</p>
Ryan, 2012 (177)	<p>The study evaluated the performance (measured with sensitivity, specificity, PPV, and AUC) of 8 different methods in 5 US health-care databases (administrative claims and EHR) participating in the OMOP network, using 53 unique drug-outcome pairs (9 PC and 44 NC drug-outcome pairs). The methods evaluated were: 1) disproportionality analysis, 2) ICTPD, 3) SCCS, 4) case-control surveillance, 5) case-crossover, 6) observational screening, 7) high-dimensional propensity score, and 8) incident user design.</p>
Schuemie, 2012 (182)	<p>The study evaluated the performance (measured with AUC) of 10 different methods to detect ADR in health-care databases participating in EU-ADR, using a set of PCs and NCs as reference standards. The methods evaluated in the study included: 1) spontaneous reporting system methods (PRR, ROR, GPS, BCPNN), 2) cohort methods (IRR, LGPS, Bayesian hierarchical model), 3) case-based methods (matched case-control, SCCS), and 4) LEOPARD.</p>
DuMouchel, 2013 (162)	<p>The study evaluated the performance (measured with AUC) of disproportionality analysis with different combinations of analytical choices (with regard to outcome definition, disproportionality metric, stratification, time at risk across) for 4 health outcomes (acute liver failure, acute MI, acute renal failure, and upper GI bleeding) in 5 US health-care databases (administrative claims and EHR) participating in the OMOP network and 6 simulated data sets. A collection of 165 PC and 234 NC drug-outcome pairs were used as reference standards for evaluation.</p>
Madigan, 2013 (168)	<p>The study evaluated the performance of case-control design for ADR risk identification and compared AUC for the design with various combinations of analytical choices with regard to: 1) number of controls, 2) minimum time before outcome, 3) time at risk, and 4) cohort nesting within indication. The reference standards included 165 PC and 234 NC drug-outcome pairs across 4 health outcomes (acute liver failure, acute MI, acute renal failure, upper GI bleeding). The performance was evaluated in 5 US health-care databases (administrative claims and EHR) participating in the OMOP network.</p>
Norén, 2013 (171)	<p>The study evaluated the performance of the calibrated self-controlled cohort analysis with temporal pattern discovery as a tool for risk identification, by comparing AUC for different combinations of the design choices with regard to: 1) surveillance period and 2) control period. The method was applied to 165 PC and 234 NC drug-outcome pairs across 4 health outcomes (acute liver failure, acute MI, acute renal failure, and upper GI bleeding). The performance was evaluated in 5 US health-care databases (administrative claims and EHR) participating in the OMOP network.</p>
Reich, 2013 (175)	<p>The study evaluated the performance of different algorithms to define health outcomes of interest for 3 health outcomes (acute kidney injury, acute liver injury, and acute MI) for risk identification methods in observational data. The algorithms were tested across an array of analytical methods and design choices (including choices for preexposure control period and time at risk) in 4 US health-care databases participating in OMOP network. The methods were executed against predefined drug-outcome pairs (test cases) that included PCs and NCs. The performance was measured with AUC, bias, and minimal detectable relative risk.</p>

Table continues

Table 3. Continued

First Author, Year (Reference No.)	Description of the Use of NCs in the Study
Reich, 2013 (176)	The study assessed the effects of drug and outcome prevalence on the feasibility and performance for risk identification methods in observational data. The methods of interest (new user cohort design, case control design, SCCS, ICC, ICTPD, disproportionality analysis, and LGPS) were tested against PCs and NCs across 4 health outcomes (acute kidney injury, acute liver injury, acute MI, and upper GI bleeding) in 3 US health-care databases participating in the OMOP network. The performance was measured with AUC.
Ryan, 2013 (178)	The study evaluated the performance (measured with AUC, bias, and coverage probabilities) of new-user cohort method for detecting safety signals in 5 US health-care databases (administrative claims and EHR) participating in the OMOP network and 6 simulated data sets. Different combinations of analytical choices were tested, including choices for: 1) required observation time prior to the exposure, 2) nesting within population with the indication of the target drug, 3) comparator population, 4) time-at-risk, 5) propensity score covariate selection strategy, 6) covariate eligibility window, 7) dimension to include as potential covariates, 8) additional covariates in the propensity score model, 9) propensity score trimming, and 10) analysis method. The method was tested against 165 PC and 234 NC drug-outcome pairs across 4 health outcomes (acute liver failure, acute MI, acute renal failure, and upper GI bleeding).
Ryan, 2013 (179)	The study evaluated the performance of SCC method for detecting safety signals in 5 US health-care databases (administrative claims and EHR) participating in the OMOP network and 6 simulated data sets. The performance metrics of AUC, bias (measured with relative risk estimates from the NC analysis) and coverage probabilities (proportion of the CIs that contained the true relative risk) for different combinations of design choices of the method were measured. Different combinations of design choices were tested, including choices for 1) exposure definition, 2) outcome definition, 3) time-at-risk, 4) control period. A set of 165 PC and 234 NC drug-outcome pairs across 4 health outcomes (acute liver failure, acute MI, acute renal failure, and upper GI bleeding) were used as reference standards.
Ryan, 2013 (180)	The study compared the performance (measured with AUC, bias, MSE, and CI coverage probability) of 7 established pharmacovigilance methods including: 1) new user cohort, 2) case-control, 3) SCCS, 4) SCC, 5) disproportionality analysis, 6) temporal pattern discovery, and 7) LGPS. The methods were tested against 165 PC and 234 NC drug-outcome pairs across 4 health outcomes (acute liver injury, acute MI, acute renal failure, and GI bleeding) in 5 US health-care databases (administrative claims and EHR) participating in the OMOP network.
Ryan, 2013 (181)	The study used simulated health-care data to evaluate the performance (measured with AUC, bias, and MSE) of 7 observational designs for ADR identification. The tested methods included: 1) case-control, 2) new user cohort method, 3) disproportionality methods, 4) ICTPD, 5) LGPS, 6) SCC, and 7) SCCS. The methods were tested against 399 reference drug-outcome pairs (including 165 PCs and 234 NCs) across 4 health outcomes (acute liver injury, acute MI, acute renal failure, and GI bleeding) in 6 simulated data sets.
Schuemie, 2013 (183)	The study replicated the OMOP experiment in European databases and compared the performance of the established methods for risk identification, including: 1) case-control, 2) new user cohort method, 3) disproportionality methods, 4) ICTPD, 5) LGPS, 6) SCC, and 7) SCCS. The methods were tested against 165 PC and 234 NC drug-outcome pairs across 4 health outcomes (acute liver injury, acute MI, acute renal failure, and GI bleeding) in 6 health-care databases (administrative claims and EHR) participating in the EU-ADR database network. The performance was measured with AUC and bias.
Schuemie, 2013 (184)	The study evaluated the performance (measured with AUC, bias, and coverage probabilities) of LGPS and LEOPARD in in 5 US health-care databases (administrative claims and EHR) participating in the OMOP network and 6 simulated data sets in safety signal detection. Different combinations of analytical choices were tested, including choices for: 1) data source, 2) exposure definition, 3) run-in period, 4) carry-over period, and 5) shrinkage application. The method was tested against 165 PC and 234 NC drug-outcome pairs across 4 health outcomes (acute liver failure, acute MI, acute renal failure, upper GI bleeding).
Suchard, 2013 (185)	The study evaluated the performance (measured with AUC, bias, and coverage probabilities) of SCCS for safety signal detection in 5 US health-care databases (administrative claims and EHR) participating in the OMOP network and 6 simulated data sets. Different combinations of design choices were tested, including choices for: 1) outcome definition, 2) multivariate adjustment, 3) at-risk window, 4) observation time-window definition, and 5) minimum observation length. The method was tested against 165 PC and 234 NC drug-outcome pairs across 4 health outcomes (acute liver failure, acute MI, acute renal failure, and upper GI bleeding).

Table continues

Table 3. Continued

First Author, Year (Reference No.)	Description of the Use of NCs in the Study
Gagne, 2014 (167)	The study developed a semiautomated process for active safety monitoring for new drugs in distributed data environment for health-care data. Five PC and 2 NC drug-outcome pairs were defined and used for testing the validity of the process, and whether the system generated an alert for each drug-outcome pair was reported.
Pratt, 2015 (174)	The study evaluated the robustness and consistency of PSSA to detect ADRs across settings in 5 different countries. The method was applied to 1 PC and 2 NCs as test cases. The ASR and 95% CIs were estimated for the associations with ADRs for each country.
Backenroth, 2016 (163)	The study compared AUCs for the case-control method with different analytical choices (including choices for: 1) covariates used (none, demographics, phenome-wide association study (PHEWAS) groups plus demographics), 2) covariate selection method (none, or 1-step LASSO), and 3) estimation method (marginal odds ratio, 1 model per drug-outcome pair, or 1 model per outcome)) across 4 health outcomes (acute kidney injury, acute liver injury, acute MI, and GI ulcer hospitalization) in a US EHR database.
Marbac, 2016 (169)	The study proposed a model-based method for analyzing spontaneous reporting databases (Bayesian model selection in logistic regression), and compared the performance (measured with number of signals, rate of positive controls, rate of negative controls, rate of unknown signals) to the well-established approaches, including 4 disproportionality-based methods (PRR, ROR, reporting Fisher exact test, FDR-based GPS, Lasso-based logistic regressions, and model-based logistic regressions) across 4 health outcomes (acute MI, acute kidney injury, acute liver injury, and upper GI bleeding) in the French pharmacovigilance database. The OMOP reference set of test cases were used to evaluate the performance, which included 165 PC and 234 NC drug-outcome pairs.
Osokogu, 2016 (172)	The study assessed the performance (AUC) of 2 established signal detection algorithms (PRR and empirical Bayes geometric mean) in the pediatric population in data from the US FAERS. Thirty-seven PC and 90 NC drug-outcome pairs from a pediatric-specific Global Research in Pediatric reference set were used.
Hauben, 2017 (164)	The study estimated the effects of database restriction for oncology drug on the signal detection performance using FAERS data. Multi-item GPS analysis, a type of disproportionality analysis, was carried out in restricted and unrestricted data sets for a predefined oncology-specific reference set (including 638 PC and 67 NC drug-outcome pairs). The method performance was measured with sensitivity, specificity, PPV, NPV, signal/noise, $F$ , and Matthews correlation coefficient.
Nishtala, 2017 (170)	The study assessed the performance of PSSA in the New Zealand prescription database for risk identification using 6 PC drug exposures and 6 NC drug exposures. The adjusted sequence ratios and 95% CIs were estimated for the associations with ADRs.
Pierce, 2017 (173)	The study assessed the potential utility of social media data for monitoring drug safety signals and examined whether a safety signal appears in social media prior to being reported to a spontaneous reporting system. An analysis of social media posts was conducted for 10 safety signals (PCs) recently identified by FDA, and 6 NC drug-outcome pairs. The number of associations appeared in the social media data was reported for each of the pairs as an outcome.
Trinh, 2018 (161)	The study evaluated the performance of combining change-point analysis method with disproportionality analysis for detecting changes in safety signals in the French National Pharmacovigilance Database (Base Nationale de Pharmacovigilance or BNPV) and EudraVigilance database. The performance was assessed by testing the change-point analysis hypothesis against the null hypothesis for 39 PC and 56 NC drug-outcome pairs selected from the OMOP reference set in EudraVigilance database, and 2 PC drug-outcome pairs were used as test cases in BNPV database.
Thurin, 2020 (186)	The study assessed the performance of 3 case-based designs (SCCS, case-control, and case-population) for identifying risk of upper GI bleeding in the French National Healthcare System Database. Different combinations of design and analytical choices for each method were tested against the reference standards (22 PC and 42 NC drugs for upper GI bleeding), and the performance (measured with AUC, MSE, and coverage probability) was compared. The best-performing method was selected for calibration using the empirical $P$ -value calibration method to improve the accuracy of the method.

Table continues



Table 3. Continued

First Author, Year (Reference No.)	Description of the Use of NCs in the Study
McGrath, 2015 (86)	<p><i>Studies That Used NCs to Compare Different Methods for Bias Mitigation in Pharmacoepidemiologic Studies</i></p> <p>The study assessed the residual confounding associated with different analytical choices (marginal structural models with different specifications and sample restriction in multiple different ways) in evaluating influenza vaccine (comparing the vaccinated vs. unvaccinated) effectiveness in adult hemodialysis patients. The study used data from the US Renal Data System linked with data from a commercial dialysis provider to compare all-cause mortality between the groups during 3 influenza seasons. The outcome measured in pre-influenza season was used as a NC outcome to measure residual confounding. The study found that marginal structural models did not improve confounding controls even after adding additional variables. Restricting the sample to more homogeneous and healthier populations (e.g., by applying stricter survival requirements) moved the results toward the null in pre-influenza seasons, indicating reduction of residual bias.</p>
Davies, 2017 (31)	<p>The study applied 3 techniques (NC outcome, NC populations, and tests of covariate balance) to assess the relative bias of the effect estimates by 2 analytical methods: 1) instrumental variable (physicians' prescribing preferences) approach, and 2) a conventional regression. The Clinical Practice Research Datalink was used to investigate the effect of smoking cessation therapies (varenicline vs. nicotine replacement products) on suicide and self-harm, depression, and urinary tract infection was used as the NC outcome. Bias component plot was applied to illustrate the relative bias of the 2 methods.</p>
Weinstein, 2017 (141)	<p>The study compared the residual bias associated with 6 different analytical methods (multivariate outcome modeling with 3 different specifications, publication variable PS adjustment, large-scale PS adjustment, and large-scale PS adjustment with large-scale outcome modeling) to adjust confounding in evaluating safety of frequently used over-the-counter medications (paracetamol vs. ibuprofen) using data from the Clinical Practice Research Datalink. The study used 31 NC outcomes to assess the bias and reported the proportion of NC outcomes that were statistically significant as an indicator of residual bias. While using multivariate outcome modeling alone and the publication variable PS adjustment resulted in residual bias (fraction significant &gt; 5%), the other 2 large-scale PS adjustment methods resulted in no significant NC results.</p>
Kaiser, 2018 (64)	<p>The study assessed potential biases associated with different analytical choices in estimating the association (HR) between statin use and incident MI using a retrospective cohort design. Six analytical variations were compared for potential bias: 1) crude analysis, 2) restriction to those eligible for statins, 3) multivariable adjusted among eligible population, 4) restriction to eligible new users, 5) multivariable adjusted among eligible new users, and 6) PS-matched among eligible new users. Noncardiovascular mortality was used as an NC outcome. HRs from negative control analysis were compared with HRs from the primary analysis to assess bias. Data from participants in the Cardiovascular Health Study from 1989 to 2004 were used.</p>
Izurieta, 2019 (60)	<p>The study assessed the impact of using multiple imputation in reducing potential bias due to residual confounding in a HZY effectiveness study (comparing the vaccinated vs. unvaccinated) using 13 NC outcomes. The authors had previously published a HZY effectiveness study using Medicare claims data and checked the residual bias using 13 NC outcomes in the original study (59). This study additionally used and linked the MCBS to the Medicare claims to impute 3 new MCBS variables, potential confounders missing in the original analysis. The same 13 NC outcomes from the previous study were used to detect residual confounding in the imputation analysis and compare the results with the original analysis. The study found similar HR estimates compared with the original study across the 13 NC outcomes (CIs overlapping). The point estimates shifted toward the null for 8 out of 13 outcomes, and the CIs were wider for the imputation analysis.</p>

Abbreviations: ADR, adverse drug reaction; ASR, adjusted sequence ratio; AUC, area under the receiver operating characteristic curve; BCPNN, Bayesian confidence propagation neural network; CI, confidence interval; EHR, electronic health records; EU-ADR, Exploring and Understanding Adverse Drug Reactions; FAERS, Food and Drug Administration's Adverse Event Reporting System; FDA, US Food and Drug Administration; FDR, false discovery rate; GI, gastrointestinal; GPS,  $\gamma$  Poisson shrinker; HMO, health maintenance organization; HR, hazard ratio; HZV, herpes zoster vaccine; ICPTD, information component temporal pattern discovery; IRR, incidence rate ratio; LASSO, least absolute shrinkage and selection operator; LEOPARD, longitudinal evaluation of observational profiles of adverse events related to drugs; LGPS, longitudinal  $\gamma$  Poisson shrinker; MCBS, Medicare Current Beneficiary Survey; MI, myocardial infarction; MSE, mean squared error; NPV, negative predictive value; OMOP, Observational Medical Outcomes Partnership; PC, positive control; PPV, positive predictive value; PRR, proportional reporting ratio; PS, propensity score; PSSA, prescription sequence symmetry analysis; ROR, reporting odds ratio; SCC, self-controlled cohort; SCCS, self-controlled case series; US, United States.

effects of gaps in antihypertensive medication adherence on injurious falls using the gaps for antithrombotic medication adherence as an NCE, checked the distributions of measured covariates between the exposure and NCE samples and found significant differences for gender, comorbidities, and rate of regular medication use between the samples. Accordingly, the authors reweighted the NCE sample using inverse probability weighting to adjust for differences in the covariate distributions (34). The details of all 8 studies are provided in Table 4.

## DISCUSSION

In this scoping review, we identified 184 pharmacoepidemiologic studies that reported the use of negative controls in their design and analysis or added novel utilities of negative control methods. We identified 4 domains for the use of negative controls: 1) detection of bias, 2) correction for bias in the effect estimates, 3) calibration of *P* value, and 4) assessing the performance of different methods in drug safety and effectiveness studies.

Among the methodological developments of negative controls, our study revealed that the methods for empirical calibration of *P* value or 95% CIs were used more than other methods in the field (4, 5). An advantage of the empirical calibration methods is that they do not require the assumption of equi-confounding, which is often unverifiable in practice. Applying this methodology requires defining multiple negative controls for the primary exposure-outcome association, which can be a challenging task for researchers. Thus the OMOP researchers provided a semi-automated process for selecting negative controls (191). One important condition for using the method is that the confounders of negative control associations should be drawn from the same distribution as confounders of the primary drug-outcome association. In addition, identifying negative control drugs is more challenging for case-control designs, as some of these drugs might cause confounding by indication (4). Also, a large amount of heterogeneity of bias estimates of negative controls may lead to increased type II error reducing the power to detect the true safety and efficacy signal (194).

As the use of observational data, for example during the coronavirus disease 2019 (COVID-19) pandemic, gains momentum, the importance of negative controls as a tool to address unmeasured confounding becomes more salient. Two recent studies during the COVID-19 pandemic investigated the effectiveness of mRNA vaccines using observational data, and both studies used negative controls to address unmeasured confounding in the data set (195, 196). These studies were published after our review's article inclusion period and hence were not included.

While use of negative controls has been primarily for addressing unmeasured confounding in pharmacoepidemiologic studies, they have been used for other types of biases, such as selection bias or measurement error. In addition, under recent developments in the proximal causal inference framework, negative controls are viewed as a version of proxy variables for sources of bias (197). The proximal

causal inference framework expands the potential utility of proxy variables in bias correction by acknowledging the presence of different sources of bias with imperfect information (e.g., unmeasured common cause or descendent, unmeasured mediator, censored data) using proxies to make inferences on these latent variables of interest (197–199).

Methods for bias correction have evolved from the bias attenuation work by Flanders et al. (3), which demonstrated the effect of including future exposure in time-series analysis in reducing the amount of bias. This work was further extended by Miao and Tchetgen Tchetgen (200) for further reduction and even elimination of bias by including more exposures in the future in time-series analysis under some untestable parametric assumptions. A recent study provided a selective review of major methodological developments of negative controls in broader epidemiology (9). Our review discovered that many of these new developments for bias correction, such as the control outcome calibration approach (COCA) (6), a method based on scale-invariant generalization of DiD (157), and methods of double negative controls (7, 8, 201), have not yet been widely adopted in pharmacoepidemiologic studies.

Despite the growing attention to negative controls, there is no standardized guidance or statistical method to test the underlying assumptions of negative controls. While identifying unmeasured confounders without any restriction or assumption may be unlikely, recent work by Kummerfeld et al. (202) provided an empirical test for validating a triplet of disconnected negative controls. In the absence of empirical tests, directed acyclic graphs and other structural models that require subject matter knowledge to justify the causal associations can provide a proxy, yet a formal justification for bias. Future research to provide a formal approach to test for the assumptions of negative controls is warranted.

Our study had some limitations. Although we conducted the literature search and selection of articles in a comprehensive and systematic manner, it is possible that some articles might have been missed given the nature of the topic. For example, it is possible that other nomenclature used interchangeably with negative controls might not have been included in our search strategy. A recent study published after our search was conducted, for instance, mentioned the use of falsification endpoint in some negative control studies (9). While we identified some studies using the term falsification endpoint, including every possible nomenclature in our search was not logistically feasible given the substantially large number of article outputs, and we do not believe addition of other terms would change the identified domains for utility of negative controls in pharmacoepidemiologic studies beyond our results. Negative controls are often used as a sensitivity or secondary analysis, which might not be reported in detail in the fields searchable in the databases. A previous study showed that many studies do not report use of a specific methodology in the title and abstract (203). We attempted to minimize the risk of missing a relevant article by manually reviewing the reference lists of the selected articles. Finally, to our knowledge, there is no standardized checklist for the use and report of negative control analyses in observational studies, thus we

**Table 4.** Pharmacoepidemiologic Studies that Provided Analysis for Validating the Assumptions of Negative Controls

First Author, Year (Reference No.)	Study Design	Type of Data Used	Hypothesized Causal Association	Primary Study Effect Measure	Type of NC	Types of Bias as Reported by Study	Statistical Approach Used to Test the Underlying Assumptions of NC
Ivers, 2015 (57)	Case-control study	Primary data collected for the study	Exposure: receipt of oral inactivated bivalent whole-cell cholera vaccine; outcome: acute watery diarrhea with a stool sample positive for cholera	Relative risk	NCO: acute watery diarrhea with a stool sample that tested negative for cholera	Not specified	Study showed the distributions of factors related to clinical presentation and treatment for cholera (time from symptom onset to admission, serotype, dehydration stage at presentation, treatment received at clinic, volume of oral rehydration solution given in clinic, volume of intravenous fluid given in clinic, admitted overnight to the cholera treatment unit, duration of stay at cholera treatment unit, and the discharge outcome type), age, gender, participation response rate, earthen floor in home, ever attended school, electricity in house, number of people in household, whether agriculture is the main source of income, whether main toilet is a latrine, ever admitted to a cholera treatment unit overnight, household member with cholera in the previous week, household member ever spent a night in a cholera treatment unit, and household member with diarrhea in the previous week across the NCO and the main outcome groups.
Lazarus, 2016 (72)	Cohort study	Main study: data from an existing cohort study; replication study: integrated health system data	Exposure: use of proton pump inhibitors; outcome: chronic kidney disease	HR	NCE: use of histamine H2-receptor antagonists	Not specified	Study provided the distributions of age, gender, race, education, health insurance, annual household income, estimated glomerular filtration rate, ratio of urinary albumin to creatinine, smoking, body mass index, systolic blood pressure, prevalent medical conditions, and concomitant medication use for the NCE group along with the primary exposure and the control group.
Han, 2017 (50)	SCCS	Administrative claims	Exposure: concomitant use of a precipitant of interest (vs. not receiving a precipitant) with secretagogues; outcome: serious hypoglycemia	Rate ratio	NCE: concomitant use of a precipitant of interest with metformin	Confounding bias by inherent hypoglycemic effects of the precipitants	Study provided the distributions of gender, median age, and proportion of cases receiving antidiabetics other than secretagogues in the past 30 days before events across the NCE and the primary exposures.

Table continues

Table 4. Continued

First Author, Year (Reference No.)	Study Design	Type of Data Used	Hypothesized Causal Association	Primary Study Effect Measure	Type of NC	Types of Bias as Reported by Study	Statistical Approach Used to Test the Underlying Assumptions of NC
Dillon, 2019 (34)	Cohort study	Pharmacy dispensing data, primary data collected for the study	Exposure: gaps in antihypertensive medication adherence; outcome: injurious fall	Relative risk	NCE: gaps in antithrombotic medication adherence	Confounding bias resulting from healthy adherer bias	Study checked for the distribution of covariates across the exposure (antihypertensive) sample and the NCE (antithrombotic) samples. They found statistically significantly different distributions for gender, comorbidities, and rate of regular medication use between the exposure and the NCE samples. Accordingly, the NCE sample was reweighted using inverse probability weighting, with weights being the ratio of the predicted probabilities using 2 probit regressions, one with and one without statistically different participant characteristic variables of gender, comorbidities, and rate of regular medication use.
Markovic, 2019 (83)	Cohort study	Data from an existing cohort study	Exposure: maternal exposure to NSAIDs during pregnancy; outcome: neurodevelopmental outcomes (attention problem score) in children	Mean difference	NCE: maternal NSAIDs use before pregnancy; NCO: somatic complaints in children	Confounding bias	Study provided the distributions of child factors (gender, birth weight, gestational age at birth, and Apgar score at 5 minutes after birth) and maternal factors (age, ethnicity, education, family income, body mass index, smoking during pregnancy, alcohol use during pregnancy, psychopathology symptoms, cognitive ability score, and concomitant-medication use) for the NCE and the exposure.
Ray, 2019 (104)	Case-control study	EHR	Exposure: receipt of inactive influenza vaccine; outcome: positive test result for any influenza virus	OR	NCO: positive test result for respiratory syncytial virus	Confounding bias	Study provided the distributions of age, gender, Diagnostic Cost Group risk score, influenza vaccination in previous season, year, number of days from September 1 to vaccination, and number of days from vaccination to influenza test for the NCO and the primary outcome. The study also mentioned that the known differential risk factors between the outcome and the NCO are related to age, and that the study adjusted for age and removed ages below 2 years old. Study also cited earlier research that found similar characteristics between respiratory syncytial virus and influenza A in relation to demographic characteristics, chronic illnesses, residence in long-term care facilities, and smoking.

Table continues

Table 4. Continued

First Author, Year (Reference No.)	Study Design	Type of Data Used	Hypothesized Causal Association	Primary Study Effect Measure	Type of NC	Types of Bias as Reported by Study	Statistical Approach Used to Test the Underlying Assumptions of NC
Tien, 2020 (128)	Cohort study	Primary data collected for the study	Exposure: chlorhexidine gluconate bathing (vs. over-the-counter non-chlorhexidine-based antibacterial soap or cleansing lotion); outcome: gram-positive cocci-related, skin flora-related, or central line-associated bloodstream infection	HR	NCO: gut-origin bacteremia	Confounding bias, participation bias	Study provided the results of regression analysis for the NCO with the same set of covariates that were used in the regression analysis for the primary outcome. The covariates included 2% chlorhexidine daily bathing, age, gender, diagnosis of hematological, relapsed/refractory diseases, postchemotherapy neutropenia, autologous peripheral blood stem cell transplantation, chemotherapy, central venous catheter use, portacath use, oral ciprofloxacin/levofloxacin for prophylaxis, oral sulfamethoxazole/trimethoprim for prophylaxis, multidrug-resistant organism carriage, diabetes, and hospital mortality. Results for both the univariate and multivariate analyses were provided across the primary outcome and the NCO. In the univariate analysis of the primary outcome, 2% chlorhexidine daily bathing, post-chemotherapy neutropenia, central venous catheter use, and oral ciprofloxacin/levofloxacin for prophylaxis were statistically significant. In the univariate analysis of the NCO, 2% chlorhexidine daily bathing, post-chemotherapy neutropenia, chemotherapy, and oral ciprofloxacin/levofloxacin for prophylaxis were statistically significant.
Zhou, 2020 (154)	SCCS	Administrative claims	Exposure: concomitant use of a precipitant of interest (vs. not receiving a precipitant) with anticoagulants (anticoagulant-precipitant pairs); outcome: thromboembolism (a composite outcome of stroke and venous thromboembolism)	Rate ratio	NCE: concomitant use of a precipitant of interest with pravastatin (pravastatin-precipitant pairs)	Confounding bias by inherent effect of the precipitant on the primary outcome	Study provided the distributions of gender, age, and baseline medical conditions (prior diagnoses of atrial fibrillation, stroke, serious bleeding, cirrhosis, chronic kidney disease, and hypertension), prior prescription of nonsteroidal antiinflammatory drugs in the 30 days prior to initiation of oral anticoagulant or pravastatin, and prior prescription of antiplatelet agents in the previous 30 days) across the NCE and the primary exposures.

Abbreviations: EHR, electronic health records; HR, hazard ratio; NC, negative control; NCE, negative control exposure; NCO, negative control outcome; NSAID, nonsteroidal antiinflammatory drug; OR, odds ratio; SCCS, self-controlled case series.

were not able to grade the study quality of the included articles.

## CONCLUSION

Our study laid out a holistic overview of the utility of negative controls to provide guidance for different applications of negative controls in future pharmacoepidemiologic studies. While the primary use of negative controls has been to detect bias, we identified other utility domains for their application, such as assessing the performance of epidemiologic methods of identifying ADRs. Despite considerable methodological advancements in the use of negative controls for bias correction or *P*-value calibration in recent years, only a few studies have applied these methods for their research questions. Finally, while the validity of negative control use is based on the underlying assumptions, we found that many studies selected their negative controls without efforts made toward validating the underlying assumption. Violation of the underlying assumptions can lead to biased estimates.

## ACKNOWLEDGMENTS

Author affiliations: Department of Practice, Sciences, and Health Outcomes Research, University of Maryland School of Pharmacy, Baltimore, Maryland, United States (Zafar Zafari, Jeong-eun Park, Chintal H. Shah, Susan dosReis, Emily F. Gorman); and Division of Epidemiology I, Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, Food and Drug Administration, Beltsville, Maryland, United States (Wei Hua, Yong Ma, Fang Tian).

This work was supported by the Food and Drug Administration (FDA) of the US Department of Health and Human Services (HHS) as part of a financial assistance award (FAIN: U01FD005946) totaling \$230,801 with 100% funded by FDA/HHS.

A version of this work was presented at International Conference on Pharmacoepidemiology & Therapeutic Risk Management (ICPE), August 24–28, 2022, Copenhagen, Denmark.

The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by FDA/HHS, or the US Government.

Conflict of interest: none declared.

## REFERENCES

- Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21(3):383–388.
- Flanders WD, Klein M, Darrow LA, et al. A method for detection of residual confounding in time-series and other observational studies. *Epidemiology*. 2011;22(1):59–67.
- Flanders WD, Strickland MJ, Klein M. A new method for partial correction of residual confounding in time-series and other observational studies. *Am J Epidemiol*. 2017;185(10):941–949.
- Schuemie MJ, Ryan PB, Dumouchel W, et al. Interpreting observational studies: why empirical calibration is needed to correct *P*-values. *Stat Med*. 2014;33(2):209–218.
- Schuemie MJ, Hripcsak G, Ryan PB, et al. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proc Natl Acad Sci U S A*. 2018;115(11):2571–2577.
- Tchetgen Tchetgen E. The control outcome calibration approach for causal inference with unobserved confounding. *Am J Epidemiol*. 2014;179(5):633–640.
- Shi X, Miao W, Nelson JC, et al. Multiply robust causal inference with double-negative control adjustment for categorical unmeasured confounding. *J R Stat Soc Ser B Stat Methodol*. 2020;82(2):521–540.
- Miao W, Shi X, Tchetgen TCHETGEN E. A confounding bridge approach for double negative control inference on causal effects [preprint]. *arXiv*. 2020. <https://doi.org/10.48550/arXiv.1808.04945>. Accessed August 18, 2022.
- Shi X, Miao W, Tchetgen TE. A selective review of negative control methods in epidemiology. *Curr Epidemiol Rep*. 2020;7(4):190–202.
- Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467–473.
- Aronson JK, Ferner RE. Analysis of reports of unintended pregnancies associated with the combined use of non-enzyme-inducing antibiotics and hormonal contraceptives. *BMJ Evid-Based Med*. 2021;26(3):112–113.
- Bedson J, Chen Y, Ashworth J, et al. Risk of adverse events in patients prescribed long-term opioids: a cohort study in the UK Clinical Practice Research Datalink. *Eur J Pain*. 2019;23(5):908–922.
- Bijlsma MJ, Vansteelandt S, Janssen F, et al. The effect of adherence to statin therapy on cardiovascular mortality: quantification of unmeasured bias using falsification end-points. *BMC Public Health*. 2016;16(1):303.
- Brassard P, Wu JW, Ernst P, et al. The effect of statins on influenza-like illness morbidity and mortality. *Pharmacoepidemiol Drug Saf*. 2017;26(1):63–70.
- Brookhart MA, Patrick AR, Dormuth C, et al. Adherence to lipid-lowering therapy and the use of preventive health services: an investigation of the healthy user effect. *Am J Epidemiol*. 2007;166(3):348–354.
- Brookhart MA, Walker AM, Lu Y, et al. Characterizing vaccine-associated risks using cubic smoothing splines. *Am J Epidemiol*. 2012;176(10):949–957.
- Burkard T, Hügle T, Layton JB, et al. Risk of incident osteoarthritis of the hand in statin initiators: a sequential cohort study. *Arthritis Care Res*. 2018;70(12):1795–1805.
- Busby J, McMenamin Ú, Spence A, et al. Angiotensin receptor blocker use and gastro-oesophageal cancer survival: a population-based cohort study. *Aliment Pharmacol Ther*. 2018;47(2):279–288.
- Busby J, Mills K, Zhang SD, et al. Selective serotonin reuptake inhibitor use and breast cancer survival: a population-based cohort study. *Breast Cancer Res*. 2018;20(1):4.
- Butler AM, Layton JB, Krueger WS, et al. Assessing residual bias in estimating influenza vaccine effectiveness. *Med Care*. 2019;57(1):73–78.

21. Casula M, Olmastroni E, Galimberti F, et al. Association between the cumulative exposure to bisphosphonates and hospitalization for atherosclerotic cardiovascular events: a population-based study. *Atherosclerosis*. 2020;301:1–7.
22. Casula M, Scotti L, Galimberti F, et al. Use of proton pump inhibitors and risk of ischemic events in the general population. *Atherosclerosis*. 2018;277:123–129.
23. Cheung KS, Chan EW, Wong AYS, et al. Long-term proton pump inhibitors and risk of gastric cancer development after treatment for *Helicobacter pylori*: a population-based study. *Gut*. 2018;67(1):28–35.
24. Chien LN, Huang YJ, Shao YHJ, et al. Proton pump inhibitors and risk of periampullary cancers—a nested case-control study. *Int J Cancer*. 2016;138(6):1401–1409.
25. Chou YT, Farley JF, Stinchcombe TE, et al. The association between Medicare low-income subsidy and anticancer treatment uptake in advanced lung cancer. *J Natl Cancer Inst*. 2020;112(6):637–646.
26. Christiansen CF, Thomsen RW, Schmidt M, et al. Influenza vaccination and 1-year risk of myocardial infarction, stroke, heart failure, pneumonia, and mortality among intensive care unit survivors aged 65 years or older: a nationwide population-based cohort study. *Intensive Care Med*. 2019;45(7):957–967.
27. Cohen JM, Wood ME, Hernández-Díaz S, et al. Paternal antidepressant use as a negative control for maternal use: assessing familial confounding on gestational length and anxiety traits in offspring. *Int J Epidemiol*. 2019;48(5):1665–1672.
28. Cohen JM, Hernández-Díaz S, Bateman BT, et al. Placental complications associated with psychostimulant use in pregnancy. *Obstet Gynecol*. 2017;130(6):1192–1201.
29. Danaei G, García Rodríguez LA, Canero OF, et al. Statins and risk of diabetes: an analysis of electronic medical records to evaluate possible bias due to differential survival. *Diabetes Care*. 2013;36(5):1236–1240.
30. Dave CV, Schneeweiss S, Paterno E. Association of sodium-glucose cotransporter 2 inhibitor treatment with risk of hospitalization for Fournier gangrene among men. *JAMA Intern Med*. 2019;179(11):1587–1590.
31. Davies NM, Thomas KH, Taylor AE, et al. How to compare instrumental variable and conventional regression analyses using negative controls and bias plots. *Int J Epidemiol*. 2017;46(6):2067–2077.
32. Desai RJ, Sarpatwari A, Dejene S, et al. Comparative effectiveness of generic and brand-name medication use: a database study of US health insurance claims. *PLoS Med*. 2019;16(3):e1002763–e1002720.
33. de Groot MCH, Klungel OH, Leufkens HGM, et al. Sources of heterogeneity in case-control studies on associations between statins, ACE-inhibitors, and proton pump inhibitors and risk of pneumonia. *Eur J Epidemiol*. 2014;29(10):767–775.
34. Dillon P, Smith SM, Gallagher PJ, et al. Association between gaps in antihypertensive medication adherence and injurious falls in older community-dwelling adults: a prospective cohort study. *BMJ Open*. 2019;9(3):e022927.
35. Dormuth CR, Patrick AR, Shrank WH, et al. Statin adherence and risk of accidents. *Circulation*. 2009;119(15):2051–2057.
36. Douros A, Dell’Aniello S, Yu OHY, et al. Sulfonylureas as second line drugs in type 2 diabetes and the risk of cardiovascular and hypoglycaemic events: population based cohort study. *BMJ*. 2018;362:k2693.
37. Douros A, Filion KB, Yin H, et al. Glucagon-like peptide 1 receptor agonists and the risk of incident diabetic retinopathy. *Diabetes Care*. 2018;41(11):2330–2338.
38. Duke JD, Ryan PB, Suchard MA, et al. Risk of angioedema associated with levetiracetam compared with phenytoin: findings of the observational health data sciences and informatics research network. *Epilepsia*. 2017;58(8):e101–e106.
39. Etmnan M, Westerberg BD, Kozak FK, et al. Risk of sensorineural hearing loss with macrolide antibiotics: a nested case-control study. *Laryngoscope*. 2017;127(1):229–232.
40. Farhat N, Birkett N, Haddad N, et al. Risk of adverse cardiovascular events following a myocardial infarction in patients receiving combined clopidogrel and proton pump inhibitor treatment: a nested case-control study. *Drugs Real World Outcomes*. 2020;7(3):191–203.
41. Gagne JJ, Choudhry NK, Kesselheim AS, et al. Comparative effectiveness of generic and brand-name statins on patient outcomes: a cohort study. *Ann Intern Med*. 2014;161(6):400–407.
42. Gandhi S, Shariff SZ, Al-Jaishi A, et al. Second-generation antidepressants and hyponatremia risk: a population-based cohort study of older adults. *Am J Kidney Dis*. 2017;69(1):87–96.
43. Gerhard T, Devanand DP, Huang C, et al. Lithium treatment and risk for dementia in adults with bipolar disorder: population-based cohort study. *Br J Psychiatry*. 2015;207(1):46–51.
44. Gidaya NB, Lee BK, Burstyn I, et al. In utero exposure to selective serotonin reuptake inhibitors and risk for autism spectrum disorder. *J Autism Dev Disord*. 2014;44(10):2558–2567.
45. Gokhale M, Girman C, Chen Y, et al. Comparison of diagnostic evaluations for cough among initiators of angiotensin converting enzyme inhibitors and angiotensin receptor blockers. *Pharmacoepidemiol Drug Saf*. 2016;25(5):512–520.
46. Gottlieb A, Yanover C, Cahan A, et al. Estimating the effects of second-line therapy for type 2 diabetes mellitus: retrospective cohort study. *BMJ Open Diabetes Res Care*. 2017;5(1):e000435.
47. Greene SK, Li L, Shay DK, et al. Risk of adverse events following oseltamivir treatment in influenza outpatients, Vaccine Safety Datalink Project, 2007–2010. *Pharmacoepidemiol Drug Saf*. 2013;22(4):335–344.
48. Gruber JF, Becker-Dreps S, Hudgens MG, et al. Timing of rotavirus vaccine doses and severe rotavirus gastroenteritis among vaccinated infants in low- and middle-income countries. *Epidemiology*. 2018;29(6):867–875.
49. Hamad AF, Alessi-Severini S, Mahmud S, et al. Prenatal antibiotic exposure and risk of attention-deficit/hyperactivity disorder: a population-based cohort study. *CMAJ*. 2020;192(20):E527–E535.
50. Han X, Chiang CW, Leonard CE, et al. Biomedical informatics approaches to identifying drug-drug interactions: application to insulin secretagogues. *Epidemiology*. 2017;28(3):459–468.
51. Harpaz R, Leung J. When zoster hits close to home: impact of personal zoster awareness on zoster vaccine uptake in the U.S. *Vaccine*. 2017;35(27):3457–3460.
52. Harrison PJ, Luciano S, Colbourne L. Rates of delirium associated with calcium channel blockers compared to diuretics, renin-angiotensin system agents and beta-blockers: an electronic health records network study. *J Psychopharmacol*. 2020;34(8):848–855.
53. Hripesak G, Suchard MA, Shea S, et al. Comparison of cardiovascular and safety outcomes of chlorthalidone vs hydrochlorothiazide to treat hypertension. *JAMA Intern Med*. 2020;180(4):542–551.

54. Huang W, Sundquist J, Sundquist K, et al. Use of phosphodiesterase 5 inhibitors is associated with lower risk of colorectal cancer in men with benign colorectal neoplasms. *Gastroenterology*. 2019;157(3):672–681.e4.
55. Imfeld P, Bodmer M, Jick SS, et al. Proton pump inhibitor use and risk of developing Alzheimer's disease or vascular dementia: a case-control analysis. *Drug Saf*. 2018;41(12):1387–1396.
56. Ing C, Ma X, Sun M, et al. Exposure to surgery and anesthesia in early childhood and subsequent use of attention deficit hyperactivity disorder medications. *Anesth Analg*. 2020;131(3):723–733.
57. Ivers LC, Hilaire JJ, Teng JE, et al. Effectiveness of reactive oral cholera vaccination in rural Haiti: a case-control study and bias-indicator analysis. *Lancet Glob Health*. 2015;3(3):e162–e168.
58. Izurieta HS, Chillarige Y, Kelman JA, et al. Statin use and risks of influenza-related outcomes among older adults receiving standard-dose or high-dose influenza vaccines through medicare during 2010–2015. *Clin Infect Dis*. 2018;67(3):378–387.
59. Izurieta HS, Werneck M, Kelman J, et al. Effectiveness and duration of protection provided by the live-attenuated herpes zoster vaccine in the Medicare population ages 65 years and older. *Clin Infect Dis*. 2017;64(6):785–793.
60. Izurieta HS, Wu X, Lu Y, et al. Zostavax vaccine effectiveness among US elderly using real-world evidence: addressing unmeasured confounders by using multiple imputation after linking beneficiary surveys with Medicare claims. *Pharmacoepidemiol Drug Saf*. 2019;28(7):993–1001.
61. Jackson LA, Jackson ML, Nelson JC, et al. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. *Int J Epidemiol*. 2006;35(2):337–344.
62. Jensen A, Andersen PK, Stensballe LG. Early childhood vaccination and subsequent mortality or morbidity: are observational studies hampered by residual confounding? A Danish register-based cohort study. *BMJ Open*. 2019;9(9):e029794.
63. Johanson CE, Arfken CL, di Menza S, et al. Diversion and abuse of buprenorphine: findings from national surveys of treatment patients and physicians. *Drug Alcohol Depend*. 2012;120(1–3):190–195.
64. Kaiser P, Arnold AM, Benkeser D, et al. Comparing methods to address bias in observational data: statin use and cardiovascular events in a US cohort. *Int J Epidemiol*. 2018;47(1):246–254.
65. Kim Y, Tian Y, Yang J, et al. Comparative safety and effectiveness of alendronate versus raloxifene in women with osteoporosis. *Sci Rep*. 2020;10(1):11115.
66. Kioumourtzoglou MA, Coull BA, O'Reilly EJ, et al. Association of exposure to diethylstilbestrol during pregnancy with multigenerational neurodevelopmental deficits. *JAMA Pediatr*. 2018;172(7):670–677.
67. Kipp R, Askari M, Fan J, et al. Real-world comparison of classes IC and III antiarrhythmic drugs as an initial rhythm control strategy in newly diagnosed atrial fibrillation: from the TREAT-AF study. *JACC Clin Electrophysiol*. 2019;5(2):231–241.
68. Kjerpeseth LJ, Selmer R, Ariansen I, et al. Comparative effectiveness of warfarin, dabigatran, rivaroxaban and apixaban in nonvalvular atrial fibrillation: a nationwide pharmacoepidemiological study. *PLoS One*. 2019;14(8):e0221500.
69. Kürüm E, Warren JL, Schuck-Paim C, et al. Bayesian model averaging with change points to assess the impact of vaccination and public health interventions. *Epidemiology*. 2017;28(6):889–897.
70. Lane J, Weaver J, Kostka K, et al. Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study. *Lancet Rheumatol*. 2020;2(11):e698–e711.
71. Lavikainen P, Helin-Salmivaara A, Eerola M, et al. Statin adherence and risk of acute cardiovascular events among women: a cohort study accounting for time-dependent confounding affected by previous adherence. *BMJ Open*. 2016;6(6):e011306.
72. Lazarus B, Chen Y, Wilson FP, et al. Proton pump inhibitor use and the risk of chronic kidney disease. *JAMA Intern Med*. 2016;176(2):238–246.
73. Leonard CE, Zhou M, Brensinger CM, et al. Clopidogrel drug interactions and serious bleeding: generating real-world evidence via automated high-throughput pharmacoepidemiologic screening. *Clin Pharmacol Ther*. 2019;106(5):1067–1075.
74. Leonard CE, Brensinger CM, Bilker WB, et al. Thromboembolic and neurologic sequelae of discontinuation of an antihyperlipidemic drug during ongoing warfarin therapy. *Sci Rep*. 2017;7(1):1–10.
75. Leung J, Harpaz R, Molinari NA, et al. Herpes zoster incidence among insured persons in the United States, 1993–2006: evaluation of impact of varicella vaccination. *Clin Infect Dis*. 2011;52(3):332–340.
76. Levine SZ, Kodesh A, Viktorin A, et al. Association of maternal use of folic acid and multivitamin supplements in the periods before and during pregnancy with the risk of autism spectrum disorder in offspring. *JAMA Psychiatry*. 2018;75(2):176–184.
77. Liew Z, Kioumourtzoglou MA, Roberts AL, et al. Use of negative control exposure analysis to evaluate confounding: an example of acetaminophen exposure and attention-deficit/hyperactivity disorder in Nurses' Health Study II. *Am J Epidemiol*. 2019;188(4):768–775.
78. Lin HW, Ho YF, Lin FJ. Statin use associated with lower risk of epilepsy after intracranial haemorrhage: a population-based cohort study. *Br J Clin Pharmacol*. 2018;84(9):1970–1979.
79. Lip GYH, Skjøth F, Nielsen PB, et al. Effectiveness and safety of standard-dose nonvitamin k antagonist oral anticoagulants and warfarin among patients with atrial fibrillation with a single stroke risk factor: a nationwide cohort study. *JAMA Cardiol*. 2017;2(8):872–881.
80. Liu CH, Yeh YC, Huang WT, et al. Assessment of pre-specified adverse events following varicella vaccine: a population-based self-controlled risk interval study. *Vaccine*. 2020;38(11):2495–2502.
81. Madigan D, Ryan PB, Schuemie M, et al. Evaluating the impact of database heterogeneity on observational study results. *Am J Epidemiol*. 2013;178(4):645–651.
82. Man KKC, Lau WCY, Coghill D, et al. Association between methylphenidate treatment and risk of seizure: a population-based, self-controlled case-series study. *Lancet Child Adolesc Health*. 2020;4(6):435–443.
83. Markovic M, Swanson SA, Stricker BH, et al. Prenatal exposure to non-steroidal anti-inflammatory drugs (NSAIDs) and neurodevelopmental outcomes in children. *Pharmacoepidemiol Drug Saf*. 2019;28(4):452–459.
84. Matthews A, Langan SM, Douglas IJ, et al. Phosphodiesterase type 5 inhibitors and risk of malignant melanoma: matched cohort study using primary care data



- from the UK Clinical Practice Research Datalink. *PLoS Med.* 2016;13(6):e1002037–e1002015.
85. McGrath LJ, Spangler L, Curtis JR, et al. Using negative control outcomes to assess the comparability of treatment groups among women with osteoporosis in the United States. *Pharmacoepidemiol Drug Saf.* 2020;29(8):854–863.
  86. McGrath LJ, Ellis AR, Brookhart MA. Controlling time-dependent confounding by health status and frailty: restriction versus statistical adjustment. *Am J Epidemiol.* 2015;182(1):17–25.
  87. Moon AM, Jiang Y, Rogal SS, et al. Opioid prescriptions are associated with hepatic encephalopathy in a national cohort of patients with compensated cirrhosis. *Aliment Pharmacol Ther.* 2020;51(6):652–660.
  88. Morales D, Pacurariu A, Slattery J, et al. Association between peripheral neuropathy and exposure to Oral fluoroquinolone or amoxicillin-clavulanate therapy. *JAMA Neurol.* 2019;76(7):827–833.
  89. Morales DR, Conover MM, You SC, et al. Renin-angiotensin system blockers and susceptibility to COVID-19: an international, open science, cohort analysis. *Lancet Digit Health.* 2021;3(2):e98–e114.
  90. Morales DR, Pacurariu A, Slattery J, et al. Association between hydrochlorothiazide exposure and different incident skin, lip and oral cavity cancers: a series of population-based nested case-control studies. *Br J Clin Pharmacol.* 2020;86(7):1336–1345.
  91. Morales DR, Slattery J, Pacurariu A, et al. Relative and absolute risk of tendon rupture with fluoroquinolone and concomitant fluoroquinolone/corticosteroid therapy: population-based nested case-control study. *Clin Drug Investig.* 2019;39(2):205–213.
  92. Moran LV, Ongur D, Hsu J, et al. Psychosis with methylphenidate or amphetamine in patients with ADHD. *N Engl J Med.* 2019;380(12):1128–1138.
  93. Moulis G, Sommet A, Lapeyre-Mestre M, et al. Is the risk of tumour necrosis factor inhibitor-induced lupus or lupus-like syndrome the same with monoclonal antibodies and soluble receptor? A case/non-case study in a nationwide pharmacovigilance database. *Rheumatology.* 2014;53(10):1864–1871.
  94. Muanda FT, Weir MA, Bathini L, et al. Association of Baclofen with encephalopathy in patients with chronic kidney disease. *JAMA.* 2019;322(20):1987–1995.
  95. O'Grady M, Clarke L, Turner G, et al. Statin use and risk of acute diverticulitis: a population-based case-control study. *Medicine (Baltimore).* 2020;99(20):e20264.
  96. Ospina-Romero M, Brenowitz WD, Glymour MM, et al. The association between cancer and spousal rate of memory decline: a negative control study to evaluate (unmeasured) social confounding of the cancer-memory relationship. *Alzheimer Dis Assoc Disord.* 2020;35(3):271–274.
  97. Ou SM, Chen HT, Kuo SC, et al. Dipeptidyl peptidase-4 inhibitors and cardiovascular risks in patients with pre-existing heart failure. *Heart.* 2017;103(6):414–420.
  98. Ou SM, Shih CJ, Chao PW, et al. Effects on clinical outcomes of adding dipeptidyl peptidase-4 inhibitors versus sulfonylureas to metformin therapy in patients with type 2 diabetes mellitus. *Ann Intern Med.* 2015;163(9):663–672.
  99. Pan SW, Yen YF, Feng JY, et al. Opposite effects of statins on the risk of tuberculosis and herpes zoster in patients with diabetes: a population-based cohort study. *Br J Clin Pharmacol.* 2020;86(3):569–579.
  100. Paul S, Zahurak M, Luznik L, et al. Non-myeloablative allogeneic transplantation with post-transplant cyclophosphamide after immune checkpoint inhibition for classic Hodgkin lymphoma: a retrospective cohort study: checkpoint inhibition before NMA alloBMT with PTCy in cHL. *Biol Blood Marrow Transpl.* 2020;26(9):1679–1688.
  101. Pottegård A, Broe A, Stage TB, et al. Use of dicloxacillin and risk of pregnancy among users of oral contraceptives. *Basic Clin Pharmacol Toxicol.* 2018;123(3):288–293.
  102. Quinn PD, Chang Z, Hur K, et al. ADHD medication and substance-related problems. *Am J Psychiatry.* 2017;174(9):877–885.
  103. Rai D, Lee BK, Dalman C, et al. Antidepressants during pregnancy and autism in offspring: population based cohort study. *BMJ.* 2017;358:j2811.
  104. Ray GT, Lewis N, Klein NP, et al. Intra-season waning of influenza vaccine effectiveness. *Clin Infect Dis.* 2019;68(10):1623–1630.
  105. Raymakers A, Sin DD, Sadatsafavi M, et al. Statin use and lung cancer risk in chronic obstructive pulmonary disease patients: a population-based cohort study. *Respir Res.* 2020;21(1):118.
  106. Ridenhour BJ, Campitelli MA, Kwong JC, et al. Effectiveness of inactivated influenza vaccines in preventing influenza-associated deaths and hospitalizations among Ontario residents aged  $\geq 65$  years: estimates with generalized linear models accounting for healthy vaccinee effects. *PLoS One.* 2013;8(10):e76318.
  107. Rodgers LR, Dennis JM, Shields BM, et al. Prior event rate ratio adjustment produced estimates consistent with randomized trial: a diabetes case study. *J Clin Epidemiol.* 2020;122:78–86.
  108. Rodríguez AJ, Ernst MT, Nybo M, et al. Oral bisphosphonate use reduces cardiovascular events in a cohort of Danish patients referred for bone mineral density. *J Clin Endocrinol Metab.* 2020;105(10):dgaa481.
  109. Roshanov PS, Rochweg B, Patel A, et al. Withholding versus continuing angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers before noncardiac surgery. *Anesthesiology.* 2017;126(1):16–27.
  110. Sarvet AL, Wall MM, Keyes KM, et al. Self-medication of mood and anxiety disorders with marijuana: higher in states with medical marijuana laws. *Drug Alcohol Depend.* 2018;186:10–15.
  111. Schuemie MJ, Ryan PB, Hripesak G, et al. Improving reproducibility by using high-throughput observational studies with empirical calibration. *Philos Trans R Soc Math Phys Eng Sci.* 2018;376(2128):20170356.
  112. Schuemie MJ, Ryan PB, Man KKC, et al. A plea to stop using the case-control design in retrospective database studies. *Stat Med.* 2019;38(22):4199–4208.
  113. Shao SC, Chang KC, Hung MJ, et al. Comparative risk evaluation for cardiovascular events associated with dapagliflozin vs. empagliflozin in real-world type 2 diabetes patients: a multi-institutional cohort study. *Cardiovasc Diabetol.* 2019;18(1):120.
  114. Simonov M, Abel EA, Skanderson M, et al. Use of proton pump inhibitors increases risk of incident kidney stones. *Clin Gastroenterol Hepatol.* 2021;19(1):72–79.e21.
  115. Simonsen L, Taylor RJ, Schuck-Paim C, et al. Effect of 13-valent pneumococcal conjugate vaccine on admissions to hospital 2 years after its introduction in the USA: a time series analysis. *Lancet Respir Med.* 2014;2(5):387–394.
  116. Sinnott SJ, Smeeth L, Williamson E, et al. The comparative effectiveness of fourth-line drugs in resistant hypertension:

- an application in electronic health record data. *Pharmacoepidemiol Drug Saf.* 2019;28(9):1267–1277.
117. Sköldbberg F, Svensson T, Olén O, et al. A population-based case-control study on statin exposure and risk of acute diverticular disease. *Scand J Gastroenterol.* 2016;51(2):203–210.
  118. Shoag JE, Patel N, Posada L, et al. Kidney stones and risk of narcotic use. *J Urol.* 2019;202(1):114–118.
  119. Sørup S, Benn CS, Poulsen A, et al. Simultaneous vaccination with MMR and DTaP-IPV-Hib and rate of hospital admissions with any infections: a nationwide register based cohort study. *Vaccine.* 2016;34(50):6172–6180.
  120. Spoenclin J, Gagne JJ, Lewey JJ, et al. Comparative effectiveness and safety of antiplatelet drugs in patients with diabetes mellitus and acute coronary syndrome. *Pharmacoepidemiol Drug Saf.* 2018;27(12):1361–1370.
  121. Stolfo D, Uijl A, Benson L, et al. Association between beta-blocker use and mortality/morbidity in older patients with heart failure with reduced ejection fraction. A propensity score-matched analysis from the Swedish Heart Failure Registry. *Eur J Heart Fail.* 2020;22(1):103–112.
  122. Sundbakk LM, Wood M, Gran JM, et al. Impact of prenatal exposure to benzodiazepines and z-hypnotics on behavioral problems at 5 years of age: a study from the Norwegian Mother and Child Cohort Study. *PLoS One.* 2019;14(6):e0217830.
  123. Symes RJ, Etminan M, Mikelberg FS. Risk of angle-closure glaucoma with bupropion and topiramate. *JAMA Ophthalmol.* 2015;133(10):1187–1189.
  124. Tate JE, Curns AT, Cortese MM, et al. Burden of acute gastroenteritis hospitalizations and emergency department visits in US children that is potentially preventable by rotavirus vaccination: a probe study using the now-withdrawn rotashield vaccine. *Pediatrics.* 2009;123(3):744–749.
  125. Thomsen RW, Öztürk B, Pedersen L, et al. Hospital records of pain, fatigue, or circulatory symptoms in girls exposed to human papillomavirus vaccination: cohort, self-controlled case series, and population time trend studies. *Am J Epidemiol.* 2020;189(4):277–285.
  126. Thorrington D, Andrews N, Stowe J, et al. Elucidating the impact of the pneumococcal conjugate vaccine programme on pneumonia, sepsis and otitis media hospital admissions in England using a composite control. *BMC Med.* 2018;16(1):13.
  127. Tielemans SMAJ, De Melker HE, Hahné SJM, et al. Non-specific effects of measles, mumps, and rubella (MMR) vaccination in high income setting: population based cohort study in the Netherlands. *BMJ.* 2017;358:j3862.
  128. Tien KL, Sheng WH, Shieh SC, et al. Chlorhexidine bathing to prevent central line-associated bloodstream infections in hematology units: a prospective, controlled cohort study. *Clin Infect Dis.* 2020;71(3):556–563.
  129. Totterdell J, Phillips A, Glover C, et al. Safety of live attenuated herpes zoster vaccine in adults 70–79 years: a self-controlled case series analysis using primary care data from Australia's MedicinesInsight program. *Vaccine.* 2020;38(23):3968–3979.
  130. Toulis KA, Willis BH, Marshall T, et al. All-cause mortality in patients with diabetes under treatment with dapagliflozin: a population-based, open-cohort study in the health improvement network database. *J Clin Endocrinol Metab.* 2017;102(5):1719–1725.
  131. Trønnes JN, Wood M, Lupattelli A, et al. Prenatal paracetamol exposure and neurodevelopmental outcomes in preschool-aged children. *Paediatr Perinat Epidemiol.* 2020;34(3):247–256.
  132. Tseng HF, Smith N, Harpaz R, et al. Herpes zoster vaccine in older adults and the risk of subsequent herpes zoster disease. *JAMA.* 2011;305(2):160–166.
  133. Tsujimoto T, Kajio H. Use of nitrates and risk of cardiovascular events in patients with heart failure with preserved ejection fraction. *Mayo Clin Proc.* 2019;94(7):1210–1220.
  134. van Rein N, Cannegieter SC, Rosendaal FR, et al. Suspected survivor bias in case-control studies: stratify on survival time and use a negative control. *J Clin Epidemiol.* 2014;67(2):232–235.
  135. Vonesh E, Gooch KL, Khangulov V, et al. Cardiovascular risk profile in individuals initiating treatment for overactive bladder—challenges and learnings for comparative analysis using linked claims and electronic medical record databases. *PLoS One.* 2018;13(10):e0205640.
  136. Vouri SM, Jiang X, Manini TM, et al. Magnitude of and characteristics associated with the treatment of calcium channel blocker-induced lower-extremity edema with loop diuretics. *JAMA Netw Open.* 2019;2(12):e1918425.
  137. Walsh LK, Donelle J, Dodds L, et al. Health outcomes of young children born to mothers who received 2009 pandemic H1N1 influenza vaccination during pregnancy: retrospective cohort study. *BMJ.* 2019;366:14151.
  138. Wei Y, Lin FJ, Lin SY, et al. Risk of hypoglycemia and concomitant use of Repaglinide and clopidogrel: a population-based nested case-control study. *Clin Pharmacol Ther.* 2019;106(6):1346–1352.
  139. Weinstein RB, Ryan PB, Berlin JA, et al. Channeling bias in the analysis of risk of myocardial infarction, stroke, gastrointestinal bleeding, and acute renal failure with the use of paracetamol compared with ibuprofen. *Drug Saf.* 2020;43(9):927–942.
  140. Weinstein RB, Schuemie MJ, Ryan PB, et al. Seasonality in acute liver injury? Findings in two health care claims databases. *Drug Healthc Patient Saf.* 2016;8:39–48.
  141. Weinstein R, Ryan P, Berlin J, et al. Channeling in the use of nonprescription paracetamol and ibuprofen in an electronic medical records database: evidence and implications. *Drug Saf.* 2017;40(12):1279–1292.
  142. Welk B, McClure JA, Clarke C, et al. An opioid prescription for men undergoing minor urologic surgery is associated with an increased risk of new persistent opioid use. *Eur Urol.* 2020;77(1):68–75.
  143. Xian Y, Xu H, O'Brien EC, et al. Clinical effectiveness of direct Oral anticoagulants vs warfarin in older patients with atrial fibrillation and ischemic stroke: findings from the Patient-Centered Research into Outcomes Stroke Patients Prefer and Effectiveness Research (PROSPER) study. *JAMA Neurol.* 2019;76(10):1192–1202.
  144. Whitlock RH, Hougen I, Komenda P, et al. A safety comparison of metformin vs sulfonyleurea initiation in patients with type 2 diabetes and chronic kidney disease: a retrospective cohort study. *Mayo Clin Proc.* 2020;95(1):90–100.
  145. Xie Y, Bowe B, Yan Y, et al. Estimates of all cause mortality and cause specific mortality associated with proton pump inhibitors among US veterans: cohort study. *BMJ.* 2019;365:11580.
  146. Yates TA, Tomlinson LA, Bhaskaran K, et al. Lansoprazole use and tuberculosis incidence in the United Kingdom Clinical Practice Research Datalink:

- a population based cohort. *PLoS Med.* 2017;14(11): e1002457.
147. Yip TCF, Wong VWS, Chan HLY, et al. Tenofovir is associated with lower risk of hepatocellular carcinoma than entecavir in patients with chronic HBV infection in China. *Gastroenterology.* 2020;158(1):215–225.e6.
  148. You SC, Jung S, Swerdel JN, et al. Comparison of first-line dual combination treatments in hypertension: real-world evidence from multinational heterogeneous cohorts. *Korean Circ J.* 2020;50(1):52–68.
  149. Ystrom E, Gustavson K, Brandlistuen RE, et al. Prenatal exposure to acetaminophen and risk of ADHD. *Pediatrics.* 2017;140(5):e20163840.
  150. Yuan J, Zhang C, Sparks JA, et al. Regular use of proton pump inhibitor and risk of rheumatoid arthritis in women: a prospective cohort study. *Aliment Pharmacol Ther.* 2020; 52(3):449–458.
  151. Yuan Z, DeFalco FJ, Ryan PB, et al. Risk of lower extremity amputations in people with type 2 diabetes mellitus treated with sodium-glucose co-transporter-2 inhibitors in the USA: a retrospective cohort study. *Diabetes Obes Metab.* 2018; 20(3):582–589.
  152. Zhang HT, McGrath LJ, Wyss R, et al. Controlling confounding by frailty when estimating influenza vaccine effectiveness using predictors of dependency in activities of daily living. *Pharmacoepidemiol Drug Saf.* 2017;26(12): 1500–1506.
  153. Zhang HT, McGrath LJ, Ellis AR, et al. Restriction of Pharmacoepidemiologic cohorts to initiators of medications in unrelated preventive drug classes to reduce confounding by frailty in older adults. *Am J Epidemiol.* 2019;188(7): 1371–1382.
  154. Zhou M, Leonard CE, Brensinger CM, et al. Pharmacoepidemiologic screening of potential Oral anticoagulant drug interactions leading to thromboembolic events. *Clin Pharmacol Ther.* 2020;108(2):377–386.
  155. Zhou Z, Ofori-Asenso R, Curtis AJ, et al. Association of Statin use with disability-free survival and cardiovascular disease among healthy older adults. *J Am Coll Cardiol.* 2020;76(1):17–27.
  156. Zullo AR, Zhang T, Lee Y, et al. Effect of bisphosphonates on fracture outcomes among frail older adults. *J Am Geriatr Soc.* 2019;67(4):768–776.
  157. Sofer T, Richardson DB, Colicino E, et al. On negative outcome control of unobserved confounding as a generalization of difference-in-differences. *Stat Sci.* 2016; 31(3):348–361.
  158. Richardson DB, Keil AP, Tchetgen Tchetgen E, et al. Negative control outcomes and the analysis of standardized mortality ratios. *Epidemiology.* 2015;26(5): 727–732.
  159. Richardson DB, Laurier D, Schubauer-Berigan MK, et al. Assessment and indirect adjustment for confounding by smoking in cohort studies using relative hazards models. *Am J Epidemiol.* 2014;180(9):933–940.
  160. Tchetgen Tchetgen EJ, Sofer T, Richardson D. Negative outcome control for unobserved confounding under a Cox proportional hazards model. *Harv Univ Biostat Work Pap Ser.* Paper 192. 2015. <https://biostats.bepress.com/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1200&context=harvardbiostat>. Accessed October 6, 2023.
  161. Trinh NTH, Solé E, Benkebil M. Benefits of combining change-point analysis with disproportionality analysis in pharmacovigilance signal detection. *Pharmacoepidemiol Drug Saf.* 2019;28(3):370–376.
  162. DuMouchel W, Ryan PB, Schuemie MJ, et al. Evaluation of disproportionality safety signaling applied to healthcare databases. *Drug Saf.* 2013;36(suppl 1):S123–S132.
  163. Backenroth D, Chase H, Friedman C, et al. Using rich data on comorbidities in case-control study design with electronic health record data improves control of confounding in the detection of adverse drug reactions. *PLoS One.* 2016;11(10):e0164304.
  164. Hauben M, Hung E, Wood J, et al. The impact of database restriction on pharmacovigilance signal detection of selected cancer therapies. *Ther Adv Drug Saf.* 2017;8(5): 145–156.
  165. Brown JS, Kulldorff M, Chan KA, et al. Early detection of adverse drug events within population-based health networks: application of sequential testing methods. *Pharmacoepidemiol Drug Saf.* 2007;16(12):1275–1284.
  166. Brown JS, Kulldorff M, Petronis KR, et al. Early adverse drug event signal detection within population-based health networks using sequential methods: key methodologic considerations. *Pharmacoepidemiol Drug Saf.* 2009;18(3): 226–234.
  167. Gagne JJ, Wang SV, Rassen JA, et al. A modular, prospective, semi-automated drug safety monitoring system for use in a distributed data environment. *Pharmacoepidemiol Drug Saf.* 2014;23(6):619–627.
  168. Madigan D, Schuemie MJ, Ryan PB. Empirical performance of the case-control method: lessons for developing a risk identification and analysis system. *Drug Saf.* 2013;36(suppl 1):S73–S82.
  169. Marbac M, Tubert-Bitter P, Sedki M. Bayesian model selection in logistic regression for the detection of adverse drug reactions. *Biom J.* 2016;58(6):1376–1389.
  170. Nishtala PS, Chyou T-Y. Exploring New Zealand prescription data using sequence symmetry analyses for predicting adverse drug reactions. *J Clin Pharm Ther.* 2017; 42(2):189–194.
  171. Norén GN, Bergvall T, Ryan PB, et al. Empirical performance of the calibrated self-controlled cohort analysis within temporal pattern discovery: lessons for developing a risk identification and analysis system. *Drug Saf.* 2013; 36(suppl 1):S107–S121.
  172. Osokogu OU, Dodd C, Pacurariu A, et al. Drug safety monitoring in children: performance of signal detection algorithms and impact of age stratification. *Drug Saf.* 2016; 39(9):873–881.
  173. Pierce C, Bouri K, Pamer C, et al. Evaluation of Facebook and Twitter monitoring to detect safety signals for medical products: an analysis of recent FDA safety alerts. *Drug Saf.* 2017;40(4):317–331.
  174. Pratt N, Chan EW, Choi NK, et al. Prescription sequence symmetry analysis: assessing risk, temporality, and consistency for adverse drug reactions across datasets in five countries. *Pharmacoepidemiol Drug Saf.* 2015;24(8): 858–864.
  175. Reich CG, Ryan PB, Schuemie MJ. Alternative outcome definitions and their effect on the performance of methods for observational outcome studies. *Drug Saf.* 2013; 36(suppl 1):S181–S193.
  176. Reich CG, Ryan PB, Suchard MA. The impact of drug and outcome prevalence on the feasibility and performance of analytical methods for a risk identification and analysis system. *Drug Saf.* 2013;36(suppl 1):S195–S204.
  177. Ryan PB, Madigan D, Stang PE, et al. Empirical assessment of methods for risk identification in healthcare data: results from the experiments of the Observational

- Medical Outcomes Partnership. *Stat Med.* 2012;31(30):4401–4415.
178. Ryan PB, Schuemie MJ, Gruber S, et al. Empirical performance of a new user cohort method: lessons for developing a risk identification and analysis system. *Drug Saf.* 2013;36(suppl 1):S59–S72.
  179. Ryan PB, Schuemie MJ, Madigan D. Empirical performance of a self-controlled cohort method: lessons for developing a risk identification and analysis system. *Drug Saf.* 2013;36(suppl 1):S95–S106.
  180. Ryan PB, Stang PE, Overhage JM, et al. A comparison of the empirical performance of methods for a risk identification system. *Drug Saf.* 2013;36(suppl 1):S143–S158.
  181. Ryan PB, Schuemie MJ. Evaluating performance of risk identification methods through a large-scale simulation of observational data. *Drug Saf.* 2013;36(suppl 1):S171–S180.
  182. Schuemie MJ, Coloma PM, Straatman H, et al. Using electronic health care records for drug safety signal detection: a comparative evaluation of statistical methods. *Med Care.* 2012;50(10):890–897.
  183. Schuemie MJ, Gini R, Coloma PM, et al. Replication of the OMOP experiment in Europe: evaluating methods for risk identification in electronic health record databases. *Drug Saf.* 2013;36(suppl 1):S159–S169.
  184. Schuemie MJ, Madigan D, Ryan PB. Empirical performance of LGPS and LEOPARD: lessons for developing a risk identification and analysis system. *Drug Saf.* 2013;36(suppl 1):S133–S142.
  185. Suchard MA, Zorych I, Simpson SE, et al. Empirical performance of the self-controlled case series design: lessons for developing a risk identification and analysis system. *Drug Saf.* 2013;36(suppl 1):S83–S93.
  186. Thurin NH, Lassalle R, Schuemie M, et al. Empirical assessment of case-based methods for identification of drugs associated with upper gastrointestinal bleeding in the French National Healthcare System database (SNDS). *Pharmacoepidemiol Drug Saf.* 2020;29(8):890–903.
  187. Coloma PM, Avillach P, Salvo F, et al. A reference standard for evaluation of methods for drug safety signal detection using electronic healthcare record databases. *Drug Saf.* 2013;36(1):13–23.
  188. Ryan PB, Schuemie MJ, Welebob E, et al. Defining a reference set to support methodological research in drug safety. *Drug Saf.* 2013;36(suppl 1):S33–S47.
  189. Stang PE, Ryan PB, Racoosin JA, et al. Advancing the science for active surveillance: rationale and Design for the Observational Medical Outcomes Partnership. *Ann Intern Med.* 2010;153(9):600–606.
  190. Tisdale JE, Miller DA. *Drug-Induced Diseases: Prevention, Detection, and Management.* 2nd ed. Bethesda, MD: American Society of Health-System Pharmacists; 2010.
  191. Voss EA, Boyce RD, Ryan PB, et al. Accuracy of an automated knowledge base for identifying drug adverse reactions. *J Biomed Inform.* 2017;66:72–81.
  192. Osokogu OU, Fregonese F, Ferrajolo C, et al. Pediatric drug safety signal detection: a new drug–event reference set for performance testing of data-mining methods and systems. *Drug Saf.* 2015;38(2):207–217.
  193. Brauchli Pernus Y, Nan C, Verstraeten T, et al. Reference set for performance testing of pediatric vaccine safety signal detection methods and systems. *Vaccine.* 2016;34(51):6626–6633.
  194. Gruber S, Tchetgen Tchetgen E. Limitations of empirical calibration of *P* values using observational data. *Stat Med.* 2016;35(22):3869–3882.
  195. Dickerman BA, Gerlovin H, Madenci AL, et al. Comparative effectiveness of BNT162b2 and mRNA-1273 vaccines in U.S. veterans. *N Engl J Med.* 2022;386(2):105–115.
  196. Dagan N, Barda N, Kepten E, et al. BNT162b2 mRNA Covid-19 vaccine in a Nationwide mass vaccination setting. *N Engl J Med.* 2021;384(15):1412–1423.
  197. Shi X, Li K, Miao W, et al. Theory for identification and inference with synthetic controls: a proximal causal inference framework [preprint]. *arXiv.* 2023. <https://doi.org/10.48550/arXiv.2108.13935>. Accessed March 23, 2023.
  198. Ying A. Proximal identification and estimation to handle dependent right censoring for survival analysis [preprint]. *arXiv.* 2022. <https://doi.org/10.48550/arXiv.2208.07014>. Accessed March 23, 2023.
  199. Ghassami A, Yang A, Shpitser I, et al. Causal inference with hidden mediators [preprint]. *arXiv.* 2023. <https://doi.org/10.48550/arXiv.2111.02927>. Accessed March 23, 2023.
  200. Miao W, Tchetgen TE. Invited commentary: bias attenuation and identification of causal effects with multiple negative controls. *Am J Epidemiol.* 2017;185(10):950–953.
  201. Miao W, Geng Z, Tchetgen Tchetgen EJ. Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika.* 2018;105(4):987–993.
  202. Kummerfeld E, Lim J, Shi X. Data-driven Automated Negative Control Estimation (DANCE): search for, validation of, and causal inference with negative controls [preprint]. *arXiv.* 2022. <https://doi.org/10.48550/arXiv.2210.00528>. Accessed March 13, 2023.
  203. Penning de Vries BBL, van Smeden M, Rosendaal FR, et al. Title, abstract, and keyword searching resulted in poor recovery of articles in systematic reviews of epidemiologic practice. *J Clin Epidemiol.* 2020;121:55–61.
  204. Bieler GS, Williams RL. Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity. *Biometrics.* 1993;49(3):793–801.