Journal of Clinical Epidemiology 148 (2022) 74−80

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Comparative effectiveness research considered methodological insights from simulation studies in physician's prescribing preference

Lisong Zhang*, Jim Lewsey, David A. McAllister

Institute of Health and Wellbeing, University of Glasgow, 1 Lilybank Gardens, Glasgow G12 8RZ, UK

Accepted 13 April 2022; Published online 18 April 2022

Abstract

Objectives: To review comparative effectiveness research (CER) using physician's prescribing preference as an instrumental variable (PPP IV) in pharmacoepidemiology and to review methodological studies that use simulation to evaluate the performance of PPP IV in CER.

Study Design and Setting: We conducted a review of CER using PPP IV and studies evaluating the use of PPP IV by using simulation methods. We searched Ovid, PubMed, and Google Scholar databases from 2005 to 2020.

Results: We identified six simulation studies and 18 CER studies. The simulation studies explored the most suitable ways for using PPP IV in different settings (outcome types, sample size, and the prevalence of outcomes) which can be useful guidance for using PPP IV in CER. The CER studies identified show heterogeneity in terms of validation assumptions, estimation methods, and sample size. Not all applied studies used the methodological insights from the simulation studies. However, they all concluded that PPP is a valid IV.

Conclusion: Future CER should consider a range of methodological issues to improve the validity of findings when using PPP IV. Specifically, studies should consider the impact of a different choice of statistical methods, forms of proxy for measuring preference, time-varying exposures, and the type of outcome.

Keywords: Review; Pharmacoepidemiology; Instrumental variable; Physician's prescribing preference; Comparative effectiveness research; Simulation studies

1. Introduction

In comparative effectiveness research (CER) using observational study designs, residual confounding is the most important challenge. There are well-established methods that focus on reducing covariate imbalance, such as multivariable-adjusted regression and propensity score methods [1]. However, these methods assume no unmeasured confounding after such adjustment. Originally from econometrics, the instrumental variable (IV) method can be used to address unmeasured confounding [2].

Physician's prescribing preference (PPP) is a commonly used instrument [3]. It exploits naturally occurring variation which makes the treatment assignment in observational studies using heath datasets closer to that in randomized controlled trials. The PPP is a latent variable and relies on proxy/surrogate measurement. The most common surrogate for PPP is the most recent prescription made by the same physician for patients with the same symptoms, also referred to as 'prior one' prescription [3]. Alternatively, PPP can be defined as the proportion of one particular drug under study prescribed among all the previous patients of the physician [4]. Furthermore, PPP can be constructed at higher levels of aggregation, such as general practice (a local grouping of doctors which is found in the United Kingdom and other similar health systems) or other regional levels [5].

In the past 15 years, there have been articles which introduce IV methods in general and the use of PPP IV specifically [1,6,7]. Chen et al. conducted a systematic review to

Author Contributions: Lisong Zhang proposed the idea, searched the literature, analysed the results, and drafted the original manuscript. Jim Lewsey and David reviewed and revised the draft. All authors approved the final manuscript.

Declaration of Interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

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

* Corresponding author. Health Economics and Health Technology Assessment, Institute of Health and Wellbeing (IHW), University of Glasgow, 203, 1 Lilybank Gardens, Glasgow G12 8RZ, UK. Tel.: +44-752-991-5445; fax.: +44-141-330-4010.

E-mail address: l.zhang.3@research.gla.ac.uk (L. Zhang).

https://doi.org/10.1016/j.jclinepi.2022.04.020

What is new?

Key findings

- There is an inconsistency between the applied studies and simulation studies that use physician's prescribing preference as instrumental variable in pharmacoepidemiology.

What this adds to what was known?

- This article reviews applied and simulation studies that use physician's prescribing preference as instrumental variables in pharmacoepidemiology between 2005 and 2020. In this review, the applied studies that use physician's prescribing preference as an instrumental variable do not always report their results to a standard that systematic reviews and simulation studies suggest.

What is the implication and what should change now?

- Applied studies using physician's prescribing preference as an instrumental variable should consider methodological insights from simulation studies to inform study designs.Applied studies using physician's prescribing preference as an instrumental variable should consider methodological insights from simulation studies to inform study designs.

synthesize drug research studies using IV to see whether it can be a valid approach for tackling unmeasured confounding bias [8]. Davies et al. also conducted a systematic review of using IV methods and reporting of IV results [9]. However, these cover common types of IV used in epidemiological research and so did not cover some issues of particular importance to PPPs. This article reviewed the scientific literature of CERs using PPP IV to compare how methods are employed in applied studies to those studies introducing important methodological considerations.

2. Method

2.1. Inclusion and exclusion criteria

To identify simulation studies, the inclusion criteria were: (1) Use PPP as IV in a hypothesized drug comparison study and (2) Published between 2005 and 2020.

To identify CER studies using PPP IV, we conducted a literature search using Ovid, PubMed, and Google Scholar with the following inclusion criteria: (1) Use prescription drugs as exposure, (2) Compare the effectiveness of two drugs, (3) Use PPP as IV, and (4) Published between 2005 and 2020. The exclusion criteria were: (1) Review article, (2) Clinical trial, and (3) Abstract or book.

2.2. Search terms

The search terms used in this review were IV and prescribing preference or medication or treatment for both applied and simulation studies. We distinguished between these two types of study as per whether they used simulated data or routinely collected data.

3. Results

Using the search terms, we found 1,192 records in Ovid, PubMed, and Google Scholar (Fig. 1).

We first excluded studies that were not comparing treatment effectiveness. Then we excluded studies that are irrelevant to IV methods. Finally, we deleted those studies that used the IV method but not specifically PPP IV. The remaining 18 applied and six simulation studies are included in the subsequent review.

We identified six simulation studies and 18 applied studies. Tables 1 and 2 in the supplementary material contain the key characteristics of these studies.

3.1. Summary of the simulation studies

All identified simulation studies aimed to get better understanding of using PPP IV as a method of reducing unmeasured confounding bias in pharmacoepidemiology (Table 1 in the supplementary material). For that, they formed their data to simulate hypothetical studies comparing the treatment effectiveness or the risk of adverse event in the context of large datasets. The PPP was built on the basis that all three assumptions are met. In general, better performance means lower variance (smaller standard deviation) and lower bias. Simulated data facilitate the comparison between the IV estimates with the 'true' estimate and use root of mean squares error, relative bias, and coverage rate to measure the bias quantitatively. Most of these studies mention the strength of the association between exposure and the instrument. They emphasize the association between instruments and exposures needs to be strong enough to implement an unbiased IV analysis. Ionescu-Ittu et al. used the proportion of exchangeable group to represent the strength of this association [4]. A further study [34] also focuses on the strength of PPP IV but in a more specific way by defining boundaries for weak instruments (e.g., Pearson's correlation coefficient <0.15 or odds ratio (OR) <2). Also, the limitations of using IV methods including the weak instrument and limited sample size can be shown in a more specific way [35].

3.2. Summary of the applied studies

3.2.1. The construction of proxy for physician's prescribing preference

In terms of constructing PPP IV, there are two major types of variables in the applied studies: binary (the most recent prescription made by the same physician, 'prior

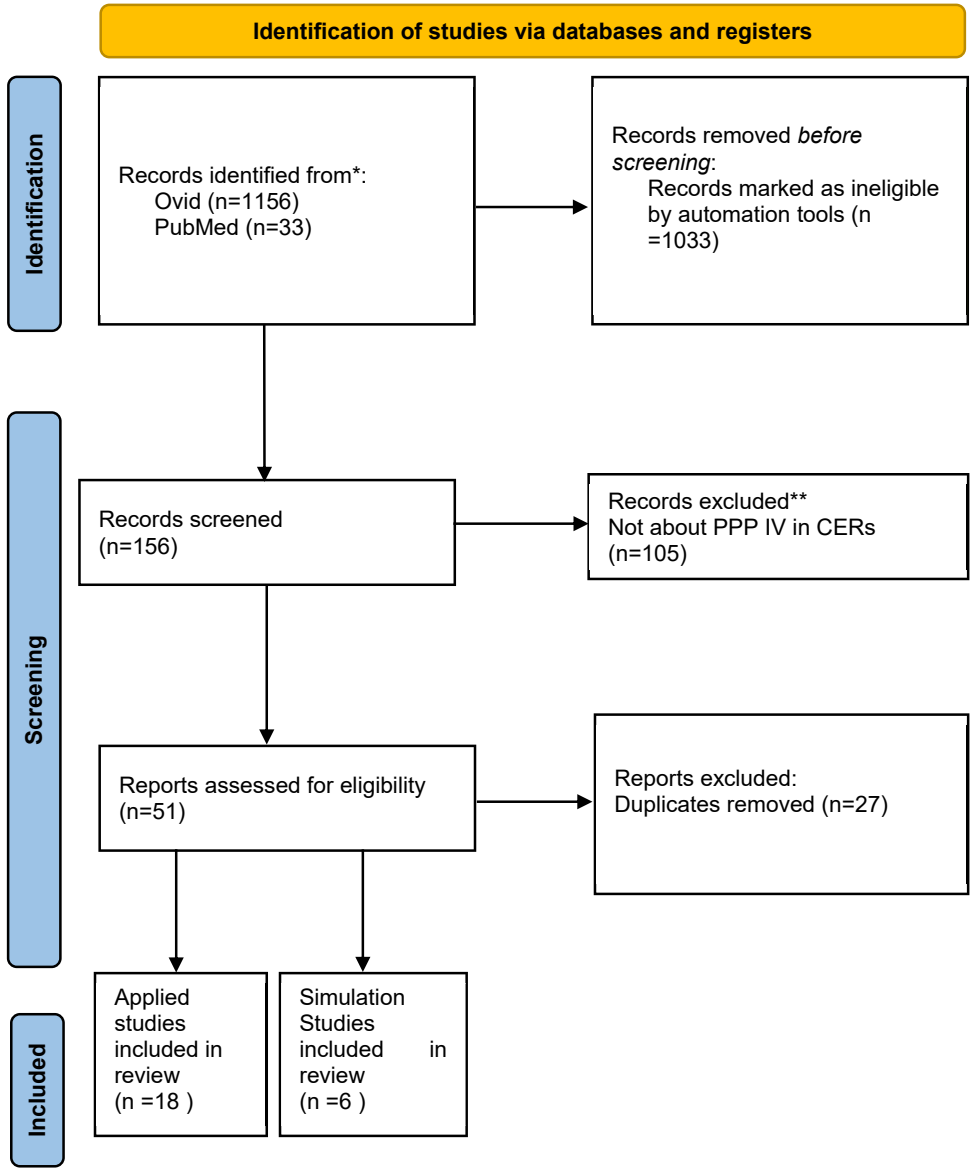

Fig. 1. PRISMA chart for a literature review. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

one') and numerical (the proportion of patients who were prescribed the drug of interest). The variance of binary outcomes is more likely to be inflated than continuous outcomes under the same settings [4]. Unlike the other studies, Koladjo et al. compared the results of generalized method of moment and two-stage residual inclusion which

Table 1. Summary of estimation method in different settings

Estimation	Suitable scenarios
Two-stage linear regression	Linear model [1]
Generalized method of moment (GMM) IV	Binary outcome [10]
Two-stage residual inclusion(2SRI)	Binary outcome [11]
	Numerical outcome [12]
	Time-to-event outcome [13]
Two-stage predictor substitution (2SPS)	Binary outcome [13], time-to-event outcome [14]
Local Average Treatment Effects (LARF)	Binary treatment and binary instrument [12]
IV Cox regression	Time-to-event outcomes [5]
GLM adaptions on IV (eg, IV Probit/IV logistic regression)	Binary outcome [15]

Table 2. Validation of assumptions. (a) The validation of relevance assumption of IV. (b) The validation of the exchangeability assumption of IV. (c) The validation of the exclusion restriction assumption of IV

Study	Validation of assumptions
[16]	(a) The probability of receiving COX-2 when the prior one prescription is COX-2.
[17]	(a) Compares the percentage of the same physician prescribed the same drug as the prior one and that of different drug. (b) The reduction of imbalance in covariates.
[18]	(a) Association between the instrument and the exposure: OR: 6.1, 95% CI (5.8–6.4).
[19]	N/A
[20]	(a) Cluster-robust F-statistics. (b) The risk difference of confounders on the level of actual treatment and on the level of IV. The reduction of Mahala nobis distance and the prevalence difference ratio.
[21]	(a) Partial F statistics.
[22]	(a) Partial F-statistic and R-square of linear regression on exposure and instrument. (b) Prevalence difference ratio.
[23]	(a) Partial F-statistic. (b) Bias component plot.
[24]	(a) Partial F-statistics.
[25]	(a) Partial F-statistic and sensitivity analysis.
[26]	(a) Partial F-statistic.
[27]	N/A
[28]	(a) Partial F-statistic and R-square, the square of the partial Spearman correlation coefficient. (b) Partial F-statistics for the regression model on instrument and three forms of IV. (c) Adjusted logistic regression on IV and outcome.
[5]	(a) Partial F-statistics and R-square.
[29]	(a) Partial F-statistics and the first stage linear probability.
[30]	(a) The percentage of actual treatment with a high preference for Bivalirudin.
[31]	(a) Point bi-serial correlation (r) for binary exposure and continuous IV; odds ratios for binary exposure and IV. (b) Standardized difference and multivariate Mahala nobis distance assess the imbalance of covariates.
[32]	(a) F-statistics of the first stage of regression. (b) Reduction of imbalance in covariates.
[33]	(a) F-statistics of first stage of regression and the mean association between exposure and instrument. (b and c) Bonet's instrumental variable inequality tests.

adds new knowledge on an IV analysis using nonlinear regression models [36]. Most of these studies used the most recent prescription or the prior one prescription of the same physician [17,20,22,27,29,37]. However, no study provided a rationale for the choice of the form of PPP.

Some articles conducted a sensitivity analysis by comparing the estimates from prior one and prior *n* prescriptions, such as prior seven and prior 20 prescriptions [20,23,24]. The other form of PPP IV is the proportion of one drug prescribed by the physician among all the previous patients which makes the IV a numerical variable [5,28]. The magnitude of a preference can also be defined using the proportion of one drug prescribed by the physician dichotomized at the median [21,38]. In higher levels of aggregation, studies do not account for the prescribing date but include the proportion of one treatment among all prescriptions [31].

3.2.2. Estimation method in different settings

Although the two-stage least square (2SLS) method is widely used in pharmacoepidemiology to account for unmeasured confounding for different types of outcome variable [1]. Literature shows that using the 2SLS estimator for binary outcomes may lead to a biased estimate [39,40]. Table 1 is a summary of the most suitable estimation methods in different settings. In the review, although all CERs have binary or time-to-event outcomes, only seven of 18 (39%) use nonlinear models instead of 2SLS. Regarding time-to-event outcomes, two-stage regression models that include the IV are proposed [41,42]. The first stage is a linear regression on the exposure and IV with measured covariates adjusted for. The second stage is a Cox proportional regression model that adjusts the survival probability from the first stage [5].

3.2.3. Validation of assumptions

In terms of validating IV assumptions (Table 2), most of the applied studies assessed the strength of the association between IV and the exposure (relevance assumption). One of the most common ways is to calculate partial F-statistics on the first stage of regression [5,20,29,22,21,24,28,26,25]. The rule of thumb for a strong enough instrument is if the

Table 3. Checklist for the researchers are interested in using PPP in CERs

1. Compare different formulations of PPP instrument.
2. Compare the IV method results with those from conventional methods, such as multivariable adjusted regression or propensity score approaches.
3. Use different estimation methods which have different underlying assumptions to explore the robustness of the study findings.
4. Consider accounting for time-varying prescribing preference (or provide a clear rationale for making the assumption that the physician's prescribing preference is time-fixed across the period of study).

F-statistic is greater than 10 [43]. Comparing the percentage of the actual treatment that match the prescribing preference is a more intuitive way to see if IV predicts the actual treatment [38,16]. There are also studies that fit regression models on exposure and IV to examine whether the association is statistically significant [29,37]. Of these different approaches, the F-statistic has the advantage that it accounts for the strength of association and being sensitive to sample size. Note, however, that a large F-statistic value only indicates that weak instrument bias is unlikely, it does not follow that there will be enough statistical power for the comparison of treatment effectiveness.

The exclusion restriction and exchangeability assumption cannot be tested empirically [1]. However, many studies compared the covariate imbalance by actual treatment and by PPP with an intuition that if PPP is less associated with measured confounders then it will also be less associated with unmeasured confounders [44]. Likewise, there are studies that reported the reduction of covariate imbalance using Mahala Nobis distance [20], bias component plot [23] and prevalence difference ratio (PDR) [20,22]. The exclusion restriction assumption which indicates that the IV does not affect the outcome directly has been overlooked in the identified studies of the review. Most studies did not mention or simply explained it using intuition that the preference of prescribing is not likely to influence the outcome [21]. Of the studies identified, only one [28] explored further and constructed an adjusted logistic regression model on IV and the outcome to examine whether PPP can affect the outcome.

4. Discussion

In this article, we have reviewed both the applied and methodological literature that uses PPP as an IV to address unmeasured confounding in CER. We found that the methodological insights from the simulation studies were not always being taking into account, or at least not reported on, in the applied studies.

Although some studies argued that 2SLS can be used as is asymptotically unbiased [4], researchers should endeavor to compare the results from 2SLS and that from other estimation methods to increase the robustness of the study. In terms of different forms of proxy for PPP, simulation studies have compared the performance of numerical and binary PPP formulations. Although some researchers concluded that the proportion form of PPP serves as a good proxy for PPP [36], others held a view that instantaneous preference is better which is consistent with the earlier definition of PPP IV [44]. Although a longer prescription history can lead to higher statistical power, there is a potential risk of overfitting to the data. Furthermore, longer prescription histories used in the PPP formulation lead to dropping those prescribers with few prescription records and dropping long 'look-back' periods from the analysis cohort, so a trade-off must be made. Given that the performance of PPP IV can depend on multiple elements, such as the rarity of outcome [4], sample size [45] and the estimation methods, a proper proxy for PPP needs more exploration that tends to occur in the applied literature to date. We suggest that researchers should present their IV estimates using more than one form of proxy as sensitivity analyses for possible violation of IV assumptions. Unlike in the simulation studies, the identified applied studies in the review did not account for or consider time-varying preferences. Abrahamowicz et al. [46] assumed the preference would switch (from preferring drug A to drug B) at a certain point in time. They include the time factor by using the prescription history as the proxy for PPP.

It should be noted that simulation studies cannot provide an insight for validation of all IV assumptions. Although the exchangeability and exclusion restriction assumptions cannot be empirically verified, researchers should give consideration to them in study reporting. Some studies have proposed a falsification strategy which may be a possible solution for assumption checking [47]. Further research may focus on explicitly how such a strategy can be used to help validate exchangeability and exclusion restriction assumptions.

Of course, simulation studies can only 'mimic' real life datasets. The simulation studies we reviewed build hypothetical studies to compare the effectiveness of two drugs. In some instances, the models parameters are set and varied without providing justification which is not consistent with the best practice for simulation studies [48]. Furthermore, the results of simulation studies often favor the new method that is being proposed/introduced which may be overstating what actually happens in real life datasets. Although there are no specific standards, some review papers have made suggestions on the reporting of IV studies. Brookhart [3] and Davies [9] proposed guidelines for reporting IV analyses results. More specifically, Swanson and Hernán [49] proposed a flowchart for reporting IV analyses of single binary non–time-varying IVs including a classification between average treatment effect and local average

treatment effect. Baiocchi [7] highlighted the exploration of concomitant treatment and sensitivity analyses. Jackson emphasized properly demonstrating the confounding bias by bias component plots [50].

Additional related issues for PPP IV are those of weighting and matching. The instrumental propensity score (IPS) is used which is the conditional probability of an IV given pretreatment covariates as weights in the regression models [51]. Cheng et al. further explored the IPS approach by implementing it in subclassifications and semiparametric models to gain treatment effects of subgroups [52]. Matching methods have been used to increase the strength of weak IVs. Baiocchi et al. [53,54] proposed a near-far approach to create pairs with similar covariate distributions and large differences in an IV value to mimic randomized trials. IPS can also be used in a full-matching approach [55]. Although not central to the design of PPP IV studies, we recommend that using these methods should be considered in CER, especially when the proportion of the complier group is relatively low.

We evaluated the 18 applied CER articles as per whether they account for insights from simulation studies by four criteria: (1) whether they compare different forms of PPP IV, (2) whether they compare IV methods with a conventional method, (3) whether they compare different estimation methods, and (4) whether they consider time-varying preferences. In summary, five of 18 studies (28%) compared different forms of PPP IV and 18 of 18 studies compared the IV methods with conventional methods. However, none of the studies considered time-varying preferences or compared different estimation methods which are often covered by the simulation studies.

5. Conclusion

In terms of exploring the suitability of the PPP IV method, simulation studies are flexible, easy to use, and have potential to make recommendations for good practice in the applied setting. We have shown that currently not all applied studies use simulation study results to guide their use of the PPP IV method. Applied studies using PPP as an IV should consider methodological insights from simulation studies to inform study designs. A brief checklist (Table 3) is presented for researchers who are interested in applying PPP IV.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2022.04.020.

References

[1] Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? Epidemiology 2006;17:360−72.
[2] Angrist JD, Imbens GW. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. J Am Stat Assoc 1995;90:431−42.
[3] Brookhart MA, Rassen JA, Wang PS, Dormuth C, Mogun H, Schneeweiss S. Evaluating the validity of an instrumental variable study of neuroleptics: can between-physician differences in prescribing patterns be used to estimate treatment effects? Med Care 2007;45: S116−22.
[4] Ionescu-Ittu R, Delaney JA, Abrahamowicz M. Bias-variance trade-off in pharmacoepidemiological studies using physician-preference-based instrumental variables: a simulation study. Pharmacoepidemiol Drug Saf 2009;18(7):562−71.
[5] Boef AG, Souverein PC, Vandenbroucke JP, van Hylckama Vlieg A, de Boer A, le Cessie S, et al. Instrumental variable analysis as a complementary analysis in studies of adverse effects: venous thromboembolism and second-generation versus third-generation oral contraceptives. Pharmacoepidemiol Drug Saf 2016;25(3):317−24.
[6] Lousdal ML. An introduction to instrumental variable assumptions, validation and estimation. Emerg Themes Epidemiol 2018;15(1): 1−7.
[7] Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. Stat Med 2014;33:2297−340.
[8] Chen Y, Briesacher BA. Use of instrumental variable in prescription drug research with observational data: a systematic review. J Clin Epidemiol 2011;64:687−700.
[9] Davies NM, Smith GD, Windmeijer F, Martin RM. Brief report: issues in the reporting and conduct of instrumental variable studies: a systematic review. Epidemiology 2013;24:363−9.
[10] Klungel O, Uddin MJ, de Boer A, Belitser SV, Groenwold RH, Roes KC. Instrumental variable analysis in epidemiologic studies: an overview of the estimation methods. Pharm Anal Acta 2015; 6(353):2.
[11] Terza JV, Bradford WD, Dismuke CE. The use of linear instrumental variables methods in health services research and health economics: a cautionary note. Health Serv Res 2008;43:1102−20.
[12] Zhang Z, Uddin MJ, Cheng J, Huang T. Instrumental variable analysis in the presence of unmeasured confounding. Ann Transl Med 2018;6(10):182.
[13] Terza JV, Basu A, Rathouz PJ. Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. J Health Econ 2008;27(3):531−43.
[14] Cai B, Small DS, Have TRT. Two-stage instrumental variable methods for estimating the causal odds ratio: analysis of bias. Stat Med 2011;30:1809−24.
[15] Rassen JA, Schneeweiss S, Glynn RJ, Mittleman MA, Brookhart MA. Instrumental variable analysis for estimation of treatment effects with dichotomous outcomes. Am J Epidemiol 2009;169:273−84.
[16] Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. Epidemiology 2006;17:268−75.
[17] Schneeweiss S, Solomon DH, Wang PS, Rassen J, Brookhart MA. Simultaneous assessment of short-term gastrointestinal benefits and cardiovascular risks of selective cyclooxygenase 2 inhibitors and nonselective nonsteroidal antiinflammatory drugs: an instrumental variable analysis. Arthritis Rheum 2006;54(11):3390−8.
[18] Schneeweiss S. Developments in post-marketing comparative effectiveness research. Clin Pharmacol Ther 2007;82:143−56.
[19] Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during coronary-artery bypass grafting and risk of death. N Engl J Med 2008;358:771−83.
[20] Davies NM, Gunnell D, Thomas KH, Metcalfe C, Windmeijer F, Martin RM. Physicians' prescribing preferences were a potential instrument for patients' actual prescriptions of antidepressants. J Clin Epidemiol 2013;66:1386−96.
[21] Kuo YF, Montie JE, Shahinian VB. Reducing bias in the assessment of treatment effectiveness: androgen deprivation therapy for prostate cancer. Med Care 2012;50:374−80.

[22] Davies NM, Smith GD, Windmeijer F, Martin RM. COX-2 selective nonsteroidal anti-inflammatory drugs and risk of gastrointestinal tract complications and myocardial infarction: an instrumental variable analysis. Epidemiology 2013;24:352−62.
[23] Davies NM, Taylor GMJ, Taylor AE, Jones T, Martin RM, Munafò MR, et al. The effects of prescribing varenicline on two-year health outcomes: an observational cohort study using electronic medical records. Addiction 2018;113(6):1105−16.
[24] Davies NM, Taylor AE, Taylor GM, Itani T, Jones T, Martin RM, et al. Varenicline versus nicotine replacement therapy for long-term smoking cessation: an observational study using the Clinical Practice Research Datalink. Health Technol Assess 2020;24(9):1−46.
[25] VanDyke RD, McPhail GL, Huang B, Fenchel MC, Amin RS, Carle AC, et al. Inhaled tobramycin effectively reduces FEV1 decline in cystic fibrosis. An instrumental variables analysis. Ann Am Thorac Soc 2013;10(3):205−12.
[26] Nelson RE, Nebeker JR, Hayden C, Reimer L, Kone K, LaFleur J. Comparing adherence to two different HIV antiretroviral regimens: an instrumental variable analysis. AIDS Behav 2013;17:160−7.
[27] Taylor GMJ, Taylor AE, Thomas KH, Jones T, Martin RM, Munafò MR, et al. The effectiveness of varenicline versus nicotine replacement therapy on long-term smoking cessation in primary care: a prospective cohort study of electronic medical records. Int J Epidemiol 2017;46:1948−57.
[28] Kollhorst B, Abrahamowicz M, Pigeot I. The proportion of all previous patients was a potential instrument for patients' actual prescriptions of nonsteroidal anti-inflammatory drugs. J Clin Epidemiol 2016;69:96−106.
[29] Chen H, Mehta S, Aparasu R, Patel A, Ochoa-Perez M. Comparative effectiveness of monotherapy with mood stabilizers versus second generation (atypical) antipsychotics for the treatment of bipolar disorder in children and adolescents. Pharmacoepidemiol Drug Saf 2014;23(3):299−308.
[30] Secemsky EA, Kirtane A, Bangalore S, Jovin IS, Patel D, Ferro EG, et al. Practice patterns and in-hospital outcomes associated with bivalirudin use among patients with non-ST-segment-elevation myocardial infarction undergoing percutaneous coronary intervention in the United States. Circ Cardiovasc Qual Outcomes 2017;10(9):e003741.
[31] Uddin MJ, Groenwold RH, de Boer A, Gardarsdottir H, Martin E, Candore G, et al. Instrumental variables analysis using multiple databases: an example of antidepressant use and risk of hip fracture. Pharmacoepidemiol Drug Saf 2016;25(Suppl 1):122−31.
[32] Suh HS, Hay JW, Johnson KA, Doctor JN. Comparative effectiveness of statin plus fibrate combination therapy and statin monotherapy in patients with type 2 diabetes: use of propensity-score and instrumental variable methods to adjust for treatment-selection bias. Pharmacoepidemiol Drug Saf 2012;21(5):470−84.
[33] Walker VM, Davies NM, Martin RM, Kehoe PG. Comparison of antihypertensive drug classes for dementia prevention. Epidemiology 2020;31:852−9.
[34] Uddin MJ, Groenwold RH, de Boer A, Belitser SV, Roes KC, Hoes AW, et al. Performance of instrumental variable methods in cohort and nested case−control studies: a simulation study. Pharmacoepidemiol Drug Saf 2014;23(2):165−77.
[35] Uddin MJ, Groenwold RH, de Boer A, Afonso AS, Primatesta P, Becker C, et al. Evaluating different physician's prescribing preference based instrumental variables in two primary care databases: a study of inhaled long-acting beta2-agonist use and the risk of myocardial infarction. Pharmacoepidemiol Drug Saf 2016;25:132−41.
[36] Koladjo BF, Escolano S, Tubert-Bitter P. Instrumental variable analysis in the context of dichotomous outcome and exposure with a numerical experiment in pharmacoepidemiology. BMC Med Res Methodol 2018;18:1−14.
[37] Schneeweiss S, Setoguchi S, Brookhart A, Dormuth C, Wang PS. Risk of death associated with the use of conventional versus atypical antipsychotic drugs among elderly patients. CMAJ 2007;176(5):627−32.
[38] Secemsky EA, Kirtane A, Bangalore S, Jovin IS, Patel D, Ferro EG, et al. Practice patterns and in-hospital outcomes associated with bivalirudin use among patients with non−ST-segment−elevation myocardial infarction undergoing percutaneous coronary intervention in the United States. Circ Cardiovasc Qual Outcomes 2017;10(9):e003741.
[39] Ertefaie A, Flory JH, Hennessy S, Small DS. Instrumental variable methods for continuous outcomes that accommodate nonignorable missing baseline values. Am J Epidemiol 2017;185:1233−9.
[40] Rassen JA, Brookhart MA, Glynn RJ, Mittleman MA, Schneeweiss S. Instrumental variables II: instrumental variable application—in 25 variations, the physician prescribing preference generally was strong and reduced covariate imbalance. J Clin Epidemiol 2009;62(12):1233−41.
[41] Tchetgen Tchetgen EJ, Walter S, Vansteelandt S, Martinussen T, Glymour M. Instrumental variable estimation in a survival context. Epidemiology 2015;26:402.
[42] Martínez-Camblor P, Mackenzie T, Staiger DO, Goodney PP, O'Malley AJ. Adjusting for bias introduced by instrumental variable estimation in the Cox proportional hazards model. Biostatistics 2019;20(1):80−96.
[43] Stock JH, Yogo M. Testing for weak instruments in linear IV regression. National Bureau of Economic Research; 2005:80−108.
[44] Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. Int J Biostat 2007;3(1):14.
[45] Boef AG, Dekkers OM, Vandenbroucke JP, le Cessie S. Sample size importantly limits the usefulness of instrumental variable methods in epidemiological studies. Obtaining causal estimates therapeutic effects observational studies. J Clin Epidemiol 2014;67(11):1258−64.
[46] Abrahamowicz M, Beauchamp ME, Ionescu-Ittu R, Delaney JA, Pilote L. Reducing the variance of the prescribing preference-based instrumental variable estimates of the treatment effect. Am J Epidemiol 2011;174:494−502.
[47] Labrecque J, Swanson SA. Understanding the assumptions underlying instrumental variable analyses: a brief review of falsification strategies and related tools. Curr Epidemiol Rep 2018;5(3):214−20.
[48] Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. Stat Med 2019;38:2074−102.
[49] Swanson SA, Hernán MA. Commentary: how to report instrumental variable analyses (suggestions welcome). Epidemiology 2013;24:370−4.
[50] Jackson JW, Swanson SA. Toward a clearer portrayal of confounding bias in instrumental variable applications. Epidemiology 2015;26:498−504.
[51] Tan Z, Regression and weighting methods for causal inference using instrumental variables. J Am Stat Assoc 2006;101(476):1607−18.
[52] Cheng J, Lin W. Understanding causal distributional and subgroup effects with the instrumental propensity score. Am J Epidemiol 2018;187(3):614−22.
[53] Baiocchi M, Small DS, Lorch S, Rosenbaum PR. Building a stronger instrument in an observational study of perinatal care for premature infants. J Am Stat Assoc 2010;105(492):1285−96.
[54] Baiocchi M, Small DS, Yang L, Polsky D, Groeneveld PW. Near/far matching: a study design approach to instrumental variables. Health Serv Outcomes Res Methodol 2012;12(4):237−53.
[55] Kang H, Kreuels B, Adjei O, Krumkamp R, May J, Small DS. The causal effect of malaria on stunting: a Mendelian randomization and matching approach. Int J Epidemiol 2013;42(5):1390−8.