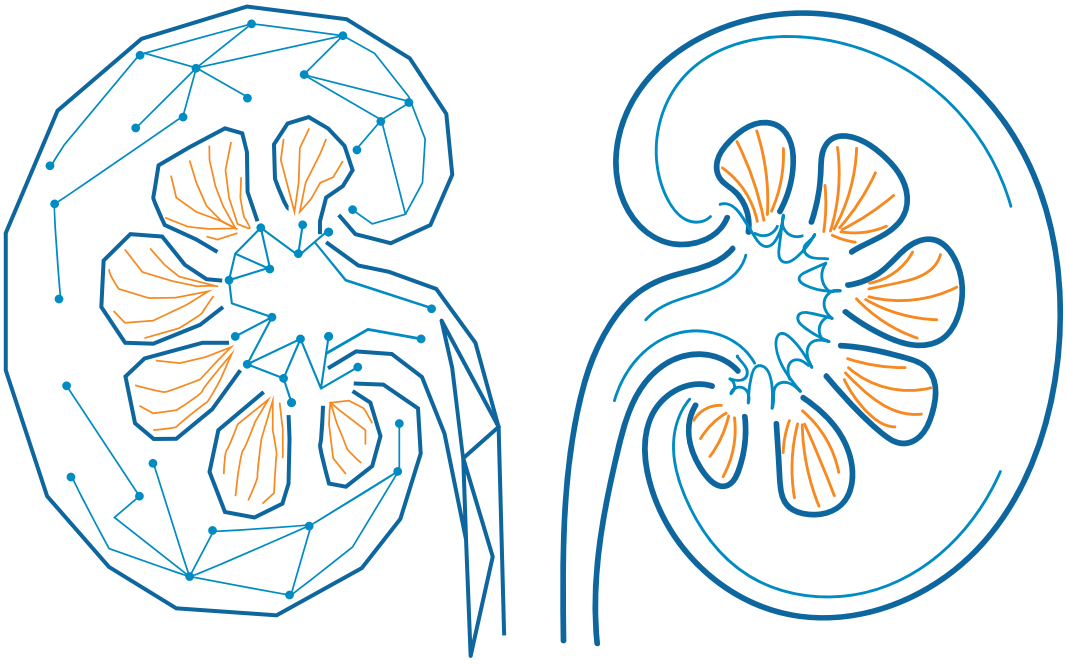
Optimal cardiovascular treatment strategies in kidney disease: casual inference from observational data

Fu, E.L.



CHAPTER 10

Summary and general discussion

Summary of main findings

The aim of this thesis was to substantially contribute to decision making in kidney disease. It answered a number of clinical questions on the effectiveness and safety of therapies by applying state-of-the-art methods to routinely collected healthcare data. In addition, we aimed to raise awareness on potential biases that could arise when using observational data for these purposes and how to mitigate them.

Observational studies with a causal aim can be plagued by a number of biases, of which some are discussed in **Chapter 2**. Although confounding by indication is a threat to any observational study assessing the causal effects of treatments, questions on unintended treatment effects and questions which involve an active comparator may be less susceptible to confounding. Remaining confounding should be addressed in the statistical analysis, where confounders to adjust for should be identified using subject matter knowledge, for example by using causal diagrams. Only measured confounding can be adjusted for, and this can be done through multivariable regression, standardization or propensity score (PS) methods. For point treatments, all methods are capable to adjust for measured confounders. However, in the setting of sustained treatments and treatment-confounder feedback, special methods based on standardization or weighting are required. The impact of residual confounding can be explored by calculating the E-value, performing quantitative bias analysis, or applying a negative control outcome or positive cohort analysis. These analyses can make the assumption of no unmeasured confounding more or less plausible. In addition to confounding, prevalent user bias and immortal time bias are important limitations in many observational studies. These biases arise whenever the start of follow-up and the start of exposure do not align. Explicit emulation of a randomized trial can eliminate these biases since it forces the alignment of follow-up and exposure. Lastly, missing data and measurement error often occur in routinely collected healthcare data. Methods such as multiple imputation and quantitative bias analysis are therefore recommended.

In **Chapter 3** we further discussed the concept, merits, and possible caveats of PS methods, a popular statistical method to control for measured confounding. Various methods that use the PS to control for confounding exist. These include PS matching, PS stratification, multivariable regression analysis including the PS as a covariate, and PS weighting. PS methods offer a number of advantages: it is possible to check for covariate balance, to choose a specific target population and to exclude individuals in non-overlapping regions of the PS distribution. Furthermore, PS methods are preferred in the setting of high-dimensional data with many confounders and relatively few outcomes.

In **Chapter 4** we highlighted immortal time bias in a published observational study aiming to estimate the causal effect of metformin in diabetic kidney disease. We propose a number of solutions that could have been applied to eliminate this bias. These include modelling metformin use as a time-varying exposure, landmarking or the use of grace periods in combination with the cloning, censoring and weighting method. Alternatively, a sequential trial approach could have been applied.

In **Chapter 5** we studied the association between acute increases of creatinine and cardiovascular and kidney outcomes. Patients with higher creatinine increase after initiation of RASi were at higher risk of death, heart failure, myocardial infarction and end-stage kidney disease. We also found that only 18% of individuals initiating RASi received the recommended creatinine monitoring, and that increases between 10-29% were common among monitored individuals. These results do not address the issue of discontinuation of RASi when creatinine increases but do suggest that patients with increases in creatinine have higher subsequent risk of cardiovascular and kidney outcomes.

Chapter 6 is a comparative effectiveness study of RASi versus calcium channel blockers on kidney replacement therapy, mortality and major adverse cardiovascular events in individuals with advanced CKD. We found that initiation of RASi was associated with a lower risk of kidney replacement therapy, and similar risks of mortality and cardiovascular events, compared with calcium channel blockers. We also performed analyses in a positive control cohort of patients with mild-to-moderate CKD, which aligned with evidence from previous randomized trials. A negative control analysis using cancer incidence did not indicate residual confounding by body mass index or smoking. These findings may inform clinical decisions on the choice of antihypertensive therapy in patients with advanced CKD.

In **Chapter 7** we found that β-blocker use at discharge was associated with a lower risk of mortality and cardiovascular mortality/heart failure hospitalization in individuals with heart failure with reduced ejection fraction and advanced CKD. A positive control analysis in individuals with heart failure with reduced ejection fraction and mild-to-moderate CKD showed a similar reduction in outcomes for β-blocker users. Such benefits were not observed in individuals with advanced CKD and midrange or preserved ejection fraction, although these results were inconclusive due to limited power. These findings suggest that β-blockers are effective in patients with heart failure with reduced ejection fraction across the spectrum of kidney function.

In **Chapter 8** we examined the effect of stopping versus continuing RASi in individuals with advanced CKD on mortality, major adverse cardiovascular events and kidney replacement therapy. We observed that individuals who stopped RASi had a higher

risk of mortality and major adverse cardiovascular events, but a lower risk of kidney replacement therapy. Similar findings were obtained when modelling RASi use as a time-varying covariate in a marginal structural model, when additionally adjusting for potassium and albumin-to-creatinine ratio and within subgroups, including when RASi was stopped at higher (eGFR 20-30 ml/min/1.73m^2) or lower eGFR (<20 ml/min/1.73m^2). These findings caution against routine discontinuation of RASi in individuals with advanced CKD, while awaiting evidence from the STOP-ACEi trial.

Chapter 9 addresses the question whether there is an optimal kidney function to start dialysis. We were able to replicate the findings of the randomized IDEAL trial using observational data. We further showed that early dialysis initiation was associated with a modest reduction in mortality and cardiovascular events (with an eGFR of 15-16 versus 6-7 ml/min/1.73m^2). This translated to a mean postponement of death of 1.6 months at the expense of starting dialysis 48 months earlier. We also show that previous observational studies suffered from lead time bias, selection bias and immortal time bias, that these biases can be avoided by applying the target trial emulation framework, and that incorrect analysis of our own data leads to similar biased results. Collectively, these findings indicate that there is little benefit of starting dialysis early based on eGFR alone. Future studies may investigate whether dialysis should be started based on symptom burden, to further improve clinical outcomes.

Future perspectives

The number of observational studies using routinely collected data is ever increasing. In this thesis we highlighted that the use of such data to inform clinical practice represents a double-edged sword: on the one hand it offers tremendous opportunities to study how treatments work in real-world practice, to study questions that are difficult to answer in randomized trials, and to study populations that were underrepresented in trials. On the other hand, several biases can invalidate the findings from observational studies: confounding bias, immortal time bias or prevalent user bias to name just a few.

How should we move forward to provide the best evidence for treating patients with kidney disease? Of course, more trials need to be conducted. In kidney disease, there have been relatively few randomized trials conducted (1), and patients with kidney disease have been largely excluded from trials in other fields, such as cardiology or oncology (2, 3). In order to solve this issue, others have called for reducing the costs and complexity of conducting trials, including the bureaucratic burden (4, 5). That this is possible has been proven by the RECOVERY trial, which randomized over 39.000 patients hospitalized for COVID-19 in less than a year (6). Additional examples include the publication of "large, simple trials" in the past decades (7) and recent innovations in

trial design such as the registry-based trial (8). Fortunately, the field of kidney disease seems to be catching up, with the publication of a number of important clinical trials (9-13) and other trials now actively recruiting patients (14-17). In addition to conducting more trials, novel approaches to generalize trial results to other populations can bridge the gap between trials and routine clinical practice (18, 19).

A fundamental question is whether it is even possible to draw causal conclusions from observational data. Indeed, some researchers are of the opinion that only randomized trials can obtain causal conclusions and observational studies cannot, and that only randomized trials are therefore useful (4, 20). However, this is a false dichotomy. Causality is not a yes/no statement and is rarely concluded on the basis of one study, since it always involves the totality of evidence, which can come from laboratory studies, observational studies, and RCTs. Furthermore, there are considerable differences in quality between observational studies, with some better able to come to causal conclusions and others less so. On the one hand, well-conducted observational studies have successfully replicated or predicted the findings of RCTs (21-31). On the other hand, numerous examples exist where observational studies have failed to do so (20, 32-35). The latter can often be explained by the fact that the these observational studies used flawed methods which introduced unnecessary biases, such as immortal time bias or prevalent user bias, rather than the presence of unmeasured confounding (36). These biases arise whenever the timing of the following three elements is not aligned at baseline: start of follow-up, start of the treatment strategies, and fulfilment of all eligibility criteria for each included patient (36). Since randomization automatically aligns the timing of these elements, trialists never have to worry about this problem. However, this is not the case in observational data where researchers must carefully think about baseline and handle this appropriately in their analyses. Examples from the literature where well-conducted observational studies were able to obtain answers similar to trials, whereas observational studies that introduced preventable biases were not, include studies on the timing of dialysis initiation and the risk of mortality (this thesis), postmenopausal hormone therapy and the risk of coronary heart disease (26, 32, 37, 38), statins and the risk of cancer (23, 33, 39, 40), timing of combined anti-retroviral therapy and risk of mortality (21), dabigatran and the risk of stroke (34, 41, 42), sodium-glucose cotransporter 2 inhibitors and the risk of mortality (35, 43, 44) and colonoscopy screening and the risk of colorectal cancer (45, 46). Using the target trial emulation framework can help to eliminate these unnecessary design flaws (47), forces the researcher to ask meaningful causal questions (48, 49), facilitates communication and guides the statistical analysis (50-52). In addition, investigators should use the analytical methods that are best suited to answer the clinical question at hand. For example, the cloning, censoring and weighting method is suitable to answer questions which 1) compare different timings of an intervention ("When should we start treatment?"); 2) compare different durations of a treatment ("How long should we

treat?") and 3) involve a grace period ("Should treatment be started within x months after event y or not?") (53). When the aim is to compare initiation of a treatment against no initiation, a random eligibility date needs to be chosen for the non-initiators or a sequential trial approach should be used to correctly handle baseline (22, 45). Researchers conducting observational studies should therefore have the appropriate methodological expertise and receive thorough training to correctly implement the methods, or involve someone with this expertise. When flawed methods are applied, flawed answers will be obtained.

Ongoing systematic replications of randomized trials using observational data, such as the RCT-DUPLICATE initiative (25, 27, 54, 55) and other efforts (24, 56-61), are therefore essential to demonstrate that observational studies can lead to the same conclusion as RCTs if done adequately. They will also provide valuable insights under which circumstances (i.e. for which study question, analytical methods, and data sources) one can come to causal conclusions in observational data and study treatment effects without randomization. These studies use principled causal inference methods and also try to emulate trial inclusion and exclusion criteria as much as possible to ensure that findings do not differ because of applying flawed methods or different patient populations (55). Such calibration studies need to be performed in the field of kidney disease as well.

Furthermore, not all observational studies are equally sensitive to confounding. Whether an observational study can come to causal conclusions greatly depends on the study question at hand, which has been referred to as an "axis of haphazardness of exposure" (62). Pharmacoepidemiological studies investigating harmful unintended effects of medications suffer less from (residual) confounding than studies investigating (un)intended beneficial effects. Indeed, regulatory agencies have a long history of using evidence from observational studies to assess the safety of medications. Furthermore, studies that involve the comparison of two drugs that are prescribed for similar indications are much less susceptible to confounding compared with studies that compare a drug against no drug (63). It is difficult to study questions comparing initiation versus no initiation in observational data, since treatment initiation is a marker of disease severity or – in the case of preventive treatments – a marker of health-seeking behaviour; both of these sources of bias may not be completely captured in observational data. In such cases, treatment selection may be so strong that baseline randomization is necessary. Nevertheless, successful applications do exist in literature (22, 26).

The ability to draw causal conclusions also greatly depends on the data that are used (64, 65), and the variables that are available in the database. Before embarking on a study, investigators should ask whether the data are of sufficient quality for the particular study. When sufficient granularity in exposures, outcomes or covariates

are missing, one may choose not to proceed with the analysis. Specifically for kidney disease, availability of routine laboratory measurements such as creatinine and albuminuria are often essential to adequately control for confounding.

Although the absence of unmeasured confounding remains an assumption that cannot be verified, researchers must carefully justify this assumption as best as possible, which is the difficult part of epidemiological research. Different analyses can be used to strengthen our confidence in the validity of findings, e.g. positive and negative control analyses, which could be either outcomes or cohorts for which we expect certain associations (either null or non-null). These analyses can be performed to explore whether it is feasible to answer a particular question even before conducting the primary analysis (66, 67). If trial results are available, the results of the observational analysis can be compared with those obtained in the trial, taking into account whether similar treatment strategies and study populations were investigated. In addition, adequate control must be made for confounding. For instance, whenever time-dependent confounding is present, G-methods such as inverse probability weighting of marginal structural models are required to obtain unbiased estimates (68). Propensity scores are a popular method to adjust for measured confounders. The many developments in propensity score methodology offer great flexibility in specification of the target population (to which population do our results apply), covariate balance, and precision (69). Importantly, covariate balance on measured and unmeasured confounders should be checked prior to analysis. In propensity score analyses, unmeasured confounders are only balanced to the extent that they are correlated with measured variables that were included in the propensity score. Therefore, a key approach to adjust for residual confounding from unobserved factors is to adjust for as many proxies of the underlying confounder as possible (e.g. diagnoses, procedures, medications, number of hospitalizations), which should be measured before the start of treatment to prevent adjusting for causal intermediates (28, 70). Whenever certain confounders are only available for a subset of the population and are not used in the adjustment set, balance in this variable after propensity score matching/weighting increases confidence that other unmeasured variables are also balanced. Besides confounding, other sources of bias should be investigated as well through sensitivity analyses, such as differential outcome ascertainment. E.g., when investigating a 30% GFR decline as outcome, one should check whether both treatment arms have the same intensity of creatinine testing during follow-up (71). Quantitative bias analyses can be used to investigate the influence of remaining biases.

Furthermore, different observational studies applying different causal methods can be used to triangulate evidence (72), since each method has its own specific assumptions. On this note, there are great opportunities for exchange of (quasi-experimental) methods from other scientific fields, such as regression discontinuity (73). Lastly, several other measures can increase the reproducibility of observational studies (74). These include preregistration of observational studies (75), the publication of codes (76), and adhering to reporting guidelines (77-80). In essence, the process of conducting observational studies should mimic the regulatory submission process of randomized trials. For example, no treatment-specific outcome analyses should be conducted until full specification and registration of the protocol (67).

In conclusion, obtaining evidence from non-experimental and experimental studies will remain important as both sources of evidence complement each other. Well-conducted observational studies can provide valuable evidence for decision-making in the field of kidney disease but also for medicine in general. All we need to do is to answer the right questions with the correct methods.

References

1. Inrig JK, Califf RM, Tasneem A, Vegunta RK, Molina C, Stanifer JW, et al. The landscape of clinical trials in nephrology: a systematic review of Clinicaltrials.gov. Am J Kidney Dis. 2014;63(5):771-80.
2. Kitchlu A, Shapiro J, Amir E, Garg AX, Kim SJ, Wald R, et al. Representation of Patients With Chronic Kidney Disease in Trials of Cancer Therapy. JAMA. 2018;319(23):2437-9.
3. Konstantinidis I, Nadkarni GN, Yacoub R, Saha A, Simoes P, Parikh CR, et al. Representation of Patients With Kidney Disease in Trials of Cardiovascular Interventions: An Updated Systematic Review. JAMA Intern Med. 2016;176(1):121-4.
4. Collins R, Bowman L, Landray M, Peto R. The Magic of Randomization versus the Myth of Real-World Evidence. N Engl J Med. 2020;382(7):674-8.
5. Eapen ZJ, Lauer MS, Temple RJ. The imperative of overcoming barriers to the conduct of large, simple trials. JAMA. 2014;311(14):1397-8.
6. Group RC. Azithromycin in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. Lancet. 2021.
7. Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? Stat Med. 1984;3(4):409-22.
8. Erlinge D, Omerovic E, Frobert O, Linder R, Danielewicz M, Hamid M, et al. Bivalirudin versus Heparin Monotherapy in Myocardial Infarction. N Engl J Med. 2017;377(12):1132-42.
9. Badve SV, Pascoe EM, Tiku A, Boudville N, Brown FG, Cass A, et al. Effects of Allopurinol on the Progression of Chronic Kidney Disease. N Engl J Med. 2020;382(26):2504-13.
10. Doria A, Galecki AT, Spino C, Pop-Busui R, Cherney DZ, Lingvay I, et al. Serum Urate Lowering with Allopurinol and Kidney Function in Type 1 Diabetes. N Engl J Med. 2020;382(26):2493-503.
11. Perkovic V, Jardine MJ, Neal B, Bompoint S, Heerspink HJL, Charytan DM, et al. Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. N Engl J Med. 2019;380(24):2295-306.
12. Heerspink HJL, Stefansson BV, Correa-Rotter R, Chertow GM, Greene T, Hou FF, et al. Dapagliflozin in Patients with Chronic Kidney Disease. N Engl J Med. 2020;383(15):1436-46.
13. Bakris GL, Agarwal R, Anker SD, Pitt B, Ruilope LM, Rossing P, et al. Effect of Finerenone on Chronic Kidney Disease Outcomes in Type 2 Diabetes. N Engl J Med. 2020;383(23):2219-29.
14. Bhandari S, Ives N, Brettell EA, Valente M, Cockwell P, Topham PS, et al. Multicentre randomized controlled trial of angiotensin-converting enzyme inhibitor/angiotensin receptor blocker withdrawal in advanced renal disease: the STOP-ACEi trial. Nephrol Dial Transplant. 2016;31(2):255-61.
15. Herrington WG, Preiss D, Haynes R, von Eynatten M, Staplin N, Hauske SJ, et al. The potential for improving cardio-renal outcomes by sodium-glucose co-transporter-2 inhibition in people with chronic kidney disease: a rationale for the EMPA-KIDNEY study. Clin Kidney J. 2018;11(6):749-61.
16. Murphy E, Burns A, Murtagh FEM, Rooshenas L, Caskey FJ. The Prepare for Kidney Care Study: prepare for renal dialysis versus responsive management in advanced chronic kidney disease. Nephrol Dial Transplant. 2020.
17. Edmonston DL, Isakova T, Dember LM, Brunelli S, Young A, Brosch R, et al. Design and Rationale of HiLo: A Pragmatic, Randomized Trial of Phosphate Management for Patients Receiving Maintenance Hemodialysis. Am J Kidney Dis. 2020.

18. Kent DM, Paulus JK, van Klaveren D, D'Agostino R, Goodman S, Hayward R, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement. Ann Intern Med. 2020;172(1):35-45.
19. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial. Am J Epidemiol. 2010;172(1):107-15.
20. Fanaroff AC, Califf RM, Harrington RA, Granger CB, McMurray JJV, Patel MR, et al. Randomized Trials Versus Common Sense and Clinical Observation: JACC Review Topic of the Week. J Am Coll Cardiol. 2020;76(5):580-9.
21. Collaboration H-C, Cain LE, Logan R, Robins JM, Sterne JA, Sabin C, et al. When to initiate combined antiretroviral therapy to reduce mortality and AIDS-defining illness in HIV-infected persons in developed countries: an observational study. Ann Intern Med. 2011;154(8):509-15.
22. Danaei G, Rodriguez LA, Cantero OF, Logan R, Hernan MA. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. Stat Methods Med Res. 2013;22(1):70-96.
23. Emilsson L, Garcia-Albeniz X, Logan RW, Caniglia EC, Kalager M, Hernan MA. Examining Bias in Studies of Statin Treatment and Survival in Patients With Cancer. JAMA Oncol. 2018;4(1):63-70.
24. Fralick M, Kesselheim AS, Avorn J, Schneeweiss S. Use of Health Care Databases to Support Supplemental Indications of Approved Medications. JAMA Intern Med. 2018;178(1):55-63.
25. Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating Randomized Clinical Trials with Nonrandomized Real-World Evidence Studies: First Results from the RCT DUPLICATE Initiative. Circulation. 2020.
26. Hernan MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology. 2008;19(6):766-79.
27. Patorno E, Schneeweiss S, Gopalakrishnan C, Martin D, Franklin JM. Using Real-World Data to Predict Findings of an Ongoing Phase IV Cardiovascular Outcome Trial: Cardiovascular Safety of Linagliptin Versus Glimepiride. Diabetes Care. 2019;42(12):2204-10.
28. Rassen JA, Murk W, Schneeweiss S. Real-world evidence of bariatric surgery and cardiovascular benefits using electronic health records data: A lesson in bias. Diabetes Obes Metab. 2021.
29. Graham DJ, Reichman ME, Wernecke M, Zhang R, Southworth MR, Levenson M, et al. Cardiovascular, bleeding, and mortality risks in elderly Medicare patients treated with dabigatran or warfarin for nonvalvular atrial fibrillation. Circulation. 2015;131(2):157-64.
30. Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during coronary-artery bypass grafting and risk of death. N Engl J Med. 2008;358(8):771-83.
31. Kim SC, Solomon DH, Rogers JR, Gale S, Klearman M, Sarsour K, et al. Cardiovascular Safety of Tocilizumab Versus Tumor Necrosis Factor Inhibitors in Patients With Rheumatoid Arthritis: A Multi-Database Cohort Study. Arthritis Rheumatol. 2017;69(6):1154-64.
32. Grodstein F, Stampfer MJ, Manson JE, Colditz GA, Willett WC, Rosner B, et al. Postmenopausal estrogen and progestin use and the risk of cardiovascular disease. N Engl J Med. 1996;335(7): 453-61.
33. Nielsen SF, Nordestgaard BG, Bojesen SE. Statin use and reduced cancer-related mortality. N Engl J Med. 2012;367(19):1792-802.

34. Sorensen R, Gislason G, Torp-Pedersen C, Olesen JB, Fosbol EL, Hvidtfeldt MW, et al. Dabigatran use in Danish atrial fibrillation patients in 2011: a nationwide study. BMJ Open. 2013;3(5).

35. Kosiborod M, Cavender MA, Fu AZ, Wilding JP, Khunti K, Holl RW, et al. Lower Risk of Heart Failure and Death in Patients Initiated on Sodium-Glucose Cotransporter-2 Inhibitors Versus Other Glucose-Lowering Drugs: The CVD-REAL Study (Comparative Effectiveness of Cardiovascular Outcomes in New Users of Sodium-Glucose Cotransporter-2 Inhibitors). Circulation. 2017;136(3):249-59.

36. Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol. 2016;79:70-5.

37. Grodstein F, Manson JE, Colditz GA, Willett WC, Speizer FE, Stampfer MJ. A prospective, observational study of postmenopausal hormone therapy and primary prevention of cardiovascular disease. Ann Intern Med. 2000;133(12):933-41.

38. Manson JE, Hsia J, Johnson KC, Rossouw JE, Assaf AR, Lasser NL, et al. Estrogen plus progestin and the risk of coronary heart disease. N Engl J Med. 2003;349(6):523-34.

39. Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. Lancet. 2016;388(10059):2532-61.

40. Danaei G, Tavakkoli M, Hernan MA. Bias in observational studies of prevalent users: lessons for comparative effectiveness research from a meta-analysis of statins. Am J Epidemiol. 2012;175(4):250-62.

41. Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med. 2009;361(12):1139-51.

42. Schneeweiss S, Gopalakrishnan C, Bartels DB, Franklin JM, Zint K, Kulldorff M, et al. Sequential Monitoring of the Comparative Effectiveness and Safety of Dabigatran in Routine Care. Circ Cardiovasc Qual Outcomes. 2019;12(2):e005173.

43. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, et al. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. N Engl J Med. 2017;377(7):644-57.

44. Pasternak B, Ueda P, Eliasson B, Svensson AM, Franzen S, Gudbjornsdottir S, et al. Use of sodium glucose cotransporter 2 inhibitors and risk of major cardiovascular events and heart failure: Scandinavian register based cohort study. BMJ. 2019;366:l4772.

45. Garcia-Albeniz X, Hsu J, Hernan MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. Eur J Epidemiol. 2017;32(6):495-500.

46. Garcia-Albeniz X, Hsu J, Bretthauer M, Hernan MA. Effectiveness of Screening Colonoscopy to Prevent Colorectal Cancer Among Medicare Beneficiaries Aged 70 to 79 Years: A Prospective Observational Study. Ann Intern Med. 2017;166(1):18-26.

47. Zhao SS, Lyu H, Solomon DH, Yoshida K. Improving rheumatoid arthritis comparative effectiveness research through causal inference principles: systematic review using a target trial emulation framework. Ann Rheum Dis. 2020;79(7):883-90.

48. Kaufman JS, Hernan MA. Epidemiologic methods are useless: they can only give you answers. Epidemiology. 2012;23(6):785-6.

49. Goetghebeur E, le Cessie S, De Stavola B, Moodie EE, Waernbaum I, on behalf of" the topic group Causal Inference of the Si. Formulating causal questions and principled statistical answers. Stat Med. 2020.

50. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64.

51. Didelez V. Commentary: Should the analysis of observational data always be preceded by specifying a target experimental trial? Int J Epidemiol. 2016;45(6):2049-51.

52. Labrecque JA, Swanson SA. Target trial emulation: teaching epidemiology and beyond. Eur J Epidemiol. 2017;32(6):473-5.

53. Zhao SS, Lyu H, Yoshida K. Versatility of the clone-censor-weight approach: response to 'trial emulation in the presence of immortal-time bias'. Int J Epidemiol. 2020.

54. Franklin JM, Pawar A, Martin D, Glynn RJ, Levenson M, Temple R, et al. Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. Clin Pharmacol Ther. 2020;107(4):817-26.

55. Franklin JM, Glynn RJ, Suissa S, Schneeweiss S. Emulation Differences vs. Biases When Calibrating Real-World Evidence Findings Against Randomized Controlled Trials. Clin Pharmacol Ther. 2020;107(4):735-7.

56. Matthews AA, Szummer K, Dahabreh IJ, Lindahl B, Erlinge D, Feychting M, et al. Comparing effect estimates in randomized trials and observational studies from the same population: an application to percutaneous coronary intervention. medRxiv. 2021:2021.02.01.21250739.

57. Lodi S, Phillips A, Lundgren J, Logan R, Sharma S, Cole SR, et al. Effect Estimates in Randomized Trials and Observational Studies: Comparing Apples With Apples. Am J Epidemiol. 2019;188(8):1569-77.

58. Webster-Clark M, Lund JL, Sturmer T, Poole C, Simpson RJ, Edwards JK. Reweighting Oranges to Apples: Transported RE-LY Trial Versus Nonexperimental Effect Estimates of Anticoagulation in Atrial Fibrillation. Epidemiology. 2020;31(5):605-13.

59. Yiu ZZN, Mason KJ, Hampton PJ, Reynolds NJ, Smith CH, Lunt M, et al. Randomized Trial Replication Using Observational Data for Comparative Effectiveness of Secukinumab and Ustekinumab in Psoriasis: A Study From the British Association of Dermatologists Biologics and Immunomodulators Register. JAMA Dermatol. 2021;157(1):66-73.

60. Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, et al. Real world effects of COPD medications: a cohort study with validation against RCT results. Eur Respir J. 2020.

61. Spoendlin J, Desai RJ, Franklin JM, Glynn RJ, Payne E, Schneeweiss S. Using Healthcare Databases to Replicate Trial Findings for Supplemental Indications: Adalimumab in Patients with Ulcerative Colitis. Clin Pharmacol Ther. 2020;108(4):874-84.

62. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. PLoS Med. 2008;5(3):e67.

63. Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. Nat Rev Rheumatol. 2015;11(7):437-41.

64. Franklin JM, Schneeweiss S. When and How Can Real World Data Analyses Substitute for Randomized Controlled Trials? Clin Pharmacol Ther. 2017;102(6):924-33.

65. Trevisan M, Fu EL, Xu Y, Jager K, Zoccali C, Dekker FW, et al. Pharmacoepidemiology for nephrologists (part 1): concept, applications and considerations for study design. Clinical Kidney Journal. 2020.

66. Franklin JM, Glynn RJ, Martin D, Schneeweiss S. Evaluating the Use of Nonrandomized Real-World Data Analyses for Regulatory Decision Making. Clin Pharmacol Ther. 2019;105(4):867-77.

67. Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. Pharmacoepidemiol Drug Saf. 2017;26(9):1033-9.

68. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000;11(5):550-60.

69. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. BMJ. 2019;367:l5657.

70. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology. 2009;20(4):512-22.

71. Iskander C, Cherney DZ, Clemens KK, Dixon SN, Harel Z, Jeyakumar N, et al. Use of sodium-glucose cotransporter-2 inhibitors and risk of acute kidney injury in older adults with diabetes: a population-based cohort study. CMAJ. 2020;192(14):E351-E60.

72. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. Int J Epidemiol. 2016;45(6):1866-86.

73. Bor J, Moscoe E, Mutevedzi P, Newell ML, Barnighausen T. Regression discontinuity designs in epidemiology: causal inference without randomized trials. Epidemiology. 2014;25(5):729-37.

74. Wang SV, Verpillat P, Rassen JA, Patrick A, Garry EM, Bartels DB. Transparency and Reproducibility of Observational Cohort Studies Using Large Healthcare Databases. Clin Pharmacol Ther. 2016;99(3):325-32.

75. Williams RJ, Tse T, Harlan WR, Zarin DA. Registration of observational studies: is it time? CMAJ. 2010;182(15):1638-42.

76. Goldacre B, Morton CE, DeVito NJ. Why researchers should share their analytic code. BMJ. 2019;367:l6365.

77. Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. BMJ. 2021;372:m4856.

78. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med. 2007;147(8):573-7.

79. Langan SM, Schmidt SA, Wing K, Ehrenstein V, Nicholls SG, Filion KB, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). BMJ. 2018;363:k3532.

80. Wang SV, Schneeweiss S, Berger ML, Brown J, de Vries F, Douglas I, et al. Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0. Pharmacoepidemiol Drug Saf. 2017;26(9):1018-32.