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

Fu, E.L.

Optimal cardiovascular treatment strategies in kidney disease

Causal inference from observational data

Edouard L. Fu

The work described in this thesis was performed at the Department of Clinical Epidemiology, Leiden University Medical Center, The Netherlands, and the Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Sweden.

E. L. Fu was supported by an MD/PhD grant from the Board of Directors of the Leiden University Medical Center. Part of the research described in this thesis was supported by grants from the Dutch Kidney Foundation (18OKK36) and Eurolife.

Cover design and layout	Publiss | www.publiss.nl
Printed by	Ridderprint | www.ridderprint.nl
ISBN:	978-94-6416-724-5

Optimal cardiovascular treatment strategies in kidney disease

Causal inference from observational data

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 28 oktober 2021
klokke 16:15 uur

door

Edouard Liang Fu
geboren te Leiden
in 1997

Promotores
Prof. dr. F.W. Dekker
Prof. dr. J.J. Carrero (Karolinska Institute, Sweden)

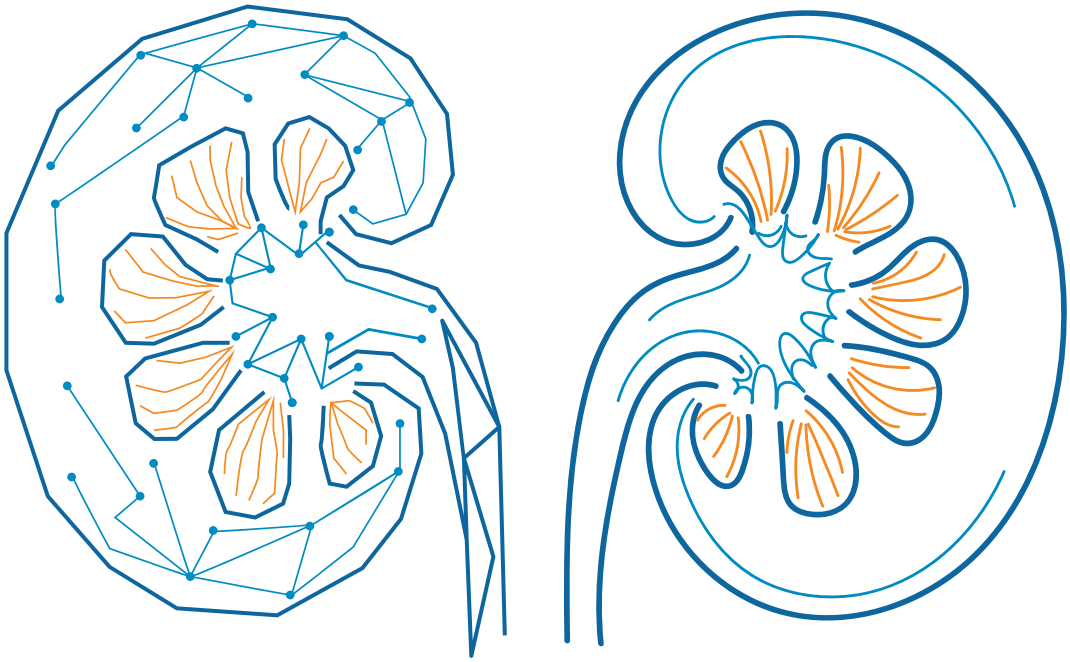
Co-promotor
Dr. M. van Diepen

Leden promotiecommissie
Prof. dr. W.J.W. Bos
Prof. dr. F.R. Rosendaal
Prof. dr. E.W. Steyerberg
Dr. S.A. Swanson (Erasmus Medical Center, The Netherlands)
Dr. L.A. Tomlinson (London School of Hygiene and Tropical Medicine, United Kingdom)

The printing of this thesis was financially supported by Leiden University Medical Center. Financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged.

Voor mijn ouders

TABLE OF CONTENTS



CHAPTER 1

General introduction and outline of the thesis

The kidneys in health and disease

The kidneys play a central role in body homeostasis. They remove waste products and toxins from the body, regulate fluid balance, acid-base homeostasis, bone mineralization and blood pressure and produce hormones such as erythropoietin and renin (1-5). Each kidney contains around 1 million nephrons, the functional units of the kidney, which consist of a glomerulus and a renal tubule. The glomerulus is a network of capillaries (small blood vessels), which are surrounded by Bowman's capsule. Blood is filtered across the capillary walls into Bowman's space, which yields a filtrate of water, small solutes and low-molecular-weight proteins (6). Next, this ultrafiltrate is modified along the tubular segments of the nephron, where reabsorption and secretion take place. Although the kidneys produce only 1.5 L of urine per day, they receive 20% of the cardiac output and form 180 L of ultrafiltrate daily.

Current international guidelines define chronic kidney disease (CKD) as a heterogeneous group of disorders leading to decreased kidney function or abnormalities in kidney structure (proteinuria or abnormalities in urinary sediment, histology or imaging) that are present for at least 3 months (7). The glomerular filtration rate is the recommended form of assessment for overall kidney function. It can be measured via exogenous markers (e.g. using iohexol), but is usually estimated based on creatinine in routine clinical practice (eGFR) (8). Proteinuria is preferably measured by the urinary albumin-to-creatinine ratio. Based on these two measurements, CKD is classified into six stages of eGFR and three stages of proteinuria (**Table 1**) (7). When eGFR is less than 15 ml/min/1.73m^2 (CKD G5), a person has kidney failure and kidney replacement therapy (either dialysis or kidney transplantation) is necessary for survival.

The classification and definition of CKD are based on large-scale epidemiological studies, which have shown strong, graded and consistent associations between eGFR or albuminuria with adverse outcomes, which are independent of age, sex, ethnicity or traditional cardiovascular risk factors (9-17). The lower the kidney function and the higher the albuminuria, the higher the risk for mortality, cardiovascular events, or end stage kidney disease. The association between CKD and cardiovascular abnormalities was first described in 1836 by the British physician Richard Bright (18). Since then, many epidemiological studies have confirmed and extended this finding. For instance, 50% of patients with CKD G4-5 have cardiovascular disease (19, 20). Compared with individuals with a normal kidney function, the risk for cardiovascular mortality is twice as high for individuals with CKD G3, and three times as high for individuals with CKD G4, even after adjustment for traditional cardiovascular risk factors and albuminuria (10, 17). Furthermore, the risks of heart failure, stroke, peripheral arterial disease, coronary artery disease and atrial fibrillation are roughly doubled in patients with an eGFR <60 ml/min/1.73m^2

compared to individuals with a preserved kidney function (21-25). Similarly, the risk of cardiovascular death is doubled in individuals with albuminuria of 300 mg/g, compared with individuals with normal albuminuria (10, 17).

Chronic kidney disease (CKD) is a significant public health problem. Estimates suggest that CKD affects around 850 million individuals (26), corresponding to 8-16% of the worldwide population (27-30). In developed countries the main causes of CKD are hypertension and diabetes, which are increasing in prevalence due to the rise in obesity (31). Like other chronic diseases, the prevalence of CKD increases with age. Among individuals above the age of 60 years the prevalence exceeds 20%, and among those older than 70 years, 35% have CKD (32). CKD was responsible for 1.2 million deaths and 35.8 million disability-adjusted life-years worldwide in 2017 (27, 28), and leads to significant healthcare costs (33, 34). Investigating the effectiveness of therapies to reduce the cardiovascular burden and prevent progression to kidney failure in CKD patients is therefore of high importance.

Table 1: CKD classification based on eGFR and albuminuria.

				Persistent albuminuria categories, description, and ACR range		
				A1	A2	A3
				Normal to mildly increased	Moderately increased	Severely increased
				<30 mg/g	30-300 mg/g	≥300 mg/g
GFR categories, description, and eGFR range (ml/min/1.73m²)	G1	Normal or high	≥90			
	G2	Mildly decreased	60-89			
	G3a	Mildly to moderately decreased	45-59			
	G3b	Moderately to severely decreased	30-44			
	G4	Severely decreased	15-29			
	G5	Kidney failure	<15			

Data from the KDIGO CKD Work Group clinical practice guidelines (7). Colors indicate the prognosis by eGFR and albuminuria category. Green, low risk; Yellow, moderately increased risk; Orange, high risk; Red, very high risk.

Management of chronic kidney disease

Currently, optimal management of CKD patients includes cardiovascular risk reduction (lipid and blood pressure management), drug dosage adjustments and avoidance of nephrotoxic agents, as well as treating complications that may arise due to CKD, such as anemia, electrolyte abnormalities and mineral and bone disorders (35). Renin-angiotensin system inhibitors (RASi; i.e. ACE inhibitors and angiotensin-II receptor blockers) are indicated in patients with diabetes and ACR >30 mg/g and in all CKD patients with ACR >300 mg/g. In patients with diabetes, glycemic control can delay CKD progression and adjustment of oral antihyperglycemic agents may be necessary. Recently, sodium-glucose cotransporter-2 inhibitors have been shown to reduce major adverse cardiovascular events and in particular CKD progression and heart failure in various patient populations, including individuals with diabetes, heart failure or CKD (36-39).

Promises and perils of "real-world evidence" in CKD

To establish the causal effects of treatments and guide decision making, we prefer data from randomized controlled trials (RCT). Due to randomization, treatment allocation depends only on chance, and not on the decision of the clinician or patient. Any differences in outcomes that are larger than we could expect from chance can then be interpreted as a causal effect of the treatment itself.

Unfortunately, evidence from randomized trials is not always available, which has led to uncertainty about the effects of therapies in patients with kidney disease. One of the reasons for this is that fewer trials are conducted in nephrology, and the quality of trials is often lower compared with other specialties (40-42). Second, patients with kidney disease have historically been excluded from many RCTs in the field of cardiovascular disease or cancer (43-45), and continue to be underrepresented in recent cardiovascular trials, especially those with CKD G4-5 (46). Besides the fact that few RCTs are performed, they also have a number of important limitations. RCTs are costly, may be unethical to conduct, and can take a long time to complete (47). Furthermore, they are often conducted in selected populations, whereas a much wider population is likely to receive these treatments in clinical practice (48, 49). For instance, a meta-analysis of 186 trials conducted in the dialysis population showed that trial participants are younger, have a different pattern of comorbidities and a lower mortality rate than the overall population of dialysis patients (48). Findings from RCTs can therefore not be readily translated to effectiveness and safety profiles of patients from routine clinical practice. Lastly, the number of clinical questions greatly outpaces the number of RCTs that can be performed.

For these reasons, there is growing interest from regulators, insurance companies and physicians in using observational data to supplement trial evidence and aid in clinical decision-making (50, 51). Vast amounts of observational data are generated each day as patients interact with the healthcare system and are stored in the form of administrative claims databases, electronic health records and registries (52). These data sources can be used relatively fast and at a fraction of the cost of RCTs to study how treatments work in routine clinical care. Indeed, many observational studies are published every year and there has been a trend towards greater reliance by regulators on such "real-world evidence" (RWE), which is defined as evidence coming from data sources other than the traditional highly controlled RCTs (53, 54). In 2019, 60% of new drug approvals to the US Food and Drug Administration (FDA) had observational evidence as part of their submission, of which 12% was considered essential evidence (55).

However, drawing causal conclusions from routinely collected healthcare data is not an easy task. Because treatment has not been randomized, there can be systematic differences between the treatment groups, known in the epidemiology literature as confounding (56). After all, in observational data we just observe what happens in every day care, where treatments are often given to individuals with a worse prognosis, or in the case of preventive treatments to health-seeking individuals. In addition to this confounding, selection bias (57-59) and information bias (60, 61) can also affect the validity of observational studies. Because data are widely available while the knowledge to properly analyse these is not, observational studies can easily yield associations that are precise (due to the large sample size) but not causal. A notable example is the debacle surrounding postmenopausal hormone therapy and coronary heart disease (62). Observational studies had shown that women using postmenopausal hormone replacement therapy had a markedly lower risk of coronary heart disease compared with women who did not use these medications (63, 64). As a consequence, this therapy was widely promoted in postmenopausal women to prevent heart disease (65, 66). However, a subsequent large randomized trial showed completely opposite results: a 24% increased risk of coronary heart disease (67). This example illustrates that making mistakes in observational data is not without costs and can have grave consequences (62).

Fortunately, in the past four decades there have been tremendous developments in the field of causal inference, the discipline that investigates the assumptions, study designs and methods that are needed to draw causal conclusions based on data (68). A theoretical framework for causality has been developed such that we are able to ask better causal questions, and novel methods increase the validity of findings from observational data (47, 69, 70). Our understanding of how the various biases arise and how they can be addressed has vastly improved. At the same time, databases are becoming larger and more detailed, including longitudinal patient-level information on medications, lab tests, procedures and diagnoses. As the flourishing of the field of causal inference coincides with the upcoming of big data, this combination provides immense opportunities to contribute valuable evidence for decision making in nephrology.

Aims and outline

The aim of this thesis is to answer a number of clinical questions on the effectiveness and safety of treatments in the field of kidney disease by using observational data from routine clinical care in combination with causal inference methods. This thesis is structured in two parts. In **part I** we explore the methodological aspects of using routinely collected observational data to establish causal effects. In **part II** we apply state-of-the-art statistical and epidemiological methods to answer a number of pressing clinical questions in nephrology.

Part I: Methodological considerations

A number of biases threaten the validity of observational studies that try to estimate the causal effects of treatments. Some of these biases can be prevented by applying a sound study design, whereas others need to be addressed in the statistical analysis. In **Chapter 2** we specifically discuss confounding, prevalent user bias, immortal time bias, missing data and measurement error and methods to handle these biases, including active comparator new user designs, target trial emulation, propensity scores, marginal structural models and multiple imputation.

Different methods exist to adjust for measured confounding in the statistical analysis. **Chapter 3** discusses the merits and caveats of propensity scores, a commonly used method to adjust for confounding. We discuss four types of propensity score methods, including propensity score matching, stratification, adjustment and weighting, and illustrate these with a clinical example. We also provide guidance when to choose propensity score methods versus conventional multivariable regression. In **Chapter 4** we point out immortal time bias in a published observational study which aimed to estimate the causal effect of metformin on kidney outcomes, and discuss how this could have been prevented with an appropriate study design or statistical analysis.

Part II: Clinical applications

RASi (i.e., ACE inhibitors and angiotensin-II receptor blockers) are widely prescribed drugs that are a cornerstone in the treatment of hypertension, heart failure and proteinuric kidney disease. Acute increases in creatinine (reflecting a drop in kidney function) are often observed after initiation of RASi, but the clinical significance of such increases is controversial. **Chapter 5** investigates the significance of this acute creatinine increase by studying the association between the magnitude of increase and the outcomes mortality, cardiovascular events and end-stage kidney disease using standard multivariable regression to adjust for confounding. Since guidelines recommend to monitor creatinine, we also investigate which proportion of patients receive this monitoring.

We do not know the best antihypertensive medications to use in patients with advanced CKD (stage 4/5), because they have not been included in many trials. In **Chapter 6** we compare the effects of RASi versus calcium channel blockers, two of the most widely used antihypertensive drugs in this population. An active comparator new user design in combination with propensity score weighting is applied to reduce confounding and other biases. Since ample evidence is available on the efficacy of RASi in moderate CKD (stage 3) from randomized trials, we replicate our analyses in a positive control cohort of patients with stage 3 CKD to compare the trial findings against our observational estimates.

Between 10-15% of patients with heart failure have advanced CKD. Although β-blockers are a cornerstone in the treatment of patients with heart failure with reduced ejection fraction, we do not know whether they are effective in individuals with advanced CKD, a population at an extremely high risk of complications and (cardiovascular) death. **Chapter 7** investigates the effect of β-blocker use in patients with heart failure and advanced CKD on cardiovascular and all-cause mortality and heart failure hospitalization. Results are reported for reduced, midrange and preserved ejection fraction. Similar to the previous chapter, a positive control cohort of heart failure patients with moderate CKD is used to benchmark the observational findings against trial evidence.

Small single-center studies have suggested that stopping RASi in patients with advanced CKD can postpone dialysis. However, few studies have addressed the cardiovascular and kidney effects of this decision. In **Chapter 8** we therefore investigate the effect of stopping versus continuing RASi in patients with advanced CKD on mortality, cardiovascular events and kidney replacement therapy using the target trial emulation framework and the cloning, censoring and weighting method. In addition, results are replicated by modelling RASi as a time-varying exposure using a marginal structural model.

Chapter 9 investigates the optimal kidney function to initiate dialysis with respect to mortality and cardiovascular outcomes using the target trial emulation framework. Although the randomized IDEAL trial showed no differences in outcomes between early and late dialysis initiation, only two treatment strategies were compared, and the realized eGFR values were within a narrow range (7.2 vs. 9.0 ml/min/1.73m^2). As a randomized trial testing many different treatment strategies is unfeasible, observational data need to be used to answer this question. Since this question involves dynamic treatment strategies a cloning, censoring and weighting approach is used. We also investigate the influence of lead time bias, selection bias and immortal time bias in previous observational studies. The last two chapters calculate restricted mean survival time to easily interpret how long an average patient would live by following a particular treatment strategy.

References

1. Nigam SK, Wu W, Bush KT, Hoenig MP, Blantz RC, Bhatnagar V. Handling of Drugs, Metabolites, and Uremic Toxins by Kidney Proximal Tubule Drug Transporters. Clin J Am Soc Nephrol. 2015;10(11):2039-49.
2. Danziger J, Zeidel ML. Osmotic homeostasis. Clin J Am Soc Nephrol. 2015;10(5):852-62.
3. Hamm LL, Nakhoul N, Hering-Smith KS. Acid-Base Homeostasis. Clin J Am Soc Nephrol. 2015;10(12):2232-42.
4. Blaine J, Chonchol M, Levi M. Renal control of calcium, phosphate, and magnesium homeostasis. Clin J Am Soc Nephrol. 2015;10(7):1257-72.
5. Hoenig MP, Zeidel ML. Homeostasis, the milieu interieur, and the wisdom of the nephron. Clin J Am Soc Nephrol. 2014;9(7):1272-81.
6. Pollak MR, Quaggin SE, Hoenig MP, Dworkin LD. The glomerulus: the sphere of influence. Clin J Am Soc Nephrol. 2014;9(8):1461-9.
7. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney inter, Suppl. 2013;3(3):1–150.
8. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12.
9. Astor BC, Matsushita K, Gansevoort RT, van der Velde M, Woodward M, Levey AS, et al. Lower estimated glomerular filtration rate and higher albuminuria are associated with mortality and end-stage renal disease. A collaborative meta-analysis of kidney disease population cohorts. Kidney Int. 2011;79(12):1331-40.
10. Chronic Kidney Disease Prognosis C, Matsushita K, van der Velde M, Astor BC, Woodward M, Levey AS, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet. 2010;375(9731):2073-81.
11. Fox CS, Matsushita K, Woodward M, Bilo HJ, Chalmers J, Heerspink HJ, et al. Associations of kidney disease measures with mortality and end-stage renal disease in individuals with and without diabetes: a meta-analysis. Lancet. 2012;380(9854):1662-73.
12. Gansevoort RT, Matsushita K, van der Velde M, Astor BC, Woodward M, Levey AS, et al. Lower estimated GFR and higher albuminuria are associated with adverse kidney outcomes. A collaborative meta-analysis of general and high-risk population cohorts. Kidney Int. 2011;80(1):93-104.
13. Hallan SI, Matsushita K, Sang Y, Mahmoodi BK, Black C, Ishani A, et al. Age and association of kidney measures with mortality and end-stage renal disease. JAMA. 2012;308(22):2349-60.
14. Mahmoodi BK, Matsushita K, Woodward M, Blankestijn PJ, Cirillo M, Ohkubo T, et al. Associations of kidney disease measures with mortality and end-stage renal disease in individuals with and without hypertension: a meta-analysis. Lancet. 2012;380(9854):1649-61.
15. Nitsch D, Grams M, Sang Y, Black C, Cirillo M, Djurdjev O, et al. Associations of estimated glomerular filtration rate and albuminuria with mortality and renal failure by sex: a meta-analysis. BMJ. 2013;346:f324.
16. Wen CP, Matsushita K, Coresh J, Iseki K, Islam M, Katz R, et al. Relative risks of chronic kidney disease for mortality and end-stage renal disease across races are similar. Kidney Int. 2014;86(4):819-27.

17. van der Velde M, Matsushita K, Coresh J, Astor BC, Woodward M, Levey A, et al. Lower estimated glomerular filtration rate and higher albuminuria are associated with all-cause and cardiovascular mortality. A collaborative meta-analysis of high-risk population cohorts. Kidney Int. 2011;79(12):1341-52.

18. Bright R. Cases and Observations Illustrative of Renal Disease, Accompanied with the Secretion of Albuminous Urine. Med Chir Rev. 1836;25(49):23-35.

19. Stevens PE, O'Donoghue DJ, de Lusignan S, Van Vlymen J, Klebe B, Middleton R, et al. Chronic kidney disease management in the United Kingdom: NEOERICA project results. Kidney Int. 2007;72(1):92-9.

20. Jankowski J, Floege J, Fliser D, Bohm M, Marx N. Cardiovascular Disease in Chronic Kidney Disease: Pathophysiological Insights and Therapeutic Options. Circulation. 2021;143(11):1157-72.

21. Kottgen A, Russell SD, Loehr LR, Crainiceanu CM, Rosamond WD, Chang PP, et al. Reduced kidney function as a risk factor for incident heart failure: the atherosclerosis risk in communities (ARIC) study. J Am Soc Nephrol. 2007;18(4):1307-15.

22. Abramson JL, Jurkovitz CT, Vaccarino V, Weintraub WS, McClellan W. Chronic kidney disease, anemia, and incident stroke in a middle-aged, community-based population: the ARIC Study. Kidney Int. 2003;64(2):610-5.

23. Wattanakit K, Folsom AR, Selvin E, Coresh J, Hirsch AT, Weatherley BD. Kidney function and risk of peripheral arterial disease: results from the Atherosclerosis Risk in Communities (ARIC) Study. J Am Soc Nephrol. 2007;18(2):629-36.

24. Astor BC, Coresh J, Heiss G, Pettitt D, Sarnak MJ. Kidney function and anemia as risk factors for coronary heart disease and mortality: the Atherosclerosis Risk in Communities (ARIC) Study. Am Heart J. 2006;151(2):492-500.

25. Alonso A, Lopez FL, Matsushita K, Loehr LR, Agarwal SK, Chen LY, et al. Chronic kidney disease is associated with the incidence of atrial fibrillation: the Atherosclerosis Risk in Communities (ARIC) study. Circulation. 2011;123(25):2946-53.

26. Jager KJ, Kovesdy C, Langham R, Rosenberg M, Jha V, Zoccali C. A single number for advocacy and communication-worldwide more than 850 million individuals have kidney diseases. Nephrol Dial Transplant. 2019;34(11):1803-5.

27. Collaboration GBDCKD. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2020;395(10225):709-33.

28. Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, et al. Chronic kidney disease: global dimension and perspectives. Lancet. 2013;382(9888):260-72.

29. Jager KJ, Fraser SDS. The ascending rank of chronic kidney disease in the global burden of disease study. Nephrol Dial Transplant. 2017;32(suppl_2):ii121-ii8.

30. Levin A, Tonelli M, Bonventre J, Coresh J, Donner JA, Fogo AB, et al. Global kidney health 2017 and beyond: a roadmap for closing gaps in care, research, and policy. Lancet. 2017;390(10105):1888-917.

31. Webster AC, Nagler EV, Morton RL, Masson P. Chronic Kidney Disease. Lancet. 2017;389(10075):1238-52.

32. Coresh J, Selvin E, Stevens LA, Manzi J, Kusek JW, Eggers P, et al. Prevalence of chronic kidney disease in the United States. JAMA. 2007;298(17):2038-47.

33. Kerr M, Bray B, Medcalf J, O'Donoghue DJ, Matthews B. Estimating the financial cost of chronic kidney disease to the NHS in England. Nephrol Dial Transplant. 2012;27 Suppl 3:iii73-80.
34. United States Renal Data System. 2020 USRDS Annual Data Report: Epidemiology of kidney disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2020.
35. Chen TK, Knicely DH, Grams ME. Chronic Kidney Disease Diagnosis and Management: A Review. JAMA. 2019;322(13):1294-304.
36. de Boer IH, Caramori ML, Chan JCN, Heerspink HJL, Hurst C, Khunti K, et al. Executive summary of the 2020 KDIGO Diabetes Management in CKD Guideline: evidence-based advances in monitoring and treatment. Kidney Int. 2020;98(4):839-48.
37. 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.
38. Zannad F, Ferreira JP, Pocock SJ, Anker SD, Butler J, Filippatos G, et al. SGLT2 inhibitors in patients with heart failure with reduced ejection fraction: a meta-analysis of the EMPEROR-Reduced and DAPA-HF trials. Lancet. 2020;396(10254):819-29.
39. 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.
40. 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.
41. Strippoli GF, Craig JC, Schena FP. The number, quality, and coverage of randomized controlled trials in nephrology. J Am Soc Nephrol. 2004;15(2):411-9.
42. Palmer SC, Sciancalepore M, Strippoli GF. Trial quality in nephrology: how are we measuring up? Am J Kidney Dis. 2011;58(3):335-7.
43. 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.
44. Charytan D, Kuntz RE. The exclusion of patients with chronic kidney disease from clinical trials in coronary artery disease. Kidney Int. 2006;70(11):2021-30.
45. Coca SG, Krumholz HM, Garg AX, Parikh CR. Underrepresentation of renal disease in randomized controlled trials of cardiovascular disease. JAMA. 2006;296(11):1377-84.
46. 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.
47. Vandenbroucke JP. When are observational studies as credible as randomised trials? Lancet. 2004;363(9422):1728-31.
48. Smyth B, Haber A, Trongtrakul K, Hawley C, Perkovic V, Woodward M, et al. Representativeness of Randomized Clinical Trial Cohorts in End-stage Kidney Disease: A Meta-analysis. JAMA Intern Med. 2019.
49. 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.
50. 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.
51. 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.

52. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol. 2005;58(4):323-37.

53. Bonamici S. H.R.34 - 114th Congress (2015-2016): 21st Century Cures Act. 2016. Available from: https://www.congress.gov/bill/114th-congress/house-bill/34. Accessed March 24, 2021.

54. U.S. Food and Drug Administration. Prescription Drug User Fee Act (PDUFA) - PDUFA VI: Fiscal Years 2018 - 2022. Available from: https://www.fda.gov/forindustry/userfees/prescriptiondruguserfee/ucm446608.htm. Accessed March 24, 2021. .

55. Aetion. What role does RWE play in FDA approvals? 2020. Available at: https://www.aetion.com/evidence-hub/infographic-what-role-does-rwe-play-in-fda-approvals. Accessed March 24, 2021.

56. Kyriacou DN, Lewis RJ. Confounding by Indication in Clinical Research. JAMA. 2016;316(17):1818-9.

57. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004;15(5):615-25.

58. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. Int J Epidemiol. 2010;39(2):417-20.

59. Elwert F, Winship C. Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. Annu Rev Sociol. 2014;40:31-53.

60. Funk MJ, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. Curr Epidemiol Rep. 2014;1(4):175-85.

61. Keogh RH, Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1-Basic theory and simple methods of adjustment. Stat Med. 2020;39(16):2197-231.

62. 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.

63. 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.

64. 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.

65. Hersh AL, Stefanick ML, Stafford RS. National use of postmenopausal hormone therapy: annual trends and response to recent evidence. JAMA. 2004;291(1):47-53.

66. Connelly MT, Richardson M, Platt R. Prevalence and duration of postmenopausal hormone replacement therapy use in a managed care organization, 1990-1995. J Gen Intern Med. 2000;15(8):542-50.

67. 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.

68. Hill J, Stuart EA. Causal Inference: Overview. In: Wright JD. International Encyclopedia of the Social & Behavioral Sciences. 2nd ed. Elsevier; 2015. p. 255-260.

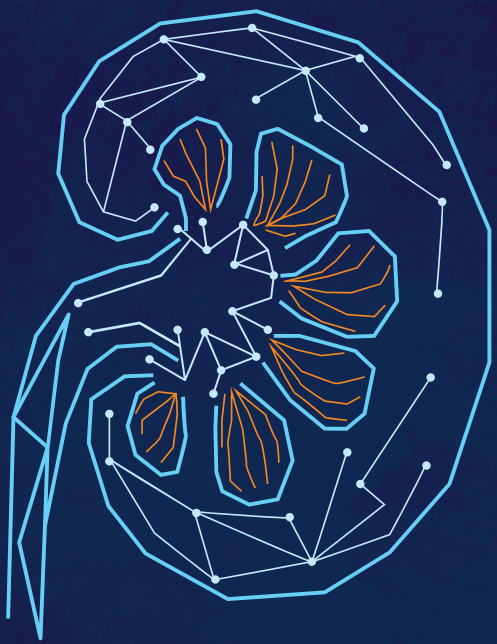
69. Imbens G, Rubin D (2015). Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge: Cambridge University Press.

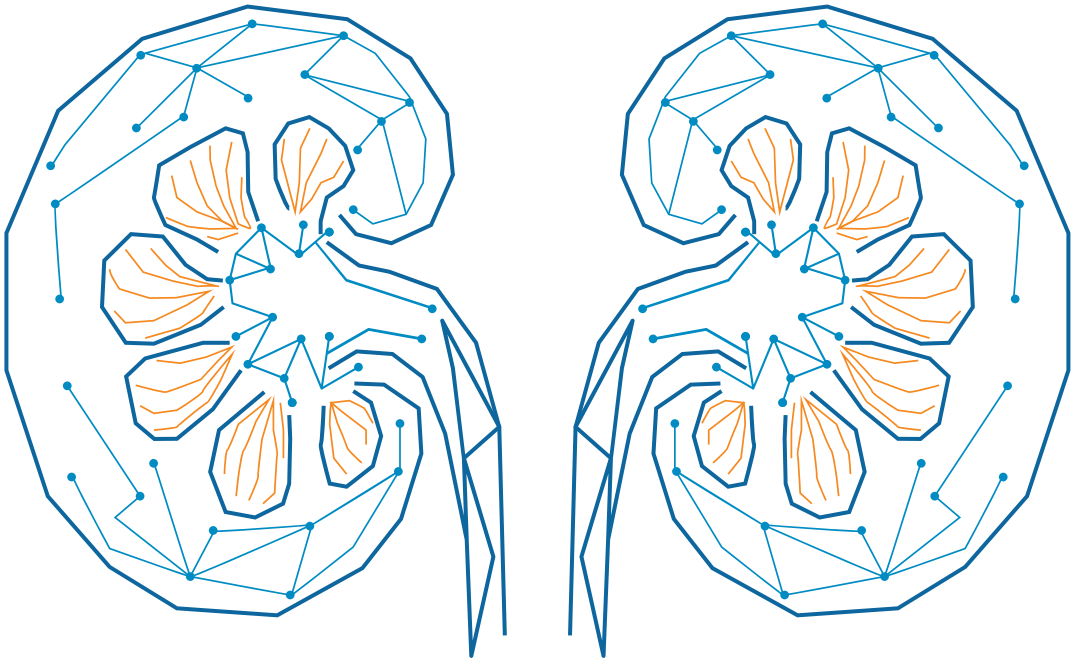
70. Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC.

1

PART I

Methodological considerations for causal inference from observational data





CHAPTER 2

Pharmacoepidemiology for nephrologists: potential biases and how to overcome them

Edouard L. Fu, Merel van Diepen, Yang Xu, Marco Trevisan, Friedo W. Dekker, Carmine Zoccali, Kitty J. Jager, Juan-Jesus Carrero

Clin Kidney J 2020; 14: 1317-1326

Abstract

Observational pharmacoepidemiological studies using routinely collected healthcare data are increasingly being used in the field of nephrology to answer questions on the effectiveness and safety of medications. This review discusses a number of biases that may arise in such studies and proposes solutions to minimize them during the design or statistical analysis phase. We first describe designs to handle confounding by indication (e.g. active comparator design) and methods to investigate the influence of unmeasured confounding, such as the E-value, the use of negative control outcomes and control cohorts. We next discuss prevalent user and immortal time biases in pharmacoepidemiology research, and how these can be prevented by focussing on incident users and applying either landmarking, using a time-varying exposure or the cloning, censoring and weighting method. Lastly, we briefly discuss the common issues with missing data and misclassification bias. When these biases are properly accounted for, pharmacoepidemiological observational studies can provide valuable information for clinical practice.

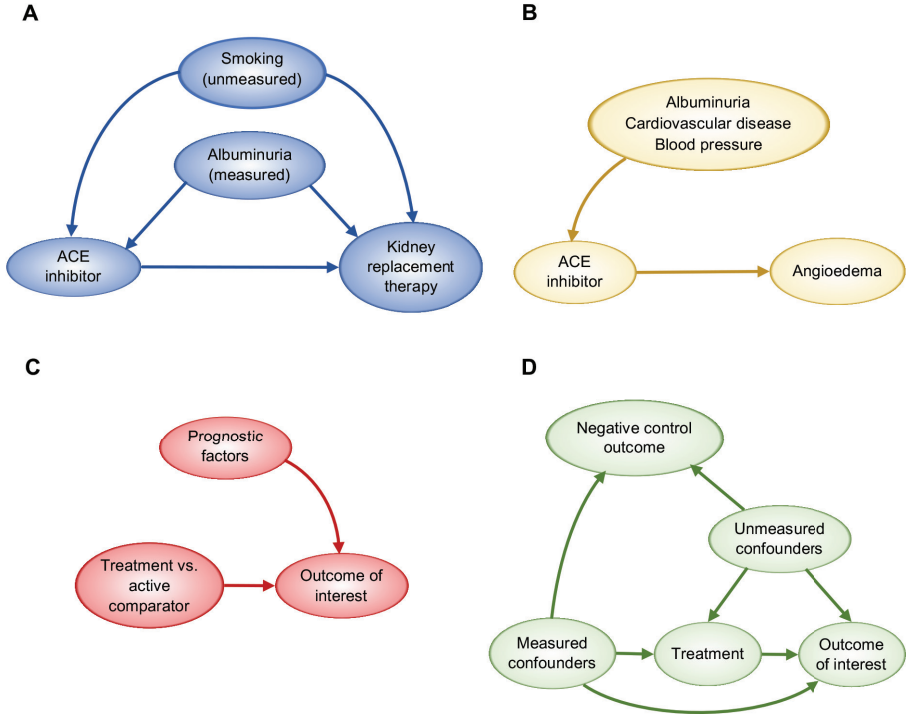
Introduction

Pharmacoepidemiology uses epidemiological methods to study the use, therapeutic effects and risks of medications in large populations (1). Due to the availability of routinely collected healthcare data from registries, electronic health records or claims databases, observational pharmacoepidemiological studies are increasingly being used to generate evidence to inform clinical practice. Our first review discussed the scope and research questions that are studied within the field of pharmacoepidemiology, and described the strengths and caveats of the most commonly used study designs to answer such questions (2). We now focus on the most common biases that may occur when using observational data to study the causal effects of medication on health outcomes. We will attempt to offer possible solutions in the design or statistical analysis to prevent or minimize such biases. The review is intended as an introduction to the field for those who wish to critically appraise pharmacoepidemiological studies or conduct such studies.

Confounding by indication

Confounding by indication is a threat to any observational study assessing the effects of medications, since treatment is not randomly assigned to patients. Confounding by indication arises when the indications for treatment, such as age and comorbidities, are also related to the outcome under study (3). For example, treatment is generally given to those with a worse prognosis. The reverse may also occur, when newly introduced drugs are first prescribed to individuals perceived as healthier and who may be more likely to tolerate them (4). Both situations lead to an uneven distribution of prognostic factors between treatment groups, which biases a direct comparison. Take the case shown in **Figure 1A**, where albuminuria is a confounder for the effect of ACE inhibitor treatment on kidney replacement therapy: individuals with albuminuria are more likely to be prescribed ACE inhibitors and albuminuria is also an independent risk factor for kidney replacement therapy. Unknown or unmeasured confounders for the treatment-outcome relationship may also be present, such as smoking (if this variable has not been measured and we would assume that smoking can affect the decision to start ACE inhibitor therapy as well as the outcome).

Figure 1. (A) Confounding by indication arises when prognostic factors for the outcome also influence the decision to start treatment. Some confounders may be measured, which can be adjusted for in the analysis, whereas others are unmeasured, leading to residual confounding. (B) When unintended outcomes are studied, less confounding by indication will be present. The indications for ACEi treatment likely do not increase the risk for the outcome angioedema. (C) An ideal active comparator has similar indications as the medication under study, thereby decreasing confounding by indication. Ideally, the active comparator should have no influence on the outcome. (D) Negative control outcomes need to have similar measured and unmeasured confounders as the treatment-outcome relationship under study. Furthermore, treatment should not have an influence on the negative control outcome.



Addressing confounding by indication when designing a pharmacoepidemiological study

When designing observational studies to investigate medication effectiveness or safety, we should be aware that some research questions will be more susceptible to confounding than others. Confounding will generally play a larger role when studying the beneficial or "intended" effects of treatments, since the indications for treatment are very likely to be related to the prognosis of the patient (5). On the other hand, if the outcome is completely unrelated to the indications for treatment, such as when studying rare side effects or "unintended" effects, no confounding would be present (6). A classic example is the relationship between ACE inhibitors and angioedema. Patient characteristics that determine treatment status (e.g. cardiovascular risk, albuminuria, blood pressure) are unlikely to be associated with the outcome angioedema. Consequently, the arrow from treatment indication to outcome will be absent, and confounding by indication will not be an issue (**Figure 1B**).

Applying an active comparator design may also decrease confounding by indication (7). In an active comparator design, the medication of interest is compared with another drug that has similar indications, instead of a non-user group. The more exchangeable the active comparator is for the medication of interest, the lower the risk for potential confounding will be. After all, if both treatment groups would have identical treatment indications (both measured and unmeasured characteristics), there would be no arrow from indication to treatment and confounding by indication is removed (**Figure 1C**) (8). A recent example applying an active comparator design investigated whether proton pump inhibitors (PPI) increased the risk of chronic kidney disease (CKD) (9). Comparing PPI users with non-users may suffer from unmeasured confounding, since non-users are generally healthier and not all confounders may have been captured in the dataset and adjusted for. Histamine-2 receptor (H2) antagonists are prescribed for similar indications as PPI. Users of these two medications may be more similar regarding comorbidities, medication use and other unmeasured variables.

Adjusting for confounding during the statistical analysis

Selecting an appropriate set of confounders to adjust for is critical when conducting pharmacoepidemiological studies (10). In general, it is not recommended to use data-driven variable selection methods to identify confounders, such as only retaining statistically significant confounders or including variables that change the regression coefficient of the treatment variable (11-13). Such data-driven approaches can lead to bias if they adjust for mediators (i.e. variables in the causal pathway between exposure and outcome) (11), colliders (e.g. a variable caused by both treatment and outcome) (14, 15) and instrumental variables (i.e. variables strongly

related to the exposure but not to the outcome) (16), although some sophisticated statistical covariate selection methods are currently under development (10, 17-19). Instead, we suggest to make directed acyclic graphs (DAGs) and select confounders based on subject-matter knowledge and biological plausibility (20, 21). Since DAGs rely on prior knowledge and assumed causal effects, they do not tell whether these assumptions are correct. Different researchers can have different views which factor causes the other and this may result in different choices regarding which confounders to adjust for. DAGs can aid in this discussion by making the causal assumptions explicit in a graphical manner.

Once the confounders have been selected, various methods can be used to adjust for confounding. These include, for instance, multivariable regression, standardization and propensity score methods (propensity score matching, weighting, stratification or adjustment). In the time-fixed setting (i.e. when treatment is only measured once), all methods generally suffice to adjust for measured confounding, although the interpretation of the effect estimate may differ depending on the method and some methods are preferred in specific settings (22-25). A thorough discussion on the merits and caveats of multivariable regression and propensity score methods can be found elsewhere (26). Nonetheless, in the setting of time-varying treatments (i.e. when treatment is received at multiple timepoints and changes over time) and time-varying confounding, methods based on weighting or standardization are required to give unbiased estimates if the confounders themselves are affected by treatment (27). It should be kept in mind that all statistical methods mentioned above can only adjust for measured confounders, but not for unmeasured confounders as is sometimes claimed (28), unless the unmeasured confounders are correlated with the variables that are adjusted for (29, 30).

Assessing the impact of unmeasured confounding

Although the possibility of residual or unmeasured confounding in observational analyses can never be fully eliminated, a number of steps can be taken to alleviate concerns and strengthen inferences. In this section we will elaborate on conducting sensitivity analyses to obtain corrected effect estimates, calculating the E-value, and conducting negative control outcome and control group analyses.

First and foremost, as many confounders as possible need to be identified and adjusted for by using appropriate statistical methods. However, if known confounders (e.g. albuminuria or smoking) have not been measured, corrected effect estimates can be calculated in quantitative bias analyses (31-34). This requires as input the assumed association between confounder and exposure and between confounder and outcome, and the prevalence of the confounder in the population. These numbers can be based on previous studies and can be varied over a range of values

to give an indication how sensitive the estimated treatment effect is to unmeasured confounding (35). If results lead to the same conclusion over a wide range of relevant scenarios, then the plausibility of the estimated treatment effect will increase.

Alternatively, one can estimate how strong the unmeasured confounding would need to be to completely explain away a certain effect estimate. The E-value has recently been introduced as an easily implemented tool for these purposes (36-39). The E-value is defined as "*the minimum strength of association … that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association, conditional on the measured covariates*" (36). As an example, researchers investigated whether sodium polystyrene sulfonate (SPS) to treat hyperkalemia, increased the risk of severe adverse gastrointestinal events in persons with CKD (40). After adjustment for measured confounding, the initiation of SPS was associated with a 1.25 (95% confidence interval 1.05-1.49) higher risk of severe gastrointestinal events. The corresponding E-value for this risk ratio was 1.80, meaning that an unmeasured confounder would need to be associated with both SPS initiation and severe gastrointestinal events by a hazard ratio of 1.80 to decrease the point estimate from 1.25 to 1.00. What constitutes a large E-value is context-specific and depends on the specific research question under study, the effect size of the exposure and the hazard ratios of the confounders that have already been adjusted for (38, 41, 42). Easily implemented online calculators are available to conduct the discussed sensitivity analyses (37, 43, 44).

For certain research questions *negative control outcomes* can be used to provide guidance about the presence and magnitude of unmeasured confounding in observational studies (45). A negative control outcome is an outcome that is not influenced by the treatment of interest but shares the same set of measured and unmeasured confounders as the treatment of interest-outcome relationship (**Figure 1D**) (46). Hence, we would not expect to find an association between the exposure of interest and the negative control outcome. As an example, one may be concerned that the unmeasured variables BMI and smoking bias the results of a study investigating the association between a cardiovascular drug and the risk of cardiovascular-related mortality. However, we would not expect the cardiovascular drug to also lower non-cardiovascular mortality. If we would unexpectedly find a lower risk of the negative control outcome non-cardiovascular mortality among treated individuals this may be an indication of residual confounding or other sources of bias. In addition to the previously mentioned assumptions, the negative control outcome should occur with a frequency similar to the primary study outcome to ensure enough power to reject the null hypothesis of no association. If such assumptions are not met, this may erroneously lead to the conclusion that no unmeasured confounding is present (47).

Similarly, one can test whether associations are as expected in a certain *control group*. The direction of the expected association (either a positive, negative or null association) can be based on physiologic mechanisms or evidence from randomized trials (48). For example, Weir *et al*. hypothesized that users of high-dialyzable β-blocker would have an increased risk of mortality compared with users of low-dialyzable β-blocker, due to loss of high-dialyzable β-blocker in the dialysate (49). To strengthen their inferences, a control group of patients with CKD G4-5 was constructed in whom a similar effectiveness of high-dialyzable and low-dialyzable beta-blockers was expected and subsequently demonstrated. Control groups can also strengthen inferences by showing similar results between observational studies and randomized trials. We recently evaluated the effectiveness of beta-blockers in patients with heart failure and advanced CKD, a population which was excluded from landmark heart failure trials (50). A positive control group including heart failure patients with moderate CKD showed a benefit similar to that observed in moderate CKD patients from randomized trials. This positive control analysis further supported a causal explanation for the results in the advanced CKD cohort.

Prevalent user and immortal time biases

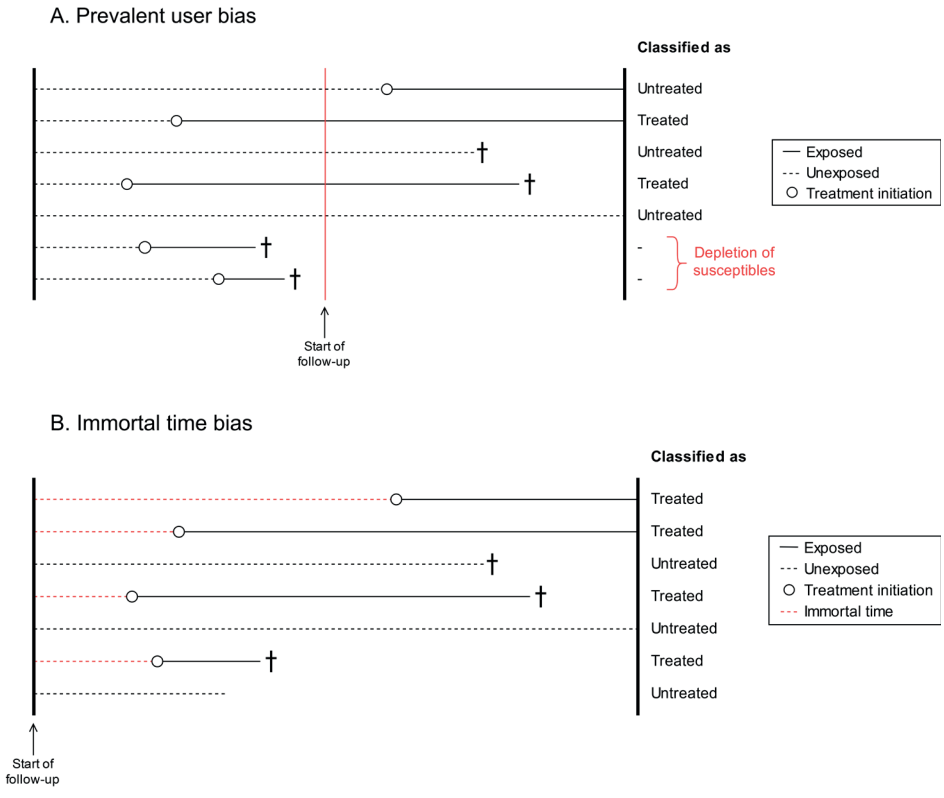
We now discuss two types of biases which often occur in pharmacoepidemiological studies, but that can and should be avoided by adhering to a simple principle: aligning the start of the follow-up with the start of exposure.

Prevalent user bias

When we want to assess the effectiveness of initiating a drug, it is recommended to include incident medication users instead of prevalent users (51, 52). In a new user or incident user design, only individuals who *initiate* the medication of interest are studied and followed from the date of treatment initiation. Therefore, the start of follow-up and start of treatment will align, and all events that occur after drug initiation are captured (53, 54). In contrast, prevalent user designs include individuals who initiated the exposure of interest some time *before* the start of follow-up (**Figure 2A**). Comparing prevalent users to non-users may introduce selection bias since individuals who died before enrolment cannot, per definition, be included in the analysis, and events occurring shortly after drug initiation are neither observed (55, 56). To better understand why this selection bias arises, we give a real-world example. Suppose we conducted a randomized trial and found that a certain medication increased the risk of myocardial infarction with a hazard ratio of 1.24. We now reanalyze the data by starting follow-up at two years after randomization. Hence, we only count the myocardial infarctions that occurred after two years of follow-up. By doing so, we also exclude all individuals who died

or experienced myocardial infarction in the first two years after randomization. This new analysis paradoxically (rather erroneously) shows that the medication *lowers* the risk of myocardial infarction. Since the medication increases the risk of myocardial infarction, the treatment arm will be progressively depleted of patients most susceptible to the event (57). After two years the treated group will only consist of survivors who likely do not have other risk factors for myocardial infarction. Therefore, comparing these survivors in the treatment group with those remaining in the control group leads to an unfair advantage for the treatment group.

Figure 2. Graphical visualization of prevalent user bias (A) and immortal time bias (B) when setting up the start of follow-up in a study. For prevalent user bias, start of follow-up occurs *after* treatment initiation whereas for immortal time bias, the start of follow-up occurs *before* treatment initiation. These biases can be prevented by aligning the start of follow-up with the start of exposure.



Prevalent user bias is one of the proposed reasons why postmenopausal hormone therapy appeared protective for coronary heart disease in observational studies, but was actually harmful when subsequent randomized trials were conducted (58, 59). Besides the fact that the effect estimates from a prevalent user design are biased, they also do not inform decision making as the decision to start the treatment was already made in the past. Studies applying a prevalent user design do not answer the relevant question whether treatment should be initiated. The results of such a study can only tell you that if a person has survived on treatment for this long, we know he is not susceptible to the event, which gives him a better prognosis than untreated individuals who are still susceptible.

Immortal time bias

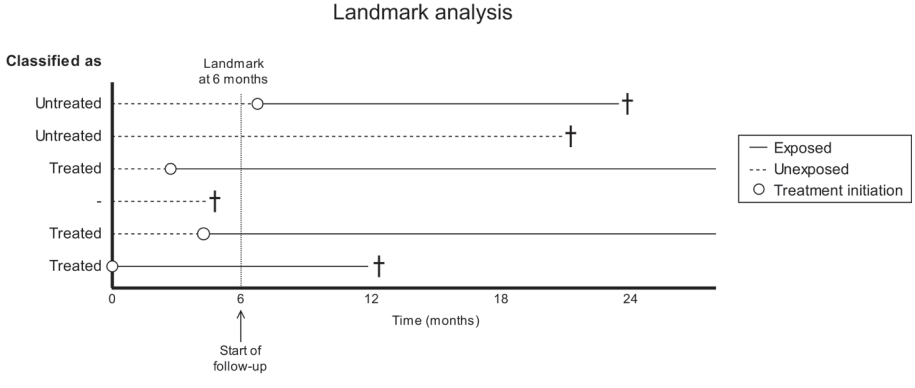
Immortal time bias occurs when patients are classified into treatment groups at baseline based on the treatment they take *after* baseline (**Figure 2B**) (60, 61). This leads to a period of time (i.e. immortal time) between baseline and start of treatment where no deaths can occur in the treatment group, thereby biasing results in favor of the treatment group. As an example, a pharmacoepidemiological study investigated the long-term effects of metformin use versus no metformin use on mortality and end-stage kidney disease (62). In this study, follow-up started when patients had a first creatinine measurement, but patients were classified as metformin users when they were prescribed metformin for more than 90 days *during* the follow-up period. Using post-baseline information on metformin use to classify patients at baseline into the metformin group leads to an unfair survival advantage for metformin users (63). Imagine that all individuals in the metformin group started medication only after 5 years of follow-up. By definition, no deaths would then occur in the metformin group during the first 5 years of follow-up. After all, individuals who have an event prior to taking up treatment would be classified as untreated. Using post-baseline information for exposure classification thus results in immortal time bias (60, 64). To what extent the effect estimate is biased depends on the total amount of follow-up that is erroneously misclassified under the metformin group. The bias will increase with a larger proportion of exposed study participants, a larger amount of time between start of follow-up and initiation of treatment and a longer duration of follow-up (65).

Potential solutions to mitigate immortal time bias

We now discuss three designs that could be applied to avoid immortal time bias: landmarking, using a time-varying exposure and using treatment strategies with grace periods. Other more complex solutions exist but are outside the scope of this review and are discussed elsewhere (66-68).

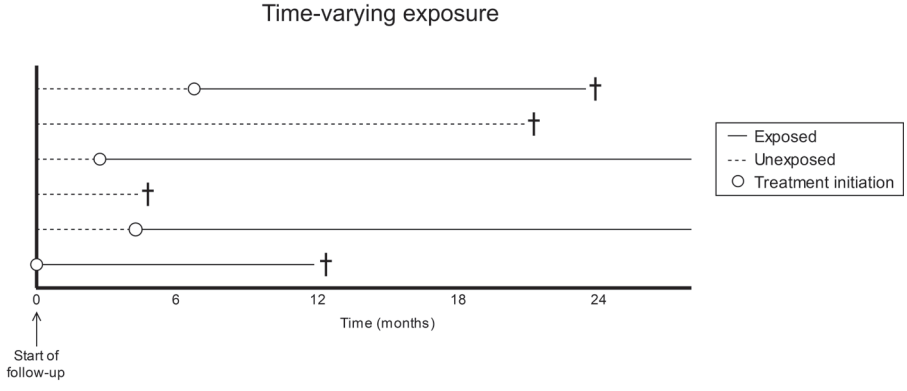
In pharmacoepidemiological studies we are often interested in the effects of initiating medication on a particular outcome after a certain event has occurred. A recent clinical example is the effect of initiating renin-angiotensin system inhibitors on mortality and recurrent AKI after acute kidney injury (AKI) (69-71). When using routinely collected healthcare data to study such questions, it is often difficult to assign individuals to the correct exposure groups: individuals become eligible for inclusion in our study immediately after the AKI event and follow-up will start at that moment. However, directly after the AKI event all individuals will likely be unexposed in our dataset, as individuals will gradually initiate therapy during follow-up. We cannot classify individuals in exposure groups based on post-baseline information as this will lead to immortal time bias. The easiest solution is then to move the baseline of our study from the date of the AKI event to a later time, e.g. 6 months after the AKI event. Our follow-up will therefore start at 6 months after the index AKI event (**Figure 3**) (72-74). This method is called landmarking and was recently applied by Brar *et al*. for this particular research question (69). In the landmarking method all individuals who died or developed the outcome between the AKI event and the newly chosen start of follow-up (i.e. 6 months after AKI) are excluded; those who initiate treatment during this period are considered exposed, and those who do not initiate treatment during this period are considered unexposed. Although landmarking prevents immortal time bias, the attentive reader will have noted that it can introduce prevalent user bias, which was discussed in the previous section.

Figure 3. Design of a landmark analysis to prevent immortal time bias. In the landmark analysis, follow-up starts at a chosen time period after a certain event, in this example at 6 months. Hence, all individuals that died before month 6 are excluded from the analysis (individual 4). Individuals are then classified according to exposure status in the first 6 months. Individuals 3, 5 and 6 are therefore considered treated, whereas individuals 1 and 2 are considered untreated.



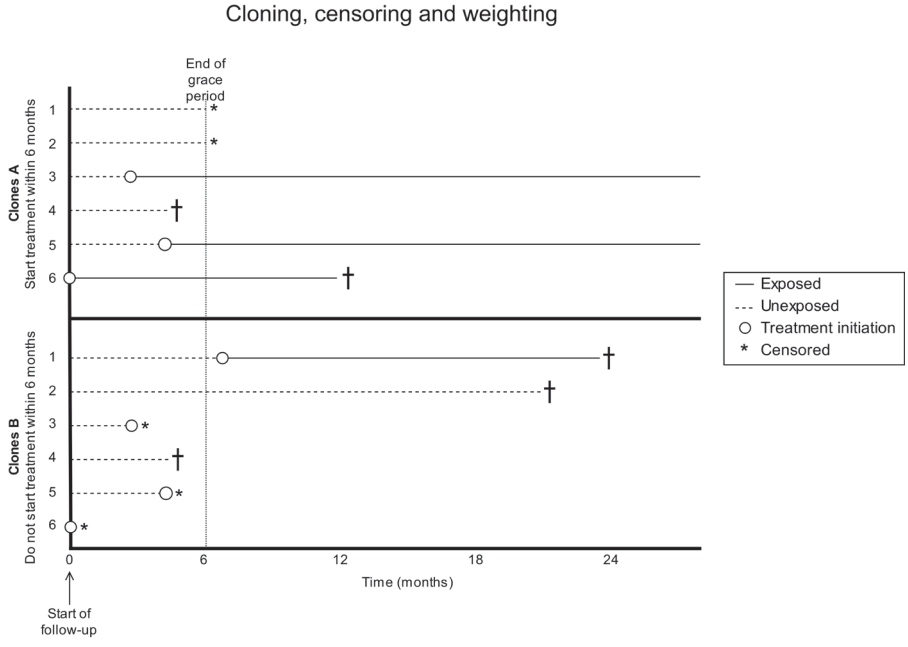
The next solution that prevents immortal time bias and allows starting follow-up immediately after the event has occurred is the use of a *time-varying exposure* (**Figure 4**). When using a time-varying exposure, individuals are allowed to switch exposure status from untreated to treated at the time of treatment initiation. Hence, individuals will contribute persontime to the unexposed group before treatment initiation, and to the exposed group after treatment initiation. This ensures that the time between start of follow-up and initiation of treatment will be correctly assigned to the non-users. For example, Hsu *et al.* used a time-varying exposure to study the effect of RASi after AKI on the risk of recurrent AKI (70). As previously mentioned, using a time-varying exposure involves time-varying confounding. When these confounders are also influenced by prior treatment, using standard methods such as multivariable regression may not be appropriate. Instead, methods such as marginal structural models that are based on inverse probability weighting can be used (27, 75). Applying these methods, the authors found that new use of RASi therapy was not associated with an increased risk of recurrent AKI.

Figure 4. Analysis using a time-varying exposure to prevent immortal time bias. In a time-varying design treatment status is allowed to change from unexposed to exposed at the moment of treatment initiation. This method allows to start of follow-up directly after the event has occurred and also does not exclude individuals. E.g., individual 1 is considered unexposed for the first 7 months of follow-up, but after 7 months will contribute to the exposed group. In the setting of time-varying exposures, time-varying confounding will be present too, which sometimes requires more advanced methods such as marginal structural models to obtain unbiased effect estimates.



Lastly, we may be interested in comparing treatment strategies that include a grace period (76). For example, we could compare the strategies "*initiate an ACE inhibitor within 6 months after the AKI event*" versus "*do not initiate an ACE inhibitor within 6 months after the AKI event*". The length of the grace period depends on what is commonly done in clinical practice. These treatment strategies with a grace period can be investigated by using a three-step method based on cloning, censoring and weighting (**Figure 5)**.

Figure 5. Design of a study using treatment strategies with a grace period based on cloning, censoring and weighting. Another method is comparing treatment strategies that include a grace period. Each individual is duplicated and assigned to one of two treatment strategies. In this example, clones 1a to 6a follow the strategy "initiate ACEi within 6 months", whereas clones 1b to 6b follow the strategy "do not initiate ACEi within 6 months". Note that copies 1a and 1b represent the same individual. Since copy 1a is assigned to initiating within 6 months, he is censored after month 6 as he did not initiate treatment. The censoring is likely to be informative and inverse probability weighting is required to adjust for this.



Briefly, each individual is duplicated so that there are two copies of each individual in the dataset. Each copy is then assigned to one of the treatment strategies. In the second step the copies are censored if and when their observed treatment does not adhere anymore to their assigned treatment strategy. Since this censoring is likely to be informative, the third step applies inverse probability weighting to correct for this. Bootstrapping can be used to take into account the cloning and weighting and obtain valid confidence intervals. An advantage of using treatment strategies with grace periods is that a wide range of questions can be answered, including questions on the duration of treatment and dynamic treatment strategies (e.g., when should treatment be initiated) (76, 77). However, this method requires that detailed longitudinal data is present to adequately adjust for the informative censoring. The three methods of landmarking, time-varying exposure and treatment strategies with grace periods are contrasted in **Table 1** and graphically depicted in **Figures 3-5**.

Table 1. Different methods to address immortal time bias in pharmacoepidemiological analyses.

	Landmark analysis	**Time-varying exposure**	**Cloning, censoring and weighting**
Immortal time bias	No	No	No
Start of follow-up	At landmark	At event	At event
Causal effect	Initiating versus not initiating at *x* months after event (landmark), conditional on having survived until landmark[†]	Initiating and always using versus never using (marginal structural model)	Initiating within *x* months versus not initiating within *x* months after event
Prevalent user bias	Possible	No	No
Results apply to	Individuals surviving until landmark	All individuals	All individuals
Baseline confounding	Yes	Yes	No[·]
Time-varying exposure	No	Yes	No
Time-varying confounding	No	Yes	No
Informative censoring	No	No	Yes
G-methods[§] required	No	Sometimes (if confounder is influenced by prior treatment)	Yes (inverse probability weighting)

[†] This is often how the effect estimate from a landmark analysis is interpreted. However, the landmark analysis conditions on surviving until a certain timepoint and classifies individuals into treatment groups based on past information, thereby possibly introducing prevalent user bias.

[·] Due to the cloning, at baseline each individual will appear in both treatment arms. Hence, no baseline confounding will be present.

[§] Methods based on standardization or inverse probability weighting, such as the G-formula or marginal structural models.

Missing data and misclassification

We now briefly discuss the implications of missing data and misclassification for bias and possible solutions. Although these two sources of bias are a common issue in pharmacoepidemiological studies, they are often less emphasized compared with confounding.

Missing data

We usually aim to adjust for as many confounders as possible in our analysis, including available laboratory tests (e.g. albuminuria, potassium) and clinical variables (e.g. blood pressure, BMI) which are indications for treatment. However, it is not unusual that a large proportion of these values are missing in routinely collected data. For example, in an analysis using data from the Swedish Renal Registry, baseline potassium and albuminuria-to-creatinine-ratio measurements were missing in 32% and 41% of patients, respectively (78). In such situations, researchers often perform a complete case analysis by restricting to individuals with both measurements available. However, this may lead to a drastic reduction in power and often also bias (79, 80). Methods such as multiple imputation are therefore recommended and are available in most software packages. These methods can reduce these biases even with large proportions of missing data (up to 90%) if data are missing at random or missing completely at random, sufficient auxiliary information is available and the imputation model is properly specified (81). It is therefore important to discuss the reasons for missingness and the plausibility of the missing at random assumption. In the above example, the researchers explained that although albuminuria and potassium values were measured in clinical practice, they were not among the list of mandatory laboratory markers that needed to be reported to the Swedish Renal Registry. Thus, some clinicians took the time to report those lab tests and others not, a decision that could be assumed to be at random. Furthermore, the authors showed that clinical characteristics were similar for individuals with and without missing data, thereby making the missing at random assumption plausible. More information on the different types of missingness (79, 80), in what situations complete case analysis leads to unbiased results (82), as well as tutorials to implement multiple imputation can be found elsewhere (83).

Misclassification

Although misclassification will be present in nearly every study, it may be especially important when using routinely collected healthcare data (84). Misclassification may for instance occur when using ICD-10 codes to ascertain the occurrence of chronic kidney disease or acute kidney injury, as these are not always coded in clinical practice and many patients are unaware of their disease (85-87). When AKI

diagnosis based on ICD coding is used as an outcome, differential misclassification will arise when doctors are more aware or more likely to encode AKI if certain drugs are prescribed. Basing kidney outcomes on biochemical criteria may sometimes mitigate such biases, but can also introduce bias when creatinine testing is more often directed towards sicker patients or patients at risk of CKD progression. Misclassification of comorbidities may be a significant concern in routinely collected data since the absence of a diagnosis (recording) is often considered to indicate absence of the comorbidity. Residual confounding may occur when confounders are misclassified and the direction can be both away or toward a null effect (84). Misclassification influences study results in ways that are often not anticipated, and simple heuristics about the impact of misclassification (towards the null or not) are often incorrect (88, 89). Many correction methods for misclassification exist, but these require information about its magnitude and structure (i.e. dependent, non-dependent, differential, non-differential) (90-93). As such information is often not available in electronic databases, sensitivity analyses similar to those for unmeasured confounding can be performed to estimate the influence of misclassification on results (33, 94).

Conclusion

Pharmacoepidemiological studies are increasingly being used to answer causal questions on the effectiveness and safety of medications in order to inform clinical decision making. In this review we discussed the most important biases that commonly occur in such studies. We also reviewed methods to account for these biases, which are summarized in **Table 2**. Researchers can and should prevent problems arising from immortal time and prevalent user biases in their study design. Confounding by indication bias can be tackled by using an active comparator design and adequately adjusting for confounders. When concerns remain about confounding or misclassification, quantifying their impact on effect estimates is recommended. When these principles are correctly applied, pharmacoepidemiological observational studies can provide valuable information for clinical practice.

Table 2. Potential biases in pharmacoepidemiological studies and proposed solutions.

Potential biases	Example of how biases may arise	Possible solutions and recommendations
Confounding by indication	• Confounding by indication arises when prognostic factors for the outcome are also an indication for initiating treatment. • Unmeasured/residual confounding arises when confounders are not adjusted for, either if they are not measured in the dataset or if they are unknown. • Time-varying confounding occurs when investigating time-varying exposures. When the confounder is influenced by past treatment, conventional methods to control for confounding will be biased.	• Research question: Unintended medication effects (e.g. rare side effects) may be less susceptible to confounding by indication than intended medication effects. • Design: Active comparator designs may decrease confounding bias if medication is given for similar indications. • Statistical methods: Multivariable regression, standardization or propensity score methods (matching, weighting, stratification, adjustment) can be used to control for measured confounding. Propensity score methods may have a number of advantages compared with regression, such as the ability to check if balance in confounders has been achieved. In the presence of time-varying confounding that are influenced by treatment, conventional methods lead to bias and the so called G-methods are required. • After analysis: The impact of unmeasured confounding on effect estimates can be investigated in simulation analyses. Negative control outcomes may investigate whether unmeasured confounders bias effect estimates.
Prevalent user bias	Comparing ever users vs. never users. Including individuals after they initiate treatment will miss early outcome events and exclude those that died (depletion of susceptibles).	• Prevalent user bias can and should be prevented by aligning initiation of treatment with start of follow-up; include new users of treatment. • Exclude prevalent users, e.g. those with drug prescription in 12 months prior to inclusion.
Immortal time bias	Classifying individuals in treatment groups based on future information not present at the start of follow-up. A period of time is created for the treated group during which the outcome cannot occur.	• Immortal time bias can and should be prevented by aligning initiation of treatment with start of follow-up. Do not use information after start of follow-up to classify individuals into exposure groups. • Landmarking, time-varying exposure, and the cloning, censoring and weighting method.
Missing data	Routinely collected healthcare data are prone to missing data. In multivariable analyses individuals with missing confounder data will be excluded. Complete case analysis often lead to bias when data is not missing completely at random, but a number of exceptions exist.	• Multiple imputation can be used to decrease bias and increase precision, even with large proportions of missing data (up to 90%) if data are missing at random or missing completely at random and the imputation model is properly specified. • Discuss the missing data mechanism and the plausibility of the missing (completely) at random assumption.
Misclassification	• Misclassification of the outcome may occur when outcomes are differentially ascertained depending on treatment status. • Misclassification of confounders may lead to residual confounding.	• The impact of misclassification on the estimated effect size can be quantified in sensitivity analyses. Online tools are available to implement these methods. • When external data are available, regression calibration, multiple imputation for measurement error or propensity score calibration can be used.

References

1. Begaud B. Dictionary of pharmacoepidemiology: Wiley, 2000.
2. Trevisan M, Fu EL, Xu Y, Jager KJ, Zoccali C, Dekker FW, et al. Pharmacoepidemiology for nephrologists (part one): Concept, applications and considerations for study design. Clinical Kidney Journal. 2020.
3. Kyriacou DN, Lewis RJ. Confounding by Indication in Clinical Research. JAMA. 2016;316(17):1818-9.
4. 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).
5. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. PLoS Med. 2008;5(3):e67.
6. Bosdriesz JR, Stel VS, van Diepen M, Meuleman Y, Dekker FW, Zoccali C, et al. Evidence-based medicine-When observational studies are better than randomized controlled trials. Nephrology (Carlton). 2020:e13742.
7. Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. Nat Rev Rheumatol. 2015;11(7):437-41.
8. Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Curr Epidemiol Rep. 2015;2(4):221-8.
9. Klatte DCF, Gasparini A, Xu H, de Deco P, Trevisan M, Johansson ALV, et al. Association Between Proton Pump Inhibitor Use and Risk of Progression of Chronic Kidney Disease. Gastroenterology. 2017;153(3):702-10.
10. VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol. 2019;34(3):211-9.
11. Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of Confounding and Reporting of Results in Causal Inference Studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. Ann Am Thorac Soc. 2019;16(1):22-8.
12. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol. 2002;155(2):176-84.
13. Robins JM. Data, design, and background knowledge in etiologic inference. Epidemiology. 2001;12(3):313-20.
14. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004;15(5):615-25.
15. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. Int J Epidemiol. 2010;39(2):417-20.
16. Myers JA, Rassen JA, Gagne JJ, Huybrechts KF, Schneeweiss S, Rothman KJ, et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. Am J Epidemiol. 2011;174(11):1213-22.
17. Blakely T, Lynch J, Simons K, Bentley R, Rose S. Reflection on modern methods: when worlds collide-prediction, machine learning and causal inference. Int J Epidemiol. 2019.
18. 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.

19. Schuler MS, Rose S. Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies. Am J Epidemiol. 2017;185(1):65-73.
20. Suttorp MM, Siegerink B, Jager KJ, Zoccali C, Dekker FW. Graphical presentation of confounding in directed acyclic graphs. Nephrol Dial Transplant. 2015;30(9):1418-23.
21. Ferguson KD, McCann M, Katikireddi SV, Thomson H, Green MJ, Smith DJ, et al. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. Int J Epidemiol. 2019.
22. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res. 2011;46(3):399-424.
23. 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.
24. Shah BR, Laupacis A, Hux JE, Austin PC. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. J Clin Epidemiol. 2005;58(6):550-9.
25. Sturmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. J Clin Epidemiol. 2006;59(5):437-47.
26. Fu EL, Groenwold RHH, Zoccali C, Jager KJ, van Diepen M, Dekker FW. Merits and caveats of propensity scores to adjust for confounding. Nephrol Dial Transplant. 2019;34(10):1629-35.
27. Williamson T, Ravani P. Marginal structural models in clinical research: when and how to use them? Nephrol Dial Transplant. 2017;32(suppl_2):ii84-ii90.
28. Dekkers IA, van der Molen AJ. Propensity Score Matching as a Substitute for Randomized Controlled Trials on Acute Kidney Injury After Contrast Media Administration: A Systematic Review. AJR Am J Roentgenol. 2018;211(4):822-6.
29. Groenwold RH, Sterne JA, Lawlor DA, Moons KG, Hoes AW, Tilling K. Sensitivity analysis for the effects of multiple unmeasured confounders. Ann Epidemiol. 2016;26(9):605-11.
30. 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.
31. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf. 2006;15(5):291-303.
32. Greenland S. Basic methods for sensitivity analysis of biases. Int J Epidemiol. 1996;25(6):1107-16.
33. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. Springer. 2009.
34. Uddin MJ, Groenwold RH, Ali MS, de Boer A, Roes KC, Chowdhury MA, et al. Methods to control for unmeasured confounding in pharmacoepidemiology: an overview. Int J Clin Pharm. 2016;38(3):714-23.
35. Groenwold RH, Nelson DB, Nichol KL, Hoes AW, Hak E. Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research. Int J Epidemiol. 2010;39(1):107-17.
36. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med. 2017;167(4):268-74.
37. https://www.evalue-calculator.com/.

38. VanderWeele TJ, Ding P, Mathur MB. Technical Considerations in the Use of the E-Value. Journal of Causal Inference. 2019;7(2).
39. Trinquart L, Erlinger AL, Petersen JM, Fox M, Galea S. Applying the E Value to Assess the Robustness of Epidemiologic Fields of Inquiry to Unmeasured Confounding. Am J Epidemiol. 2019;188(6):1174-80.
40. Laureati P, Xu Y, Trevisan M, Schalin L, Mariani I, Bellocco R, et al. Initiation of sodium polystyrene sulphonate and the risk of gastrointestinal adverse events in advanced chronic kidney disease: a nationwide study. Nephrol Dial Transplant. 2019.
41. Ioannidis JPA, Tan YJ, Blum MR. Limitations and Misinterpretations of E-Values for Sensitivity Analyses of Observational Studies. Ann Intern Med. 2019;170(2):108-11.
42. VanderWeele TJ, Mathur MB, Ding P. Correcting Misinterpretations of the E-Value. Ann Intern Med. 2019;170(2):131-2.
43. http://www.drugepi.org/dope-downloads/#Sensitivity%20Analysis.
44. https://sites.google.com/site/biasanalysis/.
45. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 2010;21(3):383-8.
46. Arnold BF, Ercumen A. Negative Control Outcomes: A Tool to Detect Bias in Randomized Trials. JAMA. 2016;316(24):2597-8.
47. Groenwold RH. Falsification end points for observational studies. JAMA. 2013;309(17):1769-70.
48. Edner M, Benson L, Dahlstrom U, Lund LH. Association between renin-angiotensin system antagonist use and mortality in heart failure with severe renal insufficiency: a prospective propensity score-matched cohort study. Eur Heart J. 2015;36(34):2318-26.
49. Weir MA, Dixon SN, Fleet JL, Roberts MA, Hackam DG, Oliver MJ, et al. beta-Blocker dialyzability and mortality in older patients receiving hemodialysis. J Am Soc Nephrol. 2015;26(4):987-96.
50. Fu EL, Uijl A, Dekker FW, Lund LH, Savarese G, Carrero JJ. Association between β-blocker use and mortality/morbidity in patients with heart failure with reduced, midrange and preserved ejection fraction and advanced chronic kidney disease. Circ Heart Fail.
51. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol. 2003;158(9):915-20.
52. Johnson ES, Bartman BA, Briesacher BA, Fleming NS, Gerhard T, Kornegay CJ, et al. The incident user design in comparative effectiveness research. Pharmacoepidemiol Drug Saf. 2013;22(1):1-6.
53. 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.
54. 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.
55. Tomlinson L, Smeeth L. Angiotensin-converting enzyme inhibitor or angiotensin receptor blocker use and renal outcomes: prevalent user designs may overestimate benefit. JAMA Intern Med. 2014;174(10):1706.
56. Stovitz SD, Banack HR, Kaufman JS. 'Depletion of the susceptibles' taught through a story, a table and basic arithmetic. BMJ Evid Based Med. 2018;23(5):199.
57. Lajous M, Banack HR, Kaufman JS, Hernan MA. Should patients with chronic disease be told to gain weight? The obesity paradox and selection bias. Am J Med. 2015;128(4):334-6.

58. 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.

59. Dickerman BA, Garcia-Albeniz X, Logan RW, Denaxas S, Hernan MA. Avoidable flaws in observational analyses: an application to statins and cancer. Nat Med. 2019;25(10):1601-6.

60. Suissa S. Immortal time bias in pharmaco-epidemiology. Am J Epidemiol. 2008;167(4):492-9.

61. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ. 2010;340:b5087.

62. Kwon S, Kim YC, Park JY, Lee J, An JN, Kim CT, et al. The Long-term Effects of Metformin on Patients With Type 2 Diabetic Kidney Disease. Diabetes Care. 2020;43(5):948-55.

63. Fu EL, Van Diepen M. Comment on Kwon et al. The Long-term Effects of Metformin on Patients With Type 2 Diabetic Kidney Disease. Diabetes Care 2020;43:948-955 Diabetes Care. 2020.

64. Suissa S. Immortal time bias in observational studies of drug effects. Pharmacoepidemiol Drug Saf. 2007;16(3):241-9.

65. Harding BN, Weiss NS. Immortal Time Bias: What Are the Determinants of Its Magnitude? Am J Epidemiol. 2019.

66. 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.

67. Karim ME, Gustafson P, Petkau J, Tremlett H, Long-Term B, Adverse Effects of Beta-Interferon for Multiple Sclerosis Study G. Comparison of Statistical Approaches for Dealing With Immortal Time Bias in Drug Effectiveness Studies. Am J Epidemiol. 2016;184(4):325-35.

68. Gran JM, Roysland K, Wolbers M, Didelez V, Sterne JA, Ledergerber B, et al. A sequential Cox approach for estimating the causal effect of treatment in the presence of time-dependent confounding applied to data from the Swiss HIV Cohort Study. Stat Med. 2010;29(26):2757-68.

69. Brar S, Ye F, James MT, Hemmelgarn B, Klarenbach S, Pannu N, et al. Association of Angiotensin-Converting Enzyme Inhibitor or Angiotensin Receptor Blocker Use With Outcomes After Acute Kidney Injury. JAMA Intern Med. 2018;178(12):1681-90.

70. Hsu CY, Liu KD, Yang J, Glidden DV, Tan TC, Pravoverov L, et al. Renin-Angiotensin System Blockade after Acute Kidney Injury (AKI) and Risk of Recurrent AKI. Clin J Am Soc Nephrol. 2020;15(1):26-34.

71. Siew ED, Parr SK, Abdel-Kader K, Perkins AM, Greevy RA, Jr., Vincz AJ, et al. Renin-angiotensin aldosterone inhibitor use at hospital discharge among patients with moderate to severe acute kidney injury and its association with recurrent acute kidney injury and mortality. Kidney Int. 2020.

72. Dafni U. Landmark analysis at the 25-year landmark point. Circ Cardiovasc Qual Outcomes. 2011;4(3):363-71.

73. Mi X, Hammill BG, Curtis LH, Lai EC, Setoguchi S. Use of the landmark method to address immortal person-time bias in comparative effectiveness research: a simulation study. Stat Med. 2016;35(26):4824-36.

74. Gleiss A, Oberbauer R, Heinze G. An unjustified benefit: immortal time bias in the analysis of time-dependent events. Transpl Int. 2018;31(2):125-30.

75. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. Stat Med. 2013;32(9):1584-618.

76. Hernan MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. Basic Clin Pharmacol Toxicol. 2006;98(3):237-42.

77. Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernan MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. Int J Biostat. 2010;6(2):Article 18.

78. Fu EL, Clase CM, Evans M, Lindholm B, Rotmans JI, van Diepen M, et al. Comparative effectiveness of renin-angiotensin system inhibitors and calcium channel blockers in individuals with advanced chronic kidney disease: a nationwide observational cohort study. Submitted. 2020.

79. de Goeij MC, van Diepen M, Jager KJ, Tripepi G, Zoccali C, Dekker FW. Multiple imputation: dealing with missing data. Nephrol Dial Transplant. 2013;28(10):2415-20.

80. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.

81. Madley-Dowd P, Hughes R, Tilling K, Heron J. The proportion of missing data should not be used to guide decisions on multiple imputation. J Clin Epidemiol. 2019;110:63-73.

82. Hughes RA, Heron J, Sterne JAC, Tilling K. Accounting for missing data in statistical analyses: multiple imputation is not always the answer. Int J Epidemiol. 2019;48(4):1294-304.

83. Blazek K, van Zwieten A, Saglimbene V, Teixeira-Pinto A. A practical guide to multiple imputation of missing data in nephrology. Kidney Int. 2020.

84. Funk MJ, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. Curr Epidemiol Rep. 2014;1(4):175-85.

85. Gasparini A, Evans M, Coresh J, Grams ME, Norin O, Qureshi AR, et al. Prevalence and recognition of chronic kidney disease in Stockholm healthcare. Nephrol Dial Transplant. 2016;31(12):2086-94.

86. McDonald HI, Shaw C, Thomas SL, Mansfield KE, Tomlinson LA, Nitsch D. Methodological challenges when carrying out research on CKD and AKI using routine electronic health records. Kidney Int. 2016;90(5):943-9.

87. Tomlinson LA, Riding AM, Payne RA, Abel GA, Tomson CR, Wilkinson IB, et al. The accuracy of diagnostic coding for acute kidney injury in England - a single centre study. BMC Nephrol. 2013;14:58.

88. van Smeden M, Lash TL, Groenwold RHH. Reflection on modern methods: five myths about measurement error in epidemiological research. Int J Epidemiol. 2020;49(1):338-47.

89. Jurek AM, Greenland S, Maldonado G, Church TR. Proper interpretation of non-differential misclassification effects: expectations vs observations. Int J Epidemiol. 2005;34(3):680-7.

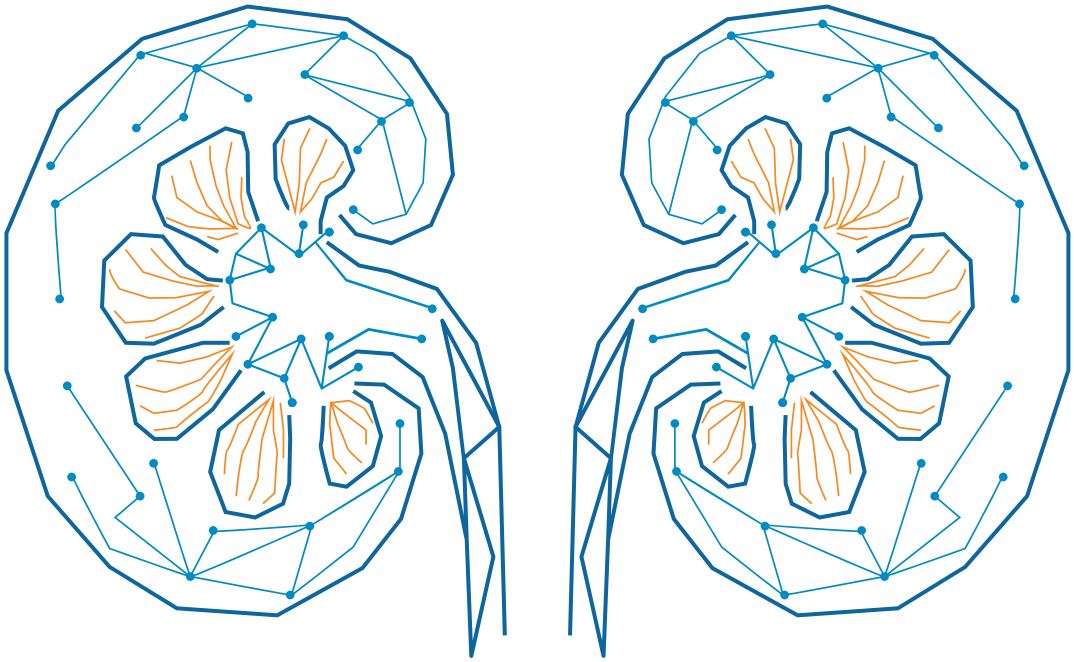
90. Sturmer T, Schneeweiss S, Avorn J, Glynn RJ. Adjusting effect estimates for unmeasured confounding with validation data using propensity score calibration. Am J Epidemiol. 2005;162(3):279-89.

91. Cole SR, Chu H, Greenland S. Multiple-imputation for measurement-error correction. Int J Epidemiol. 2006;35(4):1074-81.

92. Spiegelman D, McDermott A, Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. Am J Clin Nutr. 1997;65(4 Suppl):1179S-86S.

93. Bang H, Chiu YL, Kaufman JS, Patel MD, Heiss G, Rose KM. Bias Correction Methods for Misclassified Covariates in the Cox Model: comparison offive correction methods by simulation and data analysis. J Stat Theory Pract. 2013;7(2):381-400.

94. Fox MP, Lash TL, Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. Int J Epidemiol. 2005;34(6):1370-6.



CHAPTER 3

Merits and caveats of propensity scores to adjust for confounding

Edouard L. Fu, Rolf H.H. Groenwold, Carmine Zoccali, Kitty J. Jager, Merel van Diepen, Friedo W. Dekker

Nephrol Dial Transplant 2019; 34: 1629-1635

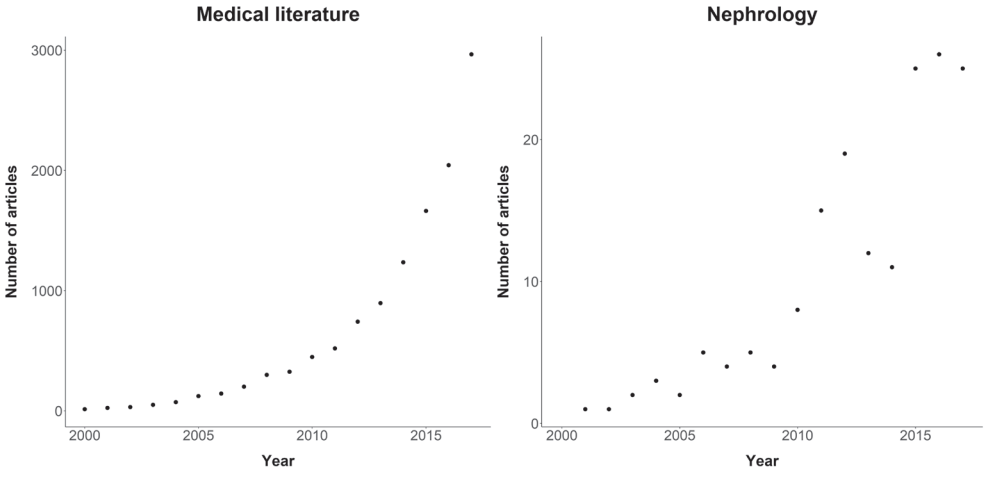
Abstract

Proper adjustment for confounding is essential when estimating the effects of treatments or risk factors on health outcomes in observational data. To this end, various statistical methods have been developed. The past couple of years the use of propensity scores to control for confounding has increased. Proper understanding of this method is necessary to critically appraise research in which it is applied. In this article we provide an overview of propensity score methods, explaining their concept, advantages, and possible disadvantages. Furthermore, the use of propensity score matching, propensity score adjustment and propensity score weighting is illustrated using data from the NECOSAD cohort of dialysis patients.

Introduction

Much research in nephrology that investigates the effects of medical treatment or risk factors on health outcomes makes use of observational data. In such observational studies prognostic factors are usually unequally distributed across different treatment groups, which may bias a direct comparison between these groups. This phenomenon is called confounding (1). To control for confounding, several statistical methods can be used, such as matching, weighting, standardization, multivariable regression analysis, and propensity score methods.

Figure 1. Total number of articles and nephrology articles indexed in Pubmed using the term "propensity score" in title or abstract. The following nephrology journals were searched on Pubmed for the term "propensity score" in title or abstract: Journal of the American Society of Nephrology, Kidney International, American Journal of Kidney Diseases, Nephrology Dialysis Transplantation, Clinical Journal of the American Society of Nephrology, BMC Nephrology and American Journal of Nephrology.



During the past decades the use of propensity scores has increased exponentially in the medical literature, and also in nephrology research (**Figure 1**). In a previous article in this series, propensity scores were briefly discussed (2). In this article we will explain what propensity scores are, to what extent they can control for confounding and their advantages and disadvantages. Furthermore, we illustrate the use of propensity score methods using data from the NECOSAD cohort of dialysis patients.

Propensity scores

In 1983, Rosenbaum and Rubin introduced the propensity score as a means to control for confounding in observational studies (3). The propensity score combines the information from all measured confounders into one score, which represents a patient's probability of receiving treatment. More formally, the propensity score is defined as the probability of treatment assignment conditional on a patient's measured confounders and can be seen as a summary measure of the confounder information. Conditional on the propensity score, the treatment and the control group are expected to have the same distribution of measured confounders and therefore allows for a direct comparison between groups (4).

Estimation of the propensity score model

For binary treatments (or exposures), the propensity score is usually estimated using a binary logistic regression analysis with the treatment as dependent variable and the measured confounders as independent variables (5). Machine learning algorithms, such as classification and regression trees, can also be used to estimate propensity scores (6). It is not advised to use data-driven approaches to select confounders for the propensity score model, e.g. by looking at the statistical significance of the relation between confounder and treatment status (7-9). Instead, directed acyclic graphs (DAGs) can be used to identify potential confounders (10, 11). Variables that are only related to the treatment and not to the outcome, should not be included in the propensity score model, since this may amplify unmeasured confounding and increase the variance of the effect estimate (12). When information on a large number of confounders is available, e.g. when performing a study using routine health care records data, high-dimensional propensity scores can be used (13). However, in previous literature this appeared to have little added value, probably because additional confounders are correlated to information already being observed (14, 15). Furthermore, the high-dimensional propensity score algorithm selects variables for the PS model in a data-driven approach, which may incorrectly select variables that are not confounders (e.g. intermediate variables or colliders) thus requiring pre-selection of possible confounders.

Using the propensity score to correct for confounding

Once the propensity score model is estimated, it can be used to compute individual probabilities of receiving treatment by entering the subject's confounder values into the estimated propensity score model. Different methods can then be used to actually control for confounding: propensity score matching (PS matching), including the propensity score as a covariate in a multivariable regression model (PS adjustment), weighting based on the propensity score (PS weighting), or

propensity score stratification. Here, we only focus on the first three methods since PS stratification may not completely remove the bias due to confounding (16).

Propensity score matching

In PS matching, each treated subject is matched to an untreated subject on the basis of similar values of their propensity scores. Different options are available for PS matching, such as one-to-one matching vs. one-to-many matching, nearest-neighbor matching vs. optimal matching, and by specifying a caliper distance (each with or without replacement). One-to-one nearest-neighbor matching is commonly used and practical to implement when sample size is large (17). In this matching method, treated subjects are picked in random order and matched to one untreated subject with the closest propensity score. When a pre-specified caliper width is used, a restriction is placed on the maximum difference between the propensity scores of the matched couple (18, 19). In one-to-many matching, a treated subject can be matched to multiple untreated subjects. Optimal matching minimizes the total within-pair difference of the propensity score, but has been shown to provide similar balance in measured confounders as nearest-neighbor matching when sample size is large (20, 21). When matching with replacement is used, an untreated subject remains available in the dataset after matching and therefore can be matched to more than one treated subject.

The assessment of balance after PS matching

It is important to assess to which extent the PS matching achieved balance with respect to the measured confounders between the matched groups of treated and untreated subjects. After all, the propensity score is a balancing score with the aim of achieving comparable groups. On a group-level, the confounders need to be equally distributed between the treatment and control group. Indeed, after PS matching the groups of treated and untreated subjects will - in expectation - be comparable in terms of the confounders included in the propensity score model, even though at an individual level, subjects with the same PS may differ on specific confounder values. A common approach to assess balance is to quantify the absolute standardized differences for each confounder in the PS matched set. The standardized difference is calculated by dividing the difference in sample means (or proportions for dichotomous variables) by the pooled standard deviation. A frequently used cut point indicating acceptable balance is a standardized difference less than 0.1 between treatment groups (19). If balance has not been achieved, the propensity score model could be modified until adequate balance is achieved, for example by including transformations of variables, higher order terms, or interactions between confounders (22, 23). Once balance is considered appropriate, treatment (or exposure) effects can be estimated directly from the PS matched set.

Propensity score adjustment

Another method to control for confounding is to include the propensity score as a covariate in a regression analysis of the treatment on the outcome. A possible disadvantage of this approach is that some (residual) confounding may remain when the relation between the propensity score and the outcome is incorrectly specified (e.g., assuming a linear relation, while in fact the relation is quadratic). PS adjustment differs from traditional multivariable regression analysis since all confounders are summarized in one score, instead of adjusting for all separate confounders.

Propensity score weighting

PS weighting makes use of weights based on each subject's propensity score in order to create a weighted dataset (pseudo-population) in which no confounding is present. A subject's weight is defined as the inverse of the probability of receiving his actual treatment, conditional on their measured confounder information. Thus, for treated subjects their weight is equal to $1/PS_i$ and for untreated subjects their weight is equal to $1/(1-PS_i)$, where PS_i indicates the propensity score of individual *i*. This method is better known as inverse probability weighting (IPW). In the weighted pseudo-population the measured confounders are expected to be distributed similarly between the treatment and control groups and hence confounding by measured confounders is eliminated. Like in the PS matched population, the pseudo-population allows for a formal assessment of balance of confounders between the treated and untreated groups. Treatment effects can be estimated directly from the PS weighted population.

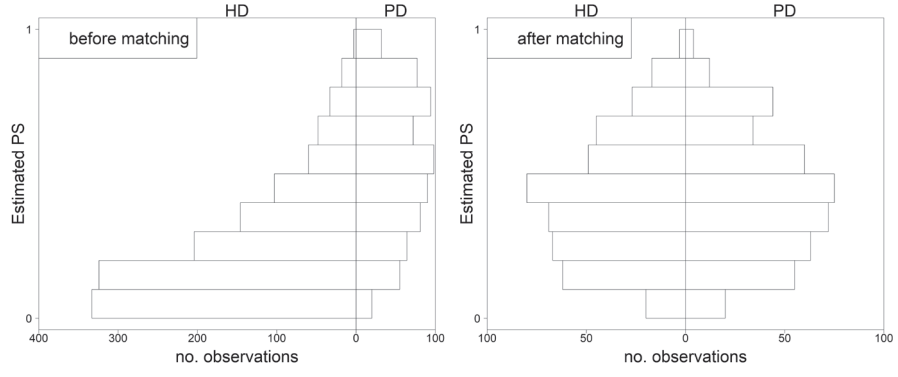
An illustration of Ps methods in the NECOSAD study

To illustrate propensity score methods, we used data of the NECOSAD study, a Dutch multicenter cohort study including 1955 incident dialysis patients. Patients older than 18 years beginning dialysis as first renal replacement therapy were included and monitored at 6-month intervals until renal transplantation or death. Further information regarding study design and patient characteristics can be found elsewhere (24). We estimated the relation between treatment modality, peritoneal dialysis (PD) vs. hemodialysis (HD), and mortality risk using a Cox proportional hazards regression analysis. Four methods to control for confounding were applied: PS matching, PS adjustment, PS weighting, and multivariable Cox regression analysis (25). Dialysis modality and confounder values measured within 3 months after renal replacement therapy was started were used. These measurements were considered proxies for the confounder values prior to the start of dialysis because of the following reasons. First, the chosen treatment modality would be more definitive and renal replacement therapy is likely to be a chronic therapy at 3

months. Second, patients who recovered or died from acute renal failure within this period would be excluded from analysis. Third, at 3 months the clinical condition of patients is more likely to have stabilized (25, 26). We used the same list of confounders that were used by Termorshuizen et al. to create the PS model using a multivariable binary logistic regression analysis (25). The confounders included were age, gender, primary kidney disease, Davies comorbidity index, SGA score, residual renal function, hemoglobin and albumin levels at baseline.

Missing confounder values were imputed using a multiple imputation by chained equations algorithm, including the confounder information, dialysis modality, follow-up time, and mortality. Ideally, multiple imputed datasets would be created, however for educational simplicity we used one imputed dataset. A detailed explanation about multiple imputation followed by propensity score methods can be found elsewhere (27, 28). Statistical analyses were performed using R version 3.4.1 (29).

Figure 2. Distribution of propensity scores before and after matching in hemodialysis (HD) and peritoneal dialysis (PD) treated patients



Patient characteristics are reported in **Table 1**. Mean age at the start of dialysis was 60 years, mean eGFR 3.8 ml/min/1.73m^2. As indicated by the absolute standardized differences, the two treatment groups differ markedly with respect to important confounders such as age. The unadjusted effect of peritoneal dialysis vs. hemodialysis on mortality risk was estimated to be HR 0.60 (95%CI 0.52-0.69) (**Table 2**). The distributions of the propensity score for the hemodialysis and for the peritoneal dialysis group are shown in **Figure 2**, showing a lower propensity score – on average – for patients who received hemodialysis, which is according to expectation since peritoneal dialysis was considered as the active treatment (coded 1), whereas hemodialysis was the control treatment (coded 0).

Table 1. Patient characteristics before and after PS matching/weighting

	Before PS matching/weighting			After PS matching			After PS weighting		
	HD (n=1272)	PD (n=683)	standardized difference*	HD (n=441)	PD (n=441)	standardized difference	HD (n=1979)	PD (n=1918)	standardized difference
Mean age in years‡ (sd)	64 (14)	53 (15)	0.721	57 (15)	57 (14)	0.008	59 (15)	59 (16)	0.002
Male sex (%)	758 (59.6%)	461 (67.5%)	0.165	285 (64.6%)	285 (64.6%)	<0.001	1224 (63.8%)	1258 (63.6%)	0.006
Primary kidney disease			0.124			0.070			0.025
Diabetes mellitus	181 (14.2%)	103 (15.1%)		66 (15.0%)	63 (14.3%)		290 (14.7%)	283 (14.8%)	
Glomerulonephritis	113 (8.9%)	127 (18.6%)		71 (16.1%)	56 (12.7%)		256 (12.9%)	260 (13.6%)	
Renal vascular	251 (19.7%)	80 (11.7%)		65 (14.7%)	69 (15.6%)		335 (16.9%)	351 (18.3%)	
All other	727 (57.2%)	373 (54.6%)		239 (54.2%)	253 (57.4%)		1098 (55.5%)	1024 (53.3%)	
Davies comorbidity index			0.328			0.074			0.024
No comorbidity	549 (43.2%)	411 (60.2%)		238 (54.0%)	255 (57.8%)		990 (50.0%)	1000 (52.1%)	
Intermediate	593 (46.6%)	228 (33.4%)		164 (37.2%)	151 (34.2%)		818 (41.3%)	743 (38.8%)	
High comorbidity	130 (10.2%)	44 (6.4%)		39 (8.8%)	35 (7.9%)		171 (8.6%)	175 (9.1%)	
SGA score			0.396			0.028			0.031
≤4	186 (14.5%)	51 (7.4%)		41 (9.3%)	25 (5.7%)		246 (12.4%)	285 (14.9%)	
5	245 (19.3%)	78 (11.4%)		59 (13.4%)	67 (15.2%)		330 (16.7%)	317 (16.5%)	
6	493 (38.8%)	234 (34.3%)		148 (33.6%)	181 (41.0%)		732 (37.0%)	656 (34.2%)	
7	348 (27.4%)	320 (46.9%)		193 (43.8%)	168 (38.1%)		671 (33.9%)	660 (33.3%)	
Residual renal function in ml/min/1.73m² (sd)	3.6 (2.9)	4.3 (2.9)	0.258	4.1 (3.1)	4.0 (2.7)	0.039	3.8 (2.8)	3.8 (3.0)	0.004
Hemoglobin in mmol/l (sd)	6.7 (0.9)	7.4 (1.0)	0.824	7.1 (0.8)	7.1 (0.9)	0.009	6.9 (1.0)	6.9 (1.0)	0.022
Albumin in g/l (sd)	3.6 (0.5)	3.6 (0.5)	0.065	3.6 (0.5)	3.6 (0.5)	0.050	3.6 (0.6)	3.6 (0.5)	0.051

Abbreviations: HD – hemodialysis; PD – peritoneal dialysis

* A standardized difference <0.1 indicates acceptable balance

‡ Patient characteristics were measured 3 months after initiation of dialysis

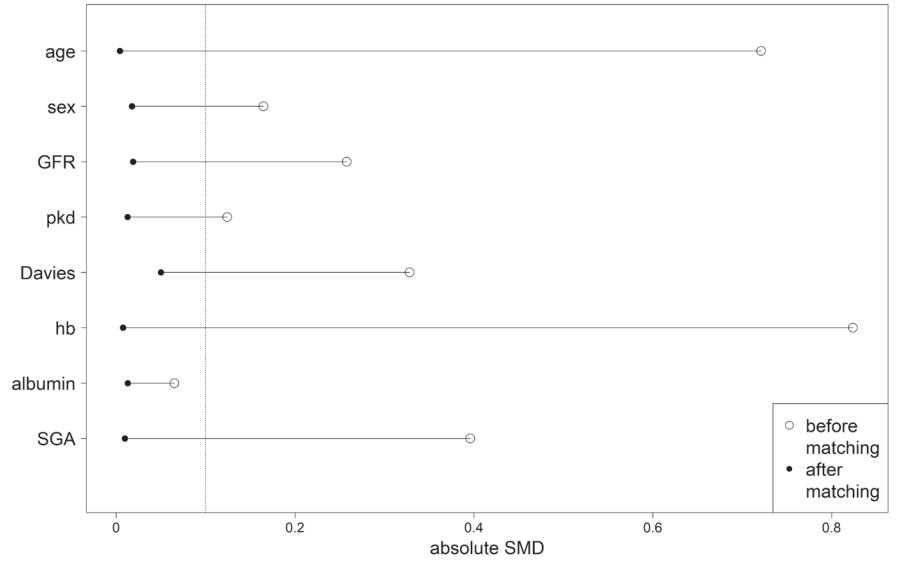
Table 2. Estimated effects of peritoneal dialysis vs. hemodialysis on mortality using different methods to control for confounding

	HR (95% CI)
Unadjusted	0.60 (0.52-0.69)
Propensity score matching[‡]	1.04 (0.85-1.28)
PS adjustment[‡]	1.00 (0.84-1.18)
PS weighting[‡]	0.93 (0.78-1.12)
Multivariable regression adjustment[*]	1.08 (0.91-1.28)

[‡]The propensity score model included the variables age, gender, primary kidney disease, Davies comorbidity index, SGA score, residual renal function, hemoglobin, and albumin levels 3 months after renal replacement therapy.
[*]The multivariable Cox regression analysis was adjusted for the variables age, gender, primary kidney disease, Davies comorbidity index, SGA score, residual renal function, hemoglobin, and albumin levels 3 months after renal replacement therapy.

First, we applied PS matching to control for confounding. We used one-to-one nearest-neighbor matching without replacement, using a caliper width equal to 0.01 on the logit of the propensity score. After PS matching, 441 matched pairs were available for the analysis. Note that the loss in number of patients included in the PS matched set can be limited by applying one-to-many matching instead of one-to-one matching (as was done here). **Figure 3** shows the absolute standardized differences before and after PS matching, indicating clear improvement in balance of confounders between treatment groups due to the matching (actual numbers shown in **Table 1**). In the PS matched set, the effect of peritoneal dialysis vs. hemodialysis on mortality was estimated to be HR 1.04 (95%CI 0.85-1.28).

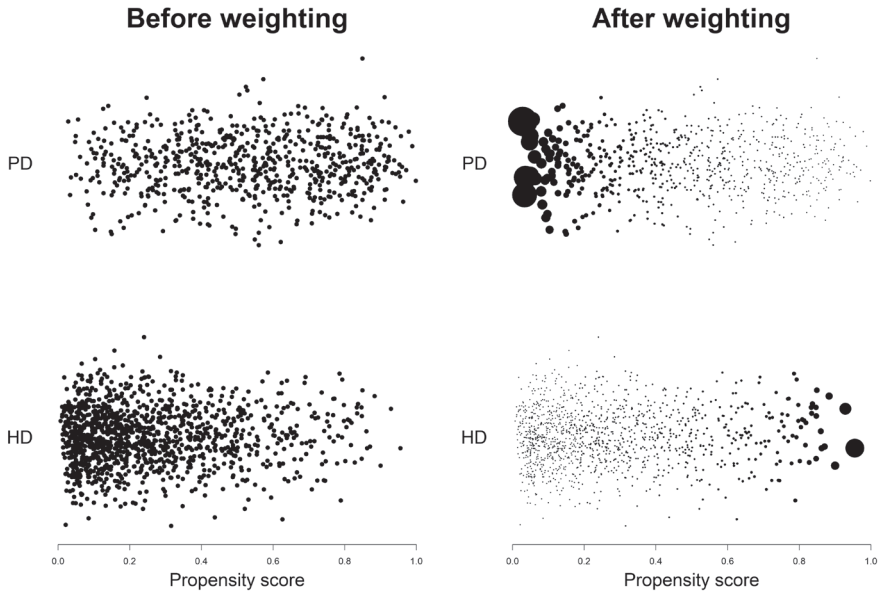
Figure 3. Absolute standardized differences for each of the confounding variables before and after PS matching. The dashed line represents the cut-off point of 0.1 indicating acceptable balance.



Second, the propensity score was included as a covariate in the Cox regression analysis (n= 1955). The relation between the propensity score and mortality was assumed to be log-linear; modelling the relation between the propensity score and mortality using splines (5 knots) yielded the same results. After PS adjustment, the effect of peritoneal dialysis vs. hemodialysis on mortality was estimated to be HR 1.00 (95%CI 0.84-1.19).

As a third approach to control for confounding, propensity score weighting was applied. The estimated PS weights ranged between 1.0 and 35.6. After weighting, the weighted pseudo-population consisted of 1918 patients in the peritoneal dialysis and 1979 patients in the hemodialysis group and confounders were balanced between treatment groups (**Table 1**). A dot plot showing the propensity score distribution of the hemodialysis and peritoneal dialysis group before and after weighting provides insight into the properties of PS weighting: subjects in the peritoneal dialysis group, who are relatively uncommon compared to subjects in the hemodialysis group with the same propensity score, receive a relatively larger weight. Likewise for subjects in the hemodialysis group, who are relatively uncommon compared to subjects in the peritoneal dialysis group with the same propensity score. By putting larger weights on relatively rare subjects, their contribution increases and as a result of that, the peritoneal and hemodialysis group are expected to have similar distributions of the propensity score in the weighted pseudo-population (**Figure 4**).

Figure 4. Dot plot showing the propensity score distributions among hemodialysis (HD) and peritoneal dialysis (PD) treated patients, before and after propensity score weighting. Each dot represents a different subject. In the right panel the dot size represents the weight of each observation in the propensity score weighted pseudo-population. Weights ranged between 1.0 and 35.6.



Consequently, the confounders are expected to be balanced between the two groups too. The effect of peritoneal dialysis vs. hemodialysis on mortality was estimated using weighted Cox regression analysis with a robust variance estimator: HR 0.93 (95%CI 0.78-1.12).

Lastly, a multivariable Cox proportional hazards model was fitted to the data, in which the confounders were included as separate covariates. This approach yielded a similar estimate as the propensity based methods. The effect of peritoneal dialysis vs. hemodialysis on mortality was estimated to be HR 1.08 (95%CI 0.91-1.28).

Choosing between the different propensity score methods

An important aspect to consider when choosing between the different PS methods is which treatment effect the researcher wants to estimate. Choosing a particular treatment effect depends on the research question and target population to whom the results need to apply (19, 30). Here, we consider three types of treatment effects: the average treatment effect (ATE), the average treatment effect for the treated (ATT) and the average treatment effect for the untreated (ATU).

The ATE represents the average effect of treatment (PD vs. HD) in the entire study population (i.e. 1955 subjects in our example study). The ATT can be interpreted as the average effect of treatment among those who were actually treated. Therefore, in our example the ATT represents the effect of peritoneal dialysis vs. hemodialysis in the subgroup of 683 subjects who actually received peritoneal dialysis. On the other hand, the ATU represents the effect of peritoneal dialysis vs. hemodialysis in the subgroup of 1272 subjects who actually received hemodialysis. Hence, the different treatment effects apply to different (sub)populations of dialysis patients. The ATE is more of interest if the goal of the researcher is to apply the treatment to all patients (e.g. when examining the effect of a public health intervention such as salt reduction in bread), whereas the ATT is more of interest when patient's characteristics determine the treatment received (e.g. when discussing dialysis options in a young patient, who is more likely to receive peritoneal dialysis because of his age) (31).

Both PS matching and PS weighting can estimate each of the different treatment effects. If the basis for PS matching is formed by the treated subjects, PS matching will estimate the ATT, whereas if untreated subjects are the basis for the matching the ATU is estimated. The ATE can be estimated by creating a matched set that represents the entire study population. In our example PS matching estimated the ATT since for each treated subject an untreated subject with the closest PS was matched. Whether PS weighting estimates an ATT, ATU or ATE depends on the method of weighting (32). PS weighting as applied in our example yields an estimate of the ATE. Therefore, the effect estimates obtained through PS matching (ATT, HR 1.04) and PS weighting (ATE, HR 0.93) have a different interpretation in our example study (**Table 2**).

PS adjustment and multivariable regression (MV) analysis both estimate a conditional effect which in our example neither can be interpreted as an ATT, ATU nor ATE (33). The effect of treatment is estimated conditional on the other variables in the regression model, in other words the effect of treatment while holding all other variables in the regression model constant.

Merits of propensity score methods

PS methods and multivariable regression have the same potential to control for confounding. However, PS methods have certain advantages compared to traditional multivariable regression. Firstly, PS matching and PS weighting clearly separate the design and analysis steps: first confounding is eliminated by balancing the confounders (design step) and afterwards the treatment effect is directly estimated (analysis step).

Secondly, by inspecting the propensity score distribution, areas of non-overlap between the treated and untreated groups can be identified. Patients in these areas of non-overlap may have an absolute indication or contra-indication for treatment (34). In PS matching, these subjects, for whom no comparison can be made, are not matched and hence excluded from the analysis. In PS weighting, these subjects may be identified since they receive very large weights. In our example data there was no clear non-overlap of the propensity score distributions (**Figure 2**).

Thirdly, propensity score methods may be preferred over multivariable regression analysis when the number of events is low relative to the number of confounders. The reason is that PS methods estimate the relationship between confounders and treatment, whereas multivariable regression estimates the relationship between confounders and outcome. In the case treatment is common but the number of events is low, there is often enough data to adequately model the relations between (many) confounders and the treatment in a propensity score model, but too little data to estimate the relations between confounders and the outcome in a traditional regression analysis in which the confounders are included as separate covariates (35). In our example, the outcome was relatively common since 963 patients (49.3%) died. Therefore, the multivariable regression analysis could easily incorporate eight confounders and the estimated HR was similar to those obtained when using the different propensity score methods.

Fourthly, in multivariable regression analysis the relation between each of the separate confounders and the outcome needs to be modelled properly in order to appropriately adjust for confounding. On the other hand, propensity score methods do not model the relations between confounders and outcome but instead model the relations between confounders and treatment. Whether this was done correctly, i.e., whether it resulted in balance of confounders between treatment groups, can be checked formally and hence provides an opportunity for correction of modeling errors.

Caveats of propensity score methods

In certain situations it is not advisable to use PS methods. In case of rare exposures or treatment, there may be insufficient data to model the relationship between confounders and treatment (i.e. in the propensity score model) and traditional regression analysis may be preferred. Also in the case of continuous treatment variables, PS methods may prove to be challenging (36). Furthermore, when the propensity score model is not correctly specified, matching (or weighting) on the propensity score may not yield the balance in the confounder distribution that is aimed for.

Lastly, although PS methods have been compared to randomized controlled trials, propensity score methods cannot control for unmeasured confounding. Like randomization is expected to do in trials, propensity score methods balance measured confounders between treatment groups. This property of the propensity score becomes clear when looking at, e.g., **Table 1**, which shows the balance achieved by PS matching and PS weighting in our example data. However, the analogy between PS methods and randomization does not apply to unmeasured confounding. For instance, unmeasured confounding may arise when clinicians use their expert knowledge and sometimes gut feeling to decide whether a patient has an indication for a certain treatment or not, and this judgment may be based on unmeasured characteristics such as severity of disease or frailty of a patient, a phenomenon which is commonly known as confounding by indication.

Key points

- The propensity score (PS) is the probability of treatment (or exposure) assignment, conditional on measured confounders.
- Propensity scores can be used to control for measured confounding in observational studies of medical treatments or risk factors, since measured confounders are balanced between treatment groups within levels of the PS.
- Methods to control for confounding using the PS include PS matching, PS stratification, multivariable regression analysis including the PS as a covariate, and PS weighting.
- The different PS methods and multivariable regression have the same potential to control for confounding. Choosing between these methods depends on the data properties and the treatment effect the researcher wants to estimate.
- PS methods as well as multivariable regression analysis cannot control for unmeasured confounding.

Conclusion

Propensity scores are often used in observational studies and it is important to understand their advantages and limitations. Traditional multivariable regression analysis and propensity score methods have been shown to give very similar results in data in which both methods could be applied, i.e., relatively large datasets (37, 38). Also in our illustrative example, similar effect estimates were found between PS matching, PS adjustment, PS weighting, and conventional multivariable Cox regression analysis. However, all these effect estimates potentially have different interpretations. Whether PS methods should be used instead of multivariable regression depends on the properties of the data at hand and which treatment effect the researcher wants to estimate. Neither of these methods, however, controls for unmeasured or unknown confounding. As a result, unmeasured confounding, for example as a consequence of confounding by indication, is not controlled for by propensity score methods and may still impact the validity of an observational study on the effects of treatments or risk factors in nephrology.

References

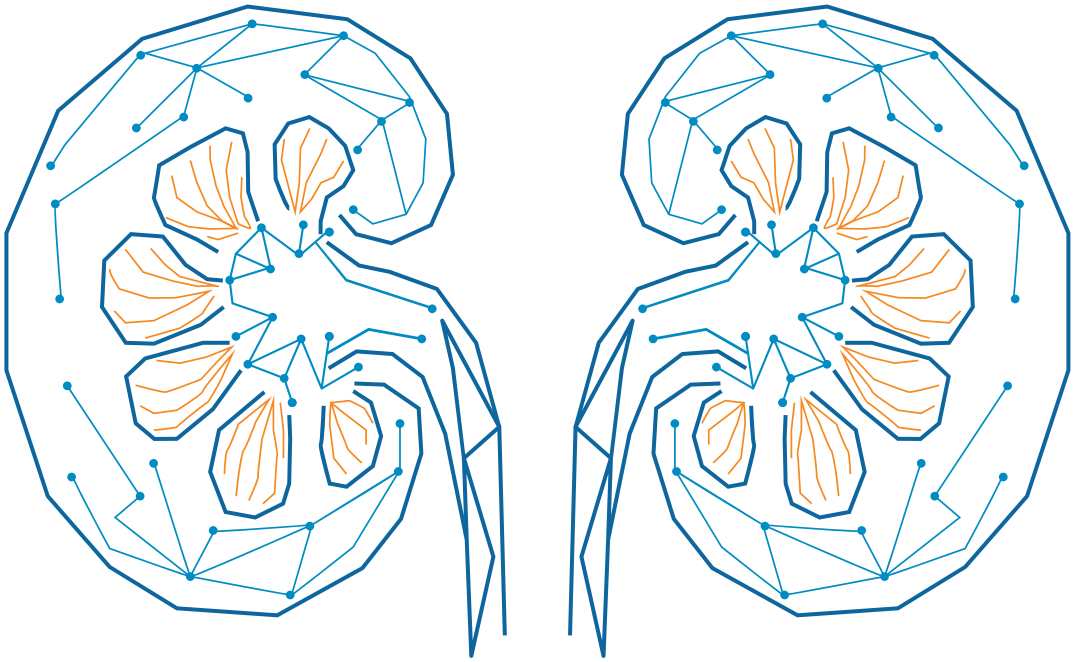
1. Jager KJ, Zoccali C, Macleod A, Dekker FW. Confounding: what it is and how to deal with it. Kidney Int. 2008;73(3):256-60.
2. Stel VS, Dekker FW, Zoccali C, Jager KJ. Instrumental variable analysis. Nephrol Dial Transplant. 2013;28(7):1694-9.
3. Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika. 1983;70(1):41-55.
4. Joffe MM, Rosenbaum PR. Invited commentary: propensity scores. Am J Epidemiol. 1999;150(4):327-33.
5. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Linear and logistic regression analysis. Kidney Int. 2008;73(7):806-10.
6. Westreich D, Lessler J, Funk MJ. Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. J Clin Epidemiol. 2010;63(8):826-33.
7. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol. 2002;155(2):176-84.
8. Robins JM. Data, design, and background knowledge in etiologic inference. Epidemiology. 2001;12(3):313-20.
9. Groenwold RH, Klungel OH, Grobbee DE, Hoes AW. Selection of confounding variables should not be based on observed associations with exposure. Eur J Epidemiol. 2011;26(8):589-93.
10. Suttorp MM, Siegerink B, Jager KJ, Zoccali C, Dekker FW. Graphical presentation of confounding in directed acyclic graphs. Nephrol Dial Transplant. 2015;30(9):1418-23.
11. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999;10(1):37-48.
12. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. Am J Epidemiol. 2006;163(12):1149-56.
13. 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.
14. Guertin JR, Rahme E, Dormuth CR, LeLorier J. Head to head comparison of the propensity score and the high-dimensional propensity score matching methods. BMC Med Res Methodol. 2016;16:22.
15. Toh S, Garcia Rodriguez LA, Hernan MA. Confounding adjustment via a semi-automated high-dimensional propensity score algorithm: an application to electronic medical records. Pharmacoepidemiol Drug Saf. 2011;20(8):849-57.
16. Cochran WG. The effectiveness of adjustment by subclassification in removing bias in observational studies. Biometrics. 1968;24(2):295-313.
17. Ali MS, Groenwold RH, Belitser SV, Pestman WR, Hoes AW, Roes KC, et al. Reporting of covariate selection and balance assessment in propensity score analysis is suboptimal: a systematic review. J Clin Epidemiol. 2015;68(2):112-21.
18. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences

in means and differences in proportions in observational studies. Pharm Stat. 2011;10(2):150-61.

19. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res. 2011;46(3):399-424.
20. Austin PC. A comparison of 12 algorithms for matching on the propensity score. Stat Med. 2014;33(6):1057-69.
21. Gu XS, Rosenbaum PR. Comparison of multivariate matching methods: Structures, distances, and algorithms. Journal of Computational and Graphical Statistics. 1993;2(4):405-20.
22. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association. 1984;79:516-24.
23. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med. 1998;17(19):2265-81.
24. Merkus MP, Jager KJ, Dekker FW, Boeschoten EW, Stevens P, Krediet RT. Quality of life in patients on chronic dialysis: self-assessment 3 months after the start of treatment. The Necosad Study Group. Am J Kidney Dis. 1997;29(4):584-92.
25. Termorshuizen F, Korevaar JC, Dekker FW, Van Manen JG, Boeschoten EW, Krediet RT, et al. Hemodialysis and peritoneal dialysis: comparison of adjusted mortality rates according to the duration of dialysis: analysis of The Netherlands Cooperative Study on the Adequacy of Dialysis 2. J Am Soc Nephrol. 2003;14(11):2851-60.
26. van Diepen M, Schroijen MA, Dekkers OM, Rotmans JI, Krediet RT, Boeschoten EW, et al. Predicting mortality in patients with diabetes starting dialysis. PLoS One. 2014;9(3):e89744.
27. Penning de Vries B, Groenwold R. A comparison of approaches to implementing propensity score methods following multiple imputation. Epidemiology Biostatistics and Public Health. 2017;14(4):e12630-1-e-21.
28. Penning de Vries B, Groenwold R. Comments on propensity score matching following multiple imputation. Stat Methods Med Res. 2016;25(6):3066-8.
29. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
30. Sturmer T, Rothman KJ, Glynn RJ. Insights into different results from different causal contrasts in the presence of effect-measure modification. Pharmacoepidemiol Drug Saf. 2006;15(10):698-709.
31. Deb S, Austin PC, Tu JV, Ko DT, Mazer CD, Kiss A, et al. A Review of Propensity-Score Methods and Their Use in Cardiovascular Research. Can J Cardiol. 2016;32(2):259-65.
32. Ali MS, Groenwold RH, Klungel OH. Best (but oft-forgotten) practices: propensity score methods in clinical nutrition research. Am J Clin Nutr. 2016;104(2):247-58.
33. Chen Q, Nian H, Zhu Y, Talbot HK, Griffin MR, Harrell FE, Jr. Too many covariates and too few cases? - a comparative study. Stat Med. 2016;35(25):4546-58.
34. Brookhart MA, Wyss R, Layton JB, Sturmer T. Propensity score methods for confounding control in nonexperimental research. Circ Cardiovasc Qual Outcomes. 2013;6(5):604-11.
35. Braitman LE, Rosenbaum PR. Rare outcomes, common treatments: analytic strategies using propensity scores. Ann Intern Med. 2002;137(8):693-5.
36. Naimi AI, Moodie EE, Auger N, Kaufman JS. Constructing inverse probability weights for continuous exposures: a comparison of methods. Epidemiology. 2014;25(2):292-9.

37. Shah BR, Laupacis A, Hux JE, Austin PC. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. J Clin Epidemiol. 2005;58(6):550-9.

38. Sturmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. J Clin Epidemiol. 2006;59(5):437-47.



CHAPTER 4

Comment on Kwon et al. The long-term effects of metformin on patients with type 2 diabetic kidney disease

Edouard L. Fu, Merel van Diepen

Diabetes Care 2020; 43: e190

We read with interest the recent study by Kwon et al. examining the long-term effects of metformin use on mortality and incident end stage kidney disease (ESKD) among 10,426 patients with type 2 diabetic kidney disease (1). The authors report that metformin use decreased the risk of all-cause mortality by 35% and ESKD progression by 33% after applying propensity score matching to adjust for a number of measured confounders.

We are concerned by the possible influence of immortal time bias on the study results. Immortal time arises when patients are classified into treatment groups at baseline based on treatment information that is only available *after* baseline (2). Since the treatment group is based on future information, by definition no deaths can occur in the treatment group between baseline and this future point in time. After all, individuals who have an event prior to taking up treatment would be classified into the untreated group. In this study, follow-up started on the date of the first creatinine measurement, but patients were classified as metformin users if they were prescribed metformin for longer than 90 days during the follow-up period (1). Such exposure classification may lead to an unfair survival advantage for the metformin users. For example, if all individuals in the metformin group started metformin treatment only after 5 years of follow-up, no deaths would occur in the metformin group during the first 5 years. The metformin group would thus be "immortal" for this time period. Due to the long-term follow-up of this study (maximum follow-up was 16 years), immortal time may have substantially biased the study results.

Immortal time bias could have been prevented by correctly assigning the persontime between start of follow-up and treatment initiation to the untreated group, e.g. by using a Cox model with a time-varying exposure (2). Individuals will then contribute persontime to the unexposed group before metformin initiation and to the exposed group after metformin initiation. When using a time dependent exposure, time-dependent confounding will also be present. If these time-dependent confounders play both the role of confounder and mediator, simply adjusting for them in a regression model will produce biased results. For example, HbA_{1c} is influenced by prior metformin treatment status, but also influences future metformin treatment status. Therefore HbA_{1c} will both confound and mediate the effect of metformin on mortality and a straightforward time-dependent Cox analysis may not suffice in this case (3). Instead, methods such as marginal structural models based on inverse probability weighting should be applied (3). Other methods that could have been used to avoid immortal time bias include landmarking (4) or the use of grace periods (5).

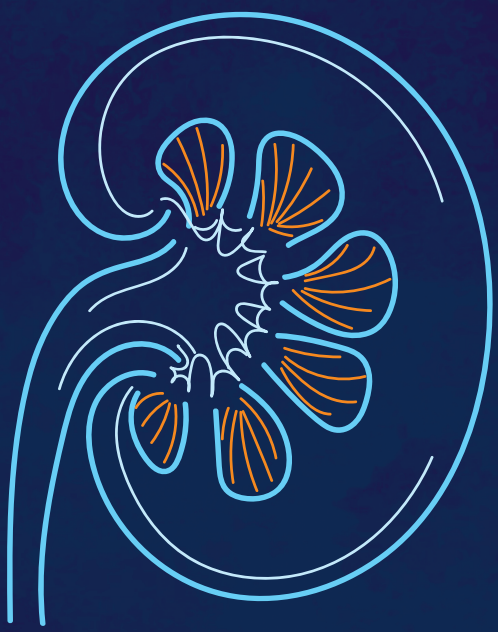
In conclusion, we feel the possibility of immortal time bias casts serious doubt on the validity of the results. Observational pharmacoepidemiologic studies must be designed and analyzed properly. Only then can the results of these studies meaningfully inform clinical practice.

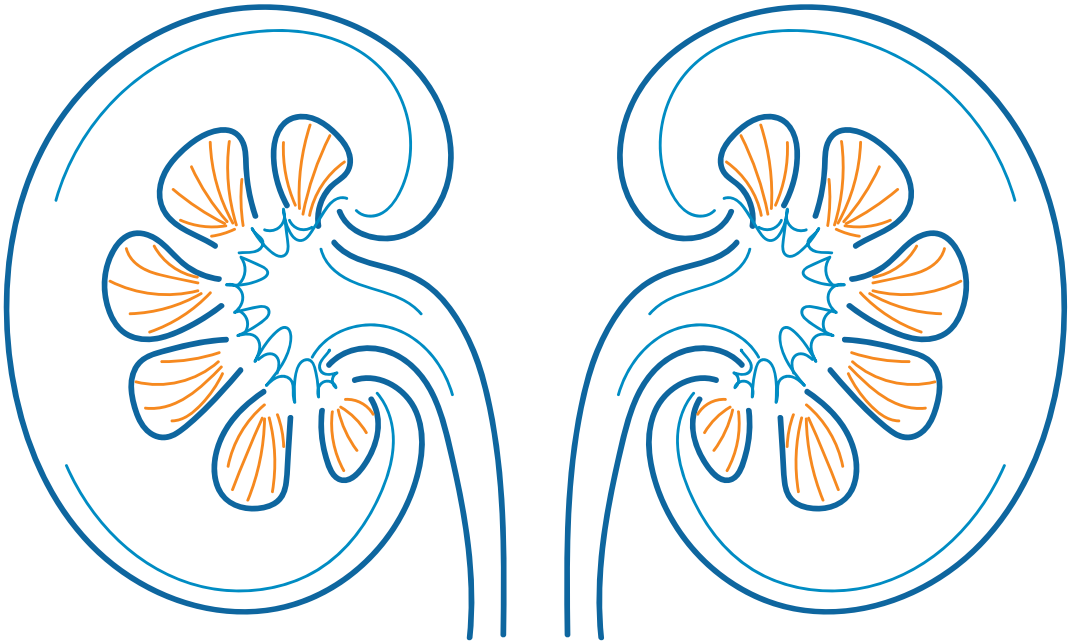
References

1. Kwon S, Kim YC, Park JY, Lee J, An JN, Kim CT, et al. The Long-term Effects of Metformin on Patients With Type 2 Diabetic Kidney Disease. Diabetes Care. 2020;43(5):948-55.
2. Suissa S. Immortal time bias in pharmaco-epidemiology. Am J Epidemiol. 2008;167(4):492-9.
3. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. BMJ. 2017;359:j4587.
4. Dafni U. Landmark analysis at the 25-year landmark point. Circ Cardiovasc Qual Outcomes. 2011;4(3):363-71.
5. Maringe C, Benitez Majano S, Exarchakou A, Smith M, Rachet B, Belot A, et al. Reflections on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. Int J Epidemiol. 2020.

PART II

Clinical applications in kidney disease





CHAPTER 5

Association of acute increases in plasma creatinine after renin-angiotensin blockade with subsequent outcomes

Edouard L. Fu, Marco Trevisan, Catherine M. Clase, Marie Evans, Bengt Lindholm, Joris I. Rotmans, Merel van Diepen, Friedo W. Dekker, Juan-Jesus Carrero

Clin J Am Soc Nephrol 2019; 14: 1336-1345

Abstract

Background and objectives. Data from observational and interventional studies provide discordant results regarding the relationship between creatinine increase following renin-angiotensin system inhibition (RASi) and adverse outcomes. We compared health outcomes among patients with different categories of increase in creatinine upon initiation of RASi in a large population-based cohort.

Design, setting, participants, and measurements. We performed a retrospective analysis of the Stockholm CREAtinine Measurements database, which contains complete information on diagnoses, medication dispensation claims, and laboratory test results for all Stockholm citizens accessing healthcare. Included were 31 951 adults initiating RASi during 2007-2011 with available pre- and post-initiation creatinine monitoring. Multivariable Cox regression was used to compare mortality, cardiovascular and end-stage kidney disease (ESKD) events among individuals with different ranges of creatinine increases within 2 months after starting treatment.

Results. In a median follow-up of 3.5 years, acute increases in creatinine were associated with mortality (3202 events) in a graded manner: compared with creatinine increases <10%, a 10-19% increase showed an adjusted HR of 1.15 (95% CI 1.05-1.27); HR 1.22 (1.07-1.40) for 20-29%; HR 1.55 (1.36-1.77) for ≥30%. Similar graded associations were present for heart failure (2275 events, p for trend <0.001) and ESKD (52 events; p for trend <0.001), and, less consistently, myocardial infarction (842 events, p for trend 0.25). Results were robust across subgroups, among continuing users, when patients with decreases in creatinine were excluded from the reference group, and after accounting for death as a competing risk.

Conclusions. Among real-world monitored adults, increases in creatinine (>10%) following initiation of RASi are associated with worse health outcomes. These results do not address the issue of discontinuation of RASi when plasma creatinine increases but do suggest that patients with increases in creatinine have higher subsequent risk of cardiovascular and kidney outcomes.

Introduction

Angiotensin-converting enzyme inhibitors (ACEi) and angiotensin receptor blockers (ARBs), known collectively as renin-angiotensin inhibitors (RASi), are widely prescribed drugs that are cornerstones in the treatment of hypertension, heart failure and proteinuric kidney disease (1, 2). Acute increases in creatinine are often observed after initiation of RASi, but the clinical significance of such increases is controversial (3-10). Current clinical guidelines recommend monitoring of creatinine during the first weeks of RASi, and discontinuing if creatinine increases exceed 30% (1, 2, 11). The rationale for the 30% threshold is unclear (12, 13).

One previous healthcare-based study, and reanalysis of 5 trials, have attempted to identify the threshold of increase in creatinine associated with increased risk of patient-important outcomes (14-17). While two reports found no outcome association for acute decrease of >20% (14) or >15% (15) in estimated glomerular filtration rate (eGFR), two other analyses suggested that increases of plasma creatinine as small as 10% were associated with worse kidney and cardiovascular outcomes (16, 17). Lack of power, limited healthcare coverage and the use of variable time windows to define the increase in creatinine may explain these differences. A report from the U.S. National Kidney Foundation recently emphasized the need to clarify this conflicting and limited evidence (13).

We used a large healthcare-based Swedish population cohort to investigate the frequency of plasma creatinine increases following RASi and whether such increases are associated with adverse health outcomes.

Material and Methods

Data sources

We used data from the Stockholm CREAtinine Measurements (SCREAM) project, a healthcare utilization cohort including all adult residents in Stockholm in whom a creatinine level was measured between 2006 and 2011 (18). SCREAM includes data from about 1.3 million adults, corresponding to 68% of the population of the region for that period (18). Laboratory results were linked to other administrative databases with complete information on demographic data, healthcare use, diagnoses, vital status, validated kidney replacement therapy endpoints, and dispensed prescriptions at Swedish pharmacies. The study utilized only de-identified data and thus was deemed not to require informed consent. It was approved by the regional ethical review boards and the Swedish National Board of Welfare.

Study design

We included all adult (>18 years old) community-dwelling patients newly initiating RASi irrespective of indication, with a creatinine measured on or within 3 months before the dispensation date, and a post-initiation creatinine within two months after. This strict window of pre- and post-initiation monitoring was chosen to align with guideline recommendations as well as previous studies (12, 13, 15, 16). We defined new users as individuals receiving a new RASi dispensation, with no dispensation of a RASi in the preceding 12 months, to ensure that the dispensation was not a continuation of an existing prescription. Additional exclusion criteria were missing age or sex, eGFR <30 ml/min/1.73m^2 or undergoing kidney replacement therapy at RASi dispensation.

Exposure

The study exposure was an increase in creatinine within the first two months of RASi, calculated as the difference between the baseline and first follow-up measurement. We only used creatinine measurements from the ambulatory setting. Creatinine tests from inpatient care, emergency room visits and taken within 24 hours before or after hospital admission were excluded. The date of the follow-up creatinine measurement was the index date of the study; the main analysis was by intention to treat. We categorized the relative increase in creatinine as follows: <10% (reference), 10-19%, 20-29% and ≥30%. In Stockholm healthcare, all laboratory tests are measured by one of three laboratories (Aleris, Unilabs and Karolinska), all of which are captured in SCREAM. Creatinine was measured in plasma, with either an enzymatic or corrected Jaffe method (alkaline picrate reaction), both methods being traceable to isotope dilution mass spectroscopy standards. Creatinine values <25 or >1500 μmol/L were considered outliers and discarded.

Time on RASi

Using information on all subsequent RASi dispensations, we defined continuous use as a refilling of prescription within the prescribed pill supply, adding 45 days to account for stockpiling and events that occur shortly after stopping drug. We quantified the proportion of patients who discontinued RASi within 14 days of the follow-up creatinine, and performed sensitivity analyses using an "as-treated" design, censoring at discontinuation.

Outcomes

Study outcomes were ascertained via linkage with the government-run National Population Registry, which registers all deaths without loss to follow up, and

the National Patient Register with codes diagnoses for essentially all (>99%) hospitalizations. The primary outcome was all-cause mortality. Secondary outcomes were hospitalization or death due to heart failure (*international classification of disease (ICD)-10* code I50); myocardial infarction (I21-I22) and end-stage kidney disease (ESKD, defined as the composite of N18.5-N18.6 codes, kidney replacement therapy initiation recorded in the validated Swedish Renal Registry, or a clinically encountered outpatient eGFR <15 ml/min/1.73m^2, whichever occurred first).

Covariates

Study covariates included age, sex, eGFR, comorbidities (hypertension, diabetes mellitus, myocardial infarction, heart failure, arrhythmia, peripheral vascular disease, cerebrovascular disease, ischemic heart disease) and medications (beta blocker, calcium channel blocker, thiazide diuretic, loop diuretic, potassium-sparing diuretic, non-steroidal anti-inflammatory drug, statin) (definitions in **Table S1**). Comorbidities identified in this study used established algorithms with an 85-95% sensitivity or positive predictive value (19). Drug dispensation data were obtained from the Dispensed Drug Registry, a nationwide register with complete information on all prescribed drugs dispensed at Swedish pharmacies. The coverage of this register is considered virtually complete, as outpatient drugs prescriptions and dispensations in Sweden are linked to the citizen's unique personal identification number. eGFR was calculated using the CKD-EPI formula (20). We defined chronic kidney disease (CKD) as eGFR below 60 ml/min/1.73m^2 on the first creatinine measurement, and categorized patients according to KDIGO criteria: category G3a (eGFR 45-59 ml/min/1.73m^2) and G3b (eGFR 30-44 ml/min/1.73m^2) (21, 22).

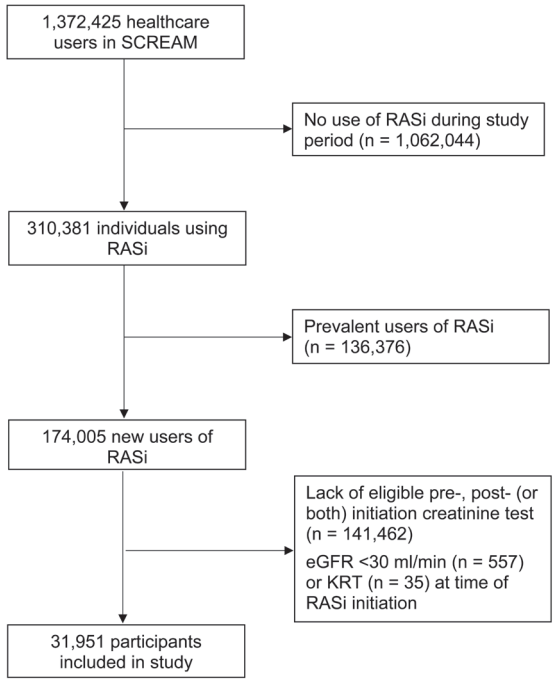
Statistical analyses

Continuous variables are presented as mean with standard deviation or median with interquartile range (IQR), depending on the distribution, and categorical variables as number and percentages. Patients were followed from dispensation of RASi until the occurrence of an event, emigration from Stockholm region or end of follow-up (2012 December 31), whichever occurred first. Cumulative incidence functions were calculated and plotted to account for the competing event mortality. Incidence rates per 1000 person years with 95% confidence intervals were calculated for each outcome. Multivariable Cox proportional hazards regression was used to calculate hazard ratios associated with creatinine increases as earlier defined. The proportional hazards assumption was checked using log-minus-log plots. Our primary analysis followed an intention-to-treat approach, assuming that RASi continued until occurrence of the first event or censoring (emigration or end of follow up). Next, we performed subgroup analyses for *a priori* defined strata: sex, comorbidities (diabetes, myocardial infarction, heart failure, hypertension, chronic

kidney disease) and treatment (ACEi, ARB or both). Finally, creatinine increase was also modelled as a continuous exposure using penalized smoothing splines. In order to elucidate short-term vs long-term risk associations, we performed time-varying Cox regression analysis splitting follow up in two intervals: <1 year and ≥1 year from baseline (23).

Sensitivity analyses included the following approaches: First, we followed an as-treated design censoring at RASi discontinuation. Second, we performed a competing risk analysis to calculate subdistribution hazards for the secondary study outcomes accounting for death as a competing risk. Third, we repeated the main analyses after excluding patients whose creatinine decreased by more than 10%. Fourth, we excluded all patients who developed hyperkalemia within the first three months of RASi (defined as an outpatient plasma potassium > 5.5 mmol/L). Lastly, we excluded patients who were hospitalized for heart failure or a myocardial infarction in the time window between the creatinine measurements. Missing data were rare, and no imputations were made. Statistical analyses were performed using R version 3.4.1 (24).

Figure 1. Flow chart of patient inclusion in the study. SCREAM = Stockholm CREAtinine Measurements project; RASi = renin-angiotensin system inhibition; eGFR = estimated glomerular filtration rate; KRT = kidney replacement therapy.



Results

A total of 174 005 new users of RASi were identified in Stockholm during 2007-2011 (**Figure 1**). Of these, 141 462 patients were excluded due to lack of an eligible pre- or post-initiation creatinine test (or both): 42 713 (30%) had a pre-initiation test, and 29 574 (21%) a post-initiation test. Of patients with chronic kidney disease (CKD; n=8273) on their pre-initiation test, 4852 (59%) had a post initiation test. An additional 592 patients were excluded for baseline eGFR <30 ml/min/1.73m^2 or kidney replacement therapy at time of RASi dispensation. The final study cohort consisted of 31 951 patients (18% of all identified new-users). For these patients, the median (interquartile range; IQR) number of days between the first creatinine measurement and start of RASi treatment was 14 (5-36), whereas median time between start of treatment and the second creatinine measurement was 19 (11-31) days.

Table 1. Baseline characteristics of new users of renin-angiotensin system inhibitors in the Stockholm CREAtinine Measurements project, overall and by increase in plasma creatinine (Cr) following drug initiation.

	Overall (n = 31 951)	**Cr increase <10% (n = 24 671)**	**Cr increase 10-19% (n = 4515)**	**Cr increase 20-29% (n = 1655)**	**Cr increase >=30% (n = 1110)**
Mean age (SD), *y*	65 (14)	64 (14)	66 (14)	66 (14)	69 (15)
Age category, *n (%)*					
<50	4344 (14)	3473 (14)	565 (13)	197 (12)	109 (10)
50-59	6353 (20)	5010 (20)	871 (19)	308 (19)	164 (15)
60-69	9101 (29)	7163 (29)	1222 (27)	456 (28)	260 (23)
70-79	7119 (22)	5429 (22)	1070 (24)	358 (22)	262 (24)
>=80	5034 (16)	3596 (15)	787 (17)	336 (20)	315 (28)
Women, *n (%)*	15768 (49)	12110 (49)	2195 (49)	845 (51)	618 (56)
Mean eGFR (SD), *ml/min/1.73m^2*	82 (19)	81 (18)	85 (19)	86 (20)	84 (22)
eGFR category, *ml/min/1.73m^2*					
30-44	1276 (4)	974 (4)	159 (4)	80 (5)	63 (6)
45-59	3017 (9)	2381 (10)	362 (8)	138 (8)	136 (12)
>=60	27658 (87)	21316 (87)	3994 (89)	1437 (87)	911 (82)
Comorbidities, *n (%)*					
Diabetes mellitus	6101 (19)	4693 (19)	854 (19)	315 (19)	239 (21)
Myocardial infarction	2140 (7)	1514 (6)	347 (8)	149 (9)	130 (12)
Heart failure	3969 (12)	2670 (11)	622 (14)	305 (18)	372 (34)
Hypertension	23374 (73)	18192 (74)	3268 (72)	1169 (71)	745 (67)

	Overall (n = 31 951)	Cr increase <10% (n = 24 671)	Cr increase 10-19% (n = 4515)	Cr increase 20-29% (n = 1655)	Cr increase >=30% (n = 1110)
Arrhythmia	4807 (15)	3500 (14)	689 (15)	315 (19)	303 (27)
Peripheral vascular disease	1362 (4)	969 (4)	214 (5)	93 (6)	86 (8)
Cerebrovascular disease	2136 (7)	1587 (6)	325 (7)	116 (7)	108 (10)
Ischemic heart disease	4587 (14)	3376 (14)	708 (16)	271 (16)	232 (21)
Chronic obstructive pulmonary disease	1826 (6)	1277 (5)	288 (6)	142 (9)	119 (11)
Medication, *n (%)*					
Beta blockers	12691 (40)	9487 (39)	1852 (41)	749 (45)	603 (54)
Calcium channel blockers	7265 (23)	5504 (22)	1053 (23)	406 (25)	302 (27)
Thiazides	2568 (8)	1965 (8)	353 (8)	153 (9)	97 (9)
Loop diuretics	4983 (16)	3305 (13)	822 (18)	400 (24)	456 (41)
Potassium-sparing diuretics	1507 (5)	976 (4)	228 (5)	108 (7)	195 (18)
NSAIDs	5951 (19)	4533 (18)	874 (19)	314 (19)	230 (21)
Statins	9666 (30)	7420 (30)	1422 (32)	484 (29)	340 (33)

eGFR = estimated glomerular filtration rate; SD = standard deviation; NSAIDs = non-steroidal anti-inflammatory drugs.

The characteristics of included patients are described in **Table 1**, overall and by increase in creatinine. Patients had a mean age of 65 years, 49% were women and 13% had CKD. Hypertension (73%), diabetes mellitus (19%), arrhythmias (15%) and ischemic heart disease (14%) were the most common comorbidities. Concurrent use of beta blockers (40%), statins (30%) and calcium-channel blockers (23%) was also common. Creatinine increases of 10-19% occurred in 4515 patients (14%), of 20-39% in 1655 (5%) and ≥30% in 1110 (4%). Patients with higher creatinine increases were on average older, had more comorbidities and a higher proportion were taking additional medications. Excluded patients (i.e., those with missing baseline or follow-up creatinine measurement) differed from those included in several ways, being in general younger, with higher GFR and lower prevalence of comorbidities (**Table S2**).

Association between creatinine increase and study outcomes

During a median follow up of 3.5 (IQR 2.1-4.7) years, there were 3202 deaths, 2275 heart failure hospitalizations, 842 myocardial infarctions and 52 ESKD events; incidence rates (95% CI) were 29.4 (28.4-30.4), 21.7 (20.8-22.6), 7.8 (7.3-8.4) and 0.5 (0.4-0.6) per 1000 person years, respectively.

Table 2. Crude and adjusted hazard ratios for the association between plasma creatinine increase category and death, cardiovascular or end-stage kidney disease outcomes.

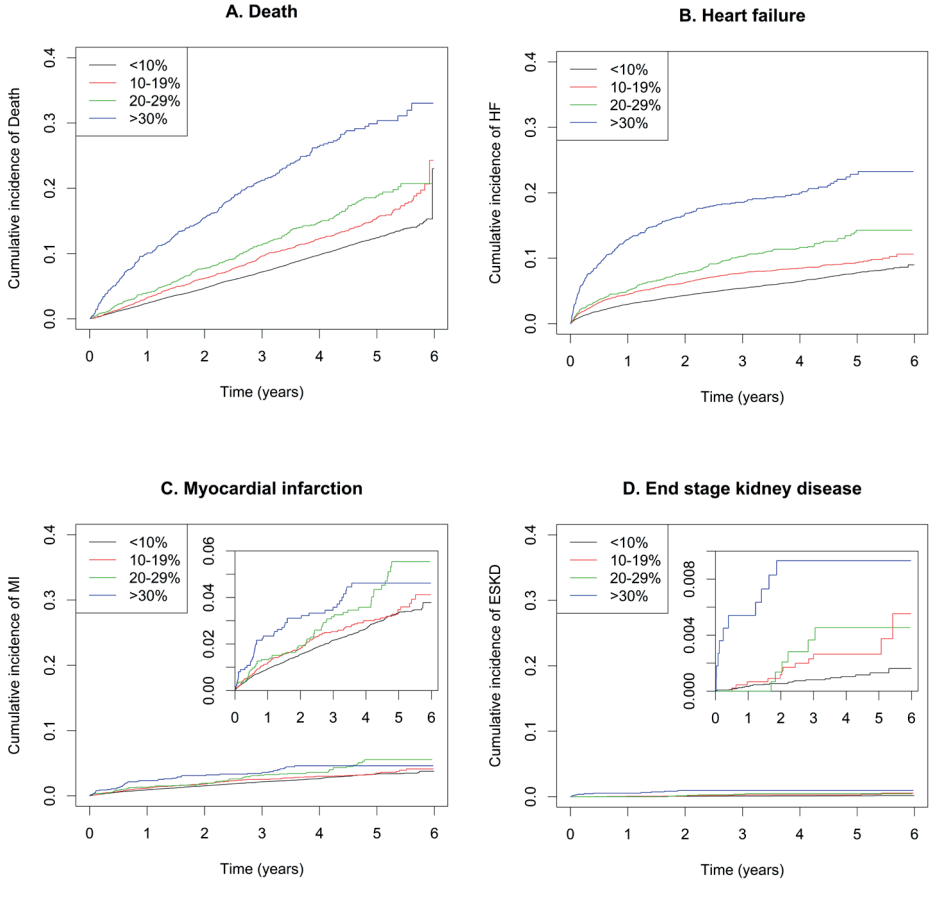
	n	Number of events	Person Years	IR per 1000PY (95% CI)	Crude HR (95% CI)	Adjusted HR (95% CI)[a]	P value for trend
Death							
Overall	31951	3202	108897	29.4 (28.4-30.4)			
<10%	24671	2183	84247	25.9 (24.8-27.0)	1	1	<0.001
10-19%	4515	522	15533	33.6 (30.8-36.6)	1.30 (1.18-1.43)	1.15 (1.05-1.27)	
20-29%	1655	231	5660	40.8 (35.7-46.4)	1.58 (1.38-1.80)	1.22 (1.07-1.40)	
>30%	1110	266	3457	76.9 (68.0-86.8)	2.98 (2.62-3.38)	1.55 (1.36-1.77)	
Heart failure							
Overall	31951	2275	104859	21.7 (20.8-22.6)			
<10%	24671	1505	81627	18.4 (17.5-19.4)	1	1	<0.001
10-19%	4515	367	14830	24.7 (22.3-27.4)	1.34 (1.20-1.51)	1.14 (1.02-1.28)	
20-29%	1655	184	5327	34.5 (29.7-39.9)	1.87 (1.60-2.18)	1.23 (1.05-1.43)	
>30%	1110	219	3075	71.2 (62.1-81.3)	3.73 (3.24-4.30)	1.41 (1.21-1.63)	
Myocardial infarction							
Overall	31951	842	107357	7.8 (7.3-8.4)			
<10%	24671	608	83126	7.3 (6.7-7.9)	1	1	0.25
10-19%	4515	127	15287	8.3 (6.9-9.9)	1.14 (0.94-1.38)	1.05 (0.86-1.27)	
20-29%	1655	62	5559	11.2 (8.6-14.3)	1.53 (1.18-1.98)	1.32 (1.02-1.72)	
>30%	1110	45	3385	13.3 (9.7-17.8)	1.80 (1.33-2.44)	1.29 (0.94-1.76)	
End-stage kidney disease							
Overall	31951	52	108815	0.5 (0.4-0.6)			
<10%	24671	24	84218	0.3 (0.2-0.4)	1	1	<0.001
10-19%	4515	12	15517	0.8 (0.4-1.4)	2.72 (1.36-5.44)	3.25 (1.61-6.53)	
20-29%	1655	6	5644	1.1 (0.4-2.3)	3.73 (1.53-9.13)	2.65 (1.05-6.72)	
>30%	1110	10	3435	2.9 (1.4-5.4)	10.13 (4.84-21.18)	8.31 (3.87-17.83)	

CI = confidence interval; IR = incidence rate; PY = person years; HR = hazard ratio.

[a] Analyses are adjusted for age, sex, diabetes mellitus, myocardial infarction, heart failure, hypertension, arrhythmia, peripheral vascular disease, eGFR, cerebrovascular disease, ischemic heart disease, use of beta blockers, calcium channel blockers, thiazide diuretics, loop diuretics, potassium-sparing diuretics, NSAIDs and statins.

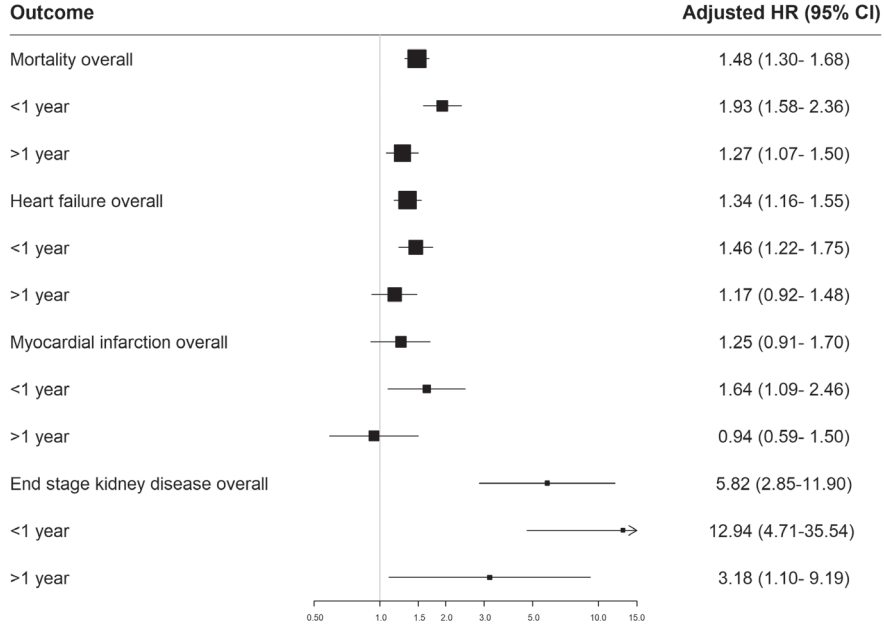
Figure 2 and **Table 2** show that in both crude and multivariable-adjusted models, there was a gradually increased risk of events with larger creatinine increases. For instance, the risk of death, HR (95% CI), was 1.15 (1.05-1.27), 1.22 (1.07-1.40) and 1.55 (1.36-1.77) times higher for increases of 10-19%, 20-29% and ≥30% respectively, compared with patients with creatinine increase <10% (p for trend < 0.001). Similar trends were observed for the outcomes heart failure and ESKD (p values for trend <0.001). The association was less robust for the outcome myocardial infarction (p for trend 0.25), but creatinine increases of 20-29% and ≥30% were associated with tendencies toward increased risk in multivariable analysis. Stratified analyses showed similar associations and tendencies, with wider confidence intervals (**Table S3**).

Figure 2. Cumulative incidence plots for death, cardiovascular and end-stage kidney disease outcomes by ranges of plasma creatinine increase during the first 2 months after RASi treatment initiation. The cumulative incidence plots for heart failure, myocardial infarction and end-stage kidney disease account for the competing risk of death.



For comparison with preceding literature and current guideline recommendations, we compared mortality, kidney and cardiovascular risks associated with increases ≥30% versus <30%. Patients with increases ≥30% were older, had more comorbidities and used more medications (**Table S4**). Creatinine increases ≥30% were associated with an increased risk for all studied outcomes, overall (**Figure 3**; **Table S5**) and across different subgroups (**Figure S1**). In restricted follow-up analyses, the associations were apparent during both short- and long-term follow up, but with consistent tendencies towards higher risk magnitude during the first year of observation (**Figure 3**).

Figure 3. Adjusted hazard ratios for the association between creatinine increase ≥30%, compared with <30%, and death, cardiovascular or end-stage kidney disease outcomes, overall, and within or beyond the first year of follow-up. HR = hazard ratio. The hazard ratio beyond the first year of follow-up is calculated conditional on surviving the first year.



When modelling creatinine change as a continuous exposure through spline curves, we observed an asymmetrical U-shaped association, with the lowest risk at decreases in creatinine of 5 to 20%, for the outcomes death and myocardial infarction. In contrast, for risk of heart failure and ESKD the association was linear (**Figure S2**). The inclusion of patients whose creatinine acutely decreased in the

reference group might alter the effect size associated with increases in creatinine. We therefore tested whether redefining the reference category by excluding patients with creatinine decreases ≥10% would modify our observations; we observed no major deviation from our main results (**Table S6**).

The median length of RASi treatment was 18 (IQR 8-33) months. In the as-treated sensitivity analysis, censoring at RASi discontinuation, effects were similar to the intention-to-treat main analysis (**Table S7**). Of note, 239 (22%) patients with creatinine increase ≥30% discontinued treatment immediately after, compared with 5477 (18%) in those with creatinine increase of <30% (crude relative risk 1.19; 95% CI 1.06-1.33). Potassium was measured at least once within the first 3 months of RASi in 29 152 patients (91%), and there were 241 patients with hyperkalemia (> 5.5 mmol/L). Hyperkalemia was more common among patients with creatinine increases ≥ 30% (5% of patients) than in those with creatinine increases <30% (0.6% of patients, relative risk of hyperkalemia 8.02; 95% CI 5.96-10.80), **Table S8**). Exclusion of patients with concurrent hyperkalemia did not modify our main observations (**Table S9**). Competing risk models accounting for death showed also similar associations regarding the risk of cardiovascular and ESKD events (**Table S10**). Finally, we compared study outcomes in included versus excluded (i.e., unmonitored) patients. Patients who were included in our analysis had a higher risk of death, but no increased risk for hospitalization for heart failure, myocardial infarction or ESKD compared with patients in whom creatinine was not monitored (**Table S11**). Exclusion of patients that were hospitalized in the time window between the creatinine measurements did also not modify the results (**Table S12**).

Discussion

In this large healthcare-based observational study, we found that i) 18% of adults initiating RASi underwent pre- and post-initiation creatinine monitoring according to current guideline recommendations; ii) creatinine increases of 10-29% within the first 2 months of RASi were common among monitored individuals, occurring in 19% of patients, and increases of 30% or more occurred in 4%; iii) acute increases in creatinine of any magnitude above 10%, relative to baseline, were consistently associated, in a graded manner, with increased subsequent risk of death, cardiovascular events and ESKD.

Clinical guidelines recommend monitoring creatinine and considering discontinuation or dose reduction of RASi if creatinine increases by 30% or more (1, 11). We found that 18% of all new users of RASi in our region underwent guideline-recommended creatinine monitoring. This is in keeping with most observations from other countries and healthcare systems (25-32), for example,

in a UK primary health care cohort, 14% of patients were monitored before and after (31). However, in a US health maintenance organization 70% of patients were monitored (33). Comparing monitoring practices between studies is problematic because of differing definitions, data collection periods, database quality and coverage. In Stockholm healthcare, laboratory tests are centrally measured by three different laboratories, all of which contribute to SCREAM, which ensures that our cohort includes all creatinine measurements. Though the proportion of patients monitored in many of these observational studies might be thought low, it is worth considering, in this context, that the recommendation for monitoring is not based on direct evidence of benefit from monitoring, but rather on extrapolation from clinical trials in which monitoring occurred. In these trials, the response to monitoring was not protocolized, and many patients with increases in creatinine likely stayed on drug (15, 17, 34). In our study, monitored patients were older, and had a higher comorbidity burden. The presence of monitoring, in adjusted analysis, was associated with outcomes that were worse than (death) or similar to (heart failure, myocardial infarction, ESKD) those in unmonitored patients, suggesting, to some extent, that its use is selective and directed at patients at higher risk.

Among those who were monitored, acute increases in creatinine were associated in a graded manner with increased subsequent mortality, kidney and cardiovascular events. Our results expand the findings of the UK primary health care cohort (16), but contrast with some analyses from other clinical trials: first, a post hoc evaluation from ONTARGET and TRANSCEND (n=9340) did not find decreases in eGFR of >= 15% to be associated with kidney or cardiovascular events, with an adjusted HR (95% CI) of 1.14 (0.93-1.39) for new micro-albuminuria and 1.17 (0.99-1.38) for the primary cardiovascular composite (15). Second, in analyses from AASK and MDRD (n=1660), acute eGFR decreases between 5-20% in the setting of intensive blood pressure control were not associated with the risk of ESKD, with an adjusted HR (95% CI) of 1.19 (0.84-1.68) for AASK and 1.08 (0.84-1.40) for the MDRD trial (14). In ADVANCE, increase in creatinine was associated with the composite outcome of mortality, major cardiovascular events, and new or worsening nephropathy, in a graded way: HR's were 1.1 (95% CI 1.0-1.3), 1.3 (1.1-1.7) and 1.4 (1.2-1.8) for increases in creatinine of 10-19%, 20-29%, and >=30% respectively, all compared with the referent category of increase <10%, with P for trend < 0.001 (17). However, it is noteworthy that half the patients contributing to these cohort analyses of ADVANCE were randomized to placebo after the active run in phase. Taken together, the tendencies and effects in these studies are in the direction of the effects that we observed, and the differences in statistical significance may reflect the greater power in the observational data sets.

Our finding that associations were stronger during the first year of follow up is a new observation. By demonstrating the asymmetry of the relationships across increase and decrease in creatinine, we have excluded the possibility that the results are

caused solely by variability itself as an adverse prognostic marker, though we recognize that variability may contribute to the magnitude of the observed effects (35, 36). We have also demonstrated that results are largely unchanged after the exclusion of patients whose creatinine significantly decreased, using patients with changes of ±10% as reference. Additional strengths include our use of a stricter definition for pre-initiation testing (a 3-month window), following criticisms that follow up after a 12-month window (13, 16) could reflect long-term progression of CKD rather than acute decreases in GFR. We based our exposure on pharmacy dispensations rather than prescriptions written, which offers better ascertainment, although we cannot ensure that the medication has been taken. We excluded patients with CKD G4-5, as the use of RASi is subject to other considerations in this patient group, and their inclusion might have driven kidney outcomes. Our work illustrates how healthcare-based analyses and clinical trials provide complementary information on benefits and harms of therapy (37).

RASi by ACEi/ARBs leads to the loss of glomerular efferent arteriolar vasoconstriction, which reduces intraglomerular pressure, resulting in an acute decrease in GFR (38, 39); mitigation of maladaptive hyperfiltration by this mechanism is thought to contribute to the kidney benefit of RASi. Studies of intensive blood pressure reduction suggest that decreases in GFR in this context reflect hemodynamic changes rather than intrinsic injury (40, 41), and after long-term empagliflozin, which is also thought to acutely decrease GFR through a hemodynamic mechanism, discontinuation of empagliflozin is followed by an acute increase in GFR (42). Acute hemodynamic change in GFR may therefore carry a different implication and prognosis than change secondary to progression.

The origin of the guideline recommendation to discontinue RASi after acute increases in creatinine >= 30% is unclear (13). It appears to have originated with an influential narrative review of 12 small trials (1102 participants), which concluded that creatinine increases of less than this magnitude were associated with more stable subsequent GFR in patients with CKD; methods and effect size for this conclusion were not shown, so it is difficult to make a direct comparison between these data and our own (12). In our larger, observational dataset, there is increased power to detect outcomes associated with more modest changes, and perhaps explains why we found that acute increases in creatinine of 10% or more were also associated with subsequent adverse events. The most significant limitation of our finding is that one cannot establish causality from this observational evidence: our results do not mean that RASi should be discontinued in any group. Instead, they are part of an emerging network of evidence that informs the decision to monitor and how to respond to monitoring, in the context of robust randomized evidence demonstrating reduction in patient-important kidney and cardiovascular outcomes with RASi (43-48). Because reanalyses of TRANSCEND and ADVANCE found no

evidence for modification of the benefit of RASi by level of creatinine increase (15, 17), we speculate that creatinine increases may therefore be a risk marker of disease rather than directly leading to adverse outcomes (49). We were unable to adjust for blood pressure and proteinuria, because blood pressure is not included in any linked database, and proteinuria data had a high degree of missingness that was unlikely to be random. In previous work, initial blood pressure is not associated with change in GFR, and for albuminuria the effect size is not strong (OR 1.2, 95% CI: 1.0-1.5) (15), so we believe they are unlikely to be important confounders. It is a limitation of our data that we were unable to comment on the persistence of the change: however, in ONTARGET and TRANSCEND, 50% of those with a decrease of GFR of ≥16 15% at 2 weeks did not have a difference of that magnitude at 8 weeks (15). For patients who are monitored and who experience an increase of 30% or more, repeating the value may therefore be helpful. Whether routinely discontinuing versus continuing RASi after a relevant creatinine increase would result in improved outcomes is outside the scope of our analysis because of the complexity of time-dependent confounding. We note that this is precisely the aim of an ongoing trial of patients with CKD G4-5 (50).

To conclude, acute increases in creatinine following initiation of RASi of 10% or more were robustly associated with increased risk of death, cardiovascular events (myocardial infarction, heart failure) and development of ESKD in an observational clinical setting. Monitoring creatinine before and after initiation of RASi identifies patients at high risk.

References

1. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, et al. 2018 ESC/ESH Guidelines for the management of arterial hypertension. Eur Heart J. 2018;39(33):3021-104.
2. Whelton PK, Carey RM, Aronow WS, Casey DE, Jr., Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2018;71(19):e127-e248.
3. Palmer BF. Angiotensin-converting enzyme inhibitors and angiotensin receptor blockers: what to do if the serum creatinine and/or serum potassium concentration rises. Nephrol Dial Transplant. 2003;18(10):1973-5.
4. Weir MR. Acute changes in glomerular filtration rate with renin-angiotensin system (RAS) inhibition: clinical implications. Kidney Int. 2017;91(3):529-31.
5. Ahmed A. Use of angiotensin-converting enzyme inhibitors in patients with heart failure and renal insufficiency: how concerned should we be by the rise in serum creatinine? J Am Geriatr Soc. 2002;50(7):1297-300.
6. Cohen DL, Townsend RR. What should the physician do when creatinine increases after starting an angiotensin-converting enzyme inhibitor or an angiotensin receptor blocker? J Clin Hypertens (Greenwich). 2008;10(10):803-4.
7. Epstein BJ. Elevations in serum creatinine concentration: concerning or reassuring? Pharmacotherapy. 2004;24(5):697-702; discussion -3.
8. Hirsch S, Hirsch J, Bhatt U, Rovin BH. Tolerating increases in the serum creatinine following aggressive treatment of chronic kidney disease, hypertension and proteinuria: pre-renal success. Am J Nephrol. 2012;36(5):430-7.
9. Navis G, Faber HJ, de Zeeuw D, de Jong PE. ACE inhibitors and the kidney. A risk-benefit assessment. Drug Saf. 1996;15(3):200-11.
10. Schoolwerth AC, Sica DA, Ballermann BJ, Wilcox CS, Council on the Kidney in Cardiovascular D, the Council for High Blood Pressure Research of the American Heart A. Renal considerations in angiotensin converting enzyme inhibitor therapy: a statement for healthcare professionals from the Council on the Kidney in Cardiovascular Disease and the Council for High Blood Pressure Research of the American Heart Association. Circulation. 2001;104(16):1985-91.
11. Kidney Disease: Improving Global Outcomes (KDIGO) Blood Pressure Work Group. KDIGO Clinical Practice Guideline for the Management of Blood Pressure in Chronic Kidney Disease. Kidney inter, Suppl. 2012;2:351.
12. Bakris GL, Weir MR. Angiotensin-converting enzyme inhibitor-associated elevations in serum creatinine: is this a cause for concern? Arch Intern Med. 2000;160(5):685-93.
13. Weir MR, Lakkis JI, Jaar B, Rocco MV, Choi MJ, Mattrix-Kramer H, et al. Use of Renin-Angiotensin System Blockade in Advanced CKD: An NKF-KDOQI Controversies Report. Am J Kidney Dis. 2018.
14. Ku E, Bakris G, Johansen KL, Lin F, Sarnak MJ, Campese VM, et al. Acute Declines in Renal Function during Intensive BP Lowering: Implications for Future ESRD Risk. J Am Soc Nephrol. 2017;28(9):2794-801.

15. Clase CM, Barzilay J, Gao P, Smyth A, Schmieder RE, Tobe S, et al. Acute change in glomerular filtration rate with inhibition of the renin-angiotensin system does not predict subsequent renal and cardiovascular outcomes. Kidney Int. 2017;91(3):683-90.
16. Schmidt M, Mansfield KE, Bhaskaran K, Nitsch D, Sorensen HT, Smeeth L, et al. Serum creatinine elevation after renin-angiotensin system blockade and long term cardiorenal risks: cohort study. BMJ. 2017;356:j791.
17. Ohkuma T, Jun M, Rodgers A, Cooper ME, Glasziou P, Hamet P, et al. Acute Increases in Serum Creatinine After Starting Angiotensin-Converting Enzyme Inhibitor-Based Therapy and Effects of its Continuation on Major Clinical Outcomes in Type 2 Diabetes Mellitus. Hypertension. 2019;73(1):84-91.
18. Runesson B, Gasparini A, Qureshi AR, Norin O, Evans M, Barany P, et al. The Stockholm CREAtinine Measurements (SCREAM) project: protocol overview and regional representativeness. Clin Kidney J. 2016;9(1):119-27.
19. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.
20. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12.
21. Kidney Disease Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney inter, Suppl. 2013;3:1-150.
22. Inker LA, Astor BC, Fox CH, Isakova T, Lash JP, Peralta CA, et al. KDOQI US commentary on the 2012 KDIGO clinical practice guideline for the evaluation and management of CKD. Am J Kidney Dis. 2014;63(5):713-35.
23. Dekker FW, de Mutsert R, van Dijk PC, Zoccali C, Jager KJ. Survival analysis: time-dependent effects and time-varying risk factors. Kidney Int. 2008;74(8):994-7.
24. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
25. Bootsma JE, Warle-van Herwaarden MF, Verbeek AL, Fussenich P, De Smet PA, Olde Rikkert MG, et al. Adherence to biochemical monitoring recommendations in patients starting with renin angiotensin system inhibitors: a retrospective cohort study in the Netherlands. Drug Saf. 2011;34(7):605-14.
26. Coleman JJ, McDowell SE, Evans SJ, Gill PS, Ferner RE. Oversight: a retrospective study of biochemical monitoring in patients beginning antihypertensive drug treatment in primary care. Br J Clin Pharmacol. 2010;70(1):109-17.
27. Hurley JS, Roberts M, Solberg LI, Gunter MJ, Nelson WW, Young L, et al. Laboratory safety monitoring of chronic medications in ambulatory care settings. J Gen Intern Med. 2005;20(4):331-3.
28. Kalra PA, Kumwenda M, MacDowall P, Roland MO. Questionnaire study and audit of use of angiotensin converting enzyme inhibitor and monitoring in general practice: the need for guidelines to prevent renal failure. BMJ. 1999;318(7178):234-7.
29. Mathieson L, Severn A, Guthrie B. Monitoring and adverse events in relation to ACE inhibitor/angiotensin receptor blocker initiation in people with diabetes in general practice: a population database study. Scott Med J. 2013;58(2):69-76.

30. McDowell SE, Coleman JJ, Evans SJ, Gill PS, Ferner RE. Laboratory monitoring and adverse patient outcomes with antihypertensive therapy in primary care. Pharmacoepidemiol Drug Saf. 2010;19(5):482-9.

31. Schmidt M, Mansfield KE, Bhaskaran K, Nitsch D, Sorensen HT, Smeeth L, et al. Adherence to guidelines for creatinine and potassium monitoring and discontinuation following renin-angiotensin system blockade: a UK general practice-based cohort study. BMJ Open. 2017;7(1):e012818.

32. van Blijderveen JC, Straus SM, de Ridder MA, Stricker BH, Sturkenboom MC, Verhamme KM. Adherence to renal function monitoring guidelines in patients starting antihypertensive therapy with diuretics and RAAS inhibitors: a retrospective cohort study. Drug Saf. 2014;37(5):369-77.

33. Simon SR, Andrade SE, Ellis JL, Nelson WW, Gurwitz JH, Lafata JE, et al. Baseline laboratory monitoring of cardiovascular medications in elderly health maintenance organization enrollees. J Am Geriatr Soc. 2005;53(12):2165-9.

34. Clark H, Krum H, Hopper I. Worsening renal function during renin-angiotensin-aldosterone system inhibitor initiation and long-term outcomes in patients with left ventricular systolic dysfunction. Eur J Heart Fail. 2014;16(1):41-8.

35. Lau KK, Wong YK, Chan YH, Teo KC, Chan KH, Wai Li LS, et al. Visit-to-visit blood pressure variability as a prognostic marker in patients with cardiovascular and cerebrovascular diseases--relationships and comparisons with vascular markers of atherosclerosis. Atherosclerosis. 2014;235(1):230-5.

36. Uehara K, Yasuda T, Shibagaki Y, Kimura K. Estimated Glomerular Filtration Rate Variability Independently Predicts Renal Prognosis in Advanced Chronic Kidney Disease Patients. Nephron. 2015;130(4):256-62.

37. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-World Evidence - What Is It and What Can It Tell Us? N Engl J Med. 2016;375(23):2293-7.

38. Denton KM, Fennessy PA, Alcorn D, Anderson WP. Morphometric analysis of the actions of angiotensin II on renal arterioles and glomeruli. Am J Physiol. 1992;262(3 Pt 2):F367-72.

39. Hall JE, Guyton AC, Jackson TE, Coleman TG, Lohmeier TE, Trippodo NC. Control of glomerular filtration rate by renin-angiotensin system. Am J Physiol. 1977;233(5):F366-72.

40. Nadkarni GN, Chauhan K, Rao V, Ix JH, Shlipak MG, Parikh CR, et al. Effect of Intensive Blood Pressure Lowering on Kidney Tubule Injury: Findings From the ACCORD Trial Study Participants. Am J Kidney Dis. 2018.

41. Zhang WR, Craven TE, Malhotra R, Cheung AK, Chonchol M, Drawz P, et al. Kidney Damage Biomarkers and Incident Chronic Kidney Disease During Blood Pressure Reduction: A Case-Control Study. Ann Intern Med. 2018;169(9):610-8.

42. Wanner C, Inzucchi SE, Lachin JM, Fitchett D, von Eynatten M, Mattheus M, et al. Empagliflozin and Progression of Kidney Disease in Type 2 Diabetes. N Engl J Med. 2016;375(4):323-34.

43. Savarese G, Costanzo P, Cleland JG, Vassallo E, Ruggiero D, Rosano G, et al. A meta-analysis reporting effects of angiotensin-converting enzyme inhibitors and angiotensin receptor blockers in patients without heart failure. J Am Coll Cardiol. 2013;61(2):131-42.

44. Casas JP, Chua W, Loukogeorgakis S, Vallance P, Smeeth L, Hingorani AD, et al. Effect of inhibitors of the renin-angiotensin system and other antihypertensive drugs on renal outcomes: systematic review and meta-analysis. Lancet. 2005;366(9502):2026-33.

45. Wu HY, Huang JW, Lin HJ, Liao WC, Peng YS, Hung KY, et al. Comparative effectiveness of renin-angiotensin system blockers and other antihypertensive drugs in patients with diabetes: systematic review and bayesian network meta-analysis. BMJ. 2013;347:f6008.

46. van Vark LC, Bertrand M, Akkerhuis KM, Brugts JJ, Fox K, Mourad JJ, et al. Angiotensin-converting enzyme inhibitors reduce mortality in hypertension: a meta-analysis of randomized clinical trials of renin-angiotensin-aldosterone system inhibitors involving 158,998 patients. Eur Heart J. 2012;33(16):2088-97.

47. Cheng J, Zhang W, Zhang X, Han F, Li X, He X, et al. Effect of angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers on all-cause mortality, cardiovascular deaths, and cardiovascular events in patients with diabetes mellitus: a meta-analysis. JAMA Intern Med. 2014;174(5):773-85.

48. Xie X, Liu Y, Perkovic V, Li X, Ninomiya T, Hou W, et al. Renin-Angiotensin System Inhibitors and Kidney and Cardiovascular Outcomes in Patients With CKD: A Bayesian Network Meta-analysis of Randomized Clinical Trials. Am J Kidney Dis. 2016;67(5):728-41.

49. Palmer BF. Renal dysfunction complicating the treatment of hypertension. N Engl J Med. 2002;347(16):1256-61.

50. 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.

Supplementary material

Supplementary Table S1. Definition of medications and comorbidities.

Supplementary Table S2. Characteristics of unmonitored and monitored patients.

Supplementary Table S3. Adjusted hazard ratios for the association between plasma creatinine increase and death, cardiovascular or end-stage kidney disease outcomes in different subgroups.

Supplementary Table S4. Characteristics at initiation of RASi, overall and according to plasma creatinine increase of < or ≥30% within the first 2 months of RASi.

Supplementary Table S5. Crude and adjusted hazard ratios for the association between plasma creatinine increases ≥30% and death, cardiovascular or end-stage kidney disease outcomes.

Supplementary Table S6. Sensitivity analysis: adjusted hazard ratios for the association between plasma creatinine increase and death, cardiovascular or end-stage kidney disease outcomes after exclusion of patients with creatinine decreases >10% following RASi.

Supplementary Table S7. Sensitivity analysis: adjusted hazard ratios for the association between plasma creatinine increase and death, cardiovascular or end-stage kidney disease outcomes censoring at time of RASi discontinuation.

Supplementary Table S8. Risk of hyperkalemia within the first 3 months of RASi overall and according to plasma creatinine increase categories.

Supplementary Table S9. Sensitivity analysis: adjusted hazard ratios for the association between plasma creatinine increase and death, cardiovascular or end-stage kidney disease outcomes after excluding patients developing hyperkalemia (plasma K^+ > 5.5 mmol/L) within the first 3 months of RASi (n = 241).

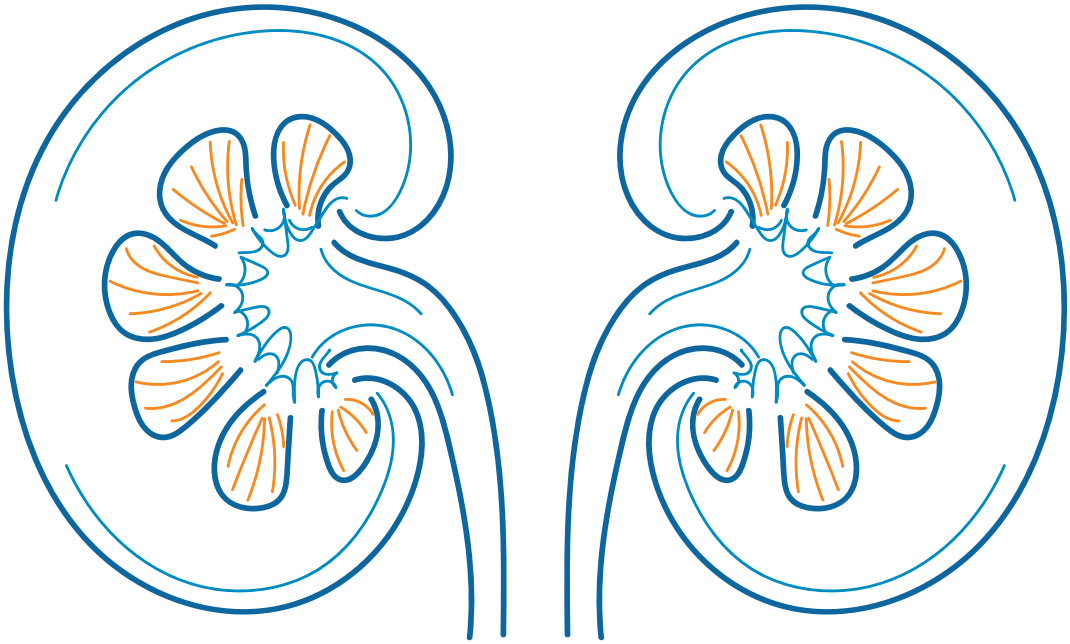
Supplementary Table S10. Sensitivity analysis: Fine and Gray competing risk analysis showing crude and adjusted subdistributional hazard ratios (sHR) for the association between plasma creatinine increase and cardiovascular or end-stage kidney disease outcomes outcomes with death (by other causes) as competing risk.

Supplementary Table S11. Crude and adjusted hazard ratios for adverse outcomes among monitored and unmonitored patients.

Supplementary Table S12. Sensitivity analysis: adjusted hazard ratios for the association between plasma creatinine increase and death, cardiovascular or end-stage kidney disease outcomes excluding patients with hospitalization between creatinine measurements.

Supplementary Figure S1. Adjusted hazard ratios for the association between plasma creatinine increase ≥30% vs. <30% in different subgroups for (**A)** mortality; (**B)** heart failure; (**C**) myocardial infarction; (**D**) end-stage kidney disease.

Supplementary Figure S2. Penalized smoothing spline curve associated with plasma creatinine increases (continuous variable) for (**A)** mortality; (**B)** heart failure; (**C**) myocardial infarction; (**D**) end-stage kidney disease.



CHAPTER 6

Comparative effectiveness of renin-angiotensin system inhibitors and calcium channel blockers in individuals with advanced CKD: a nationwide observational cohort study

Edouard L. Fu, Catherin M. Clase, Marie Evans, Bengt Lindholm, Joris I. Rotmans, Friedo W. Dekker, Merel van Diepen, Juan-Jesus Carrero

Am J Kidney Dis. 2020; 77(5): 719-729

Abstract

Rationale & Objective. It is unknown whether initiating a renin-angiotensin system inhibitor (RASi) in patients with advanced chronic kidney disease (CKD) is superior to alternative antihypertensive agents such as calcium channel blockers (CCB). We compared the risks of kidney replacement therapy (KRT), mortality and major adverse cardiovascular events (MACE) in patients with advanced CKD in routine nephrology practice who were initiating either RASi or CCB therapy.

Study Design. Observational study in the Swedish Renal Registry, 2007 to 2017.

Settings & Participants. 2458 new users of RASi and 2345 CCB users with estimated glomerular filtration rates (eGFR) <30 ml/min/1.73m^2 (CKD G4-5 without KRT) who were being followed up by a nephrologist. As a positive control cohort, new users of the same drugs in patients with CKD G3 (eGFR 30-60 ml/min/1.73m^2) were evaluated.

Exposures. RASi vs. CCB therapy initiation.

Outcome. Initiation of KRT (maintenance dialysis or transplantation), all-cause mortality and MACE (composite of cardiovascular death, myocardial infarction or stroke).

Analytical approach. Hazard ratios (HRs) with 95% CIs were estimated using propensity score-weighted Cox proportional hazards regression adjusting for demographic, clinical and laboratory covariates.

Results. Median age was 74 years, 38% were women and median follow-up was 4.1 years. After propensity score weighting, there was significantly lower risk of KRT after new use of RASi compared with new use of CCBs (adjusted HR 0.79; 95% CI 0.69-0.89), but similar risks of mortality (adjusted HR 0.97; 95% CI 0.88-1.07) and MACE (adjusted HR 1.00; 95% CI 0.88-1.15). Results were consistent across subgroups and in as-treated analyses. The positive control cohort of patients with CKD G3 showed similar KRT risk reduction (0.67; 0.56-0.80) with RASi therapy compared with CCBs.

Limitations. Potential confounding by indication.

Conclusions. Our findings provide evidence from routine care that initiation of RASi therapy compared with CCBs may confer kidney benefits among patients with advanced CKD, with similar cardiovascular protection.

Introduction

Randomized trials of angiotensin-converting enzyme inhibitors (ACEi) or angiotensin receptor blockers (ARB), collectively renin-angiotensin system inhibitors (RASi), have shown that these drugs are more effective in delaying the progression of CKD than placebo or alternative agents, such as diuretics, beta-blockers or calcium channel blockers (CCB) (1-6). Clinical guidelines recommend RASi as the first-line pharmacologic antihypertensive treatment strategy in patients with CKD G1-3 and proteinuria, with or without diabetes (7-9). There is, however, less evidence on the benefits of RASi in patients with CKD G4-5, a population that was under-represented in pivotal trials (3, 10-15). A small randomized trial (16) and various observational studies (17-20) suggest that RASi confer reno-protection compared with placebo or no use, but no data exist to inform the choice of RASi over alternative antihypertensive agents. This, together with concerns about the persistent hemodynamic effects of RASi (21, 22), may lead to underutilization of these medications in advanced CKD (23, 24). Indeed, recent studies indicate that a significant proportion of individuals with CKD G3-5 do not receive RASi therapy (23-25). A recent NKF-KDOQI controversies report (14) identified the lack of comparative effectiveness data as a critical knowledge gap, and emphasized the need of further studies to inform practice.

CCBs are also frequently prescribed to treat hypertension, especially to patients with CKD (26-28). Although CCBs were used as an active comparator to RASi in trials such as AASK or IDNT (4, 11), these trials included very few patients with advanced CKD to allow for stratification. In the absence of trial evidence, observational studies in patients cared for in routine clinical practice can provide insights into the relative efficacy of medications. To fill this knowledge gap we studied kidney and cardiovascular outcomes in patients with advanced CKD who initiated RASi or CCB therapy.

Methods

Data sources

We conducted an observational cohort study using data from the Swedish Renal Registry (SRR), a nationwide registry including patients with CKD G3–5 under nephrologist care (29, 30). The SRR includes information on outpatient visits, including laboratory tests and results from clinical examination. According to the guidelines of the registry, patients with an estimated glomerular filtration rate (eGFR) <30 mL/min/1.73m^2 should be enrolled. Registrations of subsequent outpatient visits to nephrology care are thereafter performed until death, emigration from the country or start of kidney replacement therapy (KRT). Nearly all nephrology clinics in Sweden (96%) report to the SRR-CKD and the estimated national coverage is 75-90% of nephrologist-referred patients with recognized CKD G4-5 (31).

Using each citizen's unique personal identification number, the SRR-CKD was linked to other national registries. The Swedish Prescribed Drug Register provided complete information on all prescribed drugs dispensed at Swedish pharmacies (32); the Swedish Patient Register added information on all outpatient specialist consultations and hospitalizations occurring in Swedish healthcare, and was used to obtain information on comorbidities and outcomes (33); the Swedish Cause of Death Register added information on date and causes of death (34). All these registers are run by the Swedish National Board of Welfare and are considered to have no or minimal loss to follow up. We used de-identified data, the study was approved by the regional ethical review boards and the Swedish National Board of Welfare, and was judged not to require informed consent.

Patient selection and study design

We created a cohort of all adult patients in the SRR-CKD (≥ 18 years) newly initiating a RASi or CCB between 1 January 2007 and 1 June 2017. New users were defined as individuals receiving a RASi or CCB without dispensation of either drug in the previous six months. Prevalent users of these drugs were excluded to prevent prevalent user bias (35). We further excluded all individuals with a history of kidney transplantation, an eGFR > 30 ml/min/1.73m^2, or those initiating both drugs simultaneously.

The date of initiation was defined as the index date of the study and start of follow-up. Patients were followed from index date to the first occurrence of a study outcome or end of follow-up (1 June 2017). eGFR was calculated with the CKD-EPI equation from routine plasma creatinine measurements performed by enzymatic or corrected Jaffe methods traceable to isotope dilution mass spectroscopy standards. Information on race is not available in Sweden by law; we assumed that all patients were Caucasian.

Study exposure and covariates

The exposure of interest was RASi initiation compared with initiation of a CCB. Baseline covariates included age, sex, eGFR, comorbidities (diabetes mellitus, myocardial infarction, heart failure, arrhythmia, peripheral vascular disease, cerebrovascular disease, ischemic heart disease), medications (β-blocker, thiazide diuretic, loop diuretic, potassium-sparing diuretic, potassium binder, non-steroidal anti-inflammatory drug, statin), systolic blood pressure, diastolic blood pressure, urinary albumin-to-creatinine ratio (ACR), potassium. In addition, we considered other covariates in an attempt to evaluate reasons that led to the use of either medication: the rate of kidney function decline prior to therapy initiation, the occurrence of a cardiovascular-related hospitalization in the preceding six months, the number of overall hospitalizations in the year prior and a history of hyperkalemia or AKI. Covariate definitions are detailed in **Supplemental Table S1**.

Study outcomes

The primary study outcome was initiation of KRT, defined as the date of start of maintenance dialysis or kidney transplantation, as registered in the SRR. Secondary outcomes were all-cause mortality and major adverse cardiovascular events (MACE), defined as a composite of cardiovascular death (ICD-10 code of the I family as main cause of death), hospitalization due to stroke (I63) or myocardial infarction (I21-I23). For the analysis of mortality and MACE, KRT was not considered a censoring event. In addition, we reported information about hospitalizations for hyperkalemia and acute kidney injury (AKI) after medication initiation.

Statistical analysis

We used doubly robust methods, i.e., combining outcome regression with inverse probability of treatment weighting (IPTW), to control for confounding (36). A multivariable logistic regression model was used to calculate the probability of receiving RASi (versus CCB) as a function of baseline covariates. Weighting was considered appropriate if the standardized mean difference (SMD) between treatment groups was <0.1. Weights were stabilized to increase precision by adding the marginal probability of treatment to the numerator of the weights. Robust variance estimation was used to calculate confidence intervals after weighting. We assessed the association between RASi use compared with CCB use on outcomes using multivariable cause-specific Cox proportional hazards regression in the inverse probability weighted sample, additionally adjusting for all baseline covariates. In addition, we estimated adjusted cumulative incidence curves standardized to the distribution of the baseline variables in the study population. To do so, we fitted a weighted pooled logistic model including an indicator for treatment, month and its quadratic term, all baseline confounders, and interactions between treatment and time (37). Interaction terms were included to allow for nonproportional hazards (38). The predicted probabilities from this logistic model were used to estimate the adjusted absolute risks of KRT, mortality and MACE which were then standardized to the baseline distribution of confounders. For the calculation of the cumulative incidence of KRT and MACE, we took into account the competing risk of (non-cardiovascular) death (39-41). Pointwise 95% confidence intervals for the cumulative incidence curves were calculated using nonparametric bootstrap based on 500 full samples. In primary analyses, we adopted an intention-to-treat (ITT) approach and analyzed patients according to their initially assigned treatment group irrespective of discontinuation or treatment switch. Next, we examined whether there was an interaction between treatment effect and the following variables, according to a priori defined strata: age (≥70 vs <70 years), sex, diabetes, myocardial infarction, heart failure, systolic blood pressure (≥140 vs <140 mmHg), eGFR (≥15 vs <15 mL/min/1.73m^2) and ACR (≥70 vs <70 mg/mmol). To calculate the subgroup hazard

ratios, we separately estimated the propensity score model and Cox model in each subgroup (42). Multiplicative interaction was tested by including interaction terms between treatment and the variable of interest to the Cox model.

Multiple imputation by chained equations was used to impute missing data on systolic and diastolic blood pressure (missing for 2.3% of patients). Treatment, confounding variables, outcomes and interaction terms between treatment and confounders were used in the imputation model to derive 50 imputed datasets (43). eGFR was non-normally distributed and was log-transformed before imputation. Multiple imputation was combined with IPTW using the *within* method (44). In the within method, effect estimates are obtained separately in each imputation using the propensity score, which are then combined to an overall estimate. The within method has been shown to produce unbiased estimates with appropriate confidence intervals compared with the *across* approach (44).

We performed several sensitivity analyses to test the robustness of our results. First, we additionally adjusted our analyses for plasma potassium and ACR. These variables were missing for a large proportion of patients (32% and 41%, respectively) because it was not mandatory to report these measures. Those with missing ACR measurements had similar characteristics as those without missing ACR measurements and we assumed data to be missing at random (**Supplemental Table S2**). We used multiple imputation with chained equations, a technique well suited to impute data that are missing at random. Second, we redefined new users as those not using RASi and CCB for at least 12 months. Third, we replicated our analyses in a positive control cohort of patients with CKD G3, for which we expected a reduction in kidney replacement therapy consistent with previously conducted randomized trials (3, 45-47). Fourth, we performed an as-treated analysis in which patients were censored at the time of therapy discontinuation (no dispensation for the index drug within 60 days after the estimated last day of pill supply from the previous drug dispensation), treatment switch (on the day RASi was added to CCB or vice versa) or at the end of the study period. To account for potential informative censoring due to discontinuation or treatment switch, inverse probability of censoring weighting (IPCW) was applied (see **Supplemental methods** for details). Fifth, we used incident cancer diagnosis as a negative control outcome to study the influence of potential unmeasured confounders (such as smoking and alcohol use) on our effect estimates. While unmeasured confounders may predict the risk of cancer, we did not expect initiation of RASi or CCB to cause or prevent cancer (48). For this analysis, we excluded patients with a recent cancer diagnosis (within two years from index date). Lastly, we repeated our analysis adding heart failure related hospitalization (I50) as an outcome in the composite of MACE. All analyses were performed using R version 3.6.2.

Results

Cohort characteristics

We identified 21,065 patients under nephrologist care with eGFR <30 ml/min/1.73m^2 and no history of KRT. Of these, 13 896 (66%) were prevalent users of RASi or CCB and were excluded. We further excluded 1913 patients who received neither of these drugs during observation and 453 patients who were prescribed both medications simultaneously. The final study cohort consisted of 4803 patients: 2458 (51%) who initiated RASi and 2345 (49%) who initiated CCB (**Supplemental Figure S1**). Of patients initiating RASi, the majority initiated enalapril (37.2%), candesartan (23.4%), losartan (21.4%) or ramipril (9.6%). In total, 249 of 2458 (10.1%) individuals initiating RASi had a cardiovascular hospitalization in the 6 months prior to initiation, of which 129 (5.2%) due to heart failure and 37 (1.5%) due to myocardial infarction. Five people initiated dual RAS blockade with an ACEi and ARB. The majority of patients initiating a CCB used a dihydropyridine CCB (97.7%), mainly amlodipine (55.4% of total CCB initiators) or felodipine (36.9%). In total, 231 of 2345 (9.9%) individuals initiating CCB had a cardiovascular hospitalization in the 6 months prior to initiation, of which 49 (2.1%) due to heart failure and 32 (1.4%) due to myocardial infarction.

Overall, study participants had a median (IQR) age of 74 (64-81) years and 38% were women. Median eGFR was 20 (15-21) ml/min/1.73m^2, median ACR 28 (7-108) mg/mmol, median systolic blood pressure 140 (125-153) mmHg and median diastolic blood pressure 80 (70-85) mmHg. The most common comorbidities were diabetes (34%), ischemic heart disease (26%) and heart failure (19%). Concurrent use of β-blockers (63%), loop diuretics (63%) and statins (50%) was prevalent. At baseline, patients who initiated RASi, compared with those initiating CCB, had a higher eGFR, a lower systolic blood pressure and ACR, and a higher prevalence of comorbidities such as diabetes, heart failure and arrhythmia. After weighting, all baseline covariates appeared well balanced between treatment groups (standardized differences <0.1) (**Table 1**).

Table 1. Baseline characteristics of patients with advanced CKD by RASi or CCB treatment, before and after inverse probability weighting.

	Unweighted			Weighted*		
	RASi (N = 2458)	**CCB (N = 2345)**	**Std Diff‡**	**RASi (N = 2473)**	**CCB (N = 2330)**	**Std Diff‡**
Median age (IQR)‡, *years*	73 [62, 80]	74 [66, 81]	0.22	74 [64, 80]	73 [64, 80]	0.00
Age category, *n (%)*						
<50	303 (12.3)	159 (6.8)	0.19	238 (9.6)	210 (9.0)	0.02
50-59	226 (9.2)	189 (8.1)	0.04	195 (7.9)	217 (9.3)	0.05
60-69	461 (18.8)	443 (18.9)	0.00	477 (19.3)	454 (19.5)	0.01
70-79	826 (33.6)	805 (34.3)	0.01	871 (35.2)	800 (34.4)	0.02
>=80	642 (26.1)	749 (31.9)	0.13	692 (28.0)	649 (27.8)	0.00
Women	909 (37.0)	906 (38.6)	0.03	950 (38.4)	898 (38.5)	0.00
Median eGFR (IQR)‡, *ml/min/1.73m²*	22 [17, 26]	18 [13, 24]	0.41	20 [15, 25]	20 [15, 25]	0.00
eGFR category, *n (%)*						
<15 ml/min/1.73m², *n (%)*	399 (16.2)	727 (31.0)	0.35	657 (25.4)	678 (27.0)	0.04
15-30 ml/min/1.73m², *n (%)*	2059 (83.8)	1614 (68.8)	0.36	1816 (74.6)	1652 (73.0)	0.04
Median SBP (IQR)‡, *mmHg*	133 [120, 146]	144 [130, 160]	0.51	140 [125, 155]	140 (125, 154)	0.00
SBP category, *n (%)*						
<120	486 (19.8)	161 (6.9)	0.39	333 (13.5)	304 (13.0)	0.02
120-139	934 (38.0)	689 (29.4)	0.18	842 (34.1)	801 (34.4)	0.01
140-159	661 (26.9)	804 (34.3)	0.16	774 (31.3)	740 (31.8)	0.01
>160	323 (13.1)	633 (27.0)	0.35	524 (21.2)	485 (20.8)	0.01
Missing	54 (2.2)	58 (2.5)	0.02	-	-	-
Median DBP (IQR)‡, *mmHg*	78 [70, 84]	80 [70, 89]	0.28	80 [70, 85]	80 [70, 85]	0.00
DBP category, *n (%)*						
<80	1264 (51.4)	942 (40.2)	0.23	1156 (46.7)	1077 (46.2)	0.01
80-89	776 (31.6)	783 (33.4)	0.04	847 (34.3)	772 (33.1)	0.03
90-99	260 (10.6)	380 (16.2)	0.16	323 (13.1)	330 (14.2)	0.03
>100	104 (4.2)	182 (7.8)	0.15	147 (6.0)	151 (6.5)	0.02
Missing	54 (2.2)	58 (2.5)	0.02	-	-	-
Median ACR (IQR)‡, *mg/mmol*	24 [5, 95]	33 [9, 116]	0.12	29 [7, 111]	29 [7, 113]	0.00
ACR category, *n (%)*						
A1 (<3)	276 (11.2)	150 (6.4)	0.17	373 (15.1)	342 (14.7)	0.01
A2 (3-29)	542 (22.1)	483 (20.6)	0.04	880 (35.6)	829 (35.6)	0.00
A3 (30-69)	240 (9.8)	204 (8.7)	0.04	400 (16.2)	383 (16.4)	0.01

	Unweighted			Weighted*		
	RASi (N = 2458)	CCB (N = 2345)	Std Diff‡	RASi (N = 2473)	CCB (N = 2330)	Std Diff‡
A3 (≥70)	461 (18.8)	472 (20.1)	0.03	820 (33.2)	776 (33.3)	0.00
Missing	939 (38.2)	1036 (44.2)	0.12	-	-	-
Median potassium (IQR)‡, *mmol/L**	4.4 [4.1, 4.8]	4.3 [4.0, 4.7]	0.15	4.4 [4.0, 4.7]	4.4 [4.0, 4.7]	0.00
Comorbidities, *n (%)*						
Diabetes mellitus	916 (37.3)	734 (31.3)	0.13	851 (34.4)	833 (35.8)	0.03
Myocardial infarction	423 (17.2)	353 (15.1)	0.06	398 (16.1)	361 (15.5)	0.02
Heart failure	580 (23.6)	320 (13.6)	0.26	457 (18.5)	420 (18.0)	0.01
Arrhythmia	469 (19.1)	316 (13.5)	0.15	416 (16.8)	395 (17.0)	0.00
Peripheral vascular disease	313 (12.7)	312 (13.3)	0.02	330 (13.3)	313 (13.5)	0.00
Cerebrovascular disease	294 (12.0)	327 (13.9)	0.06	321 (13.0)	311 (13.3)	0.01
Ischemic heart disease	691 (28.1)	574 (24.5)	0.08	657 (26.6)	617 (26.5)	0.00
Medication, *n (%)*						
β-blockers	1443 (58.7)	1586 (67.6)	0.19	1563 (63.2)	1486 (63.8)	0.01
Thiazides	79 (3.2)	66 (2.8)	0.02	71 (2.9)	70 (3.0)	0.01
Loop diuretics	1613 (65.6)	1395 (59.5)	0.13	1551 (62.7)	1463 (62.8)	0.00
Potassium-sparing diuretics	167 (6.8)	114 (4.9)	0.08	136 (5.5)	121 (5.2)	0.01
Potassium binders	242 (9.8)	240 (10.2)	0.01	254 (10.2)	216 (9.3)	0.03
NSAIDs	103 (4.2)	90 (3.8)	0.02	101 (4.1)	92 (4.0)	0.01
Statins	1270 (51.7)	1121 (47.8)	0.08	1232 (49.8)	1167 (50.1)	0.01
Hospitalizations, *n (%)*						
Any hospitalization in previous year	1084 (44.1)	1254 (53.5)	0.19	1210 (48.9)	1138 (48.8)	0.00
Cardiovascular hospitalization in previous 6 months	249 (10.1)	231 (9.9)	0.01	251 (10.1)	229 (9.8)	0.01
Hyperkalemia hospitalization	35 (1.4)	39 (1.7)	0.02	38 (1.5)	37 (1.6)	0.00
AKI hospitalization	125 (5.1)	213 (9.1)	0.16	187 (7.6)	169 (7.2)	0.01
Previous eGFR decline, *ml/min/1.73m² (SE)* §	-2.03 (0.08)	-1.98 (0.08)	0.02	-	-	-

* Inverse probability weighting was performed after imputation. Baseline characteristics are shown after imputation and weighting (marked with *).

‡ A standardized difference >0.1 indicates meaningful imbalance between groups.

‡ Standardized difference for the mean was calculated for age, eGFR, blood pressure, ACR and potassium.

§ Calculated in the overall population on all previous eGFR measurements with a linear mixed model containing fixed effects for time, treatment and time/treatment interaction and random intercept and slope.

Comparative effectiveness of RASi vs. CCB initiation

Median follow-up was 4.1 (95% CI 3.9-4.2) years, maximum follow-up was 10.4 years, and the total follow-up time of the cohort was 14 682 person years. During follow-up 1416 individuals initiated KRT. The absolute 5-year risk of KRT was 39.0% among CCB users and 34.8% among RASi users, with a 5-year absolute risk difference of -4.3% (-8.0 to -0.6). The KRT risk was consistently lower in RASi users compared with CCB users during the entire follow-up period. For instance, risk differences were -3.3% (-4.9 to -1.6) at 1 year and -4.4% (-7.4 to -1.6) at 3 years (**Figure 1** and **Supplemental Table S3**). For patients initiating RASi, compared with those initiating CCB, we observed a weighted hazard ratio of 0.79 (0.69-0.89), in favor of RASi initiation (**Table 2**).

Table 2. Number of events, incidence rates as well as crude and adjusted hazard ratios for the association between RASi vs. CCB initiation and all-cause mortality, MACE and kidney replacement therapy.

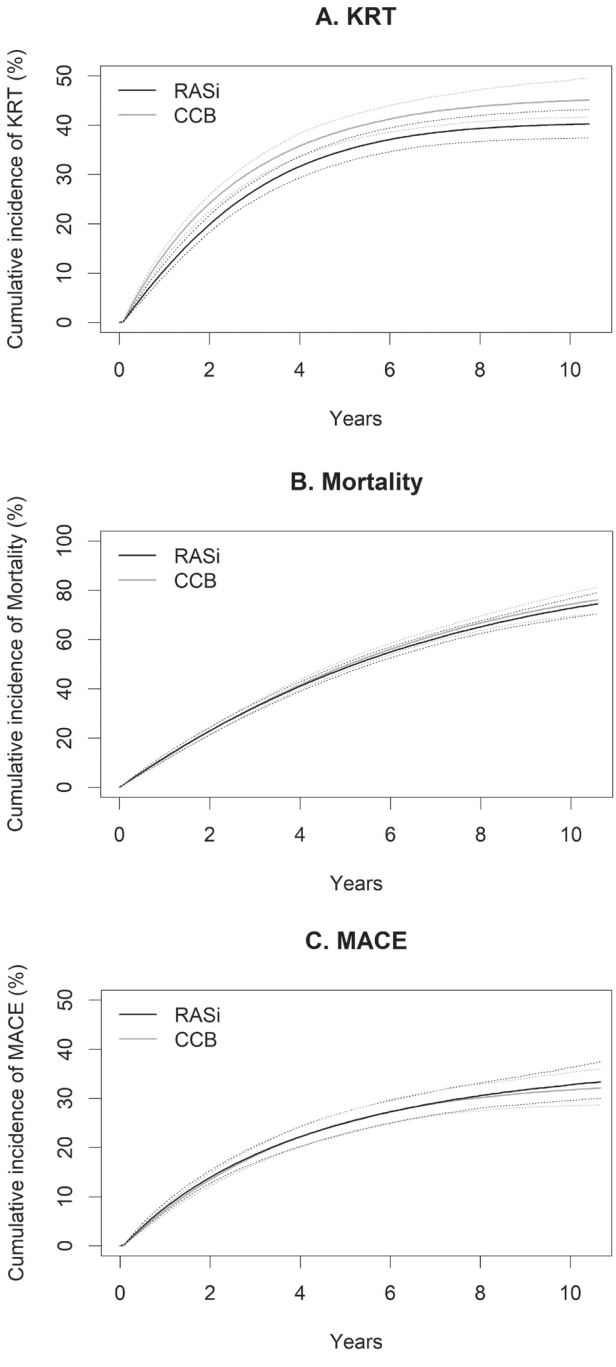
	Number of events*	Person years*	IR per 100PY (95% CI)*	Crude HR (95% CI)	Adjusted HR (95% CI)ǂ
KRT					
Overall	1416	11044	12.8 (12.2-13.5)		
CCB	753	4872	15.5 (14.4-16.6)	1 (reference)	1
RASi	663	6172	10.7 (9.9-11.6)	0.70 (0.63- 0.78)	0.79 (0.69-0.89)
All-cause mortality					
Overall	1974	14682	13.4 (12.9-14.1)		
CCB	991	6769	14.6 (13.7-15.6)	1	1
RASi	983	7912	12.4 (11.7-13.2)	0.85 (0.78- 0.93)	0.97 (0.88-1.07)
MACE					
Overall	1043	13814	7.6 (7.1-8.0)		
CCB	510	6311	8.1 (7.4-8.8)	1	1
RASi	533	7503	7.1 (6.5-7.7)	0.90 (0.80- 1.02)	1.00 (0.88-1.15)

IR = incidence rate; PY = person years; HR = hazard ratio; CI = confidence interval; MACE = major adverse cardiovascular events; RASi = renin-angiotensin system inhibitor; CCB = calcium channel blocker; KRT = kidney replacement therapy.

* Number of events, person years and incidence rates were calculated in the unweighted population.

ǂ Analyses were adjusted for age, sex, eGFR, heart failure, arrhythmia, peripheral vascular disease, cerebrovascular disease, ischemic heart disease, diabetes mellitus, systolic blood pressure, diastolic blood pressure, use of β-blocker, thiazide diuretic, potassium-sparing diuretic and statin, total number of hospitalizations in previous year, hospitalization in previous year (yes/no), history of hyperkalemia hospitalization and history of AKI hospitalization using inverse probability of treatment weighting.

Figure 1. Weighted standardized survival curves for KRT (panel A), mortality (panel B) and major adverse cardiovascular events (MACE, panel C) stratified by RASi or CCB use.

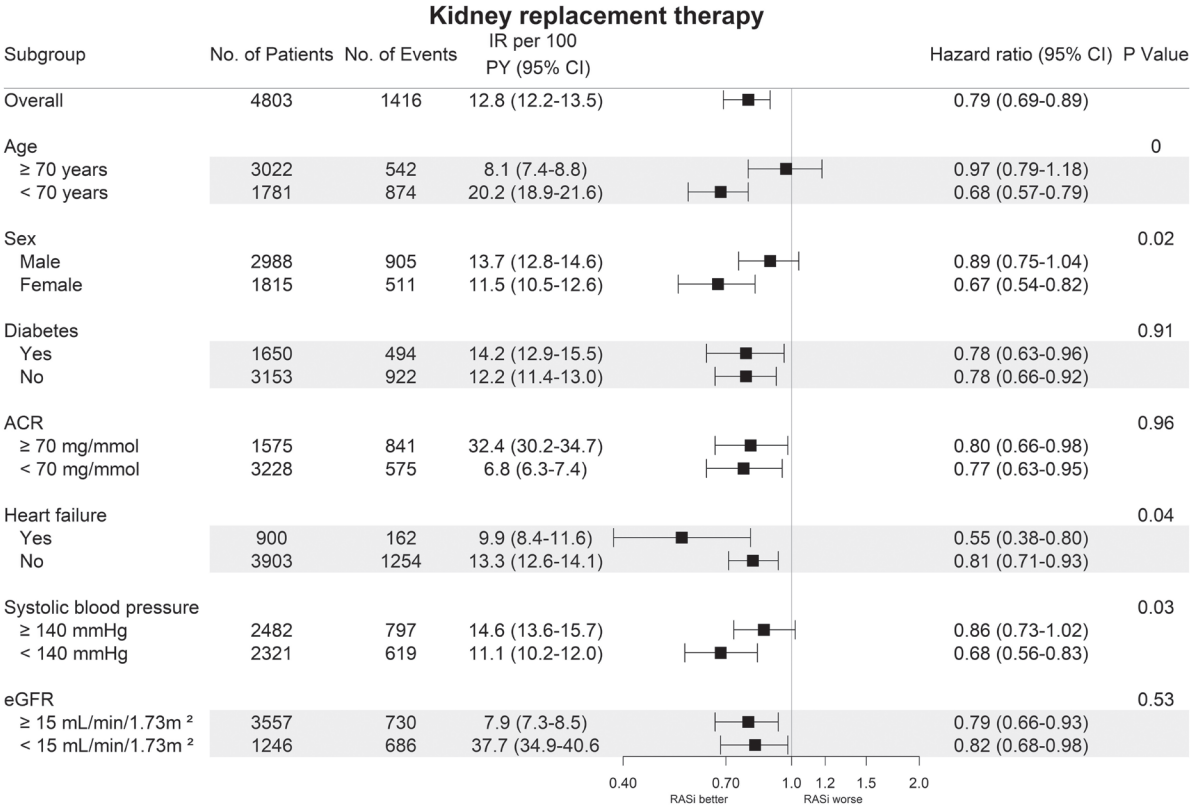


In total, 1974 individuals died, with an absolute 5-year mortality risk of 49.5% among CCB users and 48.3% among RASi users. The absolute risk difference at 5 years was -1.2% (-4.1 to 1.7), with a weighted mortality hazard ratio of 0.97 (95% CI 0.88-1.07). During follow-up, 1043 individuals experienced a MACE, with a weighted hazard ratio of 1.00 (0.88-1.15). The absolute 5-year risk of MACE was 25.1% among CCB users and 25.0% among RASi users, with a 5-year risk difference of -0.1% (-3.4 to 3.0). Among individuals initiating RASi, 18 (0.7%) experienced a hospitalization for hyperkalemia and 83 (3.4%) a hospitalization for AKI. Among those initiating CCB, 18 (0.8%) experienced a hospitalization for hyperkalemia and 72 (3.1%) individuals experienced a hospitalization for AKI.

Subgroup and sensitivity analyses

Results were robust in most subgroup analyses (**Figure 2**, **Supplemental Figures S2-S3**, **Supplemental Table S4**). A lower risk of KRT for RASi users compared with CCB was observed across strata of sex, diabetes, ACR, eGFR, heart failure and systolic blood pressure, but a significant interaction was observed for age, with benefit for initiating RASi in younger but not older patients ($p < 0.01$). An increased risk of mortality and MACE (interaction $p < 0.01$) was observed for patients with baseline heart failure and CKD G4-5 initiating RASi, compared with CCB, as well as a significant interaction for MACE according to sex ($p < 0.01$). Other than this, risks of mortality and MACE did not differ by prespecified subgroups (all interaction $p > 0.12$).

Figure 2. Number of events, incidence rates and adjusted hazard ratios for kidney replacement therapy following RASi vs. CCB initiation, according to subgroups of age, sex, diabetes, ACR, heart failure, systolic blood pressure and eGFR.



The positive control cohort included 2608 nephrologist-referred patients with CKD G3 of whom 1663 started RASi and 945 started CCB (baseline characteristics in **Supplemental Table S5**). After IPTW, the adjusted hazard ratio for RASi compared with CCB was 0.68 (0.48-0.98) for KRT, 0.97 (0.81-1.17) for mortality and 1.09 (0.85-1.40) for MACE (**Supplemental Table S6**).

In the as-treated analysis, a hazard ratio of 0.67 (0.56-0.80) was observed for KRT for RASi initiation compared with CCB initiation. The adjusted hazard ratios for mortality and MACE were 1.05 (0.87-1.26) and 1.03 (0.83-1.26), respectively (**Supplemental Table S7**). Additional adjustment for ACR and potassium or redefining new users as no recorded dispensation of either RASi or CCB for at least 12 months, produced hazard ratios consistent with the results of our main analysis (**Supplemental Table S7**). Individuals who initiated RASi had similar risks of cancer compared with CCB initiators, with a weighted HR of 1.03 (0.87-1.22). Adding heart failure-related hospitalization to the MACE outcome did not alter our results (adjusted HR 1.00; 95% CI 0.89-1.13) (**Supplemental Table S8**).

Discussion

Current clinical guidelines recommend the use of ACEi or ARBs as first-line therapy in patients with CKD and proteinuria, with or without diabetes (7-9, 49), but provide no guidance regarding eGFR thresholds for which these recommendations are valid (14, 15). In our study of a large, nationwide cohort of nephrologist-referred patients with advanced CKD, initiation of RASi compared with CCB was associated with a reduced risk of KRT, but similar risk of mortality and MACE. These findings were robust across subgroups of patients and following an as-treated design.

Our study does not evaluate the health benefits of RASi versus no use in patients with CKD G4 and 5. This has been investigated previously (17, 18, 24), including the randomized trial by Hou *et al.* (16) or the post-hoc analysis of the REIN (Ramipril Efficacy in Nephropathy) trial (10). Our goal was to inform on the choice of antihypertensive agents in the advanced CKD population by comparing outcomes associated with initiating RASi or CCB as the two most commonly used antihypertensive agents in clinical practice (28). A considerable proportion of patients reach CKD stage 4-5 without these medications. In our register this equaled to 34% of the population, a figure which agrees with other contemporary reports: in the CRIC cohort, ~30% of patients CKD G4 and about 73% of patients CKD G5 did not receive RASi, and similar proportions of non-use were reported for CCB in CKD G4 (50% not using CCB) and G5 (40% not using CCB) (24). Recent data from CKDOPPS indicates that this pattern is followed globally: for instance, only 52% of DOPPS patients in the United States and 66% in Brazil were receiving RASi (25).

We observed that RASi may be superior to CCB in delaying KRT in advanced CKD. This is consistent with a recent network meta-analysis of patients with CKD G3 showed that ACEi reduced the odds of KRT by 35% (OR 0.65; 95% credibility interval 0.51-0.80), and ARBs reduced the odds of kidney failure by 25% (0.75; 0.54-0.97), compared with other antihypertensive drugs, which included CCBs, diuretics and beta-blockers (13). Our positive control cohort of individuals with CKD G3 showed a reduction in KRT risk (HR 0.68; 95% CI 0.48-0.98) of magnitude similar to that meta-analysis, which lends reassurance to our observations. We note that 98% of our patients used dihydropyridine CCB, and the comparative effectiveness and safety of non-dihydropyridine CCB cannot be informed by our study.

We observed no differences in the risk of MACE between both therapies in persons with advanced CKD, a finding we believe is novel (7, 14) and in a magnitude similar to our control population of patients with CKD G3. Again this agrees and expands two large meta-analyses of randomized trials comparing antihypertensive agents in patients with CKD G3 (13, 50). Compared with placebo, blood-pressure-lowering regimens significantly reduced the risk of MACE in individuals with CKD G3 (HR 0.83; 95% CI 0.76-0.90), but results were similar whether the regimen was based on ACEi, CCB, diuretics or beta-blockers (50). Another Bayesian network meta-analysis found odds ratios of 0.94 (95% credibility interval 0.75-1.12) for ACEi and 0.86 (95% credibility interval 0.70-1.03) for ARB versus active controls (either CCB, diuretics or beta-blockers) on cardiovascular events (13). Collectively these findings may suggest that there is little evidence to support a particular drug class for the prevention of cardiovascular outcomes in the general population with CKD.

Finally, few studies have compared the mortality risk of RASi versus alternative antihypertensive agents in advanced CKD. In the meta-analysis by the Blood Pressure Lowering Treatment Trialists' Collaboration, both ACEi vs. placebo and CCB vs. placebo were associated with similar reductions in all-cause mortality for CKD patients (predominantly CKD G3a), with HR (95% CI) of 0.86 (0.76-0.97) and 0.83 (0.56-1.24), respectively (50). Head-to-head comparisons of RASi vs. CCB in patients with CKD yielded a hazard ratio of 1.00 (0.89-1.13) (50), which is again similar to what we observed in patients with CKD G4-5ND (0.97; 0.88-1.07) and our control cohort of patients with CKD G3 (0.97; 0.81-1.17).

We studied a unique nationwide inception cohort design of patients referred to a nephrologist in a country with universal healthcare access, with long-term follow-up data of over 10 years, assessment of multiple relevant endpoints, virtually no loss to follow-up and low likelihood of misclassification for the outcomes KRT and mortality. Furthermore, results were robust in multiple subgroup and sensitivity analyses. Our positive control analysis of persons with CKD G3 aligned with findings from two meta-analyses of trials and the patients included are representative of

routine clinical practice. In addition, the negative control analysis with cancer did not indicate that the observed associations were due to different health status. However, we recognize limitations. Despite adjustment for a wide range of potential confounders, selection of patients referred to nephrologists, and the use of an active comparator (CCB initiation), residual confounding-by-indication bias cannot be excluded in observational designs, and the reasons for the initiation of these drugs in the patients of our study remain unknown. Because only around 10% of individuals starting RASi or CCB in our study had a cardiovascular hospitalization in the 6 months prior to therapy start, we speculate that medications may have been initiated for renoprotection or as antihypertensive agents. Data were missing for ACR and potassium, but our results were similar whether these variables were included using multiple imputation or not, and those with missing measurements had similar characteristics to those without missing measurements. We recognize that it may be unusual to start RASi or CCB this late in the course of disease, and that there may be special indications for it. While we acknowledge that we do not have the precise reasons that prompted the use of these therapies, we went through a great deal of efforts to identify and control for these potential indications. Our results are likely generalizable to Swedish clinical practice during the period 2007-2017. However, extrapolations to other ethnicities, countries or periods should be done with caution. Finally, our conclusions remain observational in nature and do not substitute for randomized trials. However, until these trials are conducted they may assist in informing clinical decisions.

In conclusion, in patients with CKD G4-5ND, RASi initiation, compared with CCB initiation, was associated with a lower risk of KRT, but similar risks of MACE or mortality. These results suggest that use of RASi may confer additional renal benefits compared with CCB in patients with CKD G4-5ND. This evidence may potentially inform clinical decisions on the choice of antihypertensive therapy for this patient group, minimally included in pivotal trials.

References

1. Randomised placebo-controlled trial of effect of ramipril on decline in glomerular filtration rate and risk of terminal renal failure in proteinuric, non-diabetic nephropathy. The GISEN Group (Gruppo Italiano di Studi Epidemiologici in Nefrologia). Lancet. 1997;349(9069):1857-63.
2. Brenner BM, Cooper ME, de Zeeuw D, Keane WF, Mitch WE, Parving HH, et al. Effects of losartan on renal and cardiovascular outcomes in patients with type 2 diabetes and nephropathy. N Engl J Med. 2001;345(12):861-9.
3. Jafar TH, Schmid CH, Landa M, Giatras I, Toto R, Remuzzi G, et al. Angiotensin-converting enzyme inhibitors and progression of nondiabetic renal disease. A meta-analysis of patient-level data. Ann Intern Med. 2001;135(2):73-87.
4. Lewis EJ, Hunsicker LG, Clarke WR, Berl T, Pohl MA, Lewis JB, et al. Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. N Engl J Med. 2001;345(12):851-60.
5. Ruggenenti P, Perna A, Gherardi G, Gaspari F, Benini R, Remuzzi G. Renal function and requirement for dialysis in chronic nephropathy patients on long-term ramipril: REIN follow-up trial. Gruppo Italiano di Studi Epidemiologici in Nefrologia (GISEN). Ramipril Efficacy in Nephropathy. Lancet. 1998;352(9136):1252-6.
6. Kent DM, Jafar TH, Hayward RA, Tighiouart H, Landa M, de Jong P, et al. Progression risk, urinary protein excretion, and treatment effects of angiotensin-converting enzyme inhibitors in nondiabetic kidney disease. J Am Soc Nephrol. 2007;18(6):1959-65.
7. Kidney Disease: Improving Global Outcomes (KDIGO) Blood Pressure Work Group. KDIGO Clinical Practice Guideline for the Management of Blood Pressure in Chronic Kidney Disease. Kidney inter, Suppl. 2012;2:337-414.
8. Whelton PK, Carey RM, Aronow WS, Casey DE, Jr., Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2018;71(19):e127-e248.
9. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, et al. 2018 ESC/ESH Guidelines for the management of arterial hypertension. Eur Heart J. 2018;39(33):3021-104.
10. Ruggenenti P, Perna A, Remuzzi G, Gruppo Italiano di Studi Epidemiologici in N. ACE inhibitors to prevent end-stage renal disease: when to start and why possibly never to stop: a post hoc analysis of the REIN trial results. Ramipril Efficacy in Nephropathy. J Am Soc Nephrol. 2001;12(12):2832-7.
11. Wright JT, Jr., Bakris G, Greene T, Agodoa LY, Appel LJ, Charleston J, et al. Effect of blood pressure lowering and antihypertensive drug class on progression of hypertensive kidney disease: results from the AASK trial. JAMA. 2002;288(19):2421-31.
12. Wu HY, Huang JW, Lin HJ, Liao WC, Peng YS, Hung KY, et al. Comparative effectiveness of renin-angiotensin system blockers and other antihypertensive drugs in patients with diabetes: systematic review and bayesian network meta-analysis. BMJ. 2013;347:f6008.
13. Xie X, Liu Y, Perkovic V, Li X, Ninomiya T, Hou W, et al. Renin-Angiotensin System Inhibitors and Kidney and Cardiovascular Outcomes in Patients With CKD: A Bayesian Network Meta-analysis of Randomized Clinical Trials. Am J Kidney Dis. 2016;67(5):728-41.

14. Weir MR, Lakkis JI, Jaar B, Rocco MV, Choi MJ, Kramer HJ, et al. Use of Renin-Angiotensin System Blockade in Advanced CKD: An NKF-KDOQI Controversies Report. Am J Kidney Dis. 2018;72(6):873-84.
15. Eckardt KU, Bansal N, Coresh J, Evans M, Grams ME, Herzog CA, et al. Improving the prognosis of patients with severely decreased glomerular filtration rate (CKD G4+): conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. Kidney Int. 2018;93(6):1281-92.
16. Hou FF, Zhang X, Zhang GH, Xie D, Chen PY, Zhang WR, et al. Efficacy and safety of benazepril for advanced chronic renal insufficiency. N Engl J Med. 2006;354(2):131-40.
17. Hsu TW, Liu JS, Hung SC, Kuo KL, Chang YK, Chen YC, et al. Renoprotective effect of renin-angiotensin-aldosterone system blockade in patients with predialysis advanced chronic kidney disease, hypertension, and anemia. JAMA Intern Med. 2014;174(3):347-54.
18. Voskamp PWM, Dekker FW, van Diepen M, Hoogeveen EK, Group P-S. Effect of dual compared to no or single renin-angiotensin system blockade on risk of renal replacement therapy or death in predialysis patients: PREPARE-2 study. J Am Soc Hypertens. 2017;11(10):635-43.
19. Oh YJ, Kim SM, Shin BC, Kim HL, Chung JH, Kim AJ, et al. The Impact of Renin-Angiotensin System Blockade on Renal Outcomes and Mortality in Pre-Dialysis Patients with Advanced Chronic Kidney Disease. PLoS One. 2017;12(1):e0170874.
20. Suissa S, Hutchinson T, Brophy JM, Kezouh A. ACE-inhibitor use and the long-term risk of renal failure in diabetes. Kidney Int. 2006;69(5):913-9.
21. Fu EL, Trevisan M, Clase CM, Evans M, Lindholm B, Rotmans JI, et al. Association of Acute Increases in Plasma Creatinine after Renin-Angiotensin Blockade with Subsequent Outcomes. Clin J Am Soc Nephrol. 2019;14(9):1336-45.
22. Tomlinson LA, Abel GA, Chaudhry AN, Tomson CR, Wilkinson IB, Roland MO, et al. ACE inhibitor and angiotensin receptor-II antagonist prescribing and hospital admissions with acute kidney injury: a longitudinal ecological study. PLoS One. 2013;8(11):e78465.
23. Arora N, Katz R, Bansal N. ACE Inhibitor/Angiotensin Receptor Blocker Use Patterns in Advanced CKD and Risk of Kidney Failure and Death. Kidney Med. 2020;2(3):248-57.
24. Ku E, McCulloch CE, Vittinghoff E, Lin F, Johansen KL. Use of Antihypertensive Agents and Association With Risk of Adverse Outcomes in Chronic Kidney Disease: Focus on Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers. J Am Heart Assoc. 2018;7(19):e009992.
25. Pecoits-Filho R, Fliser D, Tu C, Zee J, Bieber B, Wong MMY, et al. Prescription of renin-angiotensin-aldosterone system inhibitors (RAASi) and its determinants in patients with advanced CKD under nephrologist care. J Clin Hypertens (Greenwich). 2019;21(7):991-1001.
26. Pugh D, Gallacher PJ, Dhaun N. Management of Hypertension in Chronic Kidney Disease. Drugs. 2019;79(4):365-79.
27. Ku E, Lee BJ, Wei J, Weir MR. Hypertension in CKD: Core Curriculum 2019. Am J Kidney Dis. 2019;74(1):120-31.
28. Sinha AD, Agarwal R. Clinical Pharmacology of Antihypertensive Therapy for the Treatment of Hypertension in CKD. Clin J Am Soc Nephrol. 2019;14(5):757-64.
29. Evans M, Suttorp MM, Bellocco R, Hoekstra T, Qureshi AR, Dekker FW, et al. Trends in haemoglobin, erythropoietin-stimulating agents and iron use in Swedish chronic kidney disease patients between 2008 and 2013. Nephrol Dial Transplant. 2016;31(4):628-35.

30. Evans M, Carrero JJ, Bellocco R, Barany P, Qureshi AR, Seeberger A, et al. Initiation of erythropoiesis-stimulating agents and outcomes: a nationwide observational cohort study in anaemic chronic kidney disease patients. Nephrol Dial Transplant. 2017;32(11):1892-901.
31. Swedish Renal Registry: Annual report 2018. www.snronline.se (2 April 2019, date last accessed).
32. Wettermark B, Hammar N, Fored CM, Leimanis A, Otterblad Olausson P, Bergman U, et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. Pharmacoepidemiol Drug Saf. 2007;16(7):726-35.
33. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.
34. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. Eur J Epidemiol. 2017;32(9):765-73.
35. 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.
36. Fu EL, Groenwold RHH, Zoccali C, Jager KJ, van Diepen M, Dekker FW. Merits and caveats of propensity scores to adjust for confounding. Nephrol Dial Transplant. 2019;34(10):1629-35.
37. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. Stat Med. 1990;9(12):1501-15.
38. Danaei G, Garcia Rodriguez LA, Cantero OF, Logan RW, Hernan MA. Electronic medical records can be used to emulate target trials of sustained treatment strategies. J Clin Epidemiol. 2018;96:12-22.
39. Noordzij M, Leffondre K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ. When do we need competing risks methods for survival analysis in nephrology? Nephrol Dial Transplant. 2013;28(11):2670-7.
40. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. Am J Epidemiol. 2009;170(2):244-56.
41. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. Circulation. 2016;133(6):601-9.
42. Izem R, Liao J, Hu M, Wei Y, Akhtar S, Wernecke M, et al. Comparison of propensity score methods for pre-specified subgroup analysis with survival data. J Biopharm Stat. 2020;30(4):734-51.
43. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.
44. Granger E, Sergeant JC, Lunt M. Avoiding pitfalls when combining multiple imputation and propensity scores. Stat Med. 2019;38(26):5120-32.
45. Casas JP, Chua W, Loukogeorgakis S, Vallance P, Smeeth L, Hingorani AD, et al. Effect of inhibitors of the renin-angiotensin system and other antihypertensive drugs on renal outcomes: systematic review and meta-analysis. Lancet. 2005;366(9502):2026-33.
46. Maione A, Navaneethan SD, Graziano G, Mitchell R, Johnson D, Mann JF, et al. Angiotensin-converting enzyme inhibitors, angiotensin receptor blockers and combined therapy in patients with micro- and macroalbuminuria and other cardiovascular risk factors: a systematic review of randomized controlled trials. Nephrol Dial Transplant. 2011;26(9):2827-47.

47. Fink HA, Ishani A, Taylor BC, Greer NL, MacDonald R, Rossini D, et al. Screening for, monitoring, and treatment of chronic kidney disease stages 1 to 3: a systematic review for the U.S. Preventive Services Task Force and for an American College of Physicians Clinical Practice Guideline. Ann Intern Med. 2012;156(8):570-81.

48. Grimaldi-Bensouda L, Klungel O, Kurz X, de Groot MC, Maciel Afonso AS, de Bruin ML, et al. Calcium channel blockers and cancer: a risk analysis using the UK Clinical Practice Research Datalink (CPRD). BMJ Open. 2016;6(1):e009147.

49. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney inter, Suppl. 2013;3:1-150.

50. Blood Pressure Lowering Treatment Trialists C, Ninomiya T, Perkovic V, Turnbull F, Neal B, Barzi F, et al. Blood pressure lowering and major cardiovascular events in people with and without chronic kidney disease: meta-analysis of randomised controlled trials. BMJ. 2013;347:f5680.

Supplementary Material

Supplemental methods.

Supplemental Table S1. Definition of medications and comorbidities.

Supplemental Table S2. Comparison of baseline characteristics for individuals with and without baseline ACR measurements.

Supplemental Table S3. Absolute risks and risk differences for KRT, mortality and MACE after 1, 2, 3, 5 and 10 years of follow-up.

Supplemental Table S4. Subgroup analyses for systolic blood pressure using 130 mmHg as cutpoint.

Supplemental Table S5. Baseline characteristics of patients in the positive control cohort with CKD G3 by RASi or CCB treatment, before and after inverse probability weighting.

Supplemental Table S6. Number of events, incidence rates and crude and adjusted hazard ratios for the association between RASi vs. CCB initiation and all-cause mortality, MACE and KRT in the positive control cohort of patients with CKD G3.

Supplemental Table S7. Adjusted hazard ratios for the sensitivity analyses.

Supplemental Table S8. Number of events, incidence rates as well as crude and adjusted hazard ratios for the association between RASi vs. CCB initiation and MACE plus (composite of cardiovascular death, hospitalization due to stroke, myocardial infarction or heart failure).

Supplemental Figure S1. Flow diagram depicting the assembly of the study cohort.

Supplemental Figure S2. Number of events, incidence rates and adjusted hazard ratios for mortality following RASi vs. CCB initiation, according to subgroups of age, sex, diabetes, ACR, heart failure, systolic blood pressure and eGFR.

Supplemental Figure S3. Number of events, incidence rates and adjusted hazard ratios for MACE following RASi vs. CCB initiation, according to subgroups of age, sex, diabetes, ACR, heart failure, systolic blood pressure and eGFR.

Supplemental Methods

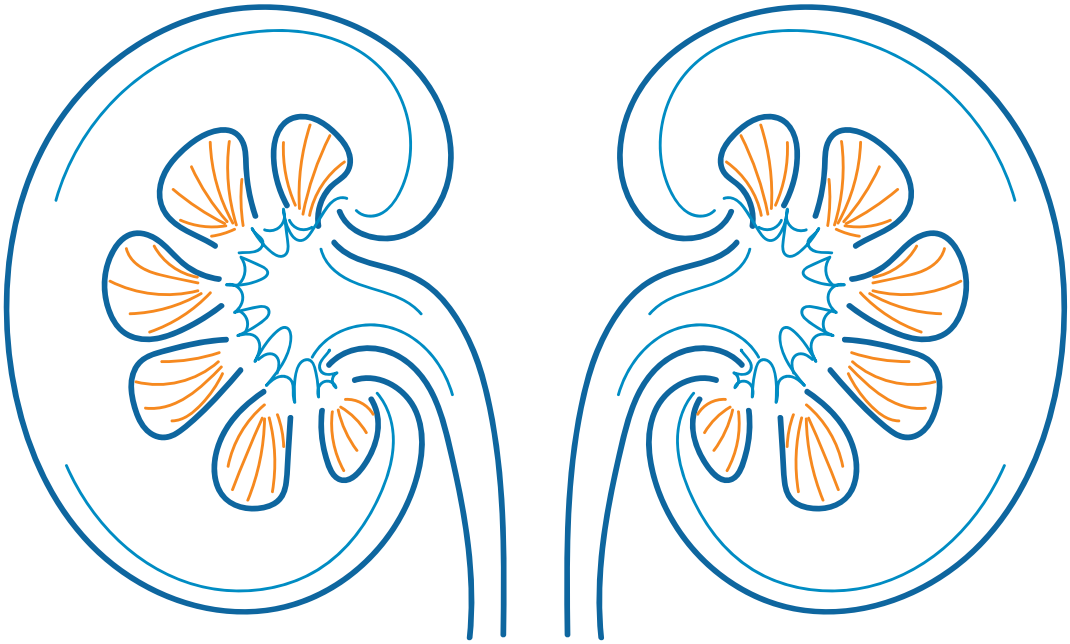
For the as-treated analysis, we censored individuals when they either discontinued therapy (no dispensation for the index drug within 60 days after the estimated last day of supply), or switched treatment (on the day of a prescription of the drug different from the index drug). To adjust for the time-varying selection bias that is introduced when censoring individuals if they deviate from the initiated medication at cohort entry, we used inverse probability of censoring weighting (IPCW). We constructed our dataset into monthly intervals and updated all comorbidities and medication use at each month. For each subject we estimated a weight that was, informally defined, proportional to the inverse of the probability of observing one's censoring history. The stabilized censoring weight at month t was calculated as

$$swC_t = \prod_{k=0}^{t} \frac{\Pr(C_{k+1} = 0 \mid \bar{C}_k = \bar{0}, \bar{A}_k, V)}{\Pr(C_{k+1} = 0 \mid \bar{C}_k = \bar{0}, \bar{A}_k, V, \bar{L}_k)}$$

where C_t indicates censoring status, A_t treatment history, V is a vector of time-fixed covariates at baseline and L_t represents the time-varying covariates. The denominator of swC_t adjusts for the informative censoring and is the probability of being uncensored in month k, conditional on past censoring history, treatment history, time-fixed covariates (measured at baseline), and time-varying covariates. The numerator of swC_t is not required for censoring adjustment but is used to stabilize the weights and improve statistical efficiency. The numerator represents the probability of remaining uncensored in month k, conditional on censoring history, treatment history, and time-fixed covariates. To estimate the weights, two separate pooled logistic models were fitted for the numerator and denominator respectively:

$$\text{logit}\,[\Pr(C_t = 0 \mid Y_t = C_{t-1} = 0, A, X, V)] = \eta_0 + \eta_1 t + \eta_2 t^2 + \eta_3 A + \eta_4 V + \eta_4 X$$
$$\text{logit}\,[\Pr(C_t = 0 \mid Y_t = C_{t-1} = 0, A, X, V)] = \eta_0 + \eta_1 t + \eta_2 t^2 + \eta_3 A + \eta_4 V + \eta_4 X + \eta_5 L_t$$

As time-fixed covariates we used age and sex, and as time-varying covariates we used all comorbidities, medications and hospitalizations listed in **Supplemental Table S1** as well as eGFR, systolic and diastolic blood pressure. The stabilized censoring weights had a mean of 1.00 and ranged from 0.30 to 5.65, indicating no violation of the positivity assumption. A weighted Cox model was then used to calculate adjusted hazard ratios for mortality, MACE and kidney replacement therapy for the as-treated analysis, using the estimated stabilized censoring weights, and additionally adjusting for all baseline covariates. Robust variance estimation was used to derive conservative 95% confidence intervals. A similar procedure was used for the competing risk analysis.



CHAPTER 7

Association between β-blocker use and mortality/morbidity in patients with heart failure with reduced, midrange, and preserved ejection fraction and advanced chronic kidney disease

Edouard L. Fu, Alicia Uijl, Friedo W. Dekker, Lars H. Lund, Gianluigi Savarese, Juan-Jesus Carrero

Circ Heart Fail 2020; 13: e007180

Abstract

Background: It is unknown if beta-blockers reduce mortality/morbidity in patients with heart failure (HF) and advanced chronic kidney disease (CKD), a population underrepresented in HF trials.

Methods: Observational cohort of HF patients with advanced CKD (eGFR [estimated glomerular filtration rate] <30 mL/min/1.73m^2) from the Swedish Heart Failure Registry between 2001-2016. We first explored associations between beta-blocker use, 5-year death, and the composite of cardiovascular (CV) death/HF hospitalization among 3,775 patients with HF with reduced ejection fraction (HFrEF) and advanced CKD. We compared observed hazards with those from a control cohort of 15,346 patients with HFrEF and moderate CKD (eGFR<60-30 mL/min/1.73m^2), for whom beta-blocker trials demonstrate benefit. Secondly, we explored outcomes associated to beta-blocker among advanced CKD participants with preserved (HFpEF; N=2,009) and midrange ejection fraction (HFmrEF; N=1,514).

Results: During median 1.3 years, 2,012 patients had a subsequent HF hospitalization, and 2,849 died in the HFrEF cohort, of which 2,016 due to cardiovascular causes. Among patients with HFrEF, beta-blocker use was associated with lower risk of death (adjusted hazard ratio 0.85; 95% confidence interval 0.75-0.96) and CV mortality/HF hospitalization (0.87; 0.77-0.98) compared to non-use. The magnitude of the associations was similar to that observed for HFrEF patients with moderate CKD. Conversely, no significant association was observed for beta-blocker users in advanced CKD with HFpEF (death: 0.88; 0.77-1.02, CV mortality/HF: 1.05; 0.90-1.23) or HFmrEF (death: 0.95; 0.79-1.14, CV mortality/HF: 1.09; 0.90-1.31).

Conclusion: In HFrEF patients with advanced CKD, use of beta-blockers was associated with lower morbidity and mortality. Although inconclusive due to limited power, these benefits were not observed in similar patients with HFpEF or HFmrEF.

Introduction

Chronic kidney disease (CKD) is highly prevalent in patients with heart failure (HF) and their coexistence is increasing due to an ageing population and shared risk factors and mechanisms (1-3). Patients with HF and CKD experience significant morbidity and mortality (4, 5), which is highest in those with advanced CKD (estimated glomerular filtration rate (eGFR) <30 ml/min/1.73m^2) (2, 6-8). Although persons with advanced CKD typically represent 10-15% of the HF population, they have been systematically excluded or underrepresented in HF clinical trials (9-12), leading to uncertainty about the effect of therapies and optimal management for them (13).

Beta-blockers are a cornerstone in the treatment of patients with HF with reduced ejection fraction (HFrEF) since they substantially reduce mortality and morbidity (13-21). A recent meta-analysis of randomized trials showed consistent benefits of beta-blockers in patients with HFrEF and moderate CKD (eGFR 30-60 ml/min/1.73m^2), but there were too few HF patients with advanced CKD (less than 3% of all patients included in the trials) to draw firm conclusions (22). Furthermore, the few observational studies conducted to date show inconsistent results, being limited by a small number of patients with advanced CKD (23-25) and/or lacking information on ejection fraction (26, 27).

We here sought to evaluate outcomes associated with the use of beta-blockers in a large, contemporary, and nationwide routine-cared cohort of patients with HFrEF and advanced CKD. As a secondary objective, we investigated whether potential benefits of beta-blockers may also extend to patients with advanced CKD and heart failure with midrange (HFmrEF) or preserved ejection fraction (HFpEF), for whom no beta-blocker trial evidence exist.

Methods

Study protocol and setting

The data that support the findings of this study are available from the corresponding author, provided that data sharing is permitted by European Union General Data Protection Regulation regulations and appropriate ethics committees. The Swedish Heart Failure Registry has been described previously (28). The inclusion criterion is clinician-judged heart failure. Approximately 80 variables are recorded at hospital discharge or after an out-patient clinic visit and entered into a web-based database managed by the Uppsala Clinical Research Center. Ejection fraction is categorized as <30%, 30-39%, 40-49% and >=50%. Deaths and causes of death are obtained from the Swedish Population Registry monthly. The National Patient Registry was

used to obtain information on additional baseline comorbidities and the outcomes hospitalization due to HF, hospitalization due to syncope and cancer. Variable definitions are reported in **Supplemental Table S1**. Linkage with Statistics Sweden provided socioeconomic characteristics. Individual patient consent is not required, but patients are informed of entry into national registries and have the opportunity to opt out. This study was approved by a multisite ethics committee and complies with the Declaration of Helsinki.

Study population

Patients registered between 11 May 2000 and 31 December 2016 with an eGFR <30 ml/min/1.73m^2 at time of registration and no missing data for beta-blocker use or ejection fraction were considered for this study. Patients receiving beta-blockers other than those recommended by HF guidelines (i.e. bisoprolol, carvedilol, or metoprolol) and those that died during the index hospitalization/outpatient visit were excluded. The index date was defined as the date of hospital discharge or date of outpatient clinic visit. If the same patient was registered more than once, we considered the first registration with eGFR <30 ml/min/1.73m^2. eGFR was calculated using the CKD Epidemiology Collaboration equation (29). Patients undergoing chronic dialysis at index date where considered to have advanced CKD. Individuals were followed from index date until occurrence of an event or end of follow-up (31 December 2016), whichever occurred first. A flow chart describing patient flow is reported in **Supplemental Figure S1**.

Outcomes

Our primary outcome was mortality due to any cause up to 5 years. Secondary outcomes included a combined endpoint of 5-year cardiovascular mortality and HF hospitalization (definitions in **Supplementary Table S1**), and each component separately. As safety outcome we considered hospitalization for syncope, as beta-blocker use is associated with increased risk of bradycardia and hypotension (10). As a negative control outcome, we used hospitalization for cancer.

Covariates

Study covariates were recorded at HF registration/discharge and were used in multivariable adjustments, and included age, sex, civil status, location (inpatient or outpatient), follow-up referral specialty, New York Heart Association (NYHA) class, left ventricular ejection fraction (LVEF) [<30 vs. 30-39% in HFrEF analyses; EF not used for adjustment in the HFpEF or HFmrEF analyses], systolic, diastolic and mean arterial pressure, heart rate, eGFR, heart failure duration, comorbidities

(hypertension, diabetes mellitus, smoking, ischemic heart disease, peripheral artery disease, stroke/transient ischemic attack, atrial fibrillation, anemia, valvular disease, lung disease, dilated cardiomyopathy), concomitant medications (renin-angiotensin-system inhibitors [RASi], mineralocorticoid receptor antagonists [MRA], digoxin, diuretic, nitrate, platelet inhibitor, oral anticoagulant, statins) and history of interventions (revascularization, valve intervention, pacemaker, cardiac resynchronization therapy, implantable cardioverter defibrillator). We further extracted information on NT-proB-type Natriuretic Peptide and body mass index but did not adjust for these variables due to a high proportion of missing values.

Statistical analysis

Continuous variables are presented as mean with SD or median with interquartile range (IQR), depending on the distribution, and categorical variables as number and percentages.

The primary analysis compared outcomes associated with beta-blocker use in patients with HFrEF (ejection fraction ≤39%). Incidence rates per 100 person-years with 95% confidence intervals (95% confidence intervals [CI]) were calculated for each outcome. We computed survival curves standardized to the distribution of the baseline variables in the study population to provide absolute survival probabilities and risk differences (30, 31). Survival probabilities were log-log transformed before pooling and combined using Rubin's rules (32). The combined estimates were back transformed onto the original scale after pooling. Multivariable Cox proportional hazards regression was used to estimate hazard ratios (HRs) for the association between beta-blocker use and outcomes. The proportional hazards assumption was verified by assessment of the Schoenfeld residuals. We performed subgroup analyses in a priori defined strata of sex, location, NYHA class (I/II vs. III/IV), ejection fraction (<30% vs. 30-39%), eGFR (<15 ml/min/1.73m^2 vs. 15-30 ml/min/1.73m^2), atrial fibrillation, diabetes, hypertension, ischemic heart disease and COPD, and non-prespecified subgroups of RASi and MRA use. In addition, we compared outcomes according to the beta-blocker dose received.

Observed estimates were contrasted with those from a positive control cohort of patients with HFrEF and moderate CKD (eGFR between 30-60 ml/min/1.73m^2), for whom a risk benefit has been observed in landmark trials (10-12, 33, 34). The positive control cohort was defined in the same way as our primary cohort. As a sensitivity analysis we repeated our analyses using a maximum follow-up of 1 year. Furthermore, to evaluate the extent of residual confounding, we used hospitalization for cancer as a negative control outcome, which is not expected to be associated with beta-blocker use.

Finally, we evaluated outcomes associated with beta-blocker use in persons with advanced CKD and HFmrEF (ejection fraction 40-49 %) or HFpEF (ejection fraction ≥50%) separately, in a manner identical to our primary analysis.

Missing confounder values were imputed using a multiple imputation by chained equations algorithm (generating 50 imputed datasets), including the confounder information, beta-blocker use, the censoring indicator of the composite outcome and the Nelson-Aalen estimate of the cumulative hazard. Missing data for each variable are reported in **Supplemental Table S2** for all cohorts separately. Statistical analyses were performed using R version 3.6.2.

Table 1. Baseline characteristics of individuals with HFrEF and advanced CKD (eGFR<30 ml/min/1.73m^2), overall and stratified by beta-blocker use.

	Beta-blocker users (N = 3371)	**Beta-blocker non-users (N = 404)**
Age, years, median (IQR)	80 [74, 85]	82 [75, 87]
Women (%)	1213 (36)	145 (36)
Location, outpatient (%)	1109 (33)	84 (21)
Follow-up location, specialty (%)	1830 (58)	168 (47)
NYHA class (%)		
I	70 (3)	7 (3)
II	670 (28)	54 (22)
III	1355 (57)	128 (53)
IV	279 (12)	54 (22)
EF (%)		
<30	1721 (51)	218 (54)
30-39	1650 (49)	186 (46)
Clinical measures		
BMI (kg/m^2), mean (SD)	27 (5)	26 (5)
SBP (mmHg), mean (SD)	122 (22)	122 (23)
DBP (mmHg), mean (SD)	70 (12)	70 (13)
MAP (mmHg), mean (SD)	87 (14)	87 (15)
Heart rate (bpm), median [IQR]	75 (16)	76 (17)
eGFR (mL/min/1.73m^2), median [IQR]	25 [20, 28]	24 [19, 28]
eGFR <15 ml/min/1.73m^2 (%)	347 (10)	57 (14)
eGFR between 15-30 ml/min/1.73m^2 (%)	3024 (90)	347 (86)

	Beta-blocker users (N = 3371)	Beta-blocker non-users (N = 404)
NT-proBNP, pg/L, median [IQR]	9176 [3914, 19894]	9950 [4241, 24107]
Smoking (%)		
Never	1100 (44)	122 (45)
Former	1176 (47)	123 (45)
Current	209 (8)	29 (11)
Medical history (%)		
Atrial fibrillation	2084 (62)	235 (58)
Anaemia	2031 (61)	254 (63)
COPD	553 (16)	73 (18)
Dilated cardiomyopathy	379 (12)	50 (13)
Diabetes	1339 (40)	147 (36)
Hypertension	2573 (76)	274 (68)
Ischemic heart disease	2542 (75)	272 (67)
Peripheral artery disease	632 (19)	89 (22)
Stroke and/or TIA	636 (19)	94 (23)
Valvular disease	1204 (36)	167 (41)
Cancer in the previous 3 years	418 (12)	66 (16)
Procedures		
Coronary revascularization	1410 (42)	138 (34)
Devices (CRT or ICD)	412 (12)	25 (6)
Pacemaker (CRT-D, CRT-P or pacemaker)	668 (20)	71 (18)
Medication use (%)		
RAS inhibitors	2320 (69)	215 (53)
MRA	827 (25)	109 (27)
Digoxin	313 (9)	36 (9)
Diuretics	3178 (95)	374 (94)
Statins	1661 (49)	141 (35)
Anticoagulants	1358 (40)	130 (32)
Antiplatelets	1798 (54)	192 (48)
Nitrates	928 (28)	93 (23)

	Beta-blocker users (N = 3371)	Beta-blocker non-users (N = 404)
Socioeconomic characteristics (%)		
Marital status		
Married	1600 (48)	191 (47)
Single	742 (22)	84 (21)
Widowed	1022 (30)	129 (32)
Education level		
Compulsory school	1751 (54)	209 (53)
Secondary school	1132 (35)	135 (34)
University	387 (12)	51 (13)
Income > median	1511 (45)	173 (43)

NYHA = New York Heart Association; EF = ejection fraction; BMI = body mass index; SBP = systolic blood pressure; DBP = diastolic blood pressure; MAP = mean arterial pressure; eGFR = estimated Glomerular Filtration Rate; NT-proBNP = N-terminal pro-B-type natriuretic peptide; CRT = cardiac resynchronization therapy; ICD = implantable cardioverter-defibrillator; CRT-D = cardiac resynchronization therapy with defibrillation; CRT-P = cardiac resynchronization therapy with pacemaker.

Results

Among a total of 76,506 patients in the Swedish Heart Failure Registry, 7,298 had advanced CKD (**Supplemental Figure S1**). Based on LVEF evaluation, 3,775 were classified as HFrEF, 2,009 as HFpEF and 1,514 as HFmrEF. Characteristics for the overall HF cohort are shown in **Supplemental Table S3**; beta-blockers were used in 6,317 (87%) individuals.

Primary analysis: Beta blockers in HFrEF with advanced CKD

Baseline characteristics for the HFrEF cohort, stratified by beta-blocker use are reported in **Table 1**. Of the 3,775 patients with HFrEF, 3,371 (89%) were treated with beta blockers and 404 were not (11%). The majority of patients received metoprolol (53%), followed by bisoprolol (41%) and carvedilol (6%). As many as 26% received target doses, 36% received 50-99% of the target dose, and the remaining 38% received <50% of target dose (**Supplemental Table S4**, **Supplemental Figure S2**). Median (IQR) age was 80 (74-85) years among beta-blocker users, compared with 82 (75-87) years among non-users, and in both groups the proportion of women was 36%. Among beta-blocker users, 51% had an ejection fraction <30%, compared with 54% among non-users. Atrial fibrillation was a common comorbidity, occurring in 62% of beta-blocker users, and 58% of non-users (**Table 1**).

The median follow-up time was 1.3 years, for a total of 6,138 person-years of follow-up. A total of 2,849 (75.5%) individuals died, of whom 2,016 (70.8% of total deaths) due to cardiovascular causes. The 5-year incidence rate of all-cause mortality was 44.8 per 100 (95% CI 43.1-46.6) person-years among beta-blocker users vs. 64.0 (57.2-71.3) for non-users (**Figure 1**). The 5-year survival was 12.9% for non-users and 16.2% for beta-blocker users (**Figure 2**, **Supplemental Table S5**). Compared to no-use, beta-blocker users had a 3.2% (95% CI 0.9%-5.6%) lower mortality risk, with an adjusted HR of 0.85 (95% CI 0.75-0.96). A total of 2,779 (73.6%) patients experienced the composite outcome of CV mortality or HF hospitalization, with again a lower incidence among beta-blocker users (incidence rate 69.8; 95% CI 67.2-72.5) than among for non-users (incidence rate 92.3; 95% CI 82.3-103.1).

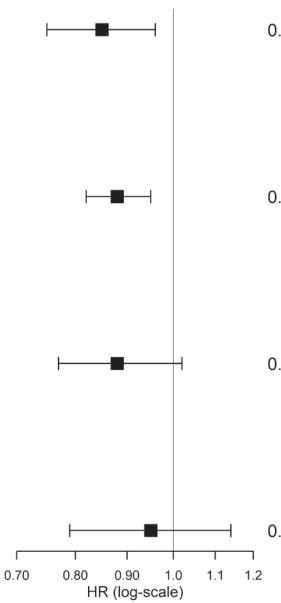
The 5-year composite-free survival was 10.3% among non-users and 12.9% for beta-blocker users (**Figure 2**, **Supplemental Table S6**). Compared to no-use, beta-blocker users had a 2.6% (95% CI 0.3%-4.8%) lower CV mortality/HF hospitalization risk (HR 0.87; 95% CI 0.77-0.98), primarily attributed to a reduction in cardiovascular death (HR 0.81; 0.71-0.93), whereas the adjusted HR for heart failure hospitalization was 0.94 (95% CI 0.81-1.10) (**Supplemental Figure S3**). Results were similar when using a shorter maximum 1-year follow-up (**Supplemental Table S7**). No differences were observed for the safety outcome, risk of syncope hospitalization, with a HR of 0.99 (95% CI 0.47-2.07) for beta-blocker users compared with no use. We also observed no association between beta-blocker use and the "negative control outcome" of cancer hospitalization, with a HR of 1.08 (0.63-1.84) (**Supplemental Table S8**).

Stratified analyses (**Figure 3**) showed significant interaction terms, with the association between beta-blocker use and mortality being stronger for inpatient than for outpatient cases, and also stronger in the absence of ischemic heart disease and those not receiving RASi. The association between beta-blocker use and the composite outcome was more favorable in patients with an eGFR <15 ml/min/1.73m^2 than in those with an eGFR between 15-<30 ml/min/1.73m^2, among those without atrial fibrillation and those not receiving RASi. Compared to non-use, the observed point estimates for benefit of beta-blocker use were similar regardless of the dose prescribed, although the confidence intervals exclude 1 only for doses that are 50% or more of target (**Supplemental Tables S9-10**).

Figure 1. Number of events, incidence rates and adjusted hazard ratios for the association between beta-blocker use and 5-year all-cause mortality, and composite outcome of CV mortality/HF hospitalization in patients with HFrEF, positive control cohort, HFpEF and HFmrEF.

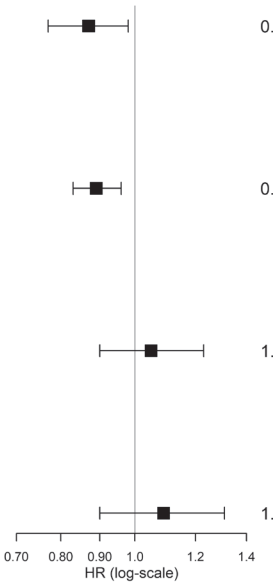
A. All-cause mortality

	No. of events	Person-years	Incidence rate (95% CI)	HR (95% CI)
HFrEF (EF ≤39) + advanced CKD				
Overall	2849	6138	46.4 (44.7-48.2)	
No BB	329	514	64 (57.2-71.3)	
BB	2520	5623	44.8 (43.1-46.6)	0.85 (0.75-0.96)
HFrEF (EF ≤39) + moderate CKD (positive control)				
Overall	7468	40121	18.6 (18.2-19)	
No BB	888	3351	26.5 (24.8-28.3)	
BB	6580	36769	17.9 (17.5-18.3)	0.88 (0.82-0.95)
HFpEF (EF ≥50) + advanced CKD				
Overall	1405	3515	40 (37.9-42.1)	
No BB	269	561	47.9 (42.4-54)	
BB	1136	2953	38.5 (36.3-40.8)	0.88 (0.77-1.02)
HFmrEF (EF 40-49) + advanced CKD				
Overall	1073	2605	41.2 (38.8-43.7)	
No BB	167	325	51.5 (43.9-59.9)	
BB	906	2280	39.7 (37.2-42.4)	0.95 (0.79-1.14)



B. Composite CV death/HF hospitalization

	No. of events	Person-years	Incidence rate (95% CI)	HR (95% CI)
HFrEF (EF ≤39) + advanced CKD				
Overall	2779	3981	69.8 (67.2-72.5)	
No BB	313	339	92.3 (82.3-103.1)	
BB	2466	3642	67.7 (65.1-70.4)	0.87 (0.77-0.98)
HFrEF (EF ≤39) + moderate CKD (positive control)				
Overall	8945	28248	31.7 (31-32.3)	
No BB	942	2387	39.5 (37-42.1)	
BB	8003	25861	30.9 (30.3-31.6)	0.89 (0.83-0.96)
HFpEF (EF ≥50) + advanced CKD				
Overall	1301	2415	53.9 (51-56.9)	
No BB	223	411	54.2 (47.3-61.8)	
BB	1078	2004	53.8 (50.6-57.1)	1.05 (0.90-1.23)
HFmrEF (EF 40-49) + advanced CKD				
Overall	1006	1803	55.8 (52.4-59.3)	
No BB	149	257	58 (49.1-68.1)	
BB	857	1547	55.4 (51.8-59.3)	1.09 (0.90-1.31)



Legend: Analyses were adjusted for age, sex, NYHA class, ejection fraction, mean arterial pressure, heart rate, eGFR, smoking, atrial fibrillation, anemia, COPD, dilated cardiomyopathy, diabetes, hypertension, ischemic heart disease, peripheral artery disease, stroke/TIA, valvular disease, cancer, dementia, coronary revascularization, devices (CRT or ICD), RAS inhibitor, MRA, digoxin, diuretics, statins, anticoagulants, antiplatelets, nitrates, marital status, education level, income, location (inpatient/outpatient), follow-up location (specialist yes/no), index year and duration of heart failure. Incidence rates are depicted per 100 person years. HFrEF was defined as ejection fraction ≤39%, HFmrEF between 40-49% and HFpEF ≥50%. EF: ejection fraction.

Figure 2. Standardized survival curves for the association between beta-blocker use and all-cause mortality and the composite outcome cardiovascular mortality or heart failure hospitalization. Legend: Panels A and B: patients with HFrEF and advanced CKD. Panels C and D: patients with HFrEF and moderate CKD (positive control analysis). HR = hazard ratio; CI = confidence interval; ARD = absolute risk difference at 5 years.

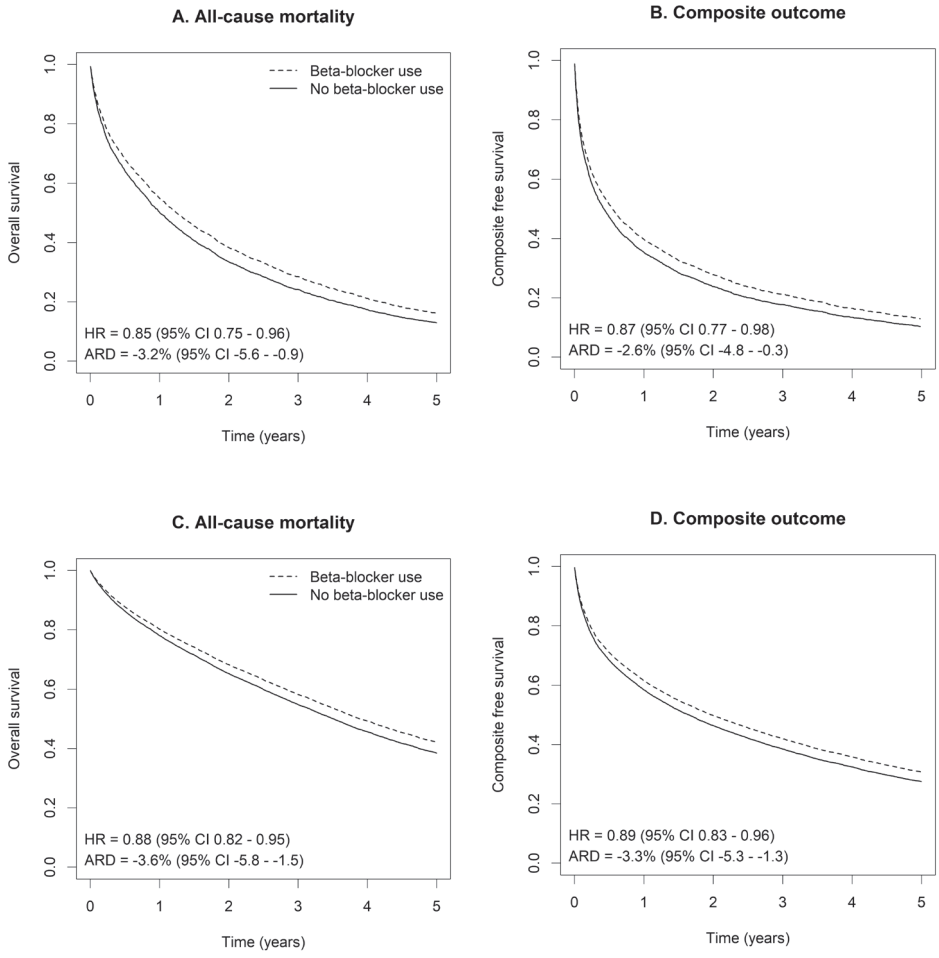
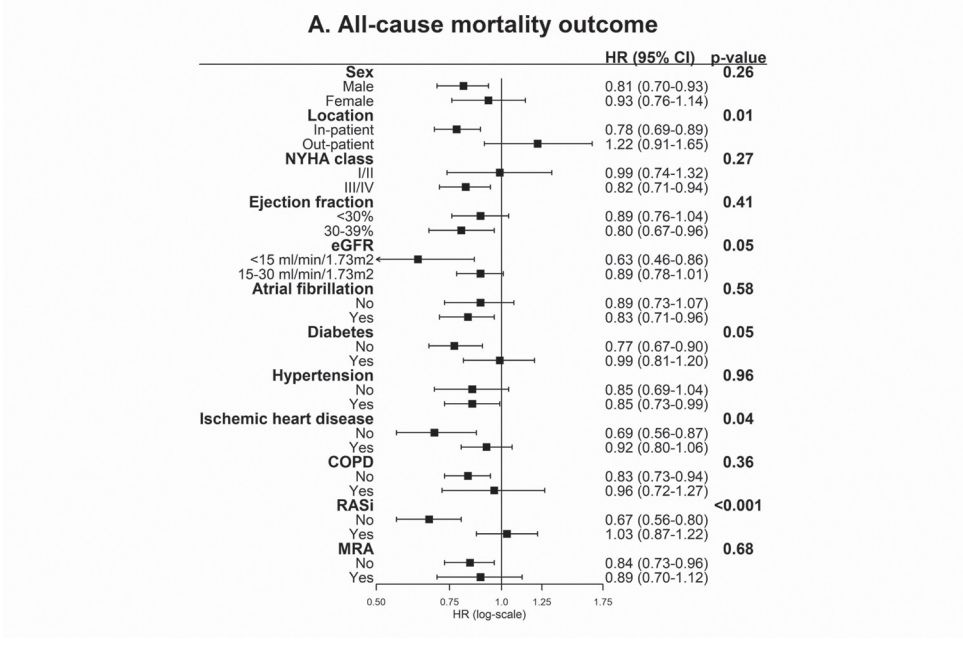
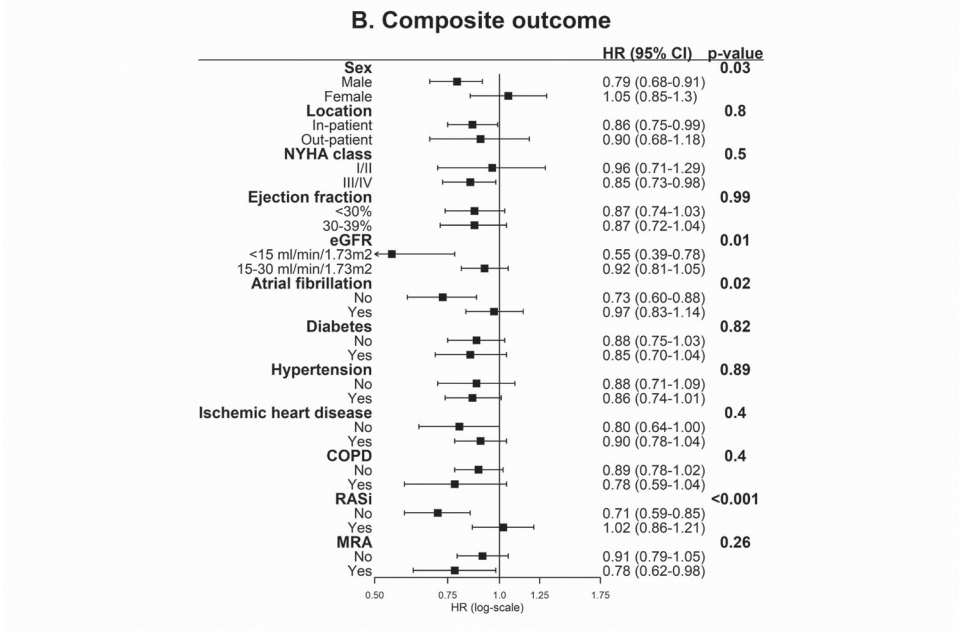


Figure 3. Association between beta-blocker use, 5-year all-cause mortality (A) and the composite of cardiovascular mortality and heart failure hospitalization (B) in subgroups of patients with HFrEF and advanced CKD.





Positive control cohort: Beta blockers in HFrEF with moderate CKD

From a total of 15,346 identified individuals with HFrEF and moderate CKD, 13,890 (90.5%) were treated with beta-blockers. Median eGFR was 48 ml/min/1.73m^2 (IQR 40-54), 60.2% had CKD G3a and 39.8% CKD G3b (**Supplemental Table S11**). The pattern of beta-blocker drug class use was similar to that observed for patients with advanced CKD (**Supplemental Table S4**). During follow up, they experienced a much lower event rate for all-cause mortality (incidence rate 18.6; 95% CI 18.2-19.0) and the composite outcome (incidence rate 31.7; 95% CI 31.0-32.3) than patients with advanced CKD (**Figure 1**). The 5-year survival was 38.4% for non-users and 42.0% for beta-blocker users (**Figure 2**, **Supplemental Table S6**). Compared to no-use, patients receiving beta-blockers had a 3.6% (95% CI 1.5%-5.8%) lower risk of death (HR 0.88; 95% CI 0.82-0.95). The risk of CV death/HF hospitalization was also lower among beta-blocker users (HR 0.89; 95% CI 0.83-0.96), attributed both to a lower cardiovascular death risk (HR 0.86; 95% CI 0.79-0.94) and a lower heart failure hospitalization risk (HR 0.88; 0.81-0.96) (**Supplemental Figure S3**).

Secondary analyses: Beta blockers in HFpEF and HFmrEF with advanced CKD

We identified 2,009 individuals with HFpEF and 1,514 individuals with HFmrEF and advanced CKD. In patients with HFpEF, 1,649 (82.1%) used beta-blockers, and 1,297 (85.7%) patients with HFmrEF used beta-blockers. Their characteristics are shown in **Supplemental Tables S12-13**, and the number of outcomes during follow up in **Figure 1** and **Supplemental Figure S3**. The pattern of specific beta-blocker class and recommended target dose within each class were similar to our primary analysis (**Supplemental Table S4**). In patients with HFpEF the use of beta-blockers did not significantly associate with the risk of death (0.88; 0.77-1.02) or CV death/HF hospitalization (1.05; 0.90-1.23) (**Figure 1**). The association was neither observed in beta-blocker users with HFmrEF (HR 0.95; 95% CI 0.79-1.14 for death and 1.09; 0.90-1.31 for CV death/HF hospitalization) (**Supplemental Figure S3**).

Discussion

This large prospective registry analysis of patients with HF and advanced CKD has the following findings: 1) Overall beta-blocker use was high despite lack of trial evidence; 2) use of beta-blocker in HFrEF and advanced CKD was associated with lower risk of all-cause mortality and the composite outcome of CV mortality/HF hospitalization. The observed risk magnitude was similar to that of patients with HFrEF and moderate CKD; (3) use of beta-blockers in HFmrEF or HFpEF and advanced CKD showed inconsistent and non-significant associations with study outcomes.

Between 10-15% of patients with HF have advanced CKD (2, 7). This population is at the highest risk of complications and (cardiovascular) death (5, 6, 8, 9, 35), attributed to the coexistence of both traditional (such as hypertension, dyslipidemia and diabetes) and nontraditional cardiovascular risk factors (inflammation, mineral and bone disorders, oxidative stress, and clinical frailty) that emerge with the failing kidney (36, 37). In our study, we indeed observed that the incidence rates of death or composite CV death/HF outcomes were doubled in those with advanced CKD compared with the moderate CKD positive control cohort. Since the event rates are much higher, the absolute risk reduction of beta-blocker use may actually be largest in individuals with the lowest kidney function, similarly to what has been observed for RASi-inhibitors in HFrEF and advanced CKD (38) or older age (39).

Beta-blockers are class I guideline-recommended therapies for patients with HFrEF (13, 14), without specifications by severity of CKD. A recent meta-analysis which pooled results of 16,740 patients from ten placebo-controlled trials, reported consistency in the death risk reduction of beta-blockers for persons with moderate CKD (eGFR 30-60 ml/min/1.73m^2), reporting a HR of 0.73 (95% CI 0.62-0.86) for patients with an eGFR of 45-59 ml/min/1.73m^2 and of 0.71 (95% CI 0.58-0.87) for patients with eGFR 30-44 ml/min/1.73m^2 (22). The results from our positive control cohort align with these findings and found a slightly lower HR of 0.88 (0.82-0.95) for mortality. However, we note that patients in our routine-care cohort were considerably older (78 vs. 68 years, respectively) and used different medications (MRA use 38% in our cohort vs. 10% in the trials, respectively) than the patients included in those trials.

Beta-blocker use in HFrEF and advanced CKD

There is a lack of evidence-based therapies for HFrEF patients with advanced CKD as they have been severely underrepresented in landmark randomized trials (9-13, 16-22). In the recent meta-analysis of 10 pooled randomized trials in HFrEF, only 448 out of 16,740 patients (2.7%) were identified to have advanced CKD at inclusion (22). Due to this low number the authors were unable to comment on the efficacy of beta-blockers in this population. Despite a lack of trial evidence, the majority (89%) of advanced CKD patients in our register used beta-blockers. However, we note that a large proportion did not receive the recommended target dose, perhaps due to fear for side effects in this vulnerable population. Our main analysis in HFrEF patients with advanced CKD suggests a possible therapeutic benefit similar to that observed for persons with moderate CKD. In support of our findings, a recent Canadian observational study (although small, with a sample size of only 200) reported a HR of 0.55 (95% CI 0.41-0.73) in the risk of death in elderly patients with HF and advanced CKD initiating beta-blockers versus no use (26). However, this study lacked information on ejection fraction. Importantly, subgroup analyses in our study showed that the benefit on all-cause mortality and CV mortality/HF hospitalization also extended to those with the lowest level of kidney function

(eGFR <15 ml/min/1.73m^2) and indicated no increased risk for syncope, although confidence intervals were wide. In addition, the negative control outcome indicated no increased risk for cancer, thereby strengthening our inferences that observed differences are not primarily explained by a worse health status. Our subgroup analyses indicated no benefit of beta-blocker use with regard to CV mortality/HF hospitalization in persons with HFrEF and atrial fibrillation, consistent with a recent meta-analysis (40). However, we observed no effect modification for all-cause mortality. Although a number of recent studies have shown absent mortality benefit for beta-blockers among patients with concomitant HF and atrial fibrillation, these analyses did not focus on patients with advanced CKD (22, 40-42). A meta-analysis specifically investigating patients with renal impairment found that beta-blockers versus placebo were associated with HRs of 0.58 (0.21-1.63; N = 72) for those with HFrEF, atrial fibrillation and an eGFR <30 ml/min/1.73m^2 and 0.83 (0.58-1.19; N = 458) for those with an eGFR between 30-44 ml/min/1.73m^2 (22). It may be that patients with advanced CKD and heart failure benefit from beta-blockers via mechanisms that are different from those with less severe renal impairment. Alternatively, residual confounding or chance may explain the benefit in individuals with HFrEF and atrial fibrillation. The larger benefit of beta-blocker use in certain subgroups such as those not receiving RASi needs replication in future studies.

Beta-blocker use in HFpEF or HFmrEF and advanced CKD

Information on ejection fraction further allowed us to evaluate the potential effectiveness of beta-blockers separately according to LVEF strata. We found that the observed benefit associated with beta-blocker use in those with HFrEF and severe renal dysfunction was not extended to those with HFmrEF (ejection fraction 40-49%) and HFpEF (ejection fraction ≥50%). A recent individual patient-level meta-analysis of randomized trials found that beta-blockers conferred similar (cardiovascular) mortality benefit in persons with LVEF between 40-49% compared to LVEF <40% (adjusted HR 0.59; 95% CI 0.34-1.03 for mortality and 0.48; 0.24-0.97 for CV mortality), although no benefit for cardiovascular hospitalization was observed (adjusted HR 0.95; 95% CI 0.68-1.32). Similar findings of a benefit in this "mildly reduced" EF range have been observed for angiotensin receptor-blockers (43), MRAs (44), and sacubitril/valsartan (45), which is also consistent with the HFmrEF resembling HFrEF in most regards, rather than being an intermediate between HFrEF and HFpEF (43, 46). In addition, this meta-analysis found no evidence of benefit from beta-blockers in the small subgroup of 244 patients with LVEF >50% in sinus rhythm. The absence of an effect of beta-blockers in persons with HFmrEF and advanced CKD in our analyses was unexpected and inconsistent with the HFrEF data in our analysis, and may be caused by effect modification according to renal function, or due to limited sample size and low event rate. Future studies should therefore confirm our findings.

Strengths and limitations

Our analysis including 3,775 patients with HFrEF and advanced CKD is the largest evaluation to date of beta-blocker effectiveness in this population. Strengths of our study include the large sample size together with detailed information available in the Swedish Heart Failure Registry, which allowed extensive adjustment for a wide range of confounders. We were also able to study multiple outcomes across the ejection fraction spectrum, and results were robust in several sensitivity analyses, including the positive control cohort and negative control outcome. However, our study also has limitations. Residual confounding by indication may be present despite adjustment for 36 variables. In addition, the cohort size was considerably smaller for those with HFmrEF and HFpEF compared with HFrEF, which may have limited power. We further defined beta-blocker use at baseline and potential cross-over may have diluted the association, although outcomes with 1-year of follow-up, for which we would expect less cross-over, showed similar results to the primary analysis with 5 years of follow-up. We did not use propensity score methods to control for confounding since there were few patients unexposed to beta-blockers at baseline (47). However, empirical studies have shown that multivariable adjusted and propensity-score adjusted studies in general do not differ much in the estimated effect size (48, 49). Our results should be considered as hypothesis generating and need confirmation in randomized trials.

In conclusion, in patients with HFrEF and advanced CKD, beta-blocker use was associated with improved survival. Our analyses support current guideline recommendations on beta-blocker therapy in HFrEF patients regardless of kidney function.

References

1. McAlister FA, Ezekowitz J, Tarantini L, Squire I, Komajda M, Bayes-Genis A, et al. Renal dysfunction in patients with heart failure with preserved versus reduced ejection fraction: impact of the new Chronic Kidney Disease-Epidemiology Collaboration Group formula. Circ Heart Fail. 2012;5(3):309-14.
2. Lofman I, Szummer K, Hagerman I, Dahlstrom U, Lund LH, Jernberg T. Prevalence and prognostic impact of kidney disease on heart failure patients. Open Heart. 2016;3(1):e000324.
3. Hillege HL, Nitsch D, Pfeffer MA, Swedberg K, McMurray JJ, Yusuf S, et al. Renal function as a predictor of outcome in a broad spectrum of patients with heart failure. Circulation. 2006;113(5):671-8.
4. Bansal N, Zelnick L, Bhat Z, Dobre M, He J, Lash J, et al. Burden and Outcomes of Heart Failure Hospitalizations in Adults With Chronic Kidney Disease. J Am Coll Cardiol. 2019;73(21):2691-700.
5. Damman K, Valente MA, Voors AA, O'Connor CM, van Veldhuisen DJ, Hillege HL. Renal impairment, worsening renal function, and outcome in patients with heart failure: an updated meta-analysis. Eur Heart J. 2014;35(7):455-69.
6. Hein AM, Scialla JJ, Edmonston D, Cooper LB, DeVore AD, Mentz RJ. Medical Management of Heart Failure With Reduced Ejection Fraction in Patients With Advanced Renal Disease. JACC Heart Fail. 2019;7(5):371-82.
7. Smith DH, Thorp ML, Gurwitz JH, McManus DD, Goldberg RJ, Allen LA, et al. Chronic kidney disease and outcomes in heart failure with preserved versus reduced ejection fraction: the Cardiovascular Research Network PRESERVE Study. Circ Cardiovasc Qual Outcomes. 2013;6(3):333-42.
8. Eckardt KU, Bansal N, Coresh J, Evans M, Grams ME, Herzog CA, et al. Improving the prognosis of patients with severely decreased glomerular filtration rate (CKD G4+): conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. Kidney Int. 2018;93(6):1281-92.
9. Damman K, Tang WH, Felker GM, Lassus J, Zannad F, Krum H, et al. Current evidence on treatment of patients with chronic systolic heart failure and renal insufficiency: practical considerations from published data. J Am Coll Cardiol. 2014;63(9):853-71.
10. Badve SV, Roberts MA, Hawley CM, Cass A, Garg AX, Krum H, et al. Effects of beta-adrenergic antagonists in patients with chronic kidney disease: a systematic review and meta-analysis. J Am Coll Cardiol. 2011;58(11):1152-61.
11. Castagno D, Jhund PS, McMurray JJ, Lewsey JD, Erdmann E, Zannad F, et al. Improved survival with bisoprolol in patients with heart failure and renal impairment: an analysis of the cardiac insufficiency bisoprolol study II (CIBIS-II) trial. Eur J Heart Fail. 2010;12(6):607-16.
12. Wali RK, Iyengar M, Beck GJ, Chartyan DM, Chonchol M, Lukas MA, et al. Efficacy and safety of carvedilol in treatment of heart failure with chronic kidney disease: a meta-analysis of randomized trials. Circ Heart Fail. 2011;4(1):18-26.
13. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J. 2016;37(27):2129-200.

14. Writing Committee M, Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Jr., et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. Circulation. 2013;128(16):e240-327.
15. The Cardiac Insufficiency Bisoprolol Study II (CIBIS-II): a randomised trial. Lancet. 1999;353 (9146):9-13.
16. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF). Lancet. 1999;353(9169):2001-7.
17. Dargie HJ. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: the CAPRICORN randomised trial. Lancet. 2001;357(9266):1385-90.
18. Flather MD, Shibata MC, Coats AJ, Van Veldhuisen DJ, Parkhomenko A, Borbola J, et al. Randomized trial to determine the effect of nebivolol on mortality and cardiovascular hospital admission in elderly patients with heart failure (SENIORS). Eur Heart J. 2005;26(3):215-25.
19. Packer M, Bristow MR, Cohn JN, Colucci WS, Fowler MB, Gilbert EM, et al. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. U.S. Carvedilol Heart Failure Study Group. N Engl J Med. 1996;334(21):1349-55.
20. Packer M, Coats AJ, Fowler MB, Katus HA, Krum H, Mohacsi P, et al. Effect of carvedilol on survival in severe chronic heart failure. N Engl J Med. 2001;344(22):1651-8.
21. Packer M, Fowler MB, Roecker EB, Coats AJ, Katus HA, Krum H, et al. Effect of carvedilol on the morbidity of patients with severe chronic heart failure: results of the carvedilol prospective randomized cumulative survival (COPERNICUS) study. Circulation. 2002;106(17):2194-9.
22. Kotecha D, Gill SK, Flather MD, Holmes J, Packer M, Rosano G, et al. Impact of Renal Impairment on Beta-Blocker Efficacy in Patients With Heart Failure. J Am Coll Cardiol. 2019;74(23):2893-904.
23. Ezekowitz J, McAlister FA, Humphries KH, Norris CM, Tonelli M, Ghali WA, et al. The association among renal insufficiency, pharmacotherapy, and outcomes in 6,427 patients with heart failure and coronary artery disease. J Am Coll Cardiol. 2004;44(8):1587-92.
24. Chang TI, Yang J, Freeman JV, Hlatky MA, Go AS. Effectiveness of beta-blockers in heart failure with left ventricular systolic dysfunction and chronic kidney disease. J Card Fail. 2013;19(3):176-82.
25. McAlister FA, Ezekowitz J, Tonelli M, Armstrong PW. Renal insufficiency and heart failure: prognostic and therapeutic implications from a prospective cohort study. Circulation. 2004;109(8):1004-9.
26. Molnar AO, Petrcich W, Weir MA, Garg AX, Walsh M, Sood MM. The association of beta-blocker use with mortality in elderly patients with congestive heart failure and advanced chronic kidney disease. Nephrol Dial Transplant. 2020;35(5):782-9.
27. Agarwal R, Rossignol P. Beta-blockers in heart failure patients with severe chronic kidney disease-time for a randomized controlled trial? Nephrol Dial Transplant. 2020;35(5):728-31.
28. Savarese G, Vasko P, Jonsson A, Edner M, Dahlstrom U, Lund LH. The Swedish Heart Failure Registry: a living, ongoing quality assurance and research in heart failure. Ups J Med Sci. 2019;124(1):65-9.
29. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12.
30. Sjolander A. Regression standardization with the R package stdReg. Eur J Epidemiol. 2016;31(6):563-74.

31. Sjolander A. Estimation of causal effect measures with the R-package stdReg. Eur J Epidemiol. 2018;33(9):847-58.

32. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med Res Methodol. 2009;9:57.

33. Ghali JK, Wikstrand J, Van Veldhuisen DJ, Fagerberg B, Goldstein S, Hjalmarson A, et al. The influence of renal function on clinical outcome and response to beta-blockade in systolic heart failure: insights from Metoprolol CR/XL Randomized Intervention Trial in Chronic HF (MERIT-HF). J Card Fail. 2009;15(4):310-8.

34. Cohen-Solal A, Kotecha D, van Veldhuisen DJ, Babalis D, Bohm M, Coats AJ, et al. Efficacy and safety of nebivolol in elderly heart failure patients with impaired renal function: insights from the SENIORS trial. Eur J Heart Fail. 2009;11(9):872-80.

35. House AA, Wanner C, Sarnak MJ, Pina IL, McIntyre CW, Komenda P, et al. Heart failure in chronic kidney disease: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. Kidney Int. 2019;95(6):1304-17.

36. Fu EL, Franko MA, Obergfell A, Dekker FW, Gabrielsen A, Jernberg T, et al. High-sensitivity C-reactive protein and the risk of chronic kidney disease progression or acute kidney injury in post-myocardial infarction patients. Am Heart J. 2019;216:20-9.

37. Muntner P, He J, Astor BC, Folsom AR, Coresh J. Traditional and nontraditional risk factors predict coronary heart disease in chronic kidney disease: results from the atherosclerosis risk in communities study. J Am Soc Nephrol. 2005;16(2):529-38.

38. Edner M, Benson L, Dahlstrom U, Lund LH. Association between renin-angiotensin system antagonist use and mortality in heart failure with severe renal insufficiency: a prospective propensity score-matched cohort study. Eur Heart J. 2015;36(34):2318-26.

39. Savarese G, Dahlstrom U, Vasko P, Pitt B, Lund LH. Association between renin-angiotensin system inhibitor use and mortality/morbidity in elderly patients with heart failure with reduced ejection fraction: a prospective propensity score-matched cohort study. Eur Heart J. 2018;39(48):4257-65.

40. Cleland JGF, Bunting KV, Flather MD, Altman DG, Holmes J, Coats AJS, et al. Beta-blockers for heart failure with reduced, mid-range, and preserved ejection fraction: an individual patient-level analysis of double-blind randomized trials. Eur Heart J. 2018;39(1):26-35.

41. Kotecha D, Flather MD, Altman DG, Holmes J, Rosano G, Wikstrand J, et al. Heart Rate and Rhythm and the Benefit of Beta-Blockers in Patients With Heart Failure. J Am Coll Cardiol. 2017;69(24):2885-96.

42. Kotecha D, Holmes J, Krum H, Altman DG, Manzano L, Cleland JG, et al. Efficacy of beta blockers in patients with heart failure plus atrial fibrillation: an individual-patient data meta-analysis. Lancet. 2014;384(9961):2235-43.

43. Lund LH, Claggett B, Liu J, Lam CS, Jhund PS, Rosano GM, et al. Heart failure with mid-range ejection fraction in CHARM: characteristics, outcomes and effect of candesartan across the entire ejection fraction spectrum. Eur J Heart Fail. 2018;20(8):1230-9.

44. Solomon SD, Claggett B, Lewis EF, Desai A, Anand I, Sweitzer NK, et al. Influence of ejection fraction on outcomes and efficacy of spironolactone in patients with heart failure with preserved ejection fraction. Eur Heart J. 2016;37(5):455-62.

45. Solomon SD, McMurray JJV, Anand IS, Ge J, Lam CSP, Maggioni AP, et al. Angiotensin-Neprilysin Inhibition in Heart Failure with Preserved Ejection Fraction. N Engl J Med. 2019;381(17):1609-20.

46. Vedin O, Lam CSP, Koh AS, Benson L, Teng THK, Tay WT, et al. Significance of Ischemic Heart Disease in Patients With Heart Failure and Preserved, Midrange, and Reduced Ejection Fraction: A Nationwide Cohort Study. Circ Heart Fail. 2017;10(6).

47. Fu EL, Groenwold RHH, Zoccali C, Jager KJ, van Diepen M, Dekker FW. Merits and caveats of propensity scores to adjust for confounding. Nephrol Dial Transplant. 2019;34(10):1629-35.

48. Shah BR, Laupacis A, Hux JE, Austin PC. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. J Clin Epidemiol. 2005;58(6):550-9.

49. Elze MC, Gregson J, Baber U, Williamson E, Sartori S, Mehran R, et al. Comparison of Propensity Score Methods and Covariate Adjustment: Evaluation in 4 Cardiovascular Studies. J Am Coll Cardiol. 2017;69(3):345-57.

Supplementary Material

Table S1. Variable definitions of comorbidities and outcomes.

Table S2. Percentage of missingness for all variables in all cohorts.

Table S3. Baseline characteristics of all individuals with advanced CKD (eGFR<30 ml/min/1.73m^2) regardless of ejection fraction, stratified by beta-blocker use.

Table S4. Guideline recommended beta-blocker agents and doses for patients with HFrEF, HFpEF and HFmrEF.

Table S5. Overall survival, composite-free survival and risk differences for all-cause mortality and the composite outcome cardiovascular mortality or heart failure hospitalization after 1, 3 and 5 years of follow-up in patients with HFrEF and advanced CKD.

Table S6. Overall survival, composite-free survival and risk differences for all-cause mortality and the composite outcome cardiovascular mortality or heart failure hospitalization after 1, 3 and 5 years of follow-up in the positive control cohort of patients with HFrEF and moderate CKD.

Table S7. Number of events, incidence rates, crude and adjusted hazard ratios for the association between beta-blocker use and 1-year all-cause mortality, CV mortality/HF hospitalization for individuals with HFrEF and advanced CKD.

Table S8. Number of events, incidence rates, crude and adjusted hazard ratios for the association between beta-blocker use and the negative control outcome cancer.

Table S9. Incidence rates and adjusted hazard ratios for mortality according to prescribed target dose of beta-blockers in patients with HFrEF and advanced CKD.

Table S10. Incidence rates and adjusted hazard ratios for the composite outcome CV death or HF hospitalization according to prescribed target dose of beta-blockers in patients with HFrEF and advanced CKD.

Table S11. Baseline characteristics in the positive control cohort of individuals with HFrEF and moderate CKD stratified by beta-blocker use.

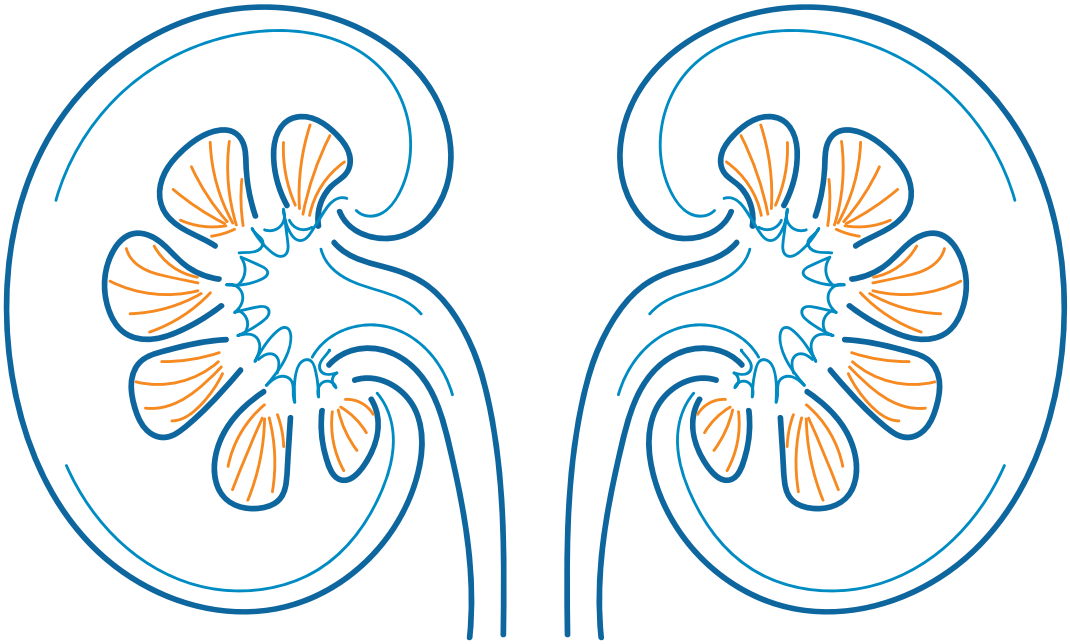
Table S12. Baseline characteristics of individuals with HFpEF and advanced CKD (eGFR<30 ml/min/1.73m^2) stratified by beta-blocker use.

Table S13. Baseline characteristics of individuals with HFmrEF and advanced CKD (eGFR<30 ml/min/1.73m^2) stratified by beta-blocker use.

Figure S1. Flow chart of patient selection from the Swedish Heart Failure Registry.

Figure S2. Density plot of beta-blocker dose levels (% of recommended dose) used in patients with HFrEF and advanced CKD from the Swedish Heart Failure Registry.

Figure S3. Number of events, incidence rates and adjusted hazard ratios for the association between beta-blocker use and 5-year cardiovascular mortality, and heart failure hospitalization in patients with HFrEF, positive control cohort, HFpEF and HFmrEF.



CHAPTER 8

Stopping renin-angiotensin system inhibitors in patients with advanced CKD and risk of adverse outcomes: a nationwide study

Edouard L. Fu, Marie Evans, Catherine M. Clase, Laurie A. Tomlinson, Merel van Diepen, Friedo W. Dekker, Juan-Jesus Carrero

J Am Soc Nephrol 2021; 32: 424-435

Abstract

Background: It is unknown whether outcomes are affected by stopping renin-angiotensin system inhibitor (RASi) therapy in patients with advanced chronic kidney disease (CKD).

Methods: We studied 10,254 nephrologist-referred patients from the Swedish Renal Registry during 2007-2017 who reached advanced CKD (glomerular filtration rate [eGFR] <30 ml/min/1.73m^2) while on RASi therapy. Target trial emulation techniques based on cloning, censoring and weighting were used to compare the risks of stopping within 6 months and remaining off treatment vs. continuing RASi on subsequent 5-year mortality, major adverse cardiovascular events (MACE) and initiation of kidney replacement therapy (KRT).

Results: Of 10,254 prevalent RASi users with new-onset eGFR <30 ml/min/1.73m^2, 1553 (15%) stopped RASi within 6 months. Median age was 72 years, 36% were women, and median eGFR was 23 mL/min/1.73m^2. Compared with the decision to continue, stopping RASi was associated with a higher absolute 5-year risk of death (40.9% vs. 54.5%) and MACE (47.6% vs. 59.5%), but lower risk of KRT (36.1% vs. 27.9%), corresponding to absolute risk differences of 13.6 (95% CI 7.0, 20.3), 11.9 (5.7, 18.6) and -8.3 (-12.8, -3.6) events per 100 patients, respectively. Results were consistent whether patients stopped at higher or lower eGFR, across pre-specified subgroups, after adjustment and stratification for albuminuria and potassium, and when modelling RASi as a time-dependent exposure using a marginal structural model.

Conclusion: In this nationwide study of people with advanced CKD, stopping RASi was associated with a higher absolute risk of mortality and MACE, but a lower absolute risk of KRT.

Introduction

Renin-angiotensin system inhibitors (RASi), that is, angiotensin-converting enzyme (ACE) inhibitors and angiotensin receptor blockers (ARB), are a cornerstone in the treatment of proteinuric chronic kidney disease (CKD), supported by trials showing their effectiveness in delaying the progression of CKD (1-7). However, evidence regarding the efficacy and safety of RASi in individuals with advanced CKD is limited to a small single-center trial (8) and post-hoc analyses of the few patients with advanced CKD who were included in the pivotal RASi trials (9, 10).

A small observational study, showing improved glomerular filtration rate (GFR) after stopping RASi (11), led to the hypothesis that continuing RASi in patients with advanced CKD might accelerate the need for kidney replacement therapy (KRT) (12). This, together with the concern that the persistent hemodynamic effects of RASi, which are manifested by an acute change in GFR at initiation (13, 14), may cause harm by chronically lowering the GFR, has led to frequently stopping RASi among patients with advanced CKD in routine clinical practice (15, 16). However, stopping RASi may also harm patients by increasing cardiovascular risk and mortality (17).

This clinical equipoise is being addressed by an ongoing randomized trial that evaluates the difference in 3-year eGFR change in patients with advanced CKD at baseline, randomized to continue or discontinue RASi, with publication anticipated in 2022 (18, 19). Recently, an observational study from a private healthcare provider in the United States (U.S.) suggested that stopping RASi in patients with advanced CKD was associated with an increased risk of major cardiovascular events (MACE) and death, but not with the risk of KRT (17). While this study has generated considerable attention, confirmation of such findings in independent and geographically diverse health systems is needed to increase generalizabilty and provide the strength of evidence needed to inform clinical practice.

We used routine-care data from patients referred to nephrologist care in Sweden, to compare the outcomes of long-term users of RASi who stopped or continued treatment after developing advanced CKD (eGFR <30 ml/min/1.73m^2). Our primary objective was to evaluate the risks of death, MACE and commencement of KRT by this treatment decision. As a secondary objective, we investigated whether observed risks and benefits differed in individuals who stopped earlier (eGFR 20-30 ml/min/1.73m^2) or later (eGFR <20 ml/min/1.73m^2) in the course of their disease progression.

Methods

Swedish Renal Registry

We used data from the Swedish Renal Registry (SRR), a nationwide registry of patients with CKD G3–5 attending routine nephrologist-specialist care in Sweden (20, 21), during the period 2007-2017. The SRR collects routine information from outpatient nephrologist visits, including CKD aetiology, laboratory tests, blood pressure and other results obtained from routine clinical examination. The registry has a mandatory enrolment policy for patients with an eGFR <30 mL/min/1.73m^2, but the registry also encourages the inclusion of patients earlier in the course of the disease (eGFR<45 mL/min/1.73m^2) provided it is done systematically by the nephrology clinic (i.e., all or none are registered from each specific clinic with eGFR<45 mL/min/1.73m^2). Registrations of subsequent outpatient visits to nephrology care (on average 2-3 per year per patient) are thereafter recorded until death, emigration from the country or start of KRT. Nearly all nephrology clinics in Sweden (96%) report to the SRR-CKD and the estimated national coverage is >75% for nephrologist-referred patients with G4-5 CKD (20).

Via each citizen's unique personal identification number, the SRR was linked to other national registries; the Swedish Prescribed Drug Registry provided complete information on all prescribed drugs dispensed at Swedish pharmacies (22), and this was used to define RASi use and changes in RASi therapy; the Swedish Patient Registry added information on all outpatient specialist consultations and hospitalizations occurring in Swedish healthcare since 1997 until end of follow up, and this was used to obtain information on comorbidities and outcomes (23); the Swedish Death Registry added information on date and causes of death (24). All these registries are run by the Swedish National Board of Welfare, a government institution, and are considered to have no, or minimal, loss to follow up. All patients are informed about their participation in the registry and have the possibility to opt out at any time. We used data linked and de-identified by the Swedish government and were judged not to require informed consent, being approved by the regional ethical review boards and the Swedish National Board of Welfare.

Patient selection and study design

This observational study emulated a pragmatic clinical trial (25) comparing the effect of stopping vs. continuing RASi on cardiovascular and renal outcomes in people with advanced CKD (19). **Supplemental Table S1** outlines the protocol of such trial, which would randomize prevalent RASi users reaching incident CKD G4-5 to either stop RASi within 6 months or to continue with the treatment.

We created a cohort of all adult (≥18 years) patients registered in the SRR after 2007 January 01, who experienced new CKD G4 (ie, whose GFR decreased to <30 ml/min/1.73m^2), and who had taken RASi for more than 80% of the two years before that date. We defined this using a medication possession ratio >80%, the proportion of the number of days of medication dispensed to total number of days of observation. Baseline (T_0) was defined as the day on which the first recorded eGFR <30 ml/min/1.73m^2 was identified. We chose to include only patients apparently adherent to RASi therapy to decrease the possibility of confounding bias due to nonadherence. We excluded patients with a history of kidney transplantation, patients with missing blood pressure measurements at the time of eGFR decrease to <30 ml/min/1.73m^2 or those who stopped RASi before the decrease in eGFR. eGFR was calculated with the CKD-EPI equation (26) from routine plasma creatinine measurements performed by enzymatic or corrected Jaffe methods traceable to isotope dilution mass spectroscopy standards. As information on race is not available in Sweden by law, we did not use the variable for African American ethnicity.

Treatment strategies

We compared the strategies "stop RASi within 6 months and remain off treatment after eGFR decrease <30 ml/min/1.73m^2" vs. "continue RASi for the whole follow-up". We chose to examine the effect of stopping and remaining off treatment because a significant proportion of individuals who discontinued RASi restarted during follow-up (57.1%). Stopping of RASi was defined as absence of a dispensation of RASi within 60 days (lag phase) after the estimated last day of pill supply from the previous dispensation, assuming the most common prescription pattern of one pill per day. When a prescription was filled before the expected end of the previous dispensation, we added the remaining pills onto the next period, for the first occurrence, but did not carry this forward. In the case of hospitalization, we added as many additional pills as days spent in the hospital.

Study outcomes

Each patient was followed until the first of: occurrence of an event, five years after baseline, or administrative censoring (June 1, 2017). The primary outcome was 5-year all-cause mortality. Secondary outcomes included MACE (defined as a composite endpoint of mortality, myocardial infarction and cerebrovascular events) and KRT (defined as undergoing kidney transplantation or initiating maintenance dialysis). ICD-10 codes for ascertainment of cardiovascular outcomes are listed in **Supplemental Table S2**. Information on date of initiation of KRT was obtained from the SRR.

Emulation of the target trial

We used the method of cloning, censoring and weighting (25, 27-29) to emulate a target trial comparing the effects of "stopping RASi within 6 months after eGFR dropped <30 ml/min/1.73m^2 and remaining off treatment" vs. "continuing RASi" (see **Supplemental Methods** and **Supplemental Figure S1** for a detailed discussion on target trial emulation). Briefly, we created a dataset with two copies of each eligible individual (cloning, or replicating) and assigned each of the replicates to one of the treatment strategies at the start of follow-up. Thereafter, we assessed at monthly intervals whether replicates adhered to their assigned treatment strategy; replicates were censored if and when their actual treatment deviated from their assigned treatment strategy, thereby ensuring that replicates followed their assigned strategy. For example, if a replicate was assigned to continuing RASi, but actually stopped RASi treatment on day 90, they would be censored at that point. A replicate that was assigned to the discontinuation arm, and discontinued within 6 months but subsequently restarted treatment would also be censored at the date of treatment restart. To adjust for the potential selection bias induced by this artificial censoring, each individual received a time-varying inverse probability weight (30). Informally, the denominator of the weights was the probability that a replicate remained uncensored (i.e., remained on the assigned treatment strategy) conditional on baseline and time-varying variables (**Supplemental Table S3**). The weights created two pseudopopulations in which treatment was independent of measured prognostic factors. We estimated the time-varying weights by fitting a pooled logistic model for the monthly probability of remaining uncensored, including variables for time and the baseline and time-varying covariates listed in **Table 1**. Models were fitted separately in both treatment arms to allow for treatment-covariate interaction (29). The variables for each model and their regression coefficients are reported in **Supplemental Tables S4-5**. To avoid undue influence of outliers, weights were truncated at the 99.5th percentile (31).

We estimated the effect of stopping RASi on 5-year all-cause mortality, MACE and KRT using weighted pooled logistic regression, including an indicator for treatment strategy, month and its quadratic term, and their interactions to allow for non-proportional hazards. The predicted probabilities from this logistic model were used to estimate the adjusted 5-year predicted probability of mortality, MACE and KRT under each treatment strategy and produce weighted cumulative incidence curves (32, 33). For the KRT curves, the competing risk of death was taken into account. Pointwise 95% confidence intervals were calculated using nonparametric bootstrap based on 500 full samples. In addition to absolute risks and risk differences, we estimated the 5-year restricted mean survival time (RMST) under each treatment strategy and the 5-year RMST difference between both strategies. The RMST is interpreted as the average survival time over a fixed follow-

up period and graphically it corresponds to the area under the survival curve (34, 35). The 5-year RMST *difference* compares the areas under the two survival curves for the intervention and control group. It is interpreted as the mean postponement of the outcome in one group compared with the other. E.g., if the 5-year RMST difference equals 6 months, then on average, patients on one strategy survive 6 months longer compared with patients on another strategy over a 5-year follow-up period. We used nonparametric bootstrapping to obtain 95% confidence intervals using the standard deviation (SD) of the bootstrap estimations as an estimation of the standard error of the RMST (36). We did not calculate hazard ratios since the proportionality of hazards assumption was not met and hazard ratios were thus difficult to interpret (29, 37, 38). R version 3.6.2 was used for all statistical analyses.

Secondary objective: stopping RASi at different eGFRs

In order to evaluate whether observed associations differed in individuals who stopped earlier or later in the course of their disease progression, we created two additional cohorts using the same methodology: we evaluated separately the outcomes associated with stopping vs. continuing RASi in a cohort of individuals on their first detected eGFR decrease to between 20-30 ml/min/1.73m^2 (higher eGFR cohort) and another cohort of individuals on their first detected eGFR below 20 ml/min/1.73m^2 (lower eGFR cohort). Note that there is some overlap of patients in these cohorts as patients progress to a lower eGFR during observation.

Supporting and sensitivity analyses

We pre-specified several analyses to test the robustness and consistency of our main results. First, we compared results when using nontruncated weights. Second, we performed stratified analyses by age (≥70 vs. < 70 years), sex, presence of diabetes, and presence of heart failure, and investigated the interaction of each of these variables with treatment on an additive scale by calculating the absolute excess risk due to interaction. Third, as a negative control analysis, we examined the association between stopping or continuing RASi and the long-term diagnosis of cancer (39). We did not expect stopping RASi to cause or prevent cancer. If we found stopping RASi to be associated with an increased risk of cancer, this would suggest that the observed effect estimate suffers from residual confounding by unmeasured clinical conditions that are associated with stopping RASi, and which are also likely to be associated with the risk of cancer, such as smoking and BMI. For this analysis, patients with a recent cancer diagnosis (within two years from the index date) were excluded from this analysis to minimise the effects of reverse causality, since people may have stopped RASi because they had been diagnosed with cancer. Fourth, we compared results from our trial emulation design with an analysis handling RASi as a time-varying covariate (40). The effect of "always using

RASi" vs. "immediately stopping and not restarting RASi" after eGFR dropped <30 ml/min/1.73m^2 was estimated using inverse probability of treatment and censoring weighted estimation of a marginal structural model (see *Supplemental Methods* for detailed explanation) (30, 41). Fifth, we additionally adjusted our analyses for time-dependent measures of urinary albumin-to-creatinine ratio (ACR) and plasma potassium. This analysis was restricted to the 3049 individuals with this data available, and evaluated consistency across baseline albuminuria (≥70 vs. <70 mg/mmol) and potassium (≥5.0 vs. <5.0 mmol/L) strata. Finally, after reviewing the results of the work above, we conducted a non-prespecified analysis, in which we examined the associations of stopping vs. continuing RASi on the combined outcome of death and KRT, as a surrogate of "net clinical benefit."

Results

Of 30,180 individuals registered in SRR during the study period, 10,254 prevalent RASi users with a medication possession ratio >80% and no history of kidney transplantation were included from the day of their first recorded eGFR below 30 ml/min/1.73m^2.

Figure 1 displays the patient selection flow chart, and **Table 1** describes their baseline characteristics. At baseline, patients had a median (IQR) age of 72 (63-79) years and 35.7% were women. Median eGFR was 23 (18-27) ml/min/1.73m^2, median ACR 35 (6-156) mg/mmol, mean (± SD) systolic blood pressure 139 (SD 22) mmHg and mean diastolic blood pressure 76 (SD 12) mmHg. Hypertension (88.7%), diabetes (49.5%), ischemic heart disease (33.1%) and heart failure (28.0%) were the most common comorbidities. Concurrent use of diuretics (79.3%), beta blockers (67.6%), statins (61.6%) and calcium channel blockers (60.5%) was also prevalent. During the first 6 months of observation 1,553 (15.1%) individuals stopped RASi. Of these, 887 (57.1%) of patients restarted RASi during follow-up.

Figure 1. Selection of study participants.
Abbreviations: RASi = Renin-angiotensin-system inhibitor; eGFR = estimated glomerular filtration rate; MPR = medication possession ratio; KRT = renal replacement therapy.

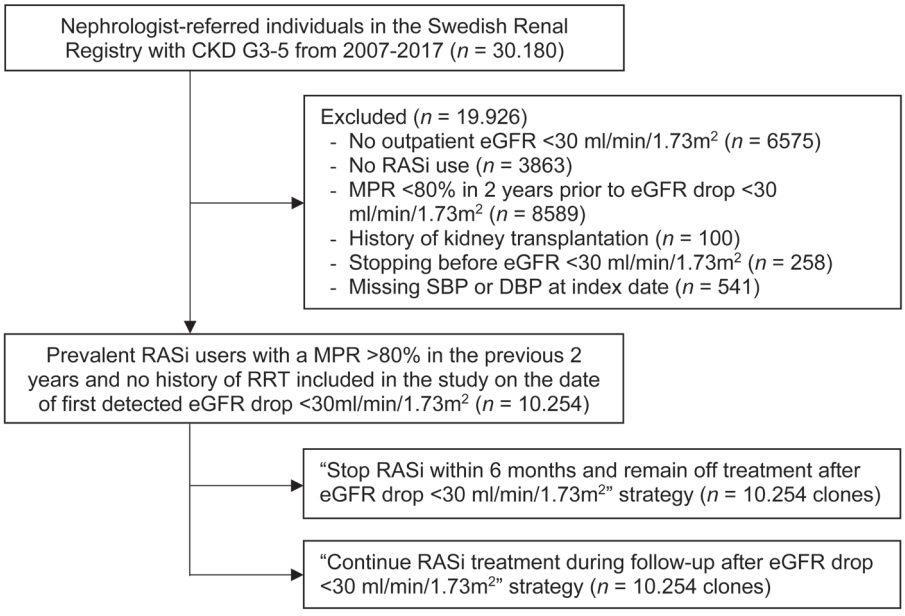


Table 1. Baseline characteristics of prevalent RASi users with eGFR <30 ml/min/1.73m² registered in the Swedish Renal Registry during 2007-2017.

	eGFR <30 ml/min/1.73m² cohort (*n* = 10,254)
Median Age (IQR)‡, *years*	72 [63, 79]
Age category, *n (%)*	
<50	848 (8.3)
50-59	1046 (10.2)
60-69	2400 (23.4)
70-79	3471 (33.9)
>=80	2489 (24.3)
Women	3662 (35.7)
Median eGFR (IQR)‡, *ml/min/1.73m²*	23 [18, 27]
eGFR category, *n (%)*	
<15 ml/min/1.73m², *n (%)*	1557 (15.2)
≥15 ml/min/1.73m², *n (%)*	8697 (84.8)
Primary kidney disease, *n* (%)	
Diabetes	2878 (28.1)
Hypertension	2512 (24.5)
Glomerulonephritis	1096 (10.7)
Polycystic kidney disease	574 (5.6)
Pyelonephritis	171 (1.7)
Other	1753 (17.1)
Missing	1270 (12.4)
Mean SBP (SD), *mmHg*	139 (22)
SBP category, *n (%)*	
<120	1430 (13.9)
120-139	3670 (35.8)
140-159	3224 (31.4)
>160	1930 (18.8)
Mean DBP (SD), *mmHg*	76 (12)
DBP category, *n (%)*	
<80	5502 (53.7)
80-89	3340 (32.6)
90-99	1066 (10.4)
>100	346 (3.4)
Median urinary ACR [IQR], *mg/mmol*	35 [6, 156]

	eGFR <30 ml/min/1.73m² cohort (*n* = 10,254)
ACR category, *n (%)*	
A1 (<3)	785 (7.7)
A2 (3-29)	1445 (14.1)
A3 (30-69)	614 (6.0)
A3 (≥70)	1835 (17.9)
Missing	5575 (54.4)
Mean serum potassium (SD), *mg/mmol**	4.5 (0.6)
Comorbidities, *n (%)*	
Hypertension	9099 (88.7)
Myocardial infarction	2212 (21.6)
Ischemic heart disease	3390 (33.1)
Arrhythmia	2302 (22.4)
Heart failure	2868 (28.0)
Peripheral vascular disease	1269 (12.4)
Cerebrovascular disease	1620 (15.8)
Diabetes mellitus	5079 (49.5)
Chronic obstructive pulmonary disease	1811 (17.7)
Cancer diagnosis in previous 2 years	1018 (9.9)
Medication, *n (%)*	
Beta blockers	6928 (67.6)
Calcium channel blockers	6202 (60.5)
Diuretics	8128 (79.3)
Statins	6312 (61.6)
Antiplatelets	4736 (46.2)
Potassium binder	941 (9.2)
Calendar year	
2007-2010	3431 (33.5)
2011-2013	3399 (33.1)
2014-2016	3424 (33.4)
Hospitalizations	
Any hospitalization in previous year, *n (%)*	4325 (42.2)
Hyperkalemia hospitalization, *n (%)*	415 (4.0)
AKI hospitalization in previous year, *n (%)*	481 (4.7)

eGFR = estimated glomerular filtration rate; SBP = systolic blood pressure; DBP = diastolic blood pressure; ACR = albumin-to-creatinine ratio; AKI = acute kidney injury.
* potassium was missing in 37% of individuals.

Stopping RASi and outcomes

After cloning, 10,254 individuals were assigned to each treatment strategy. The mean of the truncated inverse probability weights was 2.2 (maximum 35.0). The characteristics in each treatment arm at the end of the grace period (six months after baseline) before and after weighting are shown in **Supplemental Table S6**. The inverse probability weighting showed a good ability to remove covariate imbalance. The estimated 5-year mortality risk was 40.9% (95% CI 38.9, 42.8) among those who continued RASi, and 54.5% (95% CI 48.5, 61.2) among those who stopped RASi, corresponding to an absolute risk difference of 13.6 (95% CI 7.0, 20.3) deaths per 100 individuals and a 5-year RMST difference of -3.6 months (95% CI -5.4, -1.8) (**Table 2**). The 5-year risk of MACE was 47.6% (95% CI 45.9, 49.4) in the RASi continuation arm and 59.5% (95% CI 53.8, 66.1) percent in the stopping RASi arm, with an estimated 5-year absolute risk difference of 11.9 (95% CI 5.7, 18.6) events per 100 individuals and a 5-year RMST difference of -3.3 months (95% CI -5.3, -1.4) (**Figure 2, Table 2**).

Figure 2. Weighted cumulative probability curves for mortality (A), MACE (B), KRT (C) and cancer (D, negative control outcome) stratified by RASi use strategy. Thinner dotted lines represent 95% confidence intervals.

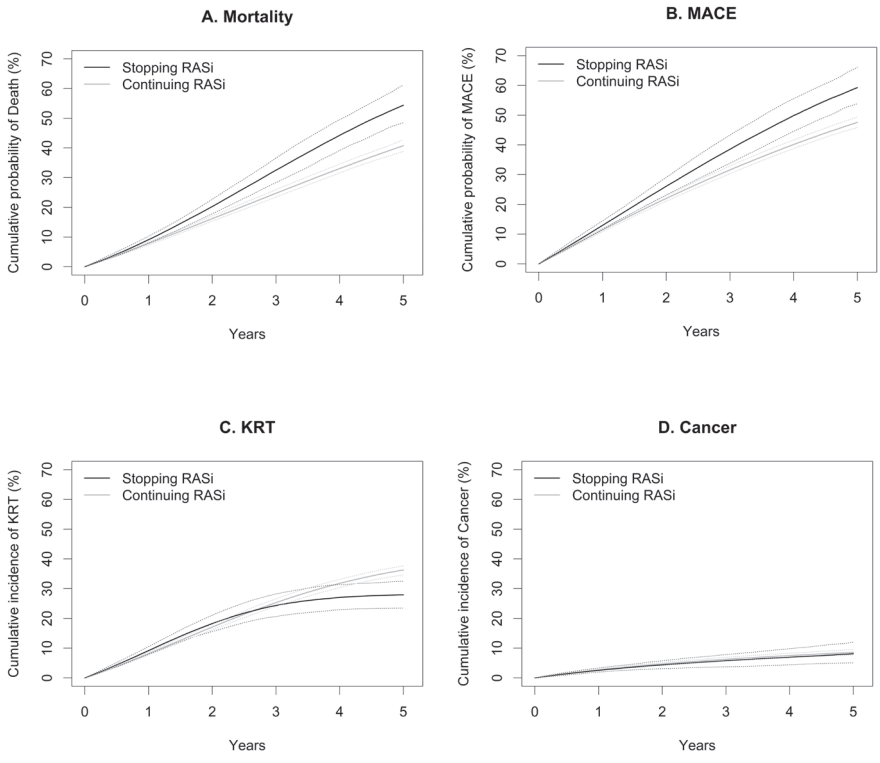


Table 2. 5-year RMST, RMST differences, absolute risks and risk differences associated with stopping RASi and continuation on mortality, MACE and KRT in advanced CKD patients with eGFR <30 ml/min/1.73m^2.

	Weighted persons, *n*	Weighted events, *n*	5-year RMST, months (95% CI)	5-year RMST difference, months (95% CI)	5-year absolute risk, % (95% CI)	5-year risk difference, % (95% CI)
All-cause mortality						
Continuing RASi	7971	3258	47.9 (46.2, 49.7)	Reference	40.9 (38.9, 42.8)	Reference
Stopping RASi	7078	3852	44.3 (43.8, 44.8)	-3.6 (-5.4, -1.8)	54.5 (48.5, 61.2)	13.6 (7.0, 20.3)
MACE						
Continuing RASi	8127	3870	44.7 (42.8, 46.5)	Reference	47.6 (45.9, 49.4)	Reference
Stopping RASi	7623	4543	41.4 (40.8, 41.9)	-3.3 (-5.3, -1.4)	59.5 (53.8, 66.1)	11.9 (5.7, 18.6)
KRT						
Continuing RASi	8329	3007	48.1 (46.5, 49.7)	Reference	36.1 (34.7, 37.7)	Reference
Stopping RASi	8808	2458	48.9 (48.3, 49.5)	0.8 (-0.8, 2.5)	27.9 (23.5, 32.5)	-8.3 (-12.8, -3.6)

N = number; CI = confidence interval; MACE = major adverse cardiovascular events; RASi = renin-angiotensin system inhibitor; KRT = renal replacement therapy; RMST = restricted mean survival time.

ǂAnalyses were adjusted through inverse probability weighting for age, sex, calendar year, eGFR, systolic and diastolic blood pressure, comorbidities (ischemic heart disease, myocardial infarction, arrhythmia, heart failure, peripheral vascular disease, cerebrovascular disease, diabetes, chronic pulmonary disease, cancer), medication use (beta blockers, calcium channel blockers, diuretic, statins, antiplatelet) and hospitalizations (total number of hospitalizations in previous year, AKI hospitalization in previous year, hyperkalaemia hospitalization). Valid 95% confidence intervals were derived using nonparametric bootstrap based on 500 samples to account for the within-subject correlation induced by weighting. Weights were truncated at the 99.5th percentile.

The 5-year estimated risk of KRT was 36.1 (95% CI 34.7, 37.7) for patients that continued with RASi and 27.9 (95% CI 23.5, 32.5) for those who stopped RASi. This corresponds to an absolute risk reduction of -8.3 (95% CI -12.8, -3.6) KRT events per 100 individuals among patients stopping RASi and a 5-year RMST difference of 0.8 months (95% CI -0.8, 2.5). **Figure 2** shows the weighted cumulative incidence curves for study outcomes stratified according to treatment strategy. The curves for mortality and MACE progressively diverged after a few months, whereas the curves for KRT crossed, and diverged after three years.

Stopping RASi and outcomes at different eGFR

The higher eGFR cohort included 7,277 individuals whose first observed eGFR was between 20-30 ml/min/1.73m² (median eGFR 25; IQR 23-28), and the lower eGFR cohort included 6,907 individuals whose first observed eGFR was below 20 ml/min/1.73m² (median eGFR 17; IQR 14-19). Baseline characteristics for both cohorts are displayed in **Supplemental Table S7**. In both cohorts an increased risk for mortality and MACE was observed when RASi was stopped (**Tables 3-4, Supplemental Figures S2-S3**). For instance, in the lower eGFR cohort, stopping RASi was associated with an increased absolute risk for mortality (17.1; 95% CI 9.9, 23.8 per 100 individuals) and MACE (12.6; 95% CI 5.8, 19.3 per 100 individuals). In both cohorts, there also was a lower absolute risk of KRT among patients stopping RASi. For instance, in the low eGFR cohort there was an absolute risk reduction of -9.6 (95% CI -15.0, -3.8) KRT events per 100 individuals among patients stopping RASi. The cumulative incidence curve showed that the risk for KRT was slightly higher in the stopping arm during the first two years of follow-up, crossed at two years, and diverged gradually (**Supplemental Figures S2-S3**).

Table 3. 5-year RMST, RMST differences, absolute risks and risk differences associated with stopping RASi and continuation on mortality, MACE and KRT in advanced CKD patients with eGFR 20-30 ml/min/1.73m².

	Weighted persons, *n*	Weighted events, *n*	5-year RMST, months (95% CI)	5-year RMST difference, months (95% CI)	5-year absolute risk, % (95% CI)	5-year risk difference, % (95% CI)
All-cause mortality						
Continuing RASi	5471	2114	48.7 (46.4, 50.9)	Reference	38.6 (36.3, 40.9)	Reference
Stopping RASi	4594	2340	46.1 (45.4, 46.8)	-2.6 (-4.9, -0.2)	50.9 (42.4, 60.1)	12.3 (3.3, 21.4)
MACE						
Continuing RASi	5634	2525	45.7 (43.3, 48.1)	Reference	44.8 (42.7, 46.9)	Reference
Stopping RASi	5005	2950	42.7 (42.0, 43.4)	-3.0 (-5.5, -0.5)	58.9 (49.2, 67.8)	14.1 (4.6, 23.5)
KRT						
Continuing RASi	5376	1360	53.3 (51.5, 55.0)	Reference	25.3 (23.4, 27.3)	Reference
Stopping RASi	5312	681	55.4 (54.9, 55.9)	2.1 (-0.3, 3.9)	12.8 (7.6, 18.6)	-12.5 (-17.8, -6.6)

N = number; CI = confidence interval; MACE = major adverse cardiovascular events; RASi = renin-angiotensin system inhibitor; KRT = renal replacement therapy; RMST = restricted mean survival time.

ǂAnalyses were adjusted through inverse probability weighting for age, sex, calendar year, eGFR, systolic and diastolic blood pressure, comorbidities (ischemic heart disease, myocardial infarction, arrhythmia, heart

failure, peripheral vascular disease, cerebrovascular disease, diabetes, chronic pulmonary disease, cancer), medication use (beta blockers, calcium channel blockers, diuretic, statins, antiplatelet) and hospitalizations (total number of hospitalizations in previous year, AKI hospitalization in previous year, hyperkalaemia hospitalization). Valid 95% confidence intervals were derived using nonparametric bootstrap based on 500 samples to account for the within-subject correlation induced by weighting. Weights were truncated at the 99.5th percentile.

Table 4. 5-year RMST, RMST differences, absolute risks and risk differences associated with stopping RASi and continuation on mortality, MACE and KRT in advanced CKD patients with eGFR <20 ml/min/1.73m^2.

	Weighted persons, *n*	Weighted events, *n*	5-year RMST, months (95% CI)	5-year RMST difference, months (95% CI)	5-year absolute risk, % (95% CI)	5-year risk difference, % (95% CI)
All-cause mortality						
Continuing RASi	5470	2401	46.4 (44.7, 48.2)	Reference	43.9 (41.3, 46.6)	Reference
Stopping RASi	5423	3309	42.0 (41.3, 42.7)	-4.4 (-6.3, -2.5)	61.0 (54.0, 67.3)	17.1 (9.9, 23.8)
MACE						
Continuing RASi	5547	2845	43.0 (41.2, 44.8)	Reference	51.3 (48.9, 53.9)	Reference
Stopping RASi	5734	3663	39.9 (39.2, 40.7)	-3.1 (-5.0, -1.1)	63.9 (57.0, 70.0)	12.6 (5.8, 19.3)
KRT						
Continuing RASi	5914	3131	40.6 (38.6, 42.6)	Reference	52.9 (50.8, 54.8)	Reference
Stopping RASi	6872	2981	42.0 (41.3, 42.7)	1.4 (-0.7, 3.5)	43.4 (38.3, 48.8)	-9.6 (-15.0, -3.8)

N = number; CI = confidence interval; MACE = major adverse cardiovascular events; RASi = renin-angiotensin system inhibitor; KRT = renal replacement therapy; RMST = restricted mean survival time.

ǂAnalyses were adjusted through inverse probability weighting for age, sex, calendar year, eGFR, systolic and diastolic blood pressure, comorbidities (ischemic heart disease, myocardial infarction, arrhythmia, heart failure, peripheral vascular disease, cerebrovascular disease, diabetes, chronic pulmonary disease, cancer), medication use (beta blockers, calcium channel blockers, diuretic, statins, antiplatelet) and hospitalizations (total number of hospitalizations in previous year, AKI hospitalization in previous year, hyperkalaemia hospitalization). Valid 95% confidence intervals were derived using nonparamtric bootstrap based on 500 samples to account for the within-subject correlation induced by weighting. Weights were truncated at the 99.5th percentile.

Supporting and sensitivity analyses

Using untruncated weights had no major influence on the point estimates (**Supplemental Table S8**). Subgroup analyses within strata of age, sex, diabetes, heart failure and ischemic heart disease showed no suggestion of heterogeneity, with higher risk differences for mortality and MACE and lower risk differences for KRT observed across all subgroups (**Supplemental Figure S4**). We did not observe an association between continuing/stopping RASi and the risk of cancer in any of the studied cohorts (**Supplemental Table S9**). In the sensitivity analysis using RASi as a time-dependent exposure through inverse probability of treatment and censoring weighted estimation of a marginal structural model, immediately stopping and not restarting RASi compared with always using RASi was associated with an 11.3% (95% CI 8.1, 14.5) higher risk for mortality, an 8.8% (95% CI 5.5, 12.5) higher risk for MACE and a -7.1% (95% CI -11.8, -3.4) lower risk for KRT (**Supplemental Table S10, Supplemental Figure S5**). In patients with available measures of ACR and potassium, additional adjustment for these covariates showed results consistent with our main analysis, although with wider confidence intervals: compared with patients continuing RASi, stopping was associated with a 9.3% (95% CI -1.1, 23.7) higher absolute risk for mortality, 7.6% (95% CI -23.6, 21.2) higher risk for MACE but -8.2% (95% CI -15.8, 5.8) lower risk for KRT (**Supplemental Table S11**). Stratified analyses by baseline ACR and potassium categories were largely consistent with the main results (**Supplemental Figure S5**). There was an increase in the magnitude of the association of stopping RASi on KRT events: risk difference of -11.4 (95% CI -19.5, -2.6) KRT events per hundred patients in patients with baseline potassium <5.0 mmol/l and -33.3 (95% CI -41.9, -25.5) in patients with potassium ≥5.0 mmol/l over a 5-year follow-up period (interaction $p < 0.001$). Finally, evaluating the composite outcome of death plus KRT favored the strategy of continuing with RASi vs. stopping, although confidence intervals were wide, with an absolute 5-year risk difference of 5.1% (95% CI -0.2, 11.3) (**Supplemental Table S12** and **Supplemental Figure S6**).

Discussion

Deciding whether and when to stop RASi in patients with advanced CKD is a frequent issue in clinical practice (15, 16). A single-center UK observational study of 52 individuals (mean eGFR of 16 ml/min/1.73m^2) reported that eGFR increased significantly after stopping RASi, leading to the idea that stopping RASi may prolong the time to KRT (11). Stopping RASi, on the other hand, may also potentially harm patients by increasing cardiovascular risk and mortality, based on generalisation from cardiovascular trials largely conducted in people with higher GFR (17). We addressed this problem by modelling the consequences of this decision in a nationwide observational study of over ten thousand individuals with advanced CKD under routine nephrological care. We found that compared with continuing RASi, stopping treatment was associated

with a higher 5-year risk of mortality and MACE, but a lower absolute KRT risk. These results appeared robust in various sensitivity and subgroup analyses, including the evaluation of stopping at a higher or lower eGFR.

Our findings of a higher absolute risk of death and MACE among patients stopping RASi confirm and expand a recent observational study of 3909 persons with advanced CKD from a single healthcare provider in the U.S. (17). Expansion of this evidence to a large, nationwide and geographically diverse cohort of patients receiving universal government-subsidized healthcare increases generalizability. Collectively, this agrees with trial evidence on the cardioprotection that RASi confers to patients with CKD (42), and with observational evidence of lower cardiovascular risk associated with RASi use at all levels of eGFR (43, 44). Our finding of a lower absolute KRT risk among patients stopping RASi differs from the previous U.S. study. Qiao *et al.* (17) observed that continuing RASi was not associated with increased risk of KRT (HR, 1.19; 95% CI, 0.86-1.65) and they summarized this as "KRT harms may not be excessive". Because the assumption of proportional hazards was not met in our study, we reported absolute risk differences, and observed an association of stopping RASi therapy with reduced risk of KRT (8.3 KRT events could have been prevented per 100 patients who continued with RASi therapy over 5 years). The composite outcome of death plus KRT, which could be considered as the overall "net-clinical benefit" of the decision strategy, favored continuing with RASi. However, this analysis assumes that death and dialysis are of equal importance, which is not the case in aggregate; individual patients may attribute different importance to these outcomes and their priorities should also be considered in decision making. Finally, individual patients may respond differently to RASi, and individualization of treatment and drug dosing are other important aspects not considered in our modelling.

We used comparable designs and analytical strategies to those used in the U.S. study (17), with one exception: we censored patients when their initial strategy was changed, in acknowledgement that patients who stopped their therapy were frequently re-started during follow-up, and thus ensuring no crossovers; we think this is a strength of this current work. However, the source and type of data differ: while our cohort is representative of the CKD population under nephrologist care in Sweden, Geisinger is a large, predominantly rural, private healthcare system in Pennsylvania that included both nephrologist-referred and non-referred patients. We believe that our selection of nephrologist-referred patients is a strength for the evaluation of KRT outcomes, because patients receive and stop or continue RASi for reasons and indications that may differ between primary care and specialist nephrology care. Both studies have a similar duration of follow-up, but a larger proportion of patients initiated KRT in our study, 35%, compared with 8% in the U.S. cohort. Between-country differences and differences between nephrologists and primary health care practitioners in clinical practice may additionally explain the divergent findings: e.g., 15% of patients stopped RASi in our study vs. 32% in the U.S. cohort.

Our study is the largest to date investigating the clinical consequences associated with this common clinical issue, whether to continue or stop RASi in patients with GFR < 30 ml/min/1.73m^2. Additional strengths are: i) the application of two complementary state-of-the-art analytic approaches (i.e. target trial emulation and marginal structural modelling) to account for time-dependent confounding of a rich range of confounders; ii) confirmation of results across risk subgroups, including those with albuminuria or elevated potassium which might have explained why drugs were stopped or continued; iii) modelling a negative control outcome to evaluate the impact of reverse-causation and unknown confounding; iv) evaluation of RASi use by pharmacy dispensations, which may be a better indicator for medication intake than prescriptions. Exclusion of patients with long-term use of RASi who did not have a high medication possession ratio reduces the likelihood that medication non-adherence was the cause of drug cessation. We acknowledge a number of limitations. We did not have information on ethnic origin. Results apply to Swedish practice and extrapolation to other populations and countries should be done with caution. Initiation of KRT is itself a treatment decision that varies by practitioner and variations in physician behavior were not captured in our study. Furthermore, the decision to stop RASi is not a random one, but the consequence of complex factors that likely herald worse outcomes. Frail patients where RASi may have been more likely to be stopped may also be more likely to be treated conservatively. Despite our sophisticated analytical design, residual confounding cannot be excluded from any observational analysis, and the precise reasons for stopping RASi remain unknown. Our conclusions remain observational in nature and therefore do not substitute for randomized trials. However, until these trials are conducted they may assist in informing clinical decisions.

To conclude, in this nationwide study, stopping RASi among patients referred to nephrologists with advanced CKD was associated with an increased absolute risk of mortality and MACE, but a lower absolute risk of KRT. To date, there is no trial evidence to inform the decision of stopping RASi therapy in these patients. Until the ongoing STOP-ACEi trial is completed (19), our analyses support current KDIGO recommendations of not routinely stopping RASi in people with advanced CKD (45, 46).

References

1. Randomised placebo-controlled trial of effect of ramipril on decline in glomerular filtration rate and risk of terminal renal failure in proteinuric, non-diabetic nephropathy. The GISEN Group (Gruppo Italiano di Studi Epidemiologici in Nefrologia). Lancet. 1997;349(9069):1857-63.
2. Brenner BM, Cooper ME, de Zeeuw D, Keane WF, Mitch WE, Parving HH, et al. Effects of losartan on renal and cardiovascular outcomes in patients with type 2 diabetes and nephropathy. N Engl J Med. 2001;345(12):861-9.
3. Jafar TH, Schmid CH, Landa M, Giatras I, Toto R, Remuzzi G, et al. Angiotensin-converting enzyme inhibitors and progression of nondiabetic renal disease. A meta-analysis of patient-level data. Ann Intern Med. 2001;135(2):73-87.
4. Kent DM, Jafar TH, Hayward RA, Tighiouart H, Landa M, de Jong P, et al. Progression risk, urinary protein excretion, and treatment effects of angiotensin-converting enzyme inhibitors in nondiabetic kidney disease. J Am Soc Nephrol. 2007;18(6):1959-65.
5. Lewis EJ, Hunsicker LG, Clarke WR, Berl T, Pohl MA, Lewis JB, et al. Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. N Engl J Med. 2001;345(12):851-60.
6. Maschio G, Alberti D, Janin G, Locatelli F, Mann JF, Motolese M, et al. Effect of the angiotensin-converting-enzyme inhibitor benazepril on the progression of chronic renal insufficiency. The Angiotensin-Converting-Enzyme Inhibition in Progressive Renal Insufficiency Study Group. N Engl J Med. 1996;334(15):939-45.
7. Ruggenenti P, Perna A, Gherardi G, Gaspari F, Benini R, Remuzzi G. Renal function and requirement for dialysis in chronic nephropathy patients on long-term ramipril: REIN follow-up trial. Gruppo Italiano di Studi Epidemiologici in Nefrologia (GISEN). Ramipril Efficacy in Nephropathy. Lancet. 1998;352(9136):1252-6.
8. Hou FF, Zhang X, Zhang GH, Xie D, Chen PY, Zhang WR, et al. Efficacy and safety of benazepril for advanced chronic renal insufficiency. N Engl J Med. 2006;354(2):131-40.
9. Ruggenenti P, Perna A, Remuzzi G, Gruppo Italiano di Studi Epidemiologici in N. ACE inhibitors to prevent end-stage renal disease: when to start and why possibly never to stop: a post hoc analysis of the REIN trial results. Ramipril Efficacy in Nephropathy. J Am Soc Nephrol. 2001;12(12):2832-7.
10. Remuzzi G, Ruggenenti P, Perna A, Dimitrov BD, de Zeeuw D, Hille DA, et al. Continuum of renoprotection with losartan at all stages of type 2 diabetic nephropathy: a post hoc analysis of the RENAAL trial results. J Am Soc Nephrol. 2004;15(12):3117-25.
11. Ahmed AK, Kamath NS, El Kossi M, El Nahas AM. The impact of stopping inhibitors of the renin-angiotensin system in patients with advanced chronic kidney disease. Nephrol Dial Transplant. 2010;25(12):3977-82.
12. Ahmed A, Jorna T, Bhandari S. Should We STOP Angiotensin Converting Enzyme Inhibitors/Angiotensin Receptor Blockers in Advanced Kidney Disease? Nephron. 2016;133(3):147-58.
13. Fu EL, Trevisan M, Clase CM, Evans M, Lindholm B, Rotmans JI, et al. Association of Acute Increases in Plasma Creatinine after Renin-Angiotensin Blockade with Subsequent Outcomes. Clin J Am Soc Nephrol. 2019;14(9):1336-45.
14. Tomlinson LA, Abel GA, Chaudhry AN, Tomson CR, Wilkinson IB, Roland MO, et al. ACE inhibitor and angiotensin receptor-II antagonist prescribing and hospital admissions with acute kidney injury: a longitudinal ecological study. PLoS One. 2013;8(11):e78465.

15. Ku E, McCulloch CE, Vittinghoff E, Lin F, Johansen KL. Use of Antihypertensive Agents and Association With Risk of Adverse Outcomes in Chronic Kidney Disease: Focus on Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers. J Am Heart Assoc. 2018;7(19):e009992.
16. Qiao Y, Shin JI, Sang Y, Inker LA, Secora A, Luo S, et al. Discontinuation of Angiotensin Converting Enzyme Inhibitors and Angiotensin Receptor Blockers in Chronic Kidney Disease. Mayo Clin Proc. 2019;94(11):2220-9.
17. Qiao Y, Shin JI, Chen TK, Inker LA, Coresh J, Alexander GC, et al. Association Between Renin-Angiotensin System Blockade Discontinuation and All-Cause Mortality Among Persons With Low Estimated Glomerular Filtration Rate. JAMA Intern Med. 2020;180(5):718-26.
18. http://www.isrctn.com/ISRCTN62869767.
19. 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.
20. Evans M, Carrero JJ, Bellocco R, Barany P, Qureshi AR, Seeberger A, et al. Initiation of erythropoiesis-stimulating agents and outcomes: a nationwide observational cohort study in anaemic chronic kidney disease patients. Nephrol Dial Transplant. 2017;32(11):1892-901.
21. Evans M, Suttorp MM, Bellocco R, Hoekstra T, Qureshi AR, Dekker FW, et al. Trends in haemoglobin, erythropoietin-stimulating agents and iron use in Swedish chronic kidney disease patients between 2008 and 2013. Nephrol Dial Transplant. 2016;31(4):628-35.
22. Wettermark B, Hammar N, Fored CM, Leimanis A, Otterblad Olausson P, Bergman U, et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. Pharmacoepidemiol Drug Saf. 2007;16(7):726-35.
23. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.
24. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. Eur J Epidemiol. 2017;32(9):765-73.
25. 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.
26. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12.
27. Hernan MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. Basic Clin Pharmacol Toxicol. 2006;98(3):237-42.
28. Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernan MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. Int J Biostat. 2010;6(2):Article 18.
29. Maringe C, Benitez Majano S, Exarchakou A, Smith M, Rachet B, Belot A, et al. Reflections on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. Int J Epidemiol. 2020.
30. Fu EL, Groenwold RHH, Zoccali C, Jager KJ, van Diepen M, Dekker FW. Merits and caveats of propensity scores to adjust for confounding. Nephrol Dial Transplant. 2019;34(10):1629-35.
31. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168(6):656-64.

32. Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. Comput Methods Programs Biomed. 2004;75(1):45-9.
33. Buchanan AL, Hudgens MG, Cole SR, Lau B, Adimora AA, Women's Interagency HIVS. Worth the weight: using inverse probability weighted Cox models in AIDS research. AIDS Res Hum Retroviruses. 2014;30(12):1170-7.
34. Kloecker DE, Davies MJ, Khunti K, Zaccardi F. Uses and Limitations of the Restricted Mean Survival Time: Illustrative Examples From Cardiovascular Outcomes and Mortality Trials in Type 2 Diabetes. Ann Intern Med. 2020;172(8):541-52.
35. McCaw ZR, Yin G, Wei LJ. Using the Restricted Mean Survival Time Difference as an Alternative to the Hazard Ratio for Analyzing Clinical Cardiovascular Studies. Circulation. 2019;140(17):1366-8.
36. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat Med. 2000;19(9):1141-64.
37. Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernan MA. A causal framework for classical statistical estimands in failure-time settings with competing events. Stat Med. 2020;39(8): 1199-236.
38. Hernan MA. The hazards of hazard ratios. Epidemiology. 2010;21(1):13-5.
39. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 2010;21(3):383-8.
40. Williamson T, Ravani P. Marginal structural models in clinical research: when and how to use them? Nephrol Dial Transplant. 2017;32(suppl_2):ii84-ii90.
41. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000;11(5):550-60.
42. Xie X, Liu Y, Perkovic V, Li X, Ninomiya T, Hou W, et al. Renin-Angiotensin System Inhibitors and Kidney and Cardiovascular Outcomes in Patients With CKD: A Bayesian Network Meta-analysis of Randomized Clinical Trials. Am J Kidney Dis. 2016;67(5):728-41.
43. Evans M, Carrero JJ, Szummer K, Akerblom A, Edfors R, Spaak J, et al. Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers in Myocardial Infarction Patients With Renal Dysfunction. J Am Coll Cardiol. 2016;67(14):1687-97.
44. Molnar MZ, Kalantar-Zadeh K, Lott EH, Lu JL, Malakauskas SM, Ma JZ, et al. Angiotensin-converting enzyme inhibitor, angiotensin receptor blocker use, and mortality in patients with chronic kidney disease. J Am Coll Cardiol. 2014;63(7):650-8.
45. Kidney Disease: Improving Global Outcomes (KDIGO) Blood Pressure Work Group. KDIGO Clinical Practice Guideline for the Management of Blood Pressure in Chronic Kidney Disease. Kidney inter, Suppl. 2012;2:337-414.
46. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney inter, Suppl. 2013;3:1-150.

Supplemental material

Supplemental Methods

Supplemental Table S1. Brief protocol of the pragmatic target trial and its emulation using data from the Swedish Renal Registry 2007-2017.

Supplemental Table S2. Definition of study outcomes and covariates.

Supplemental Table S3. Contribution to the weights at each time point by RASi treatment strategy.

Supplemental Table S4. Model coefficients for remaining uncensored in the continuation arm.

Supplemental Table S5. Model coefficients for remaining uncensored in the discontinuation arm.

Supplemental Table S6. Characteristics at six months after follow-up (end of grace period on the cloned data while accounting, or not, for informative censoring (before and after weighting, respectively).

Supplemental Table S7. Baseline characteristics of RASi users across two sub cohorts defined on their first detected eGFR drop between 20-30 ml/min/1.73m^2 or below 20 ml/min/1.73m^2.

Supplemental Table S8. Influence of weight truncation on the point estimates of risk differences comparing stopping vs. continuing (reference) RASi.

Supplemental Table S9. Sensitivity analysis: 5-year absolute risks and risk differences associated with stopping vs. continuing RASi on the negative control outcome of cancer diagnosis.

Supplemental Table S10. Sensitivity analysis: 5-year absolute risks and risk differences for always using vs. immediately stopping and not restarting RASi. RASi was modelled as a time-dependent exposure using inverse probability of treatment and censoring weighted estimation of a marginal structural model.

Supplemental Table S11. Sensitivity analysis: 5-year absolute risks and risk differences associated with stopping vs. continuing RASi among patients with ACR and potassium available (N = 3049).

Supplemental Table S12. Sensitivity analysis: 5-year absolute risks and risk differences associated with stopping vs. continuing RASi on the composite outcome of death and KRT.

Supplemental Figure S1. Schematic representation of cloning, censoring and weighting algorithm.

Supplemental Figure S2. Weighted cumulative incidence curves for mortality (A), MACE (B), KRT (C) and cancer (D) stratified by RASi use strategy in the cohort with first detected eGFR drop between 20-30 ml/min/1.73m^2. Thinner dotted lines represent 95% confidence intervals.

Supplemental Figure S3. Weighted cumulative incidence curves for mortality (A), MACE (B), KRT (C) and cancer (D) stratified by RASi use strategy in the cohort with first detected eGFR drop <20 ml/min/1.73m^2. Thinner dotted lines represent 95% confidence intervals.

Supplemental Figure S4. Weighted cumulative incidence curves for mortality (A), MACE (B) and KRT (C) standardized to the baseline distribution of confounders using a time-dependent exposure. The effect of always using vs. immediately stopping and not restarting RASi was estimated using inverse probability of treatment and censoring weighted estimation of a marginal structural model.

Supplemental Figure S5. Effect of stopping RASi on mortality (A), MACE (B) and KRT (C) across categories of age, sex, diabetes, heart failure, ischemic heart disease, ACR and potassium. Subgroup analyses for ACR and potassium were performed on the subset of individuals with these measurements available.

Supplemental Figure S6. Weighted cumulative incidence curves for the composite outcome of death or KRT by RASi strategy for the main cohort (A), cohort of individuals with first detected eGFR drop between 20-30 ml/min/1.73m^2 (B), and cohort of individuals with first detected eGFR drop <20 ml/min/1.73m^2 (C). Thinner dotted lines represent 95% confidence intervals.

Supplemental Methods

Target trial emulation using cloning, censoring and weighting

Here we describe in detail our implementation of target trial emulation and the cloning, censoring and weighting procedure. A thorough review of trial emulation can be found elsewhere (1, 2), as well as recent applications of the methodology (3-8).

Specifying details of the target trial

A simple way to structure the study design and analysis of an observational comparative effectiveness study is to use the target trial framework (1). This means that we think about a hypothetical randomized trial we would like to conduct and then use our observational data to explicitly emulate it. Explicitly emulating a randomized trial can prevent unnecessary biases such as immortal time bias and prevalent user bias (10-12), as well as making results from observational analyses more comparable to those from trials (13). Similar to a real trial, we first need to formally define the eligibility criteria of our hypothetical trial, the treatment strategies we would like to compare, how treatment is assigned to each individual, the duration of follow-up, the primary and secondary endpoints, the causal contrast of interest (intention-to-treat or per protocol effect), and the statistical analysis. Details of the target trial we wanted to emulate in our analysis are given in **Supplemental Table S1**.

In our study we were interested in comparing the treatment strategies "stop RASi within 6 months and remain off treatment" vs. "continue RASi during follow-up". We deliberately chose treatment strategies that required patients to be on or off treatment during the whole follow-up period, which ensured no cross-over between treatment arms. For example, in our study 57% of individuals who discontinued RASi within the first six months restarted treatment during follow-up. Comparing strategies such as "stop RASi within 6 months" vs. "continue RASi for 6 months" would therefore suffer from a lot of cross-over and dilution of the treatment effect.

Comparing treatment strategies that are sustained over time (as opposed to point interventions which happen only once, such as surgery or vaccination) requires methods that can appropriately adjust for time-varying confounding, such as the parametric G-formula or cloning, censoring and weighting (1, 14). We now explain in detail our implementation of the latter approach. A graphical depiction of the cloning, censoring and weighting procedure can be found in **Supplemental Figure S1**.

Step 1: Cloning and assigning replicates to the treatment strategies

The first step consists of cloning each individual into two identical replicates, each of whom is assigned to one strategy. The dataset will now be twice as large compared with the original dataset. Since each individual occurs in both strategies, no baseline confounding is present.

Step 2: Censoring replicates if and when they do not adhere to their assigned strategy

Note that there are now clones included in both strategies that do not necessarily always adhere to their assigned strategy. To estimate the effect of a particular treatment strategy, we therefore need to censor clones if and when their observed treatment does not match their assigned strategy anymore.

In our dataset, we therefore determined at each month whether a replicate was adherent to their assigned strategy and artificially censored them if they stopped adhering. Those assigned to the stopping strategy had to stop RASi within 6 months and remain off treatment for the remainder of the follow-up. Therefore, replicates in this treatment arm are censored under the following two conditions: if they had not stopped by month 6, or if they restarted treatment at any moment during follow-up after stopping. Those assigned to continuation were censored if they stopped treatment at any moment during follow-up.

Step 3: Inverse probability weighting to adjust for informative censoring

Because the artificial censoring of replicates is likely to be informative, this will lead to selection bias (collider stratification bias). We therefore need to use inverse probability weighting to adjust for this selection bias, which is the most involved step of the cloning, censoring and weighting procedure. In brief, uncensored replicates receive a weight that is equal to the inverse of the probability of remaining uncensored, conditional on their own covariate history. Intuitively, the weighting will upweight uncensored replicates who have similar characteristics as censored replicates (see also **Supplemental Figure 1**). This creates a pseudopopulation in which censoring does not depend on measured characteristics and is no longer informative.

To estimate the inverse probability of censoring weights, we first fit a pooled logistic model with being uncensored as the outcome and as independent variables an indicator for time (e.g., month and month squared [quadratic term], or more flexible functions of time such as restricted cubic splines), baseline and time-varying confounders. We fit a pooled logistic model for each arm separately for two reasons. First, the censoring pattern is likely different between both treatment strategies and secondly, this will better capture treatment by covariate interaction (2). The regression coefficients from these models are shown in **Supplemental Tables S4-5**.

Next, we used the probabilities estimated by these models to construct the inverse probability of censoring weights as shown in **Supplemental Table S3**. Weights were set to 1 during the first 5 months for replicates in the stopping arm that had not yet discontinued RASi, as their probability to remain uncensored is per definition 1. We truncated the weights at the 99.5th percentile to avoid undue influence of very large weights. Truncating the weights is a trade-off between bias and precision: truncation of large weights will lead to narrower confidence intervals at the expense of introducing some bias. The mean of the truncated weights was 2.2 and the maximum 35.0. Using untruncated weights showed virtually similar results (**Supplemental Table S8**). The weights showed good ability to remove imbalance at the end of the grace period (6 months after baseline) (**Supplemental Table S6**).

Step 4: Primary analysis

Next, we stacked the two datasets (stopping and continuing). We used a weighted pooled logistic model to estimate the per protocol effect of stopping vs. continuing. The pooled logistic model contained indicators for time (month and month squared), an indicator for treatment strategy, and interactions between time and treatment strategy, as well as the weights estimated in step 3. The pooled logistic model was used to calculate weighted cumulative incidence curves. The weighted curves were then used to calculate 5-year absolute risk differences and differences in restricted mean survival time. To account for the weighting we used nonparametric bootstrapping based on 500 samples to obtain valid 95% confidence intervals.

RASi as time-dependent exposure using inverse probability of treatment and censoring weighted estimation of a marginal structural model

We used a marginal structural model to estimate the effect of time-varying RASi use on outcomes. A marginal structural model was used because some of the time-varying confounders may also be affected by treatment itself (i.e., over time the covariate plays both the role of confounder and mediator of the effect of treatment on outcomes). Using a time-dependent regression analysis would therefore lead to biased results due to adjustment in the causal pathway and introducing collider stratification bias (15).

The method described here instead uses inverse probability weighting to appropriately adjust for time-varying confounding and censoring. Inverse probability of treatment weights (IPTW) were used to adjust for time-varying confounding, whereas inverse probability of censoring weights (IPCW) were used to adjust for informative censoring. The IPTW and IPCW were estimated using the same time-fixed and time-varying confounders that were used in the main analysis using the cloning, censoring and weighting design (see **Supplemental Table 1** for variables).

Treatment weights

The IPTW consists of a numerator and a denominator. The denominator is used to adjust for the time-varying confounding, whereas the numerator is used to stabilize the weights so that they do not become excessively large. To estimate the numerator and denominator for the IPTW, we fitted two separate pooled logistic regression models. The pooled logistic regression model for the numerator had discontinuation as the outcome and an indicator for time and all time-fixed confounders as independent variables. The pooled logistic regression model for the denominator additionally included all time-varying confounders as independent variables. Time in both models was modelled using month and month squared as predictors. The predicted values from these pooled logistic models were used to estimate the IPTW.

Censoring weights

In order to estimate the effect of "always" vs. "never" using RASi, we censored patients when they restarted RASi treatment after they had discontinued. This censoring is likely to be informative. We therefore additionally constructed IPCW to adjust for this informative censoring. The IPCW were constructed in a similar manner as the IPTW specified above, with the only difference being that the outcome was "remaining uncensored" instead of "discontinuation". Since patients who had not discontinued (yet) cannot be censored by definition, censoring weights were only calculated for the patients after they discontinued. For the other records, the IPCW were set to 1.

Outcome model

The IPTW and IPCW were multiplied to obtain the final stabilized weights used in the outcome model. We estimated the effect of RASi discontinuation vs. continuation on all-cause mortality, MACE and KRT by fitting a weighted pooled logistic model that included month, month squared, a time-dependent treatment variable, interactions between time and treatment and all baseline covariates. This model was used to estimate adjusted cumulative incidence curves. The cumulative incidence curves were standardized to the distribution of baseline variables in the study population (17). Under the assumptions of exchangeability, positivity, consistency and no model misspecification, this approach estimates the average causal effect of treatment discontinuation on outcomes in the original study population (15).

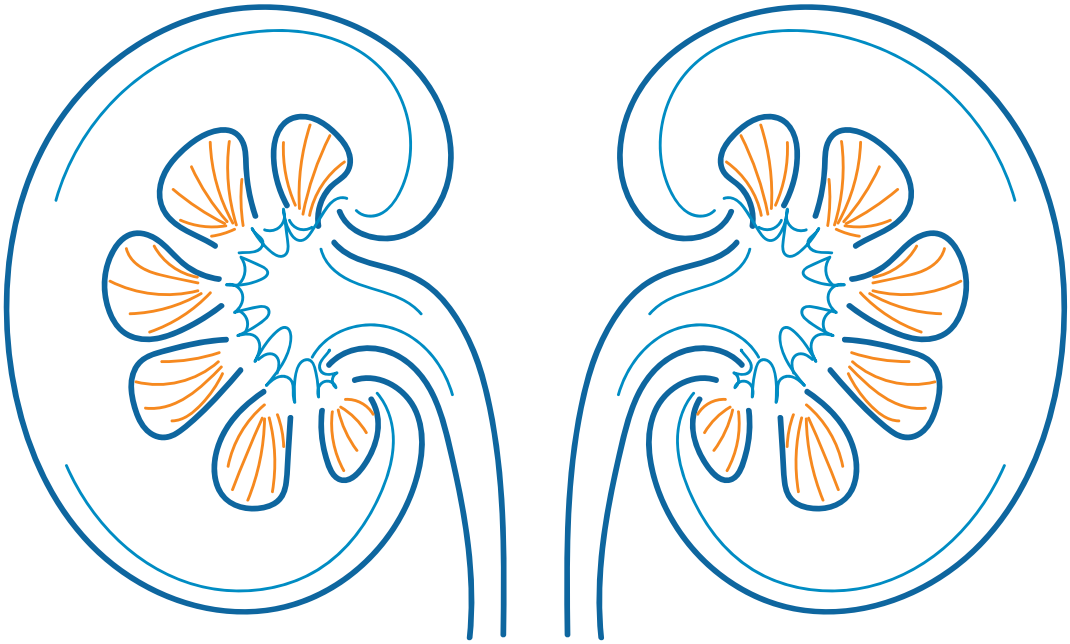
The stabilized weights had a mean of 1.0, a minimum of 0.095 and a maximum of 69.9. Weights were not truncated; truncation at the 99.5th percentile gave virtually identical results (mean of weights after truncation: 1.0; maximum: 2.4; results not shown). Nonparametric bootstrap with 500 samples was used to compute percentile-based 95% confidence intervals for the absolute estimates.

Supplemental references

1. 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.
2. Maringe C, Benitez Majano S, Exarchakou A, Smith M, Rachet B, Belot A, et al. Reflections on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. Int J Epidemiol. 2020.
3. Huitfeldt A, Kalager M, Robins JM, Hoff G, Hernan MA. Methods to Estimate the Comparative Effectiveness of Clinical Strategies that Administer the Same Intervention at Different Times. Curr Epidemiol Rep. 2015;2(3):149-61.
4. Petito LC, Garcia-Albeniz X, Logan RW, Howlader N, Mariotto AB, Dahabreh IJ, et al. Estimates of Overall Survival in Patients With Cancer Receiving Different Treatment Regimens: Emulating Hypothetical Target Trials in the Surveillance, Epidemiology, and End Results (SEER)-Medicare Linked Database. JAMA Netw Open. 2020;3(3):e200452.
5. Lyu H, Yoshida K, Zhao SS, Wei J, Zeng C, Tedeschi SK, et al. Delayed Denosumab Injections and Fracture Risk Among Patients With Osteoporosis: A Population-Based Cohort Study. Ann Intern Med. 2020.
6. Garcia-Albeniz X, Hernan MA, Logan RW, Price M, Armstrong K, Hsu J. Continuation of Annual Screening Mammography and Breast Cancer Mortality in Women Older Than 70 Years. Ann Intern Med. 2020;172(6):381-9.
7. Danaei G, Garcia Rodriguez LA, Cantero OF, Logan RW, Hernan MA. Electronic medical records can be used to emulate target trials of sustained treatment strategies. J Clin Epidemiol. 2018;96:12-22.
8. Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernan MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. Int J Biostat. 2010;6(2):Article 18.
9. Keyhani S, Cheng EM, Hoggatt KJ, Austin PC, Madden E, Hebert PL, et al. Comparative Effectiveness of Carotid Endarterectomy vs Initial Medical Therapy in Patients With Asymptomatic Carotid Stenosis. JAMA Neurol. 2020.
10. 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.
11. Dickerman BA, Garcia-Albeniz X, Logan RW, Denaxas S, Hernan MA. Avoidable flaws in observational analyses: an application to statins and cancer. Nat Med. 2019;25(10):1601-6.
12. 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.
13. 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.
14. Hernan MA. How to estimate the effect of treatment duration on survival outcomes using observational data. BMJ. 2018;360:k182.
15. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168(6):656-64.

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

17. Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. Comput Methods Programs Biomed. 2004;75(1):45-9.



CHAPTER 9

When to initiate dialysis to reduce mortality and cardiovascular events in advanced chronic kidney disease: nationwide cohort study

Edouard L. Fu, Marie Evans, Juan-Jesus Carrero, Hein Putter, Catherine M. Clase, Fergus J. Caskey, Maciej Szymczak, Claudia Torino, Nicholas C. Chesnaye, Kitty J. Jager, Christoph Wanner, Friedo W. Dekker, Merel van Diepen

Submitted

Abstract

Objectives: To identify the optimal estimated glomerular filtration rate (eGFR) to initiate dialysis in persons with advanced chronic kidney disease.

Design: Nationwide observational cohort study. We mimicked the strict design criteria of a clinical trial using the cloning, censoring, and weighting method to eliminate immortal time bias, lead time bias and survivor bias.

Setting: National Swedish Renal Registry of nephrologist-referred patients.

Participants: Individuals had a baseline eGFR between 10-20 ml/min/1.73m^2 and were included between January 1, 2007, and December 31, 2016, with follow-up until June 1, 2017.

Main outcome measures: A dynamic marginal structural model was used to estimate adjusted hazard ratios (HR) and absolute risks for 5-year all-cause mortality and major adverse cardiovascular events (MACE; composite of cardiovascular death, nonfatal myocardial infarction, or nonfatal stroke) for fifteen dialysis initiation strategies with eGFR values between 4 and 19 ml/min/1.73m^2 in increments of 1 ml/min/1.73m^2. An eGFR between 6-7 ml/min/1.73m^2 ($eGFR_{6\text{-}7}$) was taken as reference.

Results: Among 10,290 incident individuals with advanced CKD (median age 73 years; 36% women; median eGFR 16.8 ml/min/1.73m^2), 3822 individuals initiated dialysis, 4160 died and 2446 experienced a MACE. A parabolic relationship was observed for mortality, with the lowest risk for $eGFR_{15\text{-}16}$. Compared with dialysis initiation at $eGFR_{6\text{-}7}$, initiation at $eGFR_{15\text{-}16}$ was associated with a 5.1% (95% CI 2.5% to 6.9%) lower absolute 5-year mortality risk and 2.9% (95% CI 0.2% to 5.5%) lower MACE risk, corresponding with HRs of 0.89 (95% CI 0.87 to 0.92) and 0.94 (95% CI 0.91 to 0.98), respectively. This 5.1% absolute risk difference corresponded to a mean postponement of death of 1.6 months over 5 years of follow-up. However, dialysis would need to be initiated 4 years earlier. When emulating the intended strategies of the IDEAL trial ($eGFR_{10\text{-}14}$ vs. $eGFR_{5\text{-}7}$) and the achieved eGFR levels in IDEAL ($eGFR_{7\text{-}10}$ vs. $eGFR_{5\text{-}7}$), HR's for all-cause mortality were 0.96 (95% CI 0.94 to 0.99) and 0.97 (95% CI 0.94 to 1.00), respectively, which are congruent with the findings of the randomized IDEAL trial.

Conclusions: Very early dialysis initiation was associated with a modest reduction in mortality and cardiovascular events. For most individuals such a reduction may not outweigh the burden of a substantially longer period spent on dialysis.

Introduction

Worldwide, more than 3 million individuals with kidney failure require maintenance dialysis treatment for survival (1-4). These numbers are expected to double by 2030 (2). The societal and patient burden of kidney failure treated by dialysis is high: for instance, the United States Medicare fee-for-service spending for beneficiaries with kidney failure was 36.6 billion in 2018 (3). The mean annual healthcare costs per hemodialysis patient are $93,191 in the United States (3), and similar numbers are reported for European countries (5-8). Dialysis treatment also places a large burden on patients' daily lives (9, 10). Determining the optimal timing of dialysis is therefore of substantial importance.

Despite extensive previous literature, there is absence of evidence on whether an optimal GFR to start dialysis exists, and if so where it lies. Previous observational studies that attempted to investigate multiple estimated glomerular filtration rate (eGFR) strategies have been limited by insufficient power (11-13), immortal time bias (14-17) or lead time and selection biases (16-32). In 2010 the Initiating Dialysis Early and Late (IDEAL) trial (33) showed that a strategy to start dialysis at an eGFR of 10-14 ml/min/1.73m^2 was not superior to one of waiting until symptoms develop or eGFR is 5-7 ml/min/1.73m^2. This is reflected in subsequent guidelines, which recommend starting dialysis when symptoms and signs attributable to kidney failure arise rather than a specific kidney function (34-40). However, IDEAL only compared two strategies, from which an optimal GFR cannot be derived. In addition, the achieved GFR separation in IDEAL was 1.8 (9.0 vs. 7.2) ml/min/1.73m^2 by Modification of Diet in Renal Disease equation. It therefore remains possible that there is a kidney function outside this range at which starting dialysis is associated with better outcomes, and uncertainty on this issue in providers persists (41).

In the absence of evidence on an optimal GFR level, decision-making may be influenced by other factors, including potential financial incentives. Indeed, large between-country variation exists in the mean eGFR at dialysis start: from approximately 5 ml/min/1.73m^2 in Taiwan, to 8.5 in the United Kingdom and 11 ml/min/1.73m^2 in the United States (36). Some health systems in the United States (42) start at a mean eGFR of 16-17 ml/min/1.73m^2. This broad heterogeneity may lead to differences in outcomes and healthcare costs.

Ideally, this complex question would be addressed in a multi-armed randomized trial. However, such a trial is unlikely to be conducted because the required sample size is large and recruitment is problematic: IDEAL recruited 828 patients over 8 years. In the absence of trial evidence, clinical decisions could be aided by well-conducted observational studies which explicitly mimic the strict design criteria of this multi-armed trial. We therefore used novel analytical methodology to compare

different dialysis initiation strategies using data from a nationwide cohort of non-dialysis dependent patients with advanced chronic kidney disease (CKD) under nephrologist care.

Methods

This study was reported following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (43).

Data sources

We used data from the Swedish Renal Registry, a nationwide registry of patients with CKD categories G3-5 attending routine nephrologist care in Sweden (44, 45), during the period 2007-2017. The Swedish Renal Registry includes information from outpatient nephrologist visits, including CKD etiology, laboratory tests, blood pressure and other results obtained from routine clinical examination, as well as the date of kidney replacement therapy (either kidney transplantation or long-term dialysis). Registry enrolment is mandatory in Sweden when patients reach an eGFR <30 mL/min/1.73m^2, but some clinics may start reporting them earlier. Subsequent outpatient visits to nephrology care (on average 2-3 per year per patient) are registered until death or emigration. Nearly all nephrology clinics in Sweden (96%) report to the Swedish Renal Registry and the estimated national coverage is >75% for nephrologist-referred patients with CKD G4-5 (46).

Using each citizen's unique personal identification number, the Swedish Renal Registry data was linked to other national registries. The Swedish Prescribed Drug Registry provided complete information on all prescribed drugs dispensed at Swedish pharmacies (47); the Swedish Patient Registry added information on all outpatient specialist consultations and hospitalizations occurring in Swedish healthcare since 1997, and was used to obtain information on comorbidities and outcomes (48); the Swedish Death Registry added information on the date and causes of death (49). All these registries are run by the Swedish National Board of Welfare, a government institution, and are considered to have no or minimal loss to follow-up. All patients are informed about their participation in the registry and have the possibility to opt out at any time.

Study design and patient selection

This observational study emulated a pragmatic clinical trial (50) comparing the effect of initiating dialysis at various eGFR levels on mortality and cardiovascular outcomes in people with advanced CKD, and in general follows the approach

proposed by Sjölander *et al.* (51). **WebTable 1** outlines the protocol of such a trial and the emulation procedure. Explicit emulation of a trial, and in particular aligning the start of follow-up with the assignment of treatment strategies, eliminates immortal time bias, selection/survivor bias and lead time bias, which significantly affected previous observational studies (51-53). A detailed explanation of how these biases arise can be found in the **Supplemental Methods**. Our analysis included individuals who met the following eligibility criteria between January 1, 2007 and December 31, 2016: aged 18 years or older, an eGFR measurement between 10-20 ml/min/1.73m^2 with a previous eGFR measurement between 10-30 ml/min/1.73m^2 as confirmation, no history of kidney replacement therapy, and at least one available measurement of systolic blood pressure, diastolic blood pressure, total calcium, phosphate, albumin and hemoglobin. Baseline was defined as the first time when all of these eligibility criteria were met. eGFR was calculated with the CKD-EPI equation (54) from routine plasma creatinine measurements performed by enzymatic or corrected Jaffe methods traceable to isotope dilution mass spectroscopy standards. As information on ethnicity is not available in Sweden by law, we assumed all patients to be Caucasian.

Treatment strategies

We compared fifteen dialysis initiation strategies with eGFR values ranging between 4 and 19 ml/min/1.73m^2 in increments of 1 ml/min/1.73m^2. An eGFR between 6-7 ml/min/1.73m^2 ($eGFR_{6\text{-}7}$) was taken as the reference group since this is the eGFR at which most individuals initiate dialysis in Sweden.

Study outcomes

The primary outcome was 5-year all-cause mortality. The secondary outcome was MACE (defined as a composite endpoint of cardiovascular death, non-fatal myocardial infarction and non-fatal stroke). ICD-10 codes for ascertainment of cardiovascular outcomes are listed in **WebTable 2**. Each patient was followed until the first of occurrence of an event, five years after baseline, or administrative censoring (June 1, 2017).

Statistical analysis

We used the method of cloning, censoring and weighting (50, 52, 55-57) to emulate a target trial comparing the effects of different dialysis initiation strategies (see **Supplemental Methods** and **WebFigure 1** for a detailed discussion on target trial emulation and the cloning, censoring and weighting method). Explained briefly, we created a dataset with fifteen copies of each eligible individual (cloning step) and

assigned each of the replicates to one of the treatment strategies at the start of follow-up. Thereafter, we assessed at monthly intervals whether replicates adhered to their assigned treatment strategy; replicates were censored as soon as their actual treatment deviated from their assigned treatment strategy, thereby ensuring that replicates followed their assigned strategy (censoring step). To adjust for the potential selection bias induced by this artificial censoring, each individual received a time-varying inverse probability weight (58) (weighting step). Informally, the denominator of the weights was the probability that a replicate remained uncensored during follow-up (i.e., remained on the assigned treatment strategy). These weights created fifteen pseudopopulations in which censoring was independent of measured prognostic factors. We estimated the time-varying weights by fitting a pooled logistic model for the monthly probability of remaining uncensored, including variables for time and baseline plus time-varying covariates listed in **WebTable 2**. Models were fitted separately for each treatment strategy to allow for treatment-covariate interaction (57, 59). The variables for each model and their regression coefficients for the $eGFR_{6\text{-}7}$ strategy are reported in **WebTable 3**. To avoid undue influence of outliers, weights were truncated at the 99.95^{th} percentile (60).

After cloning, censoring and weighting, we estimated the effect of each dialysis initiation strategy on 5-year all-cause mortality and MACE using a weighted pooled logistic regression model, including an indicator for treatment strategy (modelled as restricted cubic spline with knots at 5, 8, 11, 14 and 17 ml/min/$1.73m^2$), month, month squared, their interactions to allow for non-proportional hazards, and all baseline covariates. This weighted model estimates the parameters of a dynamic marginal structural model when the covariates include all joint determinants of censoring and the outcome (55). The predicted probabilities from this logistic model were used to estimate the adjusted 5-year probability of mortality and MACE under each treatment strategy and to produce weighted cumulative incidence curves, which were standardized to the baseline distribution of confounders (61, 62). From these probabilities we also derived 5-year risk differences, risk ratios and hazard ratios. We estimated cause-specific cumulative incidences to account for the competing event of kidney transplantation (63, 64). In addition, we also calculated the 5-year restricted mean survival time (RMST) and the 5-year RMST differences between each dialysis initiation strategy. The RMST is interpreted as the average survival time over a fixed follow-up period. Graphically, it corresponds to the area under the survival curve (65). The 5-year RMST *difference* compares the areas under the survival curves for the different dialysis initiation strategies. It is interpreted as the mean postponement of the outcome in one group compared with the reference. Pointwise 95% percentile confidence intervals were calculated using nonparametric bootstrap based on 500 full samples. The 5-year RMST difference was compared with the postponement of dialysis initiation to provide insight into this trade-off. Postponement of dialysis initiation was determined by the average eGFR decline

before dialysis initiation using a linear mixed model (**Supplemental Methods**). R version 3.6.2 was used for all statistical analyses.

Sensitivity analyses

We pre-specified several analyses to test the robustness of our main results. First, we emulated the IDEAL trial comparing early ($eGFR_{10-14}$) versus late initiation ($eGFR_{5-7}$) on mortality and MACE to validate our analytical methods. We added a third "intermediate initiation" arm ($eGFR_{7-10}$), which includes the mean achieved eGFR in the early initiation arm in IDEAL. Second, we performed stratified analyses by age (≥70 vs. <70 years), sex, presence of diabetes, eGFR at baseline (10-15 vs. 15-20 ml/min/1.73m^2), presence of ischemic heart disease, and presence of heart failure. Third, we investigated the influence of adjustment for measured confounders on our point estimates by sequentially adjusting for baseline and time-varying confounders. Fourth, we compared results when using nontruncated weights. Fifth, we excluded individuals with cancer at baseline. Sixth, we used a different analytical method for the competing event of kidney transplantation. We modelled the direct effect of dialysis initiation strategies on mortality, not mediated through kidney transplantation, by adding additional inverse probability of censoring weights (63). Intuitively, this models the effect of dialysis initiation strategies in a hypothetical world in which no kidney transplantations occur. Seventh, we additionally adjusted for time-dependent measures of urinary albumin-to-creatinine ratio and plasma potassium in our analyses. This analysis was restricted to the 4286 individuals with these measurements available. Although these laboratory values are routinely measured in this population, reporting these to the Swedish Renal Registry was not mandatory until 2015. Because some physicians chose to report this information whereas others did not, we assumed that these data were missing completely at random (44). Eighth, we censored patients who chose conservative treatment, where patients explicitly chose treatment of kidney failure without dialysis. We used additional inverse probability of censoring weights to account for informative censoring. Intuitively, this models the effect of dialysis initiation strategies in a hypothetical world in which no patients choose conservative management. Lastly, we analyzed our data using the "from initiation" and "from threshold" method analogous to previous observational studies (14-29) to show that immortal time bias and selection/survivor bias give an artificial survival advantage to late dialysis initiation (51, 52). A detailed description of these methods and how bias arises is provided in the **Supplemental Methods**. Due to computational efficiency and lower power with fifteen strategies, subgroup and sensitivity analyses were performed using three dialysis initiation strategies only.

Patient involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of the study. No patients were asked to advise on interpretation or writing up of results. Being a study based on anonymised nationwide register data, there are no plans to disseminate the results of the research to study participants.

Results

Of 30,180 individuals registered in the Swedish Renal Registry during the study period, 10,290 individuals with an eGFR between 10-20 ml/min/1.73m^2 were eligible for inclusion in our study. **WebFigure 2** displays the patient selection flow chart, and **Table 1** describes their baseline characteristics. At baseline, individuals had a median (interquartile range; IQR) age of 73 (63-80) years, 35.7% were women and 42.1% had diabetes. The median eGFR was 16.8 (14.3-18.6) ml/min/1.73m^2 and 68.9% of the study population had an eGFR between 15-20 ml/min/1.73m^2.

Table 1. Baseline characteristics of individuals under nephrologist care with eGFR between 10-20 ml/min/1.73m^2 registered in the Swedish Renal Registry during January 2007 and December 2016.

	Overall (N = 10,290)
Age, median (IQR), y	73.0 [63.0, 80.0]
Age group, N (%)	
<50	1057 (10.3)
50-59	1030 (10.0)
60-69	2119 (20.6)
70-79	3247 (31.6)
>=80	2837 (27.6)
Female Sex	3739 (36.3)
Primary kidney disease, N (%)	
Diabetes	2427 (23.6)
Hypertension/Renovascular	2277 (22.1)
Glomerulonephritis	1066 (10.4)
Polycystic kidney disease	636 (6.2)
Pyelonephritis	313 (3.0)
Other	2083 (20.2)
Unknown	1488 (14.5)

	Overall (N = 10,290)
Clinical and laboratory values	
Previous eGFR before baseline, median (IQR), ml/min/1.73m^{2}.[b]	20.4 [16.4, 22.7]
Baseline eGFR, median (IQR), ml/min/1.73 m^{2}.[b]	16.8 [14.3, 18.6]
Baseline eGFR between 15-20 ml/min/1.73 m^{2}.[b], N (%)	7087 (68.9)
SBP, mean (SD), mmHg	139.6 (21.0)
SBP category, N (%)	
<120	1270 (12.3)
120-139	3774 (36.7)
140-159	3315 (32.2)
>160	1931 (18.8)
DBP, mean (SD), mmHg	76.6 (11.8)
DBP category, N (%)	
<80	5346 (52.0)
80-89	3354 (32.6)
90-99	1201 (11.7)
>100	389 (3.8)
BMI, mean (SD), kg/m^{2}.[c]	27.9 (5.7)
Total calcium, mean (SD), mmol/L	2.3 (0.2)
Total calcium category, N (%)	
<2.0	351 (3.4)
2.0-2.19	2156 (21.0)
2.20-2.44	6502 (63.2)
>2.45	1281 (12.4)
Phosphorus, mean (SD), mmol/L	1.4 (0.3)
Phosphorus category, N (%)	
<0.8	45 (0.4)
0.8-1.49	6628 (64.4)
1.50-1.99	3215 (31.2)
>2.0	402 (3.9)
Albumin, mean (SD), g/L	36.5 (4.7)
Albumin category, N (%)	
<25	152 (1.5)
25-29	555 (5.4)
30-39	6889 (66.9)
>40	2694 (26.2)
Hemoglobin, mean (SD), g/L	119.4 (14.1)

	Overall (N = 10,290)
Hemoglobin category, N (%)	
<90	143 (1.4)
90-99	585 (5.7)
100-114	3071 (29.8)
>115	6491 (63.1)
UACR, median (IQR), mg/mmol[c]	57.6 [11.6, 180.0]
UACR category, N (%)	
A1 (<3)	570 (9.9)
A2 (3-29)	1698 (29.4)
A3.1 (30-70)	815 (14.1)
A3.2 (>70)	2701 (46.7)
Potassium, mean, mmol/L[c]	4.5 (0.6)
C-reactive protein, median, ng/mL[c]	5.0 [2.1, 10.0]
Ferritin, median, ng/mL[c]	150.0 [77.0, 274.0]
Comorbidities, N (%)	
Hypertension	8796 (86.6)
Acute coronary syndrome	1906 (18.5)
Other ischemic heart disease	3177 (30.9)
Heart failure	2612 (25.4)
Diabetes	4329 (42.1)
Valve disorders	670 (6.5)
Stroke	1243 (12.1)
Other cerebrovascular disease	1300 (12.6)
Atrial fibrillation	1808 (17.6)
Other arrhythmia	898 (8.7)
Peripheral vascular disease	1415 (13.8)
Chronic obstructive pulmonary disease	792 (7.7)
Other lung disease	1605 (15.6)
Venous thromboembolism	816 (7.9)
Cancer in previous year	1025 (10.0)
Liver disease	368 (3.6)
Fracture in previous year	297 (2.9)
Medication use, N (%)	
Beta blocker	6736 (65.5)
Calcium channel blocker	6348 (61.7)
Diuretic	7356 (71.5)
ACEi/ARB	6971 (67.7)

	Overall (N = 10,290)
Lipid lowering drug	5610 (54.5)
Potassium binder	1270 (12.3)
Phosphate binder	1034 (10.0)
Erythropoietin-stimulating agent	3160 (30.7)
Vitamin D	5977 (58.1)
Digoxin	158 (1.5)
Nitrate	1474 (14.3)
Antiplatelet	4345 (42.2)
Anticoagulant	1214 (11.8)
Sodium bicarbonate	4381 (42.6)
Calendar Year, N (%)	
2007-2010	3211 (31.2)
2011-2013	3473 (33.8)
2014-2016	3606 (35.0)
Hospitalizations	
Number of hospital admissions in previous year, median (IQR)	0.0 [0.0, 2.0]
Any hospitalization in previous year, N (%)	4770 (46.4)
Hospital admission due to cardiovascular causes in previous year, N (%)	1614 (15.7)

eGFR = estimated glomerular filtration rate; SBP = systolic blood pressure; DBP = diastolic blood pressure; UACR = albumin-to-creatinine ratio.

[a] Due to the cloning step in the cloning, censoring and weighting method, patient characteristics are identical at baseline for the early, intermediate and late dialysis initiation groups. A detailed explanation of the cloning, censoring and weighting method can be found in the Methods and Supplemental Methods.

[b] eGFR was calculated with the CKD-EPI formula. Patients were required to have two eGFR measurements to be eligible for inclusion. The median (IQR) time between the baseline and previous eGFR measurement was 154 (93-234) days.

[c] BMI was missing in 25.8% of individuals, UACR in 43.8%, potassium in 29.1%, CRP in 15.9% and ferritin in 60.3%, because reporting these variables to the Swedish Renal registry was not mandatory (**Webtable 2**). Due to the high degree of missingness, these variables were not used in further analyses and are presented for descriptive purposes only.

During follow-up 3822 individuals started dialysis, the majority with an eGFR between 5 and 8 ml/min/1.73m^2 (**WebFigure 3**). Hemodialysis was the initial dialysis modality in 2339 individuals (61.2%) and peritoneal dialysis in 1483 individuals (38.8%).

Dialysis initiation strategies and risk of mortality or MACE

During a median (IQR) follow-up of 3.1 (1.7-5.0) years, 4160 (40.4%) individuals died. **Table 2** and **Figure 1A** show the 5-year absolute risks, risk differences, hazard ratios and cumulative incidence curves for all-cause mortality for all dialysis initiation strategies. For mortality, the absolute risk decreased from $eGFR_{18-19}$ to a nadir at $eGFR_{15-16}$ and progressively increased again between $eGFR_{15-16}$ and $eGFR_{4-5}$. Compared with $eGFR_{6-7}$, 5-year absolute risk differences varied between an increase of 0.8% (95% CI, 0.0% to 1.6%) for $eGFR_{4-5}$ and a decrease of 5.1% (95% CI, 2.5% to 6.9%) for $eGFR_{15-16}$ (**Figure 2A**), with corresponding hazard ratios of 1.01 (95% CI, 1.00 to 1.02) and 0.89 (95% CI, 0.87 to 0.92), respectively. When the mean eGFR at dialysis start in the United States was taken as reference group (i.e. $eGFR_{11-12}$), risk differences varied between an increase of 2.8% (95% CI, 0.5% to 5.3%) and a decrease of 3.1% (95% CI, 0.9% to 5.2%) (**WebTable 4**). Compared with $eGFR_{6-7}$, the maximum 5-year RMST difference was 1.6 months (95% CI, 1.0 to 2.0) for $eGFR_{15-16}$, and these patients would need to start dialysis on average 47.9 months (95% CI, 46.2 to 49.6) earlier than $eGFR_{6-7}$ (**WebTables 5-6** and **Figure 3**).

Figure 1. Weighted, standardized cumulative incidence curves for mortality (A) and MACE (B) stratified by different dialysis initiation strategies.

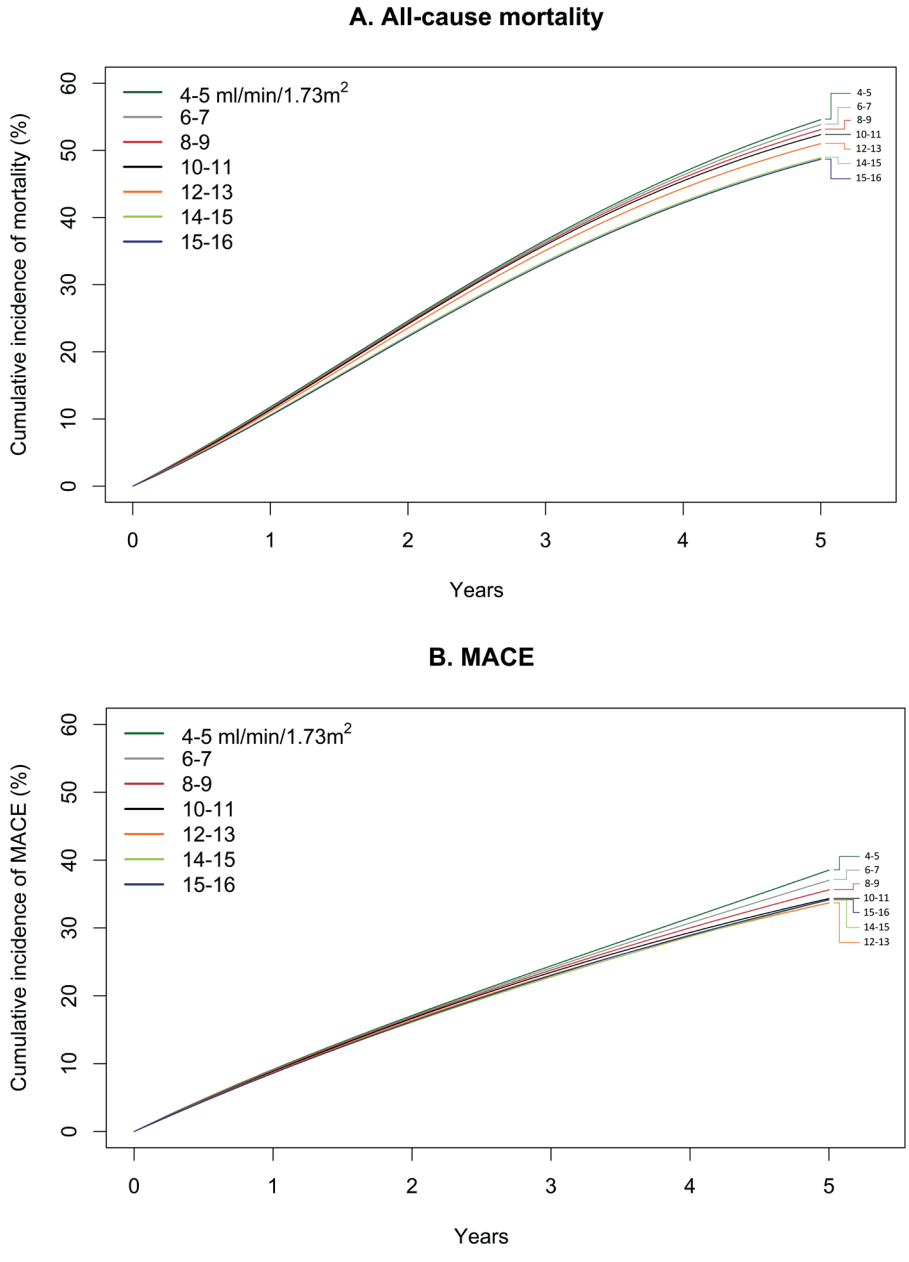


Table 2. 5-year absolute risks, risk differences, risk ratios and hazard ratios for all-cause mortality associated with initiating dialysis with eGFR values between 4 and 19 ml/min/1.73m² in increments of 1 ml/min/1.73m² with 6-7 ml/min/1.73m² as reference.

Dialysis initiation strategy	Persons[a], *n*	Outcomes, *n*	Median (IQR) eGFR at dialysis initiation[a]	5-year absolute risk (%, 95% CI)[b]	Risk difference (%, 95% CI)[b]	Risk ratio (95% CI)[b]	Hazard ratio (95% CI)[b]
18-19	3483	484	18.5 (18.2-18.7)	50.9 (44.0 to 55.3)	-2.9 (-7.2 to -0.1)	0.95 (0.86 to 1.00)	0.97 (0.87 to 1.02)
17-18	4911	742	17.6 (17.3-17.8)	50.6 (44.1 to 54.4)	-3.2 (-6.9 to -0.8)	0.94 (0.87 to 0.99)	0.93 (0.87 to 0.97)
16-17	6079	1037	16.5 (16.3-16.8)	49.5 (43.9 to 53.9)	-4.3 (-6.8 to -2.1)	0.92 (0.87 to 0.96)	0.90 (0.87 to 0.94)
15-16	7087	1312	15.5 (15.3-15.7)	48.7 (43.9 to 53.4)	-5.1 (-6.9 to -2.5)	0.90 (0.87 to 0.95)	0.89 (0.87 to 0.92)
14-15	7932	1595	14.5 (14.3-14.7)	48.9 (44.1 to 54.0)	-4.9 (-6.6 to -2.5)	0.91 (0.88 to 0.95)	0.90 (0.88 to 0.94)
13-14	8657	1888	13.5 (13.2-13.8)	49.9 (45.2 to 54.8)	-4.0 (-5.5 to -1.9)	0.93 (0.90 to 0.96)	0.92 (0.90 to 0.95)
12-13	9281	2187	12.6 (12.3-12.8)	51.0 (46.3 to 55.8)	-2.8 (-4.4 to -1.1)	0.95 (0.92 to 0.98)	0.95 (0.93 to 0.97)
11-12	9808	2426	11.5 (11.3-11.7)	51.8 (47.1 to 56.4)	-2.0 (-3.7 to -0.4)	0.96 (0.93 to 0.99)	0.96 (0.94 to 0.99)
10-11	10290	2704	10.5 (10.2-10.8)	52.4 (47.6 to 56.9)	-1.5 (-3.0 to -0.1)	0.97 (0.94 to 1.00)	0.98 (0.95 to 0.99)
9-10	10290	2839	9.5 (9.2-9.8)	52.7 (48.2 to 57.1)	-1.1 (-2.3 to 0.0)	0.98 (0.96 to 1.00)	0.98 (0.97 to 1.00)
8-9	10290	2991	8.5 (8.2-8.7)	53.1 (48.6 to 57.4)	-0.7 (-1.5 to 0.0)	0.99 (0.97 to 1.00)	0.99 (0.98 to 1.00)
7-8	10290	3088	7.5 (7.3-7.8)	53.5 (48.9 to 57.6)	-0.4 (-0.8 to 0.0)	0.99 (0.99 to 1.00)	0.99 (0.99 to 1.00)
6-7	**10290**	**3168**	**6.5 (6.2-6.7)**	**53.8 (49.2 to 58.0)**	**Reference**	**Reference**	**Reference**
5-6	10290	3196	5.5 (5.3-5.8)	54.2 (49.6 to 58.5)	0.4 (0.0 to 0.8)	1.01 (1.00 to 1.01)	1.01 (1.00 to 1.01)
4-5	10290	3188	4.6 (4.3-4.8)	54.6 (49.2 to 58.0)	0.8 (0.0 to 1.6)	1.01 (1.00 to 1.03)	1.01 (1.00 to 1.02)

Note that 5-year risk differences and risk ratios comparing any two strategies can be readily calculated from the 5-year absolute risks by subtraction or division of the absolute risks. This is not possible for the hazard ratios.

[a] Among individuals who initiate dialysis without being censored.

[b] Analyses were adjusted for all baseline and time-varying variables listed in **WebTable 3**. These are the same variables as those listed in **Table 1**, except for BMI, ACR, potassium, CRP and ferritin.

Figure 2. 5-year absolute risks and risk differences for mortality (A) and MACE (B) associated with initiating dialysis with eGFR values between 4 and 19 ml/min/1.73m² in increments of 1 ml/min/1.73m², with 6-7 ml/min/1.73m² as reference.

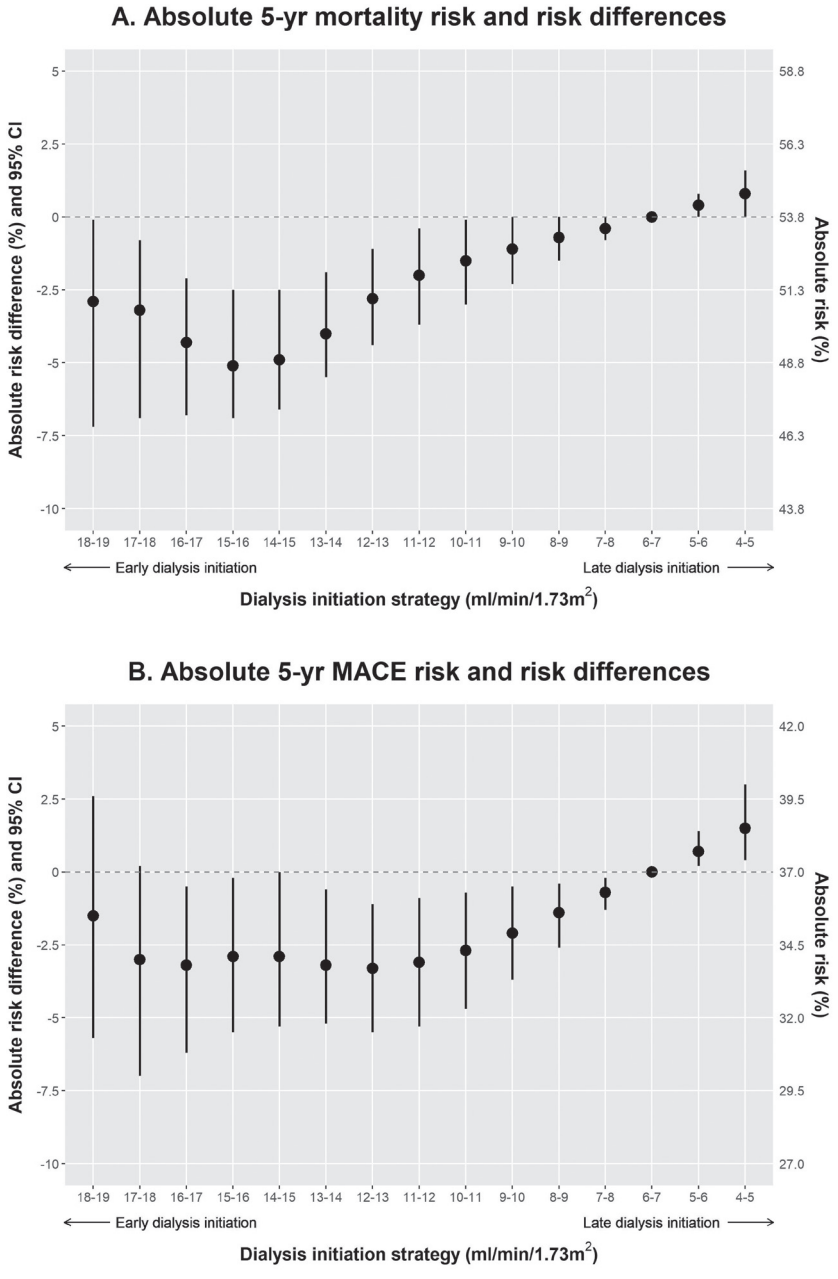
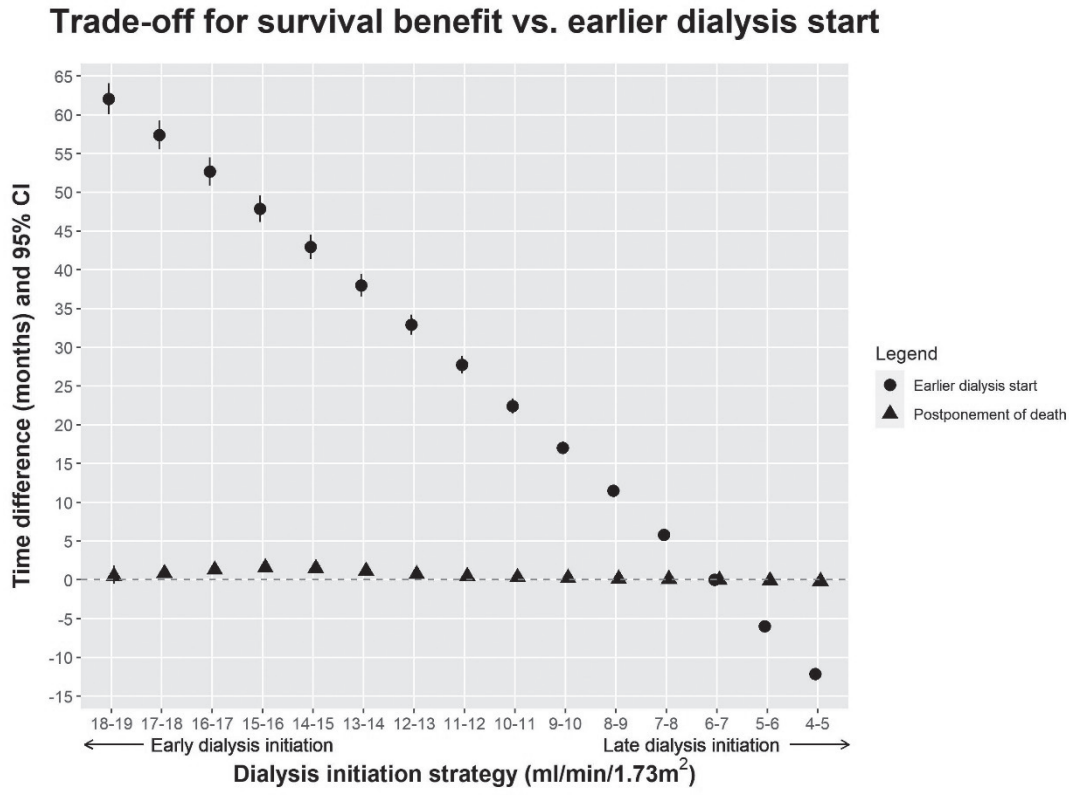


Figure 3. Trade-off between additional survival time (5-year RMST difference) and time that dialysis has to be initiated earlier, for dialysis initiation strategies with eGFR values between 4 and 19 ml/min/1.73m² in increments of 1 ml/min/1.73m², with 6-7 ml/min/1.73m² as reference. Note that a positive value indicates longer survival and an earlier dialysis start compared with the reference group. In our study population the annual eGFR decline was 2-3 ml/min/1.73m², which was estimated with a linear mixed model including linear and quadratic slope (**Supplemental Methods**). In other words, it takes ~5 months for the eGFR to decline 1 ml/min/1.73m².



For MACE the absolute risk was lowest between $eGFR_{17-18}$ and $eGFR_{11-12}$ and then progressively increased between $eGFR_{11-12}$ and $eGFR_{4-5}$ (**WebTable 7** and **Figure 2B**). Compared with $eGFR_{6-7}$, risk differences varied between an increase of 1.5% and a decrease of 3.3% (**Figure 2B**), and hazard ratios between 1.04 and 0.91, respectively. When $eGFR_{11-12}$ was taken as reference group, risk differences varied between an increase of 4.7% for $eGFR_{4-5}$ to a decrease of -0.2% for $eGFR_{12-13}$ (**WebTable 8**). The 5-year RMST differences varied between -0.3 and 0.7 months (**WebTable 5**).

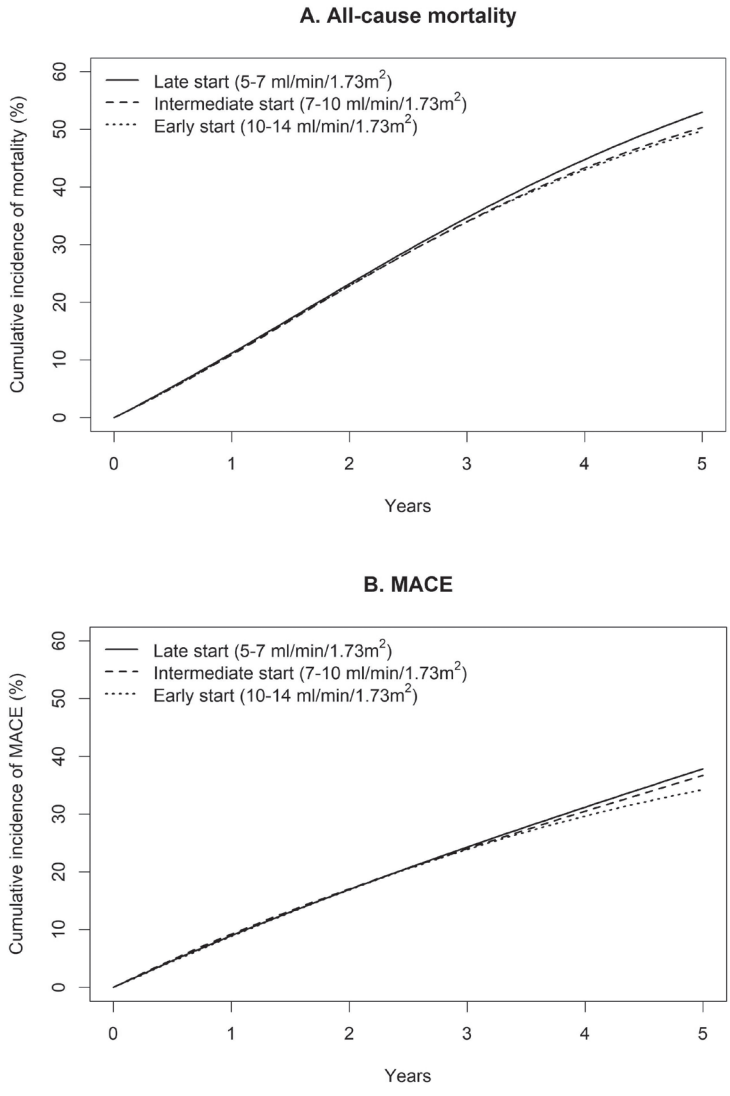
Supporting and sensitivity analyses

In our analysis mirroring the GFR thresholds from the IDEAL trial, early dialysis initiation ($eGFR_{10-14}$) was associated with a 3.3% (95% CI, 1.3% to 5.3%) lower 5-year mortality risk and 3.6% (95% CI, 1.0% to 6.0%) lower MACE risk compared with late initiation ($eGFR_{5-7}$), with hazard ratios of 0.96 (95% CI, 0.94 to 0.99) and 0.96 (95% CI, 0.93 to 1.00), respectively (**WebTable 9**, **Figure 4**). Similar results were found when comparing late versus intermediate ($eGFR_{7-10}$) dialysis initiation, in keeping with the achieved eGFR at initiation in the earlier arm of IDEAL. A lower mortality risk for early dialysis initiation was observed among all subgroups of age, sex, diabetes, eGFR, and ischemic heart disease (**WebTables 10-11**, **WebFigures 4-6**). Patients with diabetes or heart failure had a high absolute 5-year mortality and MACE risk. For instance, for the early dialysis initiation strategy the 5-year absolute mortality risk was 59.1% (95% CI, 54.9% to 65.4%) in the subgroup of patients with diabetes, and 80.5% (95% CI, 74.1% to 86.1%) in the subgroup with heart failure. Among patients with diabetes, early dialysis initiation ($eGFR_{10-14}$) was associated with a 5.4% (95% CI, 2.1% to 8.1%) lower 5-year mortality risk and 4.3% (95% CI, 0.2% to 9.1%) lower MACE risk compared with late initiation ($eGFR_{5-7}$), with hazard ratios of 0.96 (95% CI, 0.92 to 1.00) and 0.98 (95% CI, 0.93-1.04), respectively. Among patients with heart failure, early dialysis initiation was associated with a 3.3% (95% CI, -0.1% to 6.1%) lower 5-year mortality risk but no difference in MACE risk (0.3%; 95% CI, -5.2% to 5.0%) compared with late initiation, with hazard ratios of 0.95 (95% CI, 0.92 to 0.99) and 1.03 (95% CI, 0.97 to 1.08), respectively. Adjustment for confounders moved the risk difference away from the null (**WebTable 12**). As an example, the unadjusted 5-year risk difference between $eGFR_{5-7}$ and $eGFR_{10-14}$ was -0.11% and became -3.33% after full adjustment. Using untruncated weights, excluding patients with cancer, applying an alternative analytical approach for the competing risk of kidney transplantation, additionally adjusting for urinary albumin-to-creatinine ratio and potassium or censoring patients who chose conservative care did not alter our results (**WebTables 13-17**).

When we used traditional analytical approaches that introduced immortal time bias like previous observational studies (14-17) (**Supplemental Methods)**, early dialysis initiation was associated with worse outcomes, the opposite of the association we

identified in our trial emulation analysis. The hazard ratio for $eGFR_{15}$ was 1.46 (95% CI, 1.19 to 1.78) compared with $eGFR_5$ (**WebFigure 7**). In addition, when starting follow-up at dialysis initiation which introduced selection/survivor bias and lead time bias (16-31), the hazard ratio for $eGFR_{15}$ was 1.58 (95% CI, 1.37 to 1.83) compared with $eGFR_5$ (**WebFigure 8**).

Figure 4. Weighted, standardized cumulative incidence curves for mortality (A) and MACE (B) for early, intermediate and late dialysis initiation.



Discussion

In this large nationwide study of patients with advanced CKD, we estimated with novel trial emulation methodology that the maximum absolute 5-year risk reductions were 5.1% for mortality (for $eGFR_{15-16}$ vs. $eGFR_{6-7}$) and 3.3% for MACE (for $eGFR_{13-14}$ vs. $eGFR_{6-7}$). These results were robust in various sensitivity analyses and subgroups, including older patients and those with comorbidities such as diabetes, ischemic heart disease or heart failure.

Strengths and limitations of study

Strenghts of our study include its nationwide nature, large sample size, inclusion of a representative cohort of patients under routine nephrologist care, long-term follow-up and adjustment for 83 time-fixed and time-varying confounders. Furthermore, we tested the robustness of our findings in a number of supplemental analyses, and present information on absolute and relative risks, and the trade-off between restricted mean survival time and earlier dialysis start to provide a detailed picture of this issue. Our study also has limitations. First, despite adjustment for rich baseline and time-varying covariates which are used in the decision-making process (including time-varying eGFR and previous eGFR measurements), residual confounding cannot be excluded, and the precise reasons for dialysis initiation were not available in our study. Our study lacked important variables influencing this decision such as nutritional status or muscle mass stores, uremic symptoms, quality of life or physical activity. We believe however that some of these aspects were indirectly captured through adjustment for biochemical variables, hospitalizations and comorbidities. Indeed, additional adjustment for urinary albumin-to-creatinine ratio and potassium did not meaningfully alter our point estimates. Furthermore, in one of our sensitivity analyses, we sequentially adjusted for major confounder groups which are expected to induce strong confounding. However, additional adjustment resulted in at most a 1% increase in absolute risk. This, in combination with the strong probability that additional (unmeasured) confounders will be correlated with the variables we already adjusted for, reassures us that the impact from unmeasured confounders is unlikely to be large. In any case, the most compelling argument in favour of the validity of the findings is the congruence between our findings using trial emulation and those of the randomized IDEAL study. Second, the Swedish Renal Registry did not record information on symptoms or quality of life during the study period. Future studies should include symptoms in their treatment strategies and study quality of life as an outcome. Third, creatinine-based estimates of eGFR may not be an accurate reflection of true kidney function, as it may be influenced by muscle wasting or cachexia; eGFR estimated by the CKD-EPI equation is accurate within 30% to measured GFR 85% of the time (54). However, eGFR is commonly one of the factors to take into consideration by many physicians at the time of

decision-making. Lastly, as Sweden has nationwide healthcare reimbursement, and individuals in our analyses were all under nephrologist care, generalizing our results to other health systems should be done with caution.

Comparison with other studies

One randomized trial (IDEAL) and various observational studies have investigated the timing of dialysis. In a sensitivity analysis, we compared the same treatment arms as the IDEAL trial to benchmark our analytical methods (33). In IDEAL, the achieved eGFR in the early and late arms were 7.2 vs. 9.0 ml/min/1.73m^2 respectively. In our study, mean eGFR for late ($eGFR_{5-7}$) and intermediate ($eGFR_{7-10}$) start were 6.0 and 8.3 ml/min/1.73m^2, respectively. In this comparison, we observed hazard ratios of 0.97 (95% CI, 0.94 to 0.99) for mortality and 1.00 (0.97 to 1.04) for MACE. These findings are congruent with IDEAL: 1.04 (95% CI, 0.83 to 1.30) and 1.23 (95% CI 0.97 to 1.56), respectively.

Previous observational studies (14-31) investigating the timing of dialysis initiation have been criticized for the presence of immortal time, selection/survivor and lead time biases (15, 19, 51). For example, some reports found a strong protective effect of late dialysis initiation (18, 20-24, 26, 27, 29, 30), which conflicts with findings from IDEAL. In our sensitivity analyses we showed that such findings may have been attributed to either immortal time bias or selection/survivor bias. Our study design based on cloning, censoring and weighting prevents these biases by explicitly emulating a target trial, and aligning eligibility and treatment strategies at baseline. Although one previous observational study applied a similar design as ours, it did not adjust for time-varying covariates and was limited in sample size (13).

Policy implications

Our findings provide novel evidence regarding the optimal timing of dialysis initiation and show that even with maximum eGFR separations, the range of plausible effects is likely to be small. The modest increase in observed survival for initiation at higher eGFR comes at the expense of earlier dialysis initiation. Our results provide an insight into this trade-off: the maximum 5.1% absolute mortality reduction translated into a postponement of death of only 1.6 months over a 5-year follow-up period, whereas dialysis would need to be started on average 4 years earlier. For many patients this increased time on dialysis may not outweigh the modest survival benefit. Our results further suggest that in the absence of symptoms or strong indications, dialysis initiation may be postponed until lower eGFR values are reached (intent-to-defer) (40, 66), without a large increase in mortality or cardiovascular events. From a societal perspective, the elevated costs associated with earlier dialysis initiation make these strategies even less desirable. Current position papers highlight the

importance of individualized decision making in deciding whether and when to start dialysis, taking into account outcomes, quality of life and patient preferences. Our findings should not be used to suggest a single eGFR cut-off to start dialysis in all patients. Rather, our finding of similar survival across the range of eGFR where dialysis is usually considered (eGFR 5-14 ml/min/1.73m^2) should be a reassuring addition to the evidence base for clinicians: these data provide no support for any strategy other than starting dialysis based on symptoms and patient preferences, which is widespread clinical practice, recommended by guidelines, and a patient-centred approach. Our study did not address the effects of dialysis initiation versus comprehensive conservative management in patients with kidney failure. Conservative care has been proposed as a reasonable alternative to maintenance dialysis for selected older patients with comorbidities or poor functional status. Whether there are differences in survival and quality of life between dialysis and conservative management is currently unknown, and is being addressed in the ongoing randomized PREPARE for Kidney Care Study (67).

Conclusions

In conclusion, although early dialysis initiation was associated with a modest reduction in mortality and cardiovascular events, this may not outweigh the burden of a substantially longer period spent on dialysis.

References

1. G. B. D. Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2020;395(10225):709-33.
2. Liyanage T, Ninomiya T, Jha V, Neal B, Patrice HM, Okpechi I, et al. Worldwide access to treatment for end-stage kidney disease: a systematic review. Lancet. 2015;385(9981):1975-82.
3. US Renal Data System. 2020 USRDS Annual Data Report: Epidemiology of kidney disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2020.
4. Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, et al. Chronic kidney disease: global dimension and perspectives. Lancet. 2013;382(9888):260-72.
5. Mohnen SM, van Oosten MJM, Los J, Leegte MJH, Jager KJ, Hemmelder MH, et al. Healthcare costs of patients on different renal replacement modalities - Analysis of Dutch health insurance claims data. PLoS One. 2019;14(8):e0220800.
6. Mazairac AH, Blankestijn PJ, Grooteman MP, Penne EL, van der Weerd NC, den Hoedt CH, et al. The cost-utility of haemodiafiltration versus haemodialysis in the Convective Transport Study. Nephrol Dial Transplant. 2013;28(7):1865-73.
7. Eriksson JK, Neovius M, Jacobson SH, Elinder CG, Hylander B. Healthcare costs in chronic kidney disease and renal replacement therapy: a population-based cohort study in Sweden. BMJ Open. 2016;6(10):e012062.
8. Couillerot-Peyrondet AL, Sambuc C, Sainsaulieu Y, Couchoud C, Bongiovanni-Delaroziere I. A comprehensive approach to assess the costs of renal replacement therapy for end-stage renal disease in France: the importance of age, diabetes status, and clinical events. Eur J Health Econ. 2017;18(4):459-69.
9. Dabrowska-Bender M, Dykowska G, Zuk W, Milewska M, Staniszewska A. The impact on quality of life of dialysis patients with renal insufficiency. Patient Prefer Adherence. 2018;12:577-83.
10. Cohen LM, Germain MJ. Caregiver burden and hemodialysis. Clin J Am Soc Nephrol. 2014;9(5):840-2.
11. Tang SC, Ho YW, Tang AW, Cheng YY, Chiu FH, Lo WK, et al. Delaying initiation of dialysis till symptomatic uraemia--is it too late? Nephrol Dial Transplant. 2007;22(7):1926-32.
12. Evans M, Tettamanti G, Nyren O, Bellocco R, Fored CM, Elinder CG. No survival benefit from early-start dialysis in a population-based, inception cohort study of Swedish patients with chronic kidney disease. J Intern Med. 2011;269(3):289-98.
13. Crews DC, Scialla JJ, Boulware LE, Navaneethan SD, Nally JV, Jr., Liu X, et al. Comparative effectiveness of early versus conventional timing of dialysis initiation in advanced CKD. Am J Kidney Dis. 2014;63(5):806-15.
14. Traynor JP, Simpson K, Geddes CC, Deighan CJ, Fox JG. Early initiation of dialysis fails to prolong survival in patients with end-stage renal failure. J Am Soc Nephrol. 2002;13(8):2125-32.
15. Janmaat CJ, van Diepen M, Krediet RT, Hemmelder MH, Dekker FW. Effect of glomerular filtration rate at dialysis initiation on survival in patients with advanced chronic kidney disease: what is the effect of lead-time bias? Clin Epidemiol. 2017;9:217-30.

16. Susantitaphong P, Altamimi S, Ashkar M, Balk EM, Stel VS, Wright S, et al. GFR at initiation of dialysis and mortality in CKD: a meta-analysis. Am J Kidney Dis. 2012;59(6):829-40.
17. Nacak H, Bolignano D, Van Diepen M, Dekker F, Van Biesen W. Timing of start of dialysis in diabetes mellitus patients: a systematic literature review. Nephrol Dial Transplant. 2016;31(2):306-16.
18. Beddhu S, Samore MH, Roberts MS, Stoddard GJ, Ramkumar N, Pappas LM, et al. Impact of timing of initiation of dialysis on mortality. J Am Soc Nephrol. 2003;14(9):2305-12.
19. Korevaar JC, Jansen MA, Dekker FW, Jager KJ, Boeschoten EW, Krediet RT, et al. When to initiate dialysis: effect of proposed US guidelines on survival. Lancet. 2001;358(9287):1046-50.
20. Clark WF, Na Y, Rosansky SJ, Sontrop JM, Macnab JJ, Glassock RJ, et al. Association between estimated glomerular filtration rate at initiation of dialysis and mortality. CMAJ. 2011;183(1):47-53.
21. Kazmi WH, Gilbertson DT, Obrador GT, Guo H, Pereira BJ, Collins AJ, et al. Effect of comorbidity on the increased mortality associated with early initiation of dialysis. Am J Kidney Dis. 2005;46(5):887-96.
22. Stel VS, Dekker FW, Ansell D, Augustijn H, Casino FG, Collart F, et al. Residual renal function at the start of dialysis and clinical outcomes. Nephrol Dial Transplant. 2009;24(10):3175-82.
23. Lassalle M, Labeeuw M, Frimat L, Villar E, Joyeux V, Couchoud C, et al. Age and comorbidity may explain the paradoxical association of an early dialysis start with poor survival. Kidney Int. 2010;77(8):700-7.
24. Sawhney S, Djurdjev O, Simpson K, Macleod A, Levin A. Survival and dialysis initiation: comparing British Columbia and Scotland registries. Nephrol Dial Transplant. 2009;24(10):3186-92.
25. Jain AK, Sontrop JM, Perl J, Blake PG, Clark WF, Moist LM. Timing of peritoneal dialysis initiation and mortality: analysis of the Canadian Organ Replacement Registry. Am J Kidney Dis. 2014;63(5):798-805.
26. Wright S, Klausner D, Baird B, Williams ME, Steinman T, Tang H, et al. Timing of dialysis initiation and survival in ESRD. Clin J Am Soc Nephrol. 2010;5(10):1828-35.
27. Rosansky SJ, Eggers P, Jackson K, Glassock R, Clark WF. Early start of hemodialysis may be harmful. Arch Intern Med. 2011;171(5):396-403.
28. Coronel F, Cigarran S, Herrero JA. Early initiation of peritoneal dialysis in diabetic patients. Scand J Urol Nephrol. 2009;43(2):148-53.
29. Hwang SJ, Yang WC, Lin MY, Mau LW, Chen HC, Taiwan Society of N. Impact of the clinical conditions at dialysis initiation on mortality in incident haemodialysis patients: a national cohort study in Taiwan. Nephrol Dial Transplant. 2010;25(8):2616-24.
30. Shiao CC, Huang JW, Chien KL, Chuang HF, Chen YM, Wu KD. Early initiation of dialysis and late implantation of catheters adversely affect outcomes of patients on chronic peritoneal dialysis. Perit Dial Int. 2008;28(1):73-81.
31. Fink JC, Burdick RA, Kurth SJ, Blahut SA, Armistead NC, Turner MS, et al. Significance of serum creatinine values in new end-stage renal disease patients. Am J Kidney Dis. 1999;34(4):694-701.
32. Bozorgmehri S, Aboud H, Chamarthi G, Liu IC, Tezcan OB, Shukla AM, et al. Association of early initiation of dialysis with all-cause and cardiovascular mortality: A propensity score weighted analysis of the United States Renal Data System. Hemodial Int. 2021;25(2):188-97.
33. Cooper BA, Branley P, Bulfone L, Collins JF, Craig JC, Fraenkel MB, et al. A randomized, controlled trial of early versus late initiation of dialysis. N Engl J Med. 2010;363(7):609-19.

34. Kidney Disease: Improving Global Outcomes (KDIGO) 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease (CKD).

35. National Kidney Foundation. KDOQI clinical practice guideline for hemodialysis adequacy: 2015 update. Am J Kidney Dis. 2015;66(5):884-930.

36. Chan CT, Blankestijn PJ, Dember LM, Gallieni M, Harris DCH, Lok CE, et al. Dialysis initiation, modality choice, access, and prescription: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. Kidney Int. 2019;96(1):37-47.

37. Nesrallah GE, Mustafa RA, Clark WF, Bass A, Barnieh L, Hemmelgarn BR, et al. Canadian Society of Nephrology 2014 clinical practice guideline for timing the initiation of chronic dialysis. CMAJ. 2014;186(2):112-7.

38. Tattersall J, Dekker F, Heimburger O, Jager KJ, Lameire N, Lindley E, et al. When to start dialysis: updated guidance following publication of the Initiating Dialysis Early and Late (IDEAL) study. Nephrol Dial Transplant. 2011;26(7):2082-6.

39. Watanabe Y, Kawanishi H, Suzuki K, Nakai S, Tsuchida K, Tabei K, et al. Japanese society for dialysis therapy clinical guideline for "Maintenance hemodialysis: hemodialysis prescriptions". Ther Apher Dial. 2015;19 Suppl 1:67-92.

40. Tangri N, Garg AX, Ferguson TW, Dixon S, Rigatto C, Allu S, et al. Effects of a Knowledge-Translation Intervention on Early Dialysis Initiation: A Cluster Randomized Trial. J Am Soc Nephrol. 2021.

41. Mann BS, Manns BJ, Dart A, Kappel J, Molzahn A, Naimark D, et al. An assessment of dialysis provider's attitudes towards timing of dialysis initiation in Canada. Can J Kidney Health Dis. 2014;1:3.

42. Hsu CY, Parikh RV, Pravoverov LN, Zheng S, Glidden DV, Tan TC, et al. Implication of Trends in Timing of Dialysis Initiation for Incidence of End-stage Kidney Disease. JAMA Intern Med. 2020;180(12):1647-54.

43. 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.

44. Fu EL, Clase CM, Evans M, Lindholm B, Rotmans JI, Dekker FW, et al. Comparative Effectiveness of Renin-Angiotensin System Inhibitors and Calcium Channel Blockers in Individuals With Advanced CKD: A Nationwide Observational Cohort Study. Am J Kidney Dis. 2020.

45. Evans M, Carrero JJ, Bellocco R, Barany P, Qureshi AR, Seeberger A, et al. Initiation of erythropoiesis-stimulating agents and outcomes: a nationwide observational cohort study in anaemic chronic kidney disease patients. Nephrol Dial Transplant. 2017;32(11):1892-901.

46. Swedish Renal Registry: Annual report 2018. www.snronline.se (2 April 2019, date last accessed).

47. Wettermark B, Hammar N, Fored CM, Leimanis A, Otterblad Olausson P, Bergman U, et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. Pharmacoepidemiol Drug Saf. 2007;16(7):726-35.

48. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.

49. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. Eur J Epidemiol. 2017;32(9):765-73.

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. Sjolander A, Nyren O, Bellocco R, Evans M. Comparing different strategies for timing of dialysis initiation through inverse probability weighting. Am J Epidemiol. 2011;174(10):1204-10.

52. Fu EL, van Diepen M, Xu Y, Trevisan M, Dekker FW, Zoccali C, et al. Pharmacoepidemiology for nephrologists (part 2): potential biases and how to overcome them. Clinical Kidney Journal. 2020.

53. 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.

54. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12.

55. Hernan MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. Basic Clin Pharmacol Toxicol. 2006;98(3):237-42.

56. Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernan MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. Int J Biostat. 2010;6(2):Article 18.

57. Maringe C, Benitez Majano S, Exarchakou A, Smith M, Rachet B, Belot A, et al. Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. Int J Epidemiol. 2020;49(5):1719-29.

58. Fu EL, Groenwold RHH, Zoccali C, Jager KJ, van Diepen M, Dekker FW. Merits and caveats of propensity scores to adjust for confounding. Nephrol Dial Transplant. 2019;34(10):1629-35.

59. Fu EL, Evans M, Clase CM, Tomlinson LA, van Diepen M, Dekker FW, et al. Stopping Renin-Angiotensin System Inhibitors in Patients with Advanced CKD and Risk of Adverse Outcomes: A Nationwide Study. J Am Soc Nephrol. 2021;32(2):424-35.

60. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168(6):656-64.

61. Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. Comput Methods Programs Biomed. 2004;75(1):45-9.

62. Buchanan AL, Hudgens MG, Cole SR, Lau B, Adimora AA, Women's Interagency HIVS. Worth the weight: using inverse probability weighted Cox models in AIDS research. AIDS Res Hum Retroviruses. 2014;30(12):1170-7.

63. Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernan MA. A causal framework for classical statistical estimands in failure-time settings with competing events. Stat Med. 2020;39(8):1199-236.

64. van Geloven N, le Cessie S, Dekker FW, Putter H. Transplant as a competing risk in the analysis of dialysis patients. Nephrol Dial Transplant. 2017;32(suppl_2):ii53-ii9.

65. Kloecker DE, Davies MJ, Khunti K, Zaccardi F. Uses and Limitations of the Restricted Mean Survival Time: Illustrative Examples From Cardiovascular Outcomes and Mortality Trials in Type 2 Diabetes. Ann Intern Med. 2020;172(8):541-52.

66. Chau EM, Manns BJ, Garg AX, Sood MM, Kim SJ, Naimark D, et al. Knowledge Translation Interventions to Improve the Timing of Dialysis Initiation: Protocol for a Cluster Randomized Trial. Can J Kidney Health Dis. 2016;3:2054358116665257.

67. 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. 2021;36(6):975-82.

Supplemental material

Supplemental Methods. Target trial emulation using cloning, censoring and weighting.

Supplemental Methods. Why common methods introduce immortal time bias, lead-time bias or selection bias.

Supplemental Methods. A fourth bias: confounding.

Supplemental Methods. Calculation of postponement of dialysis using a linear mixed-effects regression model.

WebTable 1. Brief protocol of the pragmatic target trial and its emulation using data from the Swedish Renal Registry 2007-2017.

WebTable 2. Definition of study outcomes and covariates.

WebTable 3. Model coefficients for remaining uncensored for the $eGFR_{6-7}$ strategy.

WebTable 4. 5-year absolute risks, risks differences, risk ratios and hazard ratios for all-cause mortality associated with initiating dialysis with eGFR values between 4 and 19 ml/min/1.73m^2 in increments of 1 ml/min/1.73m^2 with 11-12 ml/min/1.73m^2 as reference (mean eGFR at dialysis initiation in United States).

WebTable 5. 5-year restricted mean survival time and 5-year restricted mean survival time differences for different dialysis initiation strategies.

WebTable 6. Difference in time until start of dialysis for various dialysis initiation strategies, compared with the $eGFR_{6-7}$ strategy. A positive number denotes an earlier dialysis start.

WebTable 7. 5-year absolute risks, risks differences, risk ratios and hazard ratios for MACE associated with initiating dialysis with eGFR values between 4 and 19 ml/min/1.73m^2 in increments of 1 ml/min/1.73m^2 with 6-7 ml/min/1.73m^2 as reference.

WebTable 8. 5-year absolute risks, risks differences, risk ratios and hazard ratios for MACE associated with initiating dialysis with eGFR values between 4 and 19 ml/min/1.73m^2 in increments of 1 ml/min/1.73m^2 with 11-12 ml/min/1.73m^2 as reference (mean eGFR at dialysis initiation in United States).

WebTable 9. 5-year absolute risks, risks differences, risk ratios and hazard ratios for all-cause mortality and MACE associated with late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation.

WebTable 10. 5-year absolute risks, risks differences, risk ratios and hazard ratios for all-cause mortality associated with late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation, stratified by subgroups of age (≥70 vs. < 70 years), sex, presence of diabetes, eGFR at baseline (10-15 vs. 15-20 ml/min/1.73m^2), presence of ischemic heart disease, and presence of heart failure.

WebTable 11. 5-year absolute risks, risks differences, risk ratios and hazard ratios for MACE associated with late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation, stratified by subgroups of age (≥70 vs. < 70 years), sex, presence of diabetes, eGFR at baseline (10-15 vs. 15-20 ml/min/1.73m^2), presence of ischemic heart disease, and presence of heart failure.

WebTable 12. Point estimates for the 5-year mortality risk differences between intermediate vs. late and early vs. late dialysis initiation with different levels of confounding adjustment.

WebTable 13. Influence of weight truncation on the point estimates of 5-year absolute risks, risks differences, risk ratios and hazard ratios for all-cause mortality and MACE for late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation.

WebTable 14. Sensitivity analysis. Influence of excluding individuals with cancer at baseline (N = 1025) on the point estimates of 5-year absolute risks, risks differences, risk ratios and hazard ratios for all-cause mortality and MACE for late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation.

WebTable 15. Sensitivity analysis. Influence of additional adjustment for urinary albumin-to-creatinine ratio and potassium in the subset of individuals with both measurements available (N = 4286) on the point estimates of 5-year absolute risks, risks differences, risk ratios and hazard ratios for all-cause mortality and MACE for late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation.

WebTable 16. Sensitivity analysis. Modelling the direct effect of late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation not mediated through kidney transplantation on all-cause mortality and MACE.

WebTable 17. Sensitivity analysis. Modelling the direct effect of late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation not mediated through conservative care on all-cause mortality and MACE.

WebFigure 1. Schematic representation of cloning, censoring and weighting algorithm for late, intermediate and early dialysis initiation.

WebFigure 2. Selection of study participants and follow-up. For simplicity, the three treatment strategies as used in the sensitivity analysis are depicted.

WebFigure 3. Histogram of eGFR at dialysis initiation in Sweden during 2007-2017.

WebFigure 4. Weighted, standardized cumulative incidence curves for mortality for late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation, stratified by subgroups of age (A), sex (B), diabetes (C), eGFR (D), ischemic heart disease (E) and heart failure (F).

WebFigure 5. Weighted, standardized cumulative incidence curves for MACE for late (eGFR 5-7 ml/min/1.73m^2), intermediate (eGFR 7-10 ml/min/1.73m^2) and early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation stratified by subgroups of age (A), sex (B), diabetes (C), eGFR (D), ischemic heart disease (E) and heart failure (F).

WebFigure 6. Forest plot depicting 5-year absolute risk differences for late (eGFR 5-7 ml/min/1.73m^2) and intermediate (eGFR 7-10 ml/min/1.73m^2) initiation compared with early (eGFR 10-14 ml/min/1.73m^2) dialysis initiation on mortality (A) and MACE (B) stratified by subgroups of age, sex, diabetes, eGFR, ischemic heart disease and heart failure.

WebFigure 7. Deliberate introduction of immortal time bias to illustrate why previous observational studies found a protective effect of late dialysis initiation on all-cause mortality. eGFR was modelled as a continuous variable using a penalized spline. The reference was set at an eGFR of 5 ml/min/1.73m^2. Dotted lines represent 95% confidence intervals.

WebFigure 8. Deliberate introduction of selection/survivor bias and lead time bias to illustrate why previous observational studies found a protective effect of late dialysis initiation on all-cause mortality. eGFR was modelled as a continuous variable using a penalized spline. The reference was set at an eGFR of 5 ml/min/1.73m^2. Dotted lines represent 95% confidence intervals.

Supplemental Methods

Target trial emulation using cloning, censoring and weighting

Here we describe in detail our implementation of target trial emulation and the cloning, censoring and weighting procedure. A thorough review of trial emulation can be found elsewhere (1, 2), as well as recent applications of the methodology (3-10).

Specifying details of the target trial

The goal of many observational studies is to compare the effects of two or more treatment strategies on a clinical outcome. A simple way to structure the study design and analysis of such a study is to use the target trial framework (1, 2). This means that we think about the hypothetical randomized trial we would like to conduct and then use our observational data to explicitly emulate it. Explicitly emulating a randomized trial can prevent unnecessary biases such as immortal time bias, selection/survivor bias and lead time bias (11-15), as well as making results from observational analyses more comparable to those from trials (16). Similar to a real trial, we first need to formally define the eligibility criteria of our hypothetical trial, the treatment strategies we would like to compare, how treatment is assigned to each individual, the duration of follow-up, the primary and secondary endpoints, the causal contrast of interest (intention-to-treat or per protocol effect), and the statistical analysis. Details of the target trial we wanted to emulate in our analysis are given in **WebTable 1**.

In our study we were interested in comparing 15 dialysis initiation strategies , with eGFR values ranging between 4-5 ml/min/1.73m^2 and 18-19 ml/min/1.73m^2. Note that it would be difficult to compare 15 strategies in a real randomized controlled trial, as this would require an extremely large sample size. The IDEAL trial required 8 years to include 828 individuals. We therefore need to rely partly on well-conducted observational studies to identify the optimal eGFR to start dialysis. We applied the same methodology when comparing three treatment strategies in our sensitivity analysis. For ease of explanation we will therefore explain the methods according to three strategies only.

Treatment strategies such as those defined above depend on the value of a time-varying individual characteristic (in this case eGFR) and are therefore called dynamic treatment strategies (5, 17). Such dynamic treatment strategies answer the question "*When* should I start a particular treatment?". Comparing the effects of dynamic treatment strategies in observational data requires methods that can appropriately adjust for time-varying confounding, such as the parametric G-formula (18, 19) or cloning, censoring and weighting (1, 14, 20). We now explain

in detail our implementation of the latter approach for three dialysis initiation strategies. A graphical depiction of the cloning, censoring and weighting procedure can be found in **WebFigure 1**.

Rationale for the cloning, censoring and weighting method

The rationale for using the cloning, censoring and weighting method is that at baseline, an individual's data is consistent with multiple strategies. For instance, an individual with an eGFR of 16 ml/min/1.73m^2 at baseline has data consistent with early (starting dialysis with an $eGFR_{10-14}$), intermediate ($eGFR_{7-10}$) or late ($eGFR_{10-14}$) dialysis initiation. This individual could be randomly assigned to one of the three strategies, similar to a real randomized trial. However, it is more statistically efficient to allocate this individual to all treatment strategies with which his/her data are consistent.

Step 1: Cloning and assigning replicates to the treatment strategies

The first step consists of cloning each individual into three identical replicates, each of whom is assigned to one strategy (either late, intermediate or early dialysis initiation). The dataset will now be three times as large as the original dataset. Since each individual occurs in all strategies, the three treatment groups will be identical in all characteristics and hence no baseline confounding is present.

Note that for the comparison of 15 dialysis initiation strategies, 15 identical replicates of each individual need to be made. At baseline some of the replicates will already have passed their assigned eGFR value to start dialysis. For example, an individual with a baseline eGFR of 13 ml/min/1.73m^2 can never comply with the strategy "initiate dialysis with an eGFR between 16-17 ml/min/1.73m^2". Such replicates that do not comply with their assigned strategy at baseline are removed from the dataset.

Step 2: Censoring replicates if and when they do not adhere to their assigned strategy

Note that there are now replicates included that do not necessarily always adhere to their assigned strategy during follow-up. To estimate the effect of a particular treatment strategy, we need to censor replicates if and when their observed treatment does not match their assigned strategy anymore.

In our dataset, we therefore determined at each month whether a replicate was adherent to their assigned strategy and artificially censored them if they stopped adhering. As an example, consider the three hypothetical persons in the **Appendix Table** on the next page. Three replicates of each person are present in the dataset (cloning step), and each replicate is assigned to a different treatment strategy (late, intermediate or early dialysis initiation).

Replicate 1.1 is assigned to the strategy "initiate dialysis with an eGFR between 5-7 ml/min/1.73m^2" (i.e. late dialysis initiation). Since his eGFR has dropped to 4.2 ml/min/1.73m^2 in month 5 and this individual has not initiated dialysis yet, he will be censored in month 5. Replicate 1.2 is assigned to the strategy "initiate dialysis with an eGFR between 7-10 ml/min/1.73m^2" (i.e. intermediate dialysis initiation). Since his eGFR has dropped to 6.5 ml/min/1.73m^2 in month 3 and this individual has not initiated dialysis yet, he will be censored in month 3. Lastly, replicate 1.3 is assigned to the strategy "initiate dialysis with an eGFR between 10-14 ml/min/1.73m^2". Since his eGFR has dropped to 7.3 ml/min/1.73m^2 in month 2 and this individual has not initiated dialysis yet, he will be censored in month 2. The first individual died in month 5. However, this death will count for none of the treatment strategies since all replicates are censored before the death is observed. Note that the three replicates represent the same person (individual 1), and that we use data from individual 1 to estimate the effect of each strategy as long as he adheres to his assigned strategy.

Person 2 is like person 1, except that dialysis is initiated in month 3 of follow-up at an eGFR of 6.5 ml/min/1.73m^2. Replicate 2.1 adheres to his assigned treatment strategy and is therefore never censored during follow-up. Replicate 2.2 and replicate 2.3 are censored in month 3 and 2, respectively, since they do not adhere to their assigned strategy anymore in those months. Note that the death is observed only for replicate 2.1 and not for replicates 2.2 or 2.3.

Person 3 dies in the first month while his eGFR was 12.0 ml/min/1.73m^2. The death will count for all three treatment strategies because the data were consistent with all strategies when it developed.

Appendix Table. Three hypothetical persons whose data are consistent with multiple dialysis initiation strategies.

Person	Replicate	Assigned strategy	Month	eGFR	Dialysis	Death	Artificial censoring
1	1.1	5-7	1	12.0	0	0	0
1	1.1	5-7	2	7.3	0	0	0
1	1.1	5-7	3	6.5	0	0	0
1	1.1	5-7	4	5.8	0	0	0
1	1.1	5-7	5	4.2	0	1	1
1	1.2	7-10	1	12.0	0	0	0
1	1.2	7-10	2	7.3	0	0	0
1	1.2	7-10	3	6.5	0	0	1
1	1.2	7-10	4	5.8	0	0	1
1	1.2	7-10	5	4.2	0	1	1
1	1.3	10-14	1	12.0	0	0	0
1	1.3	10-14	2	7.3	0	0	1
1	1.3	10-14	3	6.5	0	0	1
1	1.3	10-14	4	5.8	0	0	1
1	1.3	10-14	5	4.2	0	1	1
2	2.1	5-7	1	12.0	0	0	0
2	2.1	5-7	2	7.3	0	0	0
2	2.1	5-7	3	6.5	1	0	0
2	2.1	5-7	4	5.8	1	0	0
2	2.1	5-7	5	4.2	1	1	0
2	2.2	7-10	1	12.0	0	0	0
2	2.2	7-10	2	7.3	0	0	0
2	2.2	7-10	3	6.5	1	0	1
2	2.2	7-10	4	5.8	1	0	1
2	2.2	7-10	5	4.2	1	1	1
2	2.3	10-14	1	12.0	0	0	0
2	2.3	10-14	2	7.3	0	0	1
2	2.3	10-14	3	6.5	1	0	1
2	2.3	10-14	4	5.8	1	0	1
2	2.3	10-14	5	4.2	1	1	1
3	3.1	5-7	1	12.0	0	1	0
3	3.2	7-10	1	12.0	0	1	0
3	3.3	10-14	1	12.0	0	1	0

Step 3: Inverse probability weighting to adjust for informative censoring

Because the artificial censoring of replicates is likely to be informative, this will lead to selection bias (also called collider stratification bias in the epidemiology literature). We therefore need to use inverse probability weighting to adjust for this selection bias, which is the most involved step of the cloning, censoring and weighting procedure. In brief, uncensored replicates receive a weight that is equal to the inverse of the probability of remaining uncensored, conditional on their own covariate history. Intuitively, the weighting will upweight uncensored replicates who have similar characteristics as censored replicates (see also **WebFigure 1**). This creates a pseudopopulation in which censoring does not depend on measured characteristics and is no longer informative (21).

To estimate the inverse probability of censoring weights, we first fit a pooled logistic model with "being uncensored" as the outcome and as independent variables an indicator for time (a restricted cubic spline with prespecified knots at months 3, 7, 12, 23 and 35), baseline and time-varying confounders. We fit a pooled logistic model for each arm separately since the censoring pattern is likely to be different for each treatment strategy, and to allow for treatment-covariate interaction (2, 4). The knots for time were based on visual inspection of the censoring pattern during follow-up.

Next, we used the probabilities estimated by these models to construct the inverse probability of censoring weights. Weights were set to 1 after a replicate initiated dialysis, as their probability to remain uncensored is per definition 1. We truncated the weights at the 99.95th percentile to avoid undue influence of very large weights. Truncating the weights is a trade-off between bias and precision: truncation of large weights will lead to narrower confidence intervals at the expense of introducing some bias. The median of the truncated weights was 1.02, the mean 1.17 and the maximum 31.1. Using untruncated weights showed virtually similar results and therefore indicated that no substantial bias was introduced by truncation (**WebTable 13**).

Step 4: Primary analysis

Next, we stacked the three datasets (late, intermediate and early dialysis initiation). We used a weighted pooled logistic model to estimate the per protocol effect of late, intermediate and early dialysis initiation. The pooled logistic model contained indicators for time (month and month squared), an indicator for treatment strategy, interactions between time and treatment strategy (to allow for nonproportional hazards) and all baseline covariates, as well as the weights estimated in step 3. Treatment strategy was modelled as a factor for 3 strategies and as a restricted cubic spline with knots at 5, 8, 11, 14 and 17 for 15 strategies. This pooled logistic model was used to calculate weighted cumulative incidence curves. The weighted curves were then standardized to the baseline distribution of confounders and used to

calculate 5-year absolute risk differences and differences in restricted mean survival time. To account for the weighting and cloning procedure, we used nonparametric bootstrapping based on 500 samples to obtain valid percentile confidence intervals. From the survival curves we estimated the average hazard ratio at each month during follow-up as $\log(\text{Survival}_2)/\log(\text{Surival}_1)$. To obtain one summary hazard ratio we averaged the hazard ratio over the whole study period (22).

Why common methods introduce immortal time bias, lead-time bias or selection bias

A number of observational studies have tried to estimate the effects of dialysis timing on outcomes. Most of these studies used two methods, denoted by Sjölander et al. as the "*from initiation*" method or the "*from threshold*" method (13). Both methods introduce various biases, including lead time bias, survival/selection bias and immortal time bias.

In the from initiation method, baseline is defined as the time of dialysis initiation (**Appendix Figure 1**). All patients are included at the moment of dialysis start and eGFR levels are then compared on outcomes such as mortality. Note that the choice if baseline in the from initiation method is wrong: in a randomized trial (such as the IDEAL trial) individuals are included before dialysis. The from initiation method introduces two biases: lead time bias and survivor/selection bias. The lead time bias occurs because patients with a higher eGFR at dialysis initiation will be earlier in the course of their disease progression than individuals with a lower eGFR. This will give early starters an artificial survival advantage. It is similar to the lead time bias in observational studies investigating cancer screening. The screened group will be diagnosed with cancer earlier, and hence follow-up for this group starts earlier in the course of their disease. However, in reality patients in the screened group may not live longer than those in the non-screened group: only the diagnosis of cancer is moved earlier in time.

The second bias that is introduced by the from initiation method is selection/survival bias, also known in the epidemiology literature as collider stratification bias. This bias gives an artificial survival advantage to the late starters. Why this bias arises can be understood intuitively. Patients with a low eGFR who are included in an observational cohort must have survived long enough until sampling. As eGFR is a strong risk factor for mortality, patients who do not have other risk factors for mortality (such as diabetes) are more likely to survive until a low eGFR. After all, if the patients with a low eGFR would have had multiple other risk factors for mortality, they most likely would not have survived until sampling into the cohort. The bias can be graphically shown in a causal diagram (**Appendix Figure 2**). Conditioning

on surviving until a low eGFR value (denoted by the selection node S), induces an inverse association between eGFR and other risk factors (denoted by U). In technical terms, the conditioning on the collider S opens a backdoor path, thereby introducing collider stratification bias. To properly adjust for the selection bias, one would need to adjust for all risk factors for mortality. Failure to do so (which is very likely) will lead to biased effect estimates, e.g. if one has not measured all risk factors for mortality. It should be noted that this selection bias is distinct from confounding. Confounders are variables which influence both eGFR at dialysis initiation and mortality. Adjusting for confounders only will not be sufficient to adjust for the selection bias.

It seems that the effect of the selection/survival bias is stronger than the effect of lead time bias, since most observational studies have found a harmful effect of early dialysis initiation rather than a protective effect (23-33). When reanalyzing our data using the from initiation method, we also obtained an effect estimate favoring the late starters, with a hazard ratio of 1.58 (95% CI, 1.37 to 1.83) for $eGFR_{15}$ versus $eGFR_{5}$ (**WebFigure 8**). Even though we adjusted for a large number of confounders (similar to previous observational studies), this suggests that we – like the other observational studies – were not able to correct for all selection bias introduced by the from initiation method, since our main analysis found a completely opposite effect: a modest protective effect of early initiation.

In an attempt to mitigate lead time bias, some researchers have started follow-up at a common point in time, e.g. when eGFR drops below 20 ml/min/1.73m^2 for the first time. This method has been referred to as the from threshold method, because follow-up starts when a certain threshold is passed. However, by doing so, immortal time bias can be introduced. The problem is that at baseline it is not yet known at which eGFR dialysis will be initiated. At baseline all patients will have an eGFR around 20 ml/min/1.73m^2, and dialysis has not started yet at that moment. To overcome this problem, some researchers have classified patients into exposure groups by using future information that is not available at baseline. Whenever future information is used to classify patients into exposure groups, immortal time is introduced. All patients need to survive until dialysis start, otherwise they cannot be classified. Therefore, all included individuals will be immortal until the start of dialysis. The immortal time will be longer for individuals with a low eGFR than for those with a high eGFR, and therefore favors late dialysis initiation (**Appendix Figure 3**). When reanalyzing our data using the from threshold method and introducing immortal time bias, we obtained again an effect estimate favoring the late starters, with a hazard ratio of 1.46 (95% CI, 1.19 to 1.78) for $eGFR_{15}$ versus $eGFR_{5}$.

Both methods described above do not explicitly emulate a clinical trial. In a randomized trial we would follow patients from a common starting point (e.g. eGFR between 10-20 ml/min/1.73m^2) and randomize them at that moment to treatment groups. Therefore, the moment of start of follow-up and the assignment of treatment strategies coincide at baseline. The from initiation method does not adhere to this important principle since it starts follow-up at dialysis initiation. The from threshold method as applied by previous researchers (34-37) also does not properly emulate a randomized trial since the start of follow-up happens before the assignment of treatment strategies. The cloning step forces the alignment of the start of follow-up and assignment of treatment strategies and thereby automatically eliminates immortal time bias, lead time bias and selection/survivor bias. This cloning, censoring and weighting approach was used in an earlier analysis by Crews *et al* (38). However, their analysis was limited by a small sample size and by the fact that analyses were not adjusted for time-varying confounders. **WebTable 13** in which we sequentially adjusted for more baseline and time-varying confounders shows the importance of adjusting for time-varying confounders when applying this analytical method.

Another recent study compared dialysis initiation versus no initiation stratified by eGFR levels (39). However, this analytical approach does not answer the question *when* to initiate dialysis. Rather, it compares the effectiveness of dialysis vs. no dialysis for various levels of eGFR (i.e.: "given that my patient has survived until an eGFR of *x* ml/min/1.73m^2, what is the effect of dialysis vs. no dialysis on mortality?"). The authors found that dialysis initiation compared with no initiation was associated with an adjusted HR of 0.28 (95% CI, 0.16 to 0.45) in individuals with an eGFR <6 ml/min/1.73m^2. The hazard ratio was 0.41 for $eGFR_{6-9}$, 0.83 for $eGFR_{9-12}$, 0.88 for $eGFR_{12-15}$, 1.50 for $eGFR_{15-29}$ and 3.70 for eGFR >29 ml/min/1.73m^2. Looking at these numbers, it is tempting to compare the different hazard ratios and conclude that initiating dialysis at an eGFR <6 is associated with the best survival (since it has the lowest hazard ratio). However, we cannot compare the different hazard ratios with each other, since each hazard ratio is calculated conditional on surviving until a certain eGFR level. Therefore, the patients that contribute to the eGFR <6 analysis are only a subset of the patients that contribute to the analysis of dialysis effectiveness in individuals with an eGFR between 12-15. Naturally, the authors found that the effectiveness of dialysis was stronger in individuals with a low eGFR. These results tell you that if you do not initiate dialysis when you reach an eGFR of 6, you will die quickly. It does not tell you that initiating at an eGFR <6 is better than initiating at an eGFR between 9 and 12.

Lastly, Scialla *et al*. elegantly applied an instrumental variables approach using geographic variation as an instrument (40). Similar to conventional observational analyses, instrumental variable analyses also rely on untestable assumptions

which are difficult to verify, e.g. that there are no confounders for the instrument-outcome relationship (exogeneity assumption) and that the instrument influences the outcome mortality only through eGFR levels at dialysis initiation (exclusion restriction assumption). Secondly, it is difficult to interpret the effect estimate obtained from an instrumental variables analysis. Under additional assumptions (e.g. the monotonicity assumption) the effect estimate can be interpreted as the average causal effect of treatment in the subpopulation of compliers. However, it is not possible to identify this subpopulation of compliers, which makes it difficult to apply these findings for decision making. A more detailed discussion of merits and caveats of instrumental variable analysis can be found elsewhere (41-43).

A fourth bias: confounding

All observational studies are limited by confounding. However, published results show that confounding may not be the biggest problem in observational analyses. Rather, the preventable biases explained in the previous section are an important reason why observational analyses and randomized trials have led to different conclusions, e.g. in the case of statins and decreased cancer risk, the effect of hormonal replacement treatment on cardiovascular events in postmenopausal women, or the effect of timing of dialysis on outcomes (12, 15, 44). There are a number of recent analyses showing that properly conducted observational studies, in particular those explicitly emulating a trial, can in certain situations obtain similar estimates as randomized trials (e.g. if we apply similar inclusion/exclusion criteria, have enough data to emulate the treatment strategies, etc.) (16, 45). When data from randomized trials are available, it can help to compare the results obtained from the observational analysis with those from the trial. If these results align, this can add further validity to the methods and data used in the observational approach.

To avoid confounding as much as possible, we adjusted for a wide range of baseline and time-varying covariates, including demographic variables, laboratory measurements, medication use, medical history, and prior hospitalizations, many of which are used in the decision-making process to start dialysis. Sequential adjustment for confounding can also give an indication how large confounding bias is likely to be, and whether any additional adjustment would significantly affect the point estimate. This seemed not to be the case. Additional adjustment for urinary albumin-to-creatinine ratio and potassium measurements also did not suggest major residual confounding bias.

Calculation of postponement of dialysis using a linear mixed-effects regression model.

To estimate the time from baseline until start of dialysis for various dialysis initiation strategies, we first fit the a linear mixed-effects regression model that describes the eGFR decline of the population over time. This model estimates the coefficients β, b and ε, as previously described by Crews et al. (38):

$$eGFR_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 (t_{ij})^2 + b_{0i} + b_{1i} t_{ij} + b_{2i} (t_{ij})^2 + \varepsilon_{ij} \text{ (eq. 1)}$$

where persons i = 1, …, n have eGFR measurements at occasions j=1, …, m_i and t_{ij} = time in years after baseline. β terms represent fixed effects describing the population-average eGFR decline over time, b terms are random effects describing the patient-specific deviation from this population average, and ε terms represent the patient- and occasion-specific residuals. All eGFR measurements until the start of dialysis were used for the estimation of this model.

Using the coefficients of the fitted model, we solved the quadratic equation for *t* to obtain time until dialysis for various eGFR levels.

$$\hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2 = eGFR \text{ (eq. 2)}$$

where $\hat{\beta}_0 = 21.0$, $\hat{\beta}_1 = -2.7$ + $\hat{\beta}_2 = 0.05$.

Next, dialysis times were subtracted from the reference value of dialysis initiation at $eGFR_{6-7}$. To obtain 95% confidence intervals around these differences, parametric bootstrapping based on 10.000 samples was used.

Supplemental References

1. 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):758-764.
2. Maringe C, Benitez Majano S, Exarchakou A, et al. Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. *Int J Epidemiol.* 2020;49(5):1719-1729.
3. Huitfeldt A, Kalager M, Robins JM, Hoff G, Hernan MA. Methods to Estimate the Comparative Effectiveness of Clinical Strategies that Administer the Same Intervention at Different Times. *Curr Epidemiol Rep.* 2015;2(3):149-161.
4. Fu EL, Evans M, Clase CM, et al. Stopping Renin-Angiotensin System Inhibitors in Patients with Advanced CKD and Risk of Adverse Outcomes: A Nationwide Study. *J Am Soc Nephrol.* 2021;32(2):424-435.
5. Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernan MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. *Int J Biostat.* 2010;6(2):Article 18.
6. Danaei G, Garcia Rodriguez LA, Cantero OF, Logan RW, Hernan MA. Electronic medical records can be used to emulate target trials of sustained treatment strategies. *J Clin Epidemiol.* 2018;96:12-22.
7. Garcia-Albeniz X, Hernan MA, Logan RW, Price M, Armstrong K, Hsu J. Continuation of Annual Screening Mammography and Breast Cancer Mortality in Women Older Than 70 Years. *Ann Intern Med.* 2020;172(6):381-389.
8. Keyhani S, Cheng EM, Hoggatt KJ, et al. Comparative Effectiveness of Carotid Endarterectomy vs Initial Medical Therapy in Patients With Asymptomatic Carotid Stenosis. *JAMA Neurol.* 2020;77(9):1110-1121.
9. Lyu H, Yoshida K, Zhao SS, et al. Delayed Denosumab Injections and Fracture Risk Among Patients With Osteoporosis : A Population-Based Cohort Study. *Ann Intern Med.* 2020;173(7):516-526.
10. Petito LC, Garcia-Albeniz X, Logan RW, et al. Estimates of Overall Survival in Patients With Cancer Receiving Different Treatment Regimens: Emulating Hypothetical Target Trials in the Surveillance, Epidemiology, and End Results (SEER)-Medicare Linked Database. *JAMA Netw Open.* 2020;3(3):e200452.
11. Dickerman BA, Garcia-Albeniz X, Logan RW, Denaxas S, Hernan MA. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med.* 2019;25(10):1601-1606.
12. 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;79:70-75.
13. Sjolander A, Nyren O, Bellocco R, Evans M. Comparing different strategies for timing of dialysis initiation through inverse probability weighting. *Am J Epidemiol.* 2011;174(10):1204-1210.
14. Fu EL, van Diepen M, Xu Y, et al. Pharmacoepidemiology for nephrologists (part 2): potential biases and how to overcome them. *Clinical Kidney Journal.* 2020.
15. 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.

16. Lodi S, Phillips A, Lundgren J, et al. Effect Estimates in Randomized Trials and Observational Studies: Comparing Apples With Apples. *Am J Epidemiol.* 2019;188(8):1569-1577.
17. Hernan MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol.* 2006;98(3):237-242.
18. Wen L, Young JG, Robins JM, Hernan MA. Parametric g-formula implementations for causal survival analyses. *Biometrics.* 2020.
19. Westreich D, Cole SR, Young JG, et al. The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident AIDS or death. *Stat Med.* 2012;31(18):2000-2009.
20. Hernan MA. How to estimate the effect of treatment duration on survival outcomes using observational data. *BMJ.* 2018;360:k182.
21. Fu EL, Groenwold RHH, Zoccali C, Jager KJ, van Diepen M, Dekker FW. Merits and caveats of propensity scores to adjust for confounding. *Nephrol Dial Transplant.* 2019;34(10):1629-1635.
22. Murray EJ, Caniglia EC, Petito LC. Causal survival analysis: A guide to estimating intention-to-treat and per-protocol effects from randomized clinical trials with non-adherence. *Research Methods in Medicine & Health Sciences.* 2020;2(1):39-49.
23. Kazmi WH, Gilbertson DT, Obrador GT, et al. Effect of comorbidity on the increased mortality associated with early initiation of dialysis. *Am J Kidney Dis.* 2005;46(5):887-896.
24. Beddhu S, Samore MH, Roberts MS, et al. Impact of timing of initiation of dialysis on mortality. *J Am Soc Nephrol.* 2003;14(9):2305-2312.
25. Clark WF, Na Y, Rosansky SJ, et al. Association between estimated glomerular filtration rate at initiation of dialysis and mortality. *CMAJ.* 2011;183(1):47-53.
26. Hwang SJ, Yang WC, Lin MY, Mau LW, Chen HC, Taiwan Society of N. Impact of the clinical conditions at dialysis initiation on mortality in incident haemodialysis patients: a national cohort study in Taiwan. *Nephrol Dial Transplant.* 2010;25(8):2616-2624.
27. Lassalle M, Labeeuw M, Frimat L, et al. Age and comorbidity may explain the paradoxical association of an early dialysis start with poor survival. *Kidney Int.* 2010;77(8):700-707.
28. Rosansky SJ, Eggers P, Jackson K, Glassock R, Clark WF. Early start of hemodialysis may be harmful. *Arch Intern Med.* 2011;171(5):396-403.
29. Sawhney S, Djurdjev O, Simpson K, Macleod A, Levin A. Survival and dialysis initiation: comparing British Columbia and Scotland registries. *Nephrol Dial Transplant.* 2009;24(10):3186-3192.
30. Stel VS, Dekker FW, Ansell D, et al. Residual renal function at the start of dialysis and clinical outcomes. *Nephrol Dial Transplant.* 2009;24(10):3175-3182.
31. Wright S, Klausner D, Baird B, et al. Timing of dialysis initiation and survival in ESRD. *Clin J Am Soc Nephrol.* 2010;5(10):1828-1835.
32. Shiao CC, Huang JW, Chien KL, Chuang HF, Chen YM, Wu KD. Early initiation of dialysis and late implantation of catheters adversely affect outcomes of patients on chronic peritoneal dialysis. *Perit Dial Int.* 2008;28(1):73-81.
33. Bozorgmehri S, Aboud H, Chamarthi G, et al. Association of early initiation of dialysis with all-cause and cardiovascular mortality: A propensity score weighted analysis of the United States Renal Data System. *Hemodial Int.* 2021;25(2):188-197.
34. Janmaat CJ, van Diepen M, Krediet RT, Hemmelder MH, Dekker FW. Effect of glomerular filtration rate at dialysis initiation on survival in patients with advanced chronic kidney disease: what is the effect of lead-time bias? *Clin Epidemiol.* 2017;9:217-230.

35. Nacak H, Bolignano D, Van Diepen M, Dekker F, Van Biesen W. Timing of start of dialysis in diabetes mellitus patients: a systematic literature review. *Nephrol Dial Transplant.* 2016;31(2):306-316.

36. Susantitaphong P, Altamimi S, Ashkar M, et al. GFR at initiation of dialysis and mortality in CKD: a meta-analysis. *Am J Kidney Dis.* 2012;59(6):829-840.

37. Traynor JP, Simpson K, Geddes CC, Deighan CJ, Fox JG. Early initiation of dialysis fails to prolong survival in patients with end-stage renal failure. *J Am Soc Nephrol.* 2002;13(8):2125-2132.

38. Crews DC, Scialla JJ, Boulware LE, et al. Comparative effectiveness of early versus conventional timing of dialysis initiation in advanced CKD. *Am J Kidney Dis.* 2014;63(5):806-815.

39. Kurella Tamura M, Desai M, Kapphahn KI, Thomas IC, Asch SM, Chertow GM. Dialysis versus Medical Management at Different Ages and Levels of Kidney Function in Veterans with Advanced CKD. *J Am Soc Nephrol.* 2018;29(8):2169-2177.

40. Scialla JJ, Liu J, Crews DC, et al. An instrumental variable approach finds no associated harm or benefit with early dialysis initiation in the United States. *Kidney Int.* 2014;86(4):798-809.

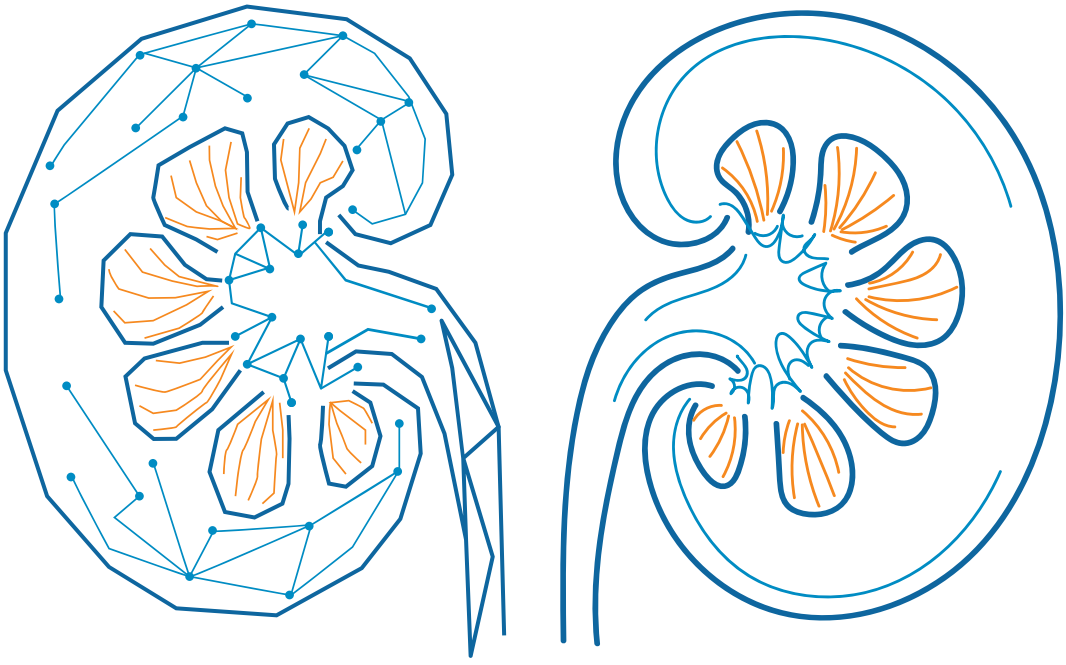
41. Hernan MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology.* 2006;17(4):360-372.

42. Swanson SA, Hernan MA. The challenging interpretation of instrumental variable estimates under monotonicity. *Int J Epidemiol.* 2018;47(4):1289-1297.

43. Martens EP, Pestman WR, de Boer A, Belitser SV, Klungel OH. Instrumental variables: application and limitations. *Epidemiology.* 2006;17(3):260-267.

44. Hernan MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology.* 2008;19(6):766-779.

45. Franklin JM, Patorno E, Desai RJ, et al. Emulating Randomized Clinical Trials with Nonrandomized Real-World Evidence Studies: First Results from the RCT DUPLICATE Initiative. *Circulation.* 2020.



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.

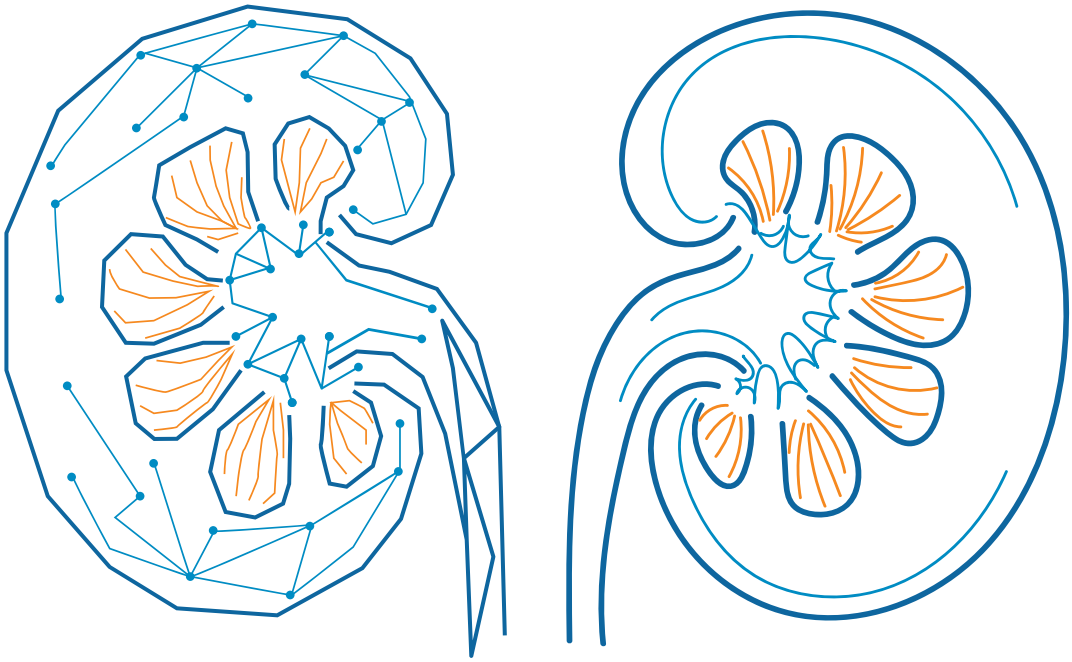
76. Goldacre B, Morton CE, DeVito NJ. Why researchers should share their analytic code. BMJ. 2019;367:l6365.

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.



APPENDICES

Nederlandse samenvatting (Dutch summary)

Samenvatting van belangrijkste bevindingen

Het doel van dit proefschrift was om een substantiële bijdrage te leveren aan de besluitvorming bij patiënten met nierziekten. In dit proefschrift zijn een aantal klinische vragen beantwoord over de effectiviteit en veiligheid van therapieën door gebruik te maken van state-of-the-art methoden en routinematig verzamelde zorgdata. Daarnaast is getracht het bewustzijn te vergroten voor potentiële systematische vertekeningen die kunnen optreden bij het gebruik van observationele data, en hoe deze biases verminderd kunnen worden.

Observationele studies met een causaal doel kunnen geplaagd worden door verschillende biases, waarvan enkele worden beschreven in **hoofdstuk 2**. Alhoewel confounding door indicatie een gevaar vormt bij alle observationele studies die als doel hebben de causale effecten van behandelingen te schatten, kunnen vragen over onbedoelde behandeleffecten en vragen met een actieve controlegroep minder gevoelig zijn voor confounding. De resterende confounding moet geadresseerd worden in de statistische analyse, waarbij confounders geïdentificeerd moeten worden op basis van vakinhoudelijke kennis, bijvoorbeeld door gebruik te maken van causale diagrammen. Er kan alleen gecorrigeerd worden voor gemeten confounding en dit kan gedaan worden met multivariabele regressie, standaardisatie of propensity score methoden. Voor eenmalige behandelingen kunnen alle methoden even goed corrigeren voor gemeten confounders. In de setting van langdurige behandelingen en behandeling-confounder feedback zijn speciale methoden gebaseerd op standaardisatie of wegen nodig. De impact van residuele confounding kan worden onderzocht door een E-waarde te berekenen, door middel van een kwantitatieve bias analyse, of door het uitvoeren van een negatieve controle uitkomst of positieve cohort analyse. Naast confounding zijn prevalent user bias en immortal time bias belangrijke limitaties in veel observationele studies. Deze biases ontstaan wanneer de start van follow-up en de start van de blootstelling niet overeenkomen. Het expliciet emuleren van een gerandomiseerde trial kan deze biases elimineren aangezien dit forceert dat de start van follow-up en blootstelling overeenkomen. Ten slotte komen missende waarden en meetfouten veelvuldig voor in routinematig verzamelde zorgdata. Het wordt daarom aanbevolen om methoden zoals multipele imputatie of kwantitatieve bias analyse toe te passen.

In **hoofdstuk 3** bediscussiëren we het concept en de voor- en nadelen van propensity score (PS) methoden, een populaire statistische methode om te corrigeren voor gemeten confounding. Er bestaan verschillende soorten PS methoden. Dit zijn PS matching, PS stratificatie, multivariabele regressie waarbij de

PS als variabele wordt toegevoegd, en PS wegen. PS methoden hebben een aantal voordelen ten opzichte van conventionele multivariabele regressie: het is mogelijk om de balans van variabelen te checken, om een specifieke target populatie te kiezen, en om individuen te excluderen die zich in de niet-overlappende delen van de PS distributie bevinden. Daarnaast hebben PS methoden de voorkeur in de setting van hoog-dimensionele data waarbij er veel confounders zijn en relatief weinig uitkomsten.

In **hoofdstuk 4** besteden we aandacht aan immortal time bias in een gepubliceerde observationele studie die als doel heeft om het causale effect van metformine te schatten in patiënten met diabetische nefropathie. We stellen een aantal oplossingen voor die toegepast hadden kunnen worden om deze bias te elimineren. Dit zijn onder andere het modelleren van metformine als tijdsafhankelijke blootstelling, landmarken of het gebruik van grace periodes in combinatie met de cloning-censoring-weighting methode. Een andere optie is het emuleren van een trial met de sequentiële trial methode.

In **hoofdstuk 5** bestuderen we de associatie tussen acute stijgingen in creatinine na het starten van RAS-blokkers en cardiovasculaire en nieruitkomsten. Patiënten met een hogere creatininestijging na de start van RAS-blokkers hadden een hoger risico op sterfte, hartfalen, myocardinfarcten en nierfalen. We vonden ook dat slechts 18% van de individuen die RAS blokkade startten de aanbevolen creatinine monitoring kregen, en dat creatininestijgingen tussen de 10-29% regelmatig voorkwamen bij gemonitorde individuen. Deze studie adresseert niet de kwestie van het stoppen van RAS-blokkers bij creatininestijgingen, maar suggereert wel dat patiënten met een creatininestijging een hoger risico hebben op cardiovasculaire en nieruitkomsten.

Hoofdstuk 6 is een vergelijkend effectiviteitsonderzoek tussen RAS-blokkers en calciumantagonisten op het starten van nierfunctie-vervangende therapie (dialyse of niertransplantatie), sterfte en cardiovasculaire uitkomsten (een combinatie van cardiovasculaire sterfte, myocardinfarct of beroerte) in patiënten met gevorderde chronische nierziekten. We vonden dat het starten van RAS blokkade was geassocieerd met een lager risico op nierfunctie-vervangende therapie en vergelijkbare risico's op sterfte en cardiovasculaire uitkomsten vergeleken met calciumantagonisten. Daarnaast is een analyse uitgevoerd in een positief controle cohort van patiënten met milde tot matige chronische nierziekten, waarvan de resultaten overeenkwamen met bewijs uit eerdere gerandomiseerde onderzoeken. Een negatieve controle analyse met kanker als uitkomst suggereerde geen residuele confounding door body mass index of roken. Deze bevindingen zouden de klinische besluitvorming rond de keuze van antihypertensiva kunnen informeren in patiënten met gevorderde chronische nierziekten.

In **hoofdstuk 7** vonden we dat het gebruik van β-blokkers bij ziekenhuisontslag geassocieerd was met een lager risico op sterfte, en een gecombineerd eindpunt van cardiovasculaire sterfte en ziekenhuisopnames voor hartfalen in individuen met hartfalen met een verminderde ejectiefractie en gevorderde chronische nierziekten. Een positieve controle analyse in individuen met hartfalen met een verminderde ejectiefractie en milde tot matige chronische nierziekten lieten vergelijkbare reducties in uitkomsten zien voor gebruikers van β-blokkers. Deze voordelen werden niet gezien in individuen met gevorderde chronische nierziekten en matige of behouden ejectiefractie, alhoewel deze analyses niet eenduidig waren door lage power. Deze bevindingen suggereren dat β-blokkers effectief zijn in patiënten met hartfalen met een verminderde ejectiefractie over het gehele spectrum van nierfunctie.

In **hoofdstuk 8** is het effect van stoppen versus doorgaan van RAS-blokkers in individuen met gevorderde chronische nierziekte onderzocht op de uitkomsten sterfte, cardiovasculaire uitkomsten, en nierfunctie-vervangende therapie. We observeerden dat individuen die stopten met RAS blokkade een hoger risico hadden op sterfte en cardiovasculaire uitkomsten, maar een lager risico hadden op nierfunctie-vervangende therapie. Vergelijkbare bevindingen werden verkregen wanneer gebruik van RAS-blokkers werd gemodelleerd als tijdsafhankelijke variabele in een marginal structural model, wanneer daarnaast gecorrigeerd werd voor kalium en albumine-creatinine ratio, en in subgroepen, inclusief het stoppen van RAS-blokkers bij een hogere (nierfunctie tussen 20-30%) of lagere eGFR (nierfunctie <20%). Deze bevindingen suggereren dat, terwijl de resultaten van de gerandomiseerde STOP-ACEi trial worden afgewacht, niet routinematig moet worden gestopt met RAS-blokkers in patiënten met gevorderde chronische nierziekten.

Hoofdstuk 9 adresseert de vraag of er een optimale nierfunctie is om te starten met dialyse. In dit onderzoek konden we met observationele data de bevindingen van de gerandomiseerde IDEAL trial repliceren. Daarnaast toonden we aan dat vroeg starten met dialyse geassocieerd was met een bescheiden vermindering in sterfte en cardiovasculaire uitkomsten (met een nierfunctie tussen de 15-16% versus een nierfunctie tussen 6-7%). Dit vertaalde zich naar een hogere levensverwachting van 1,6 maanden, terwijl dialyse gemiddeld 48 maanden eerder gestart zou moeten worden. We tonen ook aan dat eerdere observationele studies leden aan lead time bias, selectiebias en immortal time bias, dat deze systematische vertekeningen voorkomen hadden kunnen worden als het target trial emulatie framework was toegepast, en dat incorrecte analyse van onze eigen data tot dezelfde verkeerde resultaten leidt. Als geheel wijzen deze resultaten erop dat er weinig voordeel is van het vroeg starten van dialyse op basis van nierfunctie alleen. In toekomstige studies kan worden onderzocht of dialyse gestart moet worden op basis van symptomen om klinische uitkomsten verder te verbeteren.

Toekomstperspectieven

Het aantal observationele studies dat gebruik maakt van routinematig verzamelde zorgdata neemt almaar toe. In dit proefschrift is toegelicht dat het gebruik van dit soort data om de klinische praktijk te informeren een tweesnijdend zwaard betreft: aan de ene kant biedt het enorme kansen om te bestuderen hoe behandelingen in de routine klinische praktijk werken, om vragen te bestuderen die moeilijk te beantwoorden zijn in gerandomiseerde trials, en om populaties te bestuderen die zijn ondervertegenwoordigd in trials. Aan de andere kant kunnen verscheidene biases de bevindingen uit observationele studies vertekenen: confounding bias, immortal time bias en prevalent user bias om er slechts een paar te noemen.

Hoe moeten we verder om het beste bewijs te leveren voor de behandeling van patiënten met nierziekten? Natuurlijk zullen er meer gerandomiseerde trials moeten komen. In nierziekten zijn relatief weinig gerandomiseerde trials gedaan, en patiënten met nierziekten zijn overwegend geëxcludeerd uit trials in andere vakgebieden zoals de cardiologie en oncologie. Om dit probleem op te lossen hebben andere onderzoekers opgeroepen tot vermindering van de kosten en complexiteit van trials, inclusief de bureaucratische last. Dat dit mogelijk is, is recent aangetoond door de RECOVERY trial, die binnen een jaar meer dan 39.000 gehospitaliseerde COVID-19 patiënten heeft gerandomiseerd. Andere voorbeelden zijn onder andere de publicatie van zogeheten "grote, simpele trials" in de afgelopen decennia en recente innovaties in trial design zoals de registratie-gebaseerde trial. Gelukkig lijkt het vakgebied van nierziekten een inhaalslag te maken, met de publicaties van een aantal belangrijke klinische trials, terwijl andere trials nu actief bezig zijn met het includeren van patiënten. Naast het uitvoeren van meer trials kunnen nieuwe benaderingen om trialresultaten te generaliseren naar andere populaties de kloof tussen trials en de dagelijkse klinische praktijk overbruggen.

Een fundamentele vraag is of het überhaupt mogelijk is causale conclusies te trekken uit observationele data. Sommige onderzoekers zijn inderdaad van mening dat causale conclusies alleen uit gerandomiseerde trials kunnen worden getrokken, maar niet uit observationele studies, en dat alleen gerandomiseerde onderzoeken nuttig zijn. Dit is echter een valse tweedeling. Causaliteit is niet een ja/nee statement en wordt zelden geconcludeerd op basis van één studie, aangezien het altijd om de totaliteit van bewijs gaat, wat kan komen uit laboratorium onderzoek, observationele studies en gerandomiseerde trials. Daarnaast zijn er substantiële verschillen in kwaliteit tussen observationele studies, waarbij sommigen beter tot causale conclusies kunnen komen dan anderen. Aan de ene kant is het mogelijk met goed uitgevoerde observationele studies de resultaten van gerandomiseerde trials te repliceren of zelfs te voorspellen. Aan de andere kant bestaan er legio voorbeelden van observationele studies waarbij dit niet is gelukt. Dit laatste kan

vaak verklaard worden door het feit dat deze observationele studies gebrekkige methoden hebben gebruikt die onnodige systematische vertekeningen introduceerden, zoals immortal time bias en prevalent user bias, en werd in deze gevallen niet veroorzaakt door de aanwezigheid van ongemeten confounding. Deze biases ontstaan als de timing van de volgende drie elementen niet overeenkomt op baseline: de start van follow-up, de start van de behandelstrategieën en het voldoen aan alle inclusiecriteria. Omdat door randomisatie de timing van deze elementen automatisch overeenkomt, hoeven onderzoekers van gerandomiseerde trials zich hier nooit zorgen over te maken. Dit is echter niet het geval voor observationele data, waar de onderzoeker zorgvuldig moet nadenken over de baseline en hier adequaat mee om moet gaan in de analyses. Voorbeelden uit de literatuur waarbij goed uitgevoerde observationele studies in staat waren dezelfde antwoorden te krijgen als trials, terwijl observationele studies die vermijdbare biases introduceerden dit niet konden bereiken, zijn onder andere onderzoeken naar de timing van dialyse en het risico op sterfte (dit proefschrift), postmenopauzale hormoontherapie en het risico op coronaire hartziekten, statines en het risico op kanker, de timing van antiretrovirale behandeling en het risico op sterfte, dabigatran en het risico op beroertes, sodium-glucose cotransporter 2 remmers en het risico op sterfte, en colonoscopie screening en het risico op darmkanker. Het gebruik van het target trial emulatie framework kan helpen om deze onnodige studiedesign gebreken te elimineren, dwingt de onderzoeker om betekenisvolle causale vragen te stellen, faciliteert de communicatie en is leidend voor de statistische analyse. Daarnaast moeten onderzoekers de analytische methoden gebruiken die het best passen bij de klinische onderzoeksvraag. De cloning-censoring-weighting methode is bijvoorbeeld geschikt om vragen te beantwoorden die 1) verschillende initiatiemomenten van een interventie vergelijken ("Wanneer moeten we starten met de behandeling?"); 2) verschillende behandelduur vergelijken ("Hoe lang moeten we behandelen?") en 3) een grace periode bevatten ("Moet de behandeling binnen *x* maanden gestart worden na event *y* of niet?"). Wanneer het doel is om het effect van het starten van een behandeling te vergelijken met het niet starten van deze behandeling kan een random inclusiedatum gekozen worden voor de niet-starters, of zal een sequentiële trial methode gebruikt moeten worden om correct om te gaan met baseline. Onderzoekers die observationele studies uitvoeren zullen daarom moeten beschikken over geschikte methodologische expertise en een gedegen opleiding om deze methoden correct toe te passen, of iemand moeten inschakelen die deze expertise heeft. Wanneer gebrekkige methoden worden toegepast, zullen gebrekkige antwoorden worden verkregen.

Lopende systematische replicaties van gerandomiseerde trials met observationele data, zoals het RCT-DUPLICATE initiatief en andere inspanningen zijn daarom essentieel om aan te tonen dat observationele onderzoeken tot dezelfde conclusies leiden als gerandomiseerde onderzoeken, mits adequaat uitgevoerd.

Deze inspanningen zullen ook waardevolle inzichten geven onder welke omstandigheden (dat wil zeggen voor welke onderzoeksvragen, analytische methoden en databronnen) men zonder randomisatie tot causale conclusies kan komen in observationele data, en de effecten van behandelingen kan bestuderen zonder randomisatie. Deze initiatieven maken gebruik van principiële causal inference methoden en proberen ook de in- en exclusiecriteria van trials zo nauwkeurig mogelijk na te bootsen, om ervoor te zorgen dat de bevindingen niet verschillen door toepassing van gebrekkige methoden of verschillende patiëntpopulaties. Dergelijke calibratiestudies zullen ook op het gebied van nierziekten moeten worden uitgevoerd.

Bovendien zijn niet alle observationele studies even gevoelig voor confounding. Of een observationele studie tot causale conclusies kan komen hangt sterk af van de onderzoeksvraag, die wel eens de "axis of haphazardness" is genoemd. Farmaco-epidemiologische studies die schadelijke, onbedoelde effecten van medicijnen onderzoeken hebben minder last van (residuele) confounding dan studies die (on)bedoelde gunstige effecten onderzoeken. Regelgevende instanties maken inderdaad al langdurig gebruik van bewijs verkregen uit observationele studies om de veiligheid van medicijnen te beoordelen. Daarnaast zijn onderzoeken waarin twee geneesmiddelen met elkaar worden vergeleken die voor vergelijkbare indicaties worden voorgeschreven minder vatbaar voor confounding dan onderzoeken waarin een geneesmiddel wordt vergeleken met geen geneesmiddel. Het is moeilijk om in observationele data vragen te bestuderen waarin het starten van een medicijn wordt vergeleken met het niet starten, aangezien het starten van de behandeling een marker is voor de ernst van de ziekte of - in het geval van preventieve behandelingen - een marker van gezondheidszoekend gedrag; beide bronnen van bias zijn mogelijk niet volledig vastgelegd in de observationele data. In deze gevallen kan de selectie voor behandeling zo sterk zijn dat randomisatie noodzakelijk is. Desalniettemin bestaan er succesvolle toepassingen in de literatuur.

Het vermogen om causale conclusies te trekken hangt ook sterk af van de data die worden gebruikt en de variabelen die in de database beschikbaar zijn. Alvorens aan een onderzoek te beginnen, dienen onderzoekers zich af te vragen of de data van voldoende kwaliteit zijn voor het betreffende onderzoek. Wanneer voldoende gedetailleerdheid in blootstelling, uitkomsten of covariabelen ontbreekt, kan men ervoor kiezen om niet verder te gaan met de analyse. Specifiek voor nierziekten is de beschikbaarheid van routinematige laboratoriummetingen zoals creatinine en albuminurie vaak essentieel om adequaat te corrigeren voor confounding.

Hoewel de afwezigheid van ongemeten confounding een assumptie blijft die niet kan worden geverifieerd, moeten onderzoekers deze assumptie zorgvuldig en zo goed mogelijk rechtvaardigen, wat natuurlijk het moeilijke gedeelte van

epidemiologisch onderzoek is. Verschillende analyses kunnen worden gebruikt om het vertrouwen in de validiteit van bevindingen te versterken, zoals positieve en negatieve controle analyses, die zowel uitkomsten als cohorten zouden kunnen zijn waarvoor we bepaalde associaties verwachten. Deze analyses kunnen worden uitgevoerd om te onderzoeken of het haalbaar is om een bepaalde vraag te beantwoorden, zelfs voordat de primaire analyse is uitgevoerd. Indien resultaten uit gerandomiseerde trials beschikbaar zijn, kunnen de resultaten van de observationele analyse worden vergeleken met de resultaten van de trial, in acht nemend of dezelfde behandelstrategieën en studiepopulaties zijn onderzocht. Daarnaast moet er adequaat worden gecorrigeerd voor confounding. Wanneer bijvoorbeeld tijdsafhankelijke confounding aanwezig is, zijn G-methods zoals inverse probability weighting van marginal structural models vereist om juiste schattingen te krijgen. Propensity scores zijn een populaire methode om te corrigeren voor gemeten confounders. De vele ontwikkelingen in de propensity score-methodologie bieden een grote flexibiliteit voor het specificeren van de doelpopulatie (op welke populatie zijn onze resultaten van toepassing), balans in confounders, en precisie. Belangrijk is dat de balans op gemeten en ongemeten confounders voorafgaand aan de analyse moet worden gecontroleerd. In propensity score analyses worden ongemeten confounders alleen gebalanceerd voor zover ze gecorreleerd zijn met gemeten variabelen die zijn geïncludeerd in de propensity score. Daarom is een belangrijke aanpak om te corrigeren voor residuele confounding door niet-geobserveerde factoren om te corrigeren voor zoveel mogelijk proxy's van de onderliggende confounder (bijvoorbeeld diagnoses, procedures, medicijnen, aantal ziekenhuisopnames), die vóór de start van de behandeling moeten worden gemeten om correctie voor causale mediatoren te voorkomen. Wanneer bepaalde confounders alleen beschikbaar zijn voor een subset van de populatie en niet kunnen worden gebruikt om voor te corrigeren, kan de balans in deze variabele na het matchen/wegen van de propensity score het vertrouwen verhogen dat andere niet-gemeten variabelen ook in evenwicht zijn. Naast confounding moeten ook andere bronnen van vertekening worden onderzocht door middel van sensitiviteitsanalyses, zoals differentiële misclassificatie. Bij het onderzoeken van een 30% GFR-afname als uitkomst, moet men bijvoorbeeld controleren of beide behandelarmen dezelfde intensiteit van creatininetesten hebben tijdens follow-up. Kwantitatieve bias-analyses kunnen worden gebruikt om de invloed van resterende systematische vertekeningen te onderzoeken.

Bovendien kunnen verschillende observationele studies die verschillende causale methoden toepassen, worden gebruikt om bewijs te trianguleren, aangezien elke methode zijn eigen specifieke assumpties heeft. Op dit punt zijn er grote kansen voor uitwisseling van (quasi-experimentele) methoden uit andere wetenschapsgebieden, zoals regressiediscontinuïteit. Ten slotte kunnen verschillende andere maatregelen de reproduceerbaarheid van observationele

studies vergroten. Dit omvat preregistratie van observationele studies, de publicatie van codes en het volgen van rapportagerichtlijnen. In wezen zou het proces van het uitvoeren van observationele onderzoeken het regelgevende indieningsproces van gerandomiseerde onderzoeken moeten nabootsen. Er mogen bijvoorbeeld geen behandeling-uitkomstanalyses worden uitgevoerd totdat het protocol volledig is gespecificeerd en geregistreerd.

Concluderend zal het verkrijgen van bewijs uit niet-experimentele en experimentele studies belangrijk blijven, aangezien beide bronnen van bewijs elkaar aanvullen. Goed uitgevoerde observationele studies kunnen waardevol bewijs leveren voor de besluitvorming op het gebied van nierziekten, maar ook voor de geneeskunde in het algemeen. Het enige dat we moeten doen is het beantwoorden van de juiste vragen met de juiste methoden.

Dankwoord (Acknowledgments)

Dit proefschrift had niet tot stand kunnen komen zonder de hulp en steun van een groot aantal mensen. Ik wil daarom graag mijn dank uitspreken aan eenieder die heeft bijgedragen aan dit proefschrift, waarvan een aantal mensen in het bijzonder.

Friedo, als eerstejaars geneeskunde student heb je me al onder je hoede genomen. In de afgelopen vijf jaar heb je me niet alleen veel over onderzoek geleerd, maar ook veel daarbuiten. Ik ben je zeer dankbaar voor de mogelijkheden die je me hebt gegeven om mezelf te ontwikkelen en de vrijheid die je mij gunde bij de invulling van dit promotietraject.

Merel, je begeleiding, steun en alle gezelligheid de afgelopen jaren waren van onschatbare waarde. In het bijzonder dank voor je extra steun tijdens de coronaperiode. Stressvolle situaties wist jij goed te relativeren, waar ik veel aan heb gehad.

Juan, we connected from the very first day and I enjoy our efficient, fast and fun working ethic. Even if separated by a thousand kilometers, we form a great team together. I cherish great memories from our collaboration and I am grateful for the opportunities you have given me in the past years.

Alle collega's van de afdeling Klinische Epidemiologie wil ik bedanken voor de fijne werksfeer. Roommates of C7-82: our chit-chat and your support made me come to work with much joy. Leden van het nierclubje, de afgelopen jaren heb ik veel geleerd tijdens de leerzame journal clubs. Rolf, ik ben je dankbaar voor wat je me allemaal hebt geleerd over causaliteit. Maarten, je deur stond altijd open voor een goed gesprek of waardevol advies, veel dank daarvoor. Tamara, door jou voelde ik me meteen thuis op de afdeling als eerstejaars geneeskunde student.

I would like to thank Catherine and Marie for their great collaboration and valuable input. Catherine, you always took the time to review my work with much care and detail. Marie, thank you for your input on my work and help with the Swedish Renal Registry.

De Raad van Bestuur, de MD/PhD selectiecommissie en in het bijzonder Professor Pancras Hogendoorn wil ik bedanken voor de kans die zij mij hebben gegeven om dit promotietraject te starten.

Marco, my good Italian friend, thank you for teaching me so much about R and routinely collected healthcare data. You made me feel at home at Karolinska and I am grateful for our friendship.

Bas, Fas, Floris, Michael, Niels, Tjerk, dank voor jullie steun de afgelopen jaren. Ik zie jullie gauw in de kliniek!

Layla en Sophie, wat begon als de CAT is uitgemond in een ware vriendschap. Ik heb genoten van de taart en etentjes de afgelopen jaren. Sophie, dank dat je mijn paranimf wil zijn op deze bijzondere dag.

Maxim, via de GDT, het SGL en de uni zijn we al 17 jaar goede vrienden. Ik koester bijzondere herinneringen aan onze vriendschap, waaronder Hei- en Boeicop, onze vakanties en de aanval van de Schotse Hooglander. Ik prijs mij gelukkig deze langdurige vriendschap te hebben.

Papa en mama, jullie hebben mij geleerd om hard te werken om iets te bereiken in het leven. Ik realiseer me hoeveel beter de kansen zijn die jullie mij hebben gegeven dan de kansen die jullie zelf hadden. Zonder jullie steun had ik hier niet gestaan. Dit proefschrift is dan ook aan jullie opgedragen.

Victor, quel bonheur om jou als broer te hebben. Ik ben dankbaar dat je altijd voor je kleine broertje klaar staat, of dat nu urenlange telefoongesprekken, broodnodige ontspanmomenten of een reddingsmissie naar Stockholm zijn. Je goede adviezen zijn van onschatbare waarde geweest de afgelopen jaren.

Tot slot mijn lieve Anne. Zonder jou was dit proefschrift niet tot stand gekomen. Dank voor alle liefde en steun de afgelopen jaren, waarbij je me stimuleert om het beste uit mezelf te halen en ook zorgt voor balans in mijn leven. Ik ben gelukkig dat ik deze mooie momenten met jou kan delen en dat je aan mijn zijde staat op deze bijzondere dag.

List of publications

Published articles included in this thesis

1. **Fu EL**, van Diepen M, Xu Y, Trevisan M, Dekker FW, Zoccali C, Jager KJ and Carrero JJ. Pharmacoepidemiology for nephrologists (part 2): potential biases and how to overcome them. *Clin Kidney J*. 2020;14:1317-1326.
2. **Fu EL**, Groenwold RHH, Zoccali C, Jager KJ, van Diepen M and Dekker FW. Merits and caveats of propensity scores to adjust for confounding. *Nephrol Dial Transplant*. 2019;34:1629-1635.
3. **Fu EL**, van Diepen M. Comment on Kwon et al. The Long-term Effects of Metformin on Patients With Type 2 Diabetic Kidney Disease. Diabetes Care 2020;43:948-955. *Diabetes Care*. 2020;43:e190.
4. **Fu EL**, Trevisan M, Clase CM, Evans M, Lindholm B, Rotmans JI, van Diepen M, Dekker FW and Carrero JJ. Association of Acute Increases in Plasma Creatinine after Renin-Angiotensin Blockade with Subsequent Outcomes. *Clin J Am Soc Nephrol*. 2019;14:1336-1345.
5. **Fu EL**, Uijl A, Dekker FW, Lund LH, Savarese G and Carrero JJ. Association Between beta-Blocker Use and Mortality/Morbidity in Patients With Heart Failure With Reduced, Midrange, and Preserved Ejection Fraction and Advanced Chronic Kidney Disease. *Circ Heart Fail*. 2020;13:e007180.
6. **Fu EL**, Clase CM, Evans M, Lindholm B, Rotmans JI, Dekker FW, van Diepen M and Carrero JJ. Comparative Effectiveness of Renin-Angiotensin System Inhibitors and Calcium Channel Blockers in Individuals With Advanced CKD: A Nationwide Observational Cohort Study. *Am J Kidney Dis. 2020;77:719-729*.
7. **Fu EL**, Evans M, Clase CM, Tomlinson LA, van Diepen M, Dekker FW and Carrero JJ. Stopping Renin-Angiotensin System Inhibitors in Patients with Advanced CKD and Risk of Adverse Outcomes: A Nationwide Study. *J Am Soc Nephrol*. 2021;32:424-435.

Other publications

8. **Fu EL**, Franko MA, Obergfell A, Dekker FW, Gabrielsen A, Jernberg T and Carrero JJ. High-sensitivity C-reactive protein and the risk of chronic kidney disease progression or acute kidney injury in post-myocardial infarction patients. *Am Heart J*. 2019;216:20-29.
9. **Fu EL**, Janse RJ, de Jong Y, van der Endt VHW, Milders J, van der Willik EM, de Rooij ENM, Dekkers OM, Rotmans JI and van Diepen M. Acute kidney injury and kidney replacement therapy in COVID-19: a systematic review and meta-analysis. *Clin Kidney J*. 2020;13:550-563.
10. Clase CM, **Fu EL**, Joseph M, Beale RCL, Dolovich MB, Jardine MJ, Mann JFE, Pecoits-Filho R, Winkelmayer WC and Carrero JJ. Cloth Masks May Prevent Transmission of COVID-19: An Evidence-Based, Risk-Based Approach. *Ann Intern Med*. 2020;173:489-491.
11. Clase CM, **Fu EL**, Ashur A, Beale RCL, Clase IA, Dolovich MB, Jardine MJ, Joseph M, Kansiime G, Mann JFE, Pecoits-Filho R, Winkelmayer WC and Carrero JJ. Forgotten Technology in the COVID-19 Pandemic: Filtration Properties of Cloth and Cloth Masks-A Narrative Review. *Mayo Clin Proc*. 2020;95:2204-2224.

12. Bidulka P, **Fu EL**, Leyrat C, Kalogirou F, McAllister KSL, Kingdon EJ, Mansfield KE, Iwagami M, Smeeth L, Clase CM, Bhaskaran K, van Diepen M, Carrero JJ, Nitsch D and Tomlinson LA. Stopping renin-angiotensin system blockers after acute kidney injury and risk of adverse outcomes: parallel population-based cohort studies in English and Swedish routine care. *BMC Med*. 2020;18:195.
13. Trevisan M, **Fu EL**, Xu Y, Jager KJ, Zoccali C, Dekker FW and Carrero JJ. Pharmacoepidemiology for nephrologists (part 1): concept, applications and considerations for study design. *Clin Kidney J*. 2020;14:1307-1316.
14. Trevisan M, **Fu EL**, Szummer K, Norhammar A, Lundman P, Wanner C, Sjolander A, Jernberg T and Carrero JJ. Glucagon-like peptide-1 receptor agonists and the risk of cardiovascular events in diabetes patients surviving an acute myocardial infarction. *Eur Heart J Cardiovasc Pharmacother*. 2021;7:104-111.
15. de Jong Y, **Fu EL**, van Diepen M, Trevisan M, Szummer K, Dekker FW, Carrero JJ and Ocak G. Validation of risk scores for ischaemic stroke in atrial fibrillation across the spectrum of kidney function. *Eur Heart J*. 2021;42:1476-1485.
16. Windahl KM, Faxén Irving G, Almquist T, Lidén MK, Stenvinkel P, Chesnaye NC, Drechsler C, Szymczak M, Krajewska M, **Fu EL**, Torino C, Porto G, Roderick PJ, Caskey FJ, Wanner C, Dekker FW, Jager KJ and Evans M. Patient-reported measures and life-style are associated with deterioration in nutritional status in CKD stage 4-5: the EQUAL cohort study. *J Ren Nutr*. 2021.
17. Zhang JLH, Appelman-Dijkstra N, **Fu EL**, Rotmans JI and Schepers A. Practice variation in treatment of patients with renal hyperparathyroidism: a survey-based study in the Netherlands. *BMC Nephrol.* 2021;22:150.
18. Trevisan M, **Fu EL**, Xu Y, Savarese G, Dekker FW, Lund LH, Sjolander A and Carrero JJ. Stopping mineralocorticoid receptor antagonists after hyperkalemia is associated with lower recurrence of hyperkalemia but greater risk of cardiovascular events and death. *Eur J of Heart Fail*. 2021.
19. Clase CM, **Fu EL**, Jardine M, Mann JFE, Carrero JJ. Cloth Masks May Prevent Transmission of COVID-19. *Ann Intern Med. 2021;174:580.*
20. Chesnaye NC, Stel VS, Tripepi G, Dekker FW, **Fu EL**, Zoccali C, Jager KJ. An introduction to inverse probability of treatment weighting in observational research. *Clin Kidney J. 2021.*

Submitted

21. **Fu EL**, Evans M, Carrero JJ, Putter H, Clase CM, Caskey FJ, Szymczak M, Torino C, Chesnaye NC, Jager KJ, Wanner C, Dekker FW and Van Diepen M. When to initiate dialysis to reduce mortality and cardiovascular events in advanced chronic kidney disease: nationwide cohort study.
22. **Fu EL**, Clase CM, Janse RJ, Lindholm B, Dekker FW, Jardine MJ, Carrero JJ. Comparative effectiveness of SGLT2i versus GLP1-RA on cardiovascular outcomes in routine clinical practice.
23. **Fu EL**, Trevisan M, Lanka V, Clase CM, Xu Y, Van Diepen M, Dekker FW, Jardine MJ, Carrero JJ. Comparative effectiveness of SGLT2i versus DPP4i on cardiovascular, kidney and hyperkalemia outcomes in individuals from routine clinical practice: observational cohort study.
24. Xu Y, **Fu EL**, Trevisan M, Jernberg T, Sjolander A, Clase CM and Carrero JJ. Stopping renin-angiotensin system inhibitors after hyperkalemia and risk of adverse outcomes.

PhD Portfolio

Name PhD student:	E.L. Fu
LUMC department:	Clinical Epidemiology
PhD period:	September 2018 - August 2021
Title thesis:	Optimal cardiovascular treatment strategies in kidney disease – causal inference from observational data
Promotors:	Prof. dr. F.W. Dekker
	Prof. dr. J.J. Carrero
Copromotor:	Dr. M. van Diepen

PhD Training	Year	Hours
Courses		
Survival analysis	2017	42
Regression analysis	2018	42
Causal inference	2018	84
PhD Introductory Meeting	2019	5
Basic Methods and Reasoning in Biostatistics	2019	42
BROK Course	2019	42
Meta analysis	2019	42
Advanced survival analysis	2019	40
Analysis of repeated measurements	2019	42
ESP48 Causal inference Rotterdam	2019	40
Prediction modelling and intervention research	2019	84
Winterschool Dutch Kidney Foundation	2020	42
Causal inference for multiple time-point (longitudinal) exposures (SER workshop)	2020	4
An introduction to transporting treatment effects from randomized clinical trials to clinical practice (SER workshop)	2021	4
Regression discontinuity designs	2021	14
Instrumental variables	2021	14
Statistical aspects of clinical trials	2021	42

PhD Training	Year	Hours
Congress attendance and poster or oral presentations		
Poster presentation ERA-EDTA congress	2018	8
Oral presentation Dutch Nephrology Days	2019	12
Oral presentation & 2 poster presentations ERA-EDTA congress	2019	28
Oral presentation & poster presentation ERA-EDTA congress	2020	30
Oral presentation ESC congress	2020	12
Oral presentation & moderated poster presentation Dutch Nephrology Days	2020	20
Oral presentation & moderated poster presentation ERA-EDTA congress	2021	20
Poster presentation Dutch Epidemiological Conference	2021	8
Teaching activities		
Academic and Scientific Training Year 1 (BSc Medicine)	2018	6
CRiP - Advanced Concepts in Epidemiology (MSc Biomedical Sciences)	2018	3
Academic and Scientific Training Year 1 (BSc Medicine)	2019	3
Academic and Scientific Training Year 2 (BSc Medicine)	2019	6
Design and Analysis of Biomedical Studies (BSc Biomedical Sciences)	2019	3
Critical Appraisal of a Topic (BSc Medicine)	2019	28
Lecture Series (Honours College Medicine)	2019	8
Lecture Series (Honours College Medicine)	2020	8
Medicine in Numbers (Pre-University College Leiden), course coordinator	2020	66
Epidemiological Data Analysis and Critical Review of Scientific Literature (Honours College Medicine)	2020	34
Academic and Scientific Training Year 1 (BSc Medicine)	2021	19
Critical Appraisal of a Topic (BSc Medicine)	2021	28
Academic and Scientific Training Year 3 (BSc Medicine)	2021	3

PhD Training	Year	Hours
Awards and prizes		
Face of Science, Royal Academy of Sciences	2019	
ERA-EDTA 2019 Travel Grant	2019	
ERA-EDTA 2019 Best Abstract presented by Young Author	2019	
ERA-EDTA 2020 Social Media Selection Award for Best Abstract	2020	
ERA-EDTA 2021 Travel Grant	2021	
ERA-EDTA 2021 Best Abstract presented by Young Author	2021	
ERA-EDTA 2021 Best Abstract of Congress	2021	
Personal grants		
Personal MD/PhD Grant, Board of Directors LUMC	2018	
Eurolife Scholarship for Early Career Researchers	2018	
Dutch Kidney Foundation Kolff PhD Fellow Abroad Grant	2018	

Curriculum vitae

Edouard Liang Fu was born on the 4th of February 1997 in Leiden, The Netherlands. In 2015 he obtained his gymnasium diploma cum laude at the Stedelijk Gymnasium Leiden and started medical school at Leiden University. As a first year medical student he developed an early interest in clinical research, and therefore started training under supervision of Prof. dr. F.W. Dekker at the department of Clinical Epidemiology. A year later he was given the opportunity to formally continue his research as part of the MD/PhD track, preparing him to combine a PhD with his medical studies. He received further training in epidemiology and statistics by following the clinical epidemiology track of the Honours College Medicine. Besides his research, Edouard was an active member of the Dutch Surgical Society for Medical Students Leiden, of which he later became a board member. Furthermore, he taught dissection courses, was year representative and participated in the Leiden Oxford Transplantation Summer School. In 2018 he received a highly competitive MD/PhD grant from the LUMC Board of Directors, allowing him to do 3-year fulltime PhD research.

After obtaining both his bachelor's degree in medicine and Honours College degree cum laude, Edouard started as a PhD student at the department of Clinical Epidemiology, under supervision of Prof. dr. F.W. Dekker and Dr. M. van Diepen. He spent part of his PhD abroad in Sweden at the department of Medical Epidemiology and Biostatistics, Karolinska Institutet, under supervision of Prof. dr. J.J. Carrero. This was supported by a grant from the Dutch Kidney Foundation and a Eurolife scholarship.

During his PhD, he has addressed clinically important questions in the fields of cardiology, nephrology and internal medicine, and has published in journals such as the Journal of the American Society of Nephrology, Annals of Internal Medicine and the European Heart Journal. He also presented his research at multiple national and international congresses where he was awarded with various prizes. Furthermore, the Royal Dutch Academy of Arts and Sciences (KNAW) gave him the opportunity to write about his research for the lay public when he was elected as a Face of Science in 2019. With the completion of his PhD he also has completed his training as clinical epidemiologist (Epidemiologist B).

Before finishing his training to become a medical doctor, Edouard will spend two years as postdoctoral researcher at the Division of Pharmacoepidemiology and Pharmaco-economics at Brigham and Women's Hospital and Harvard Medical School. In addition, he will serve as an editorial fellow at the leading kidney journal JASN.