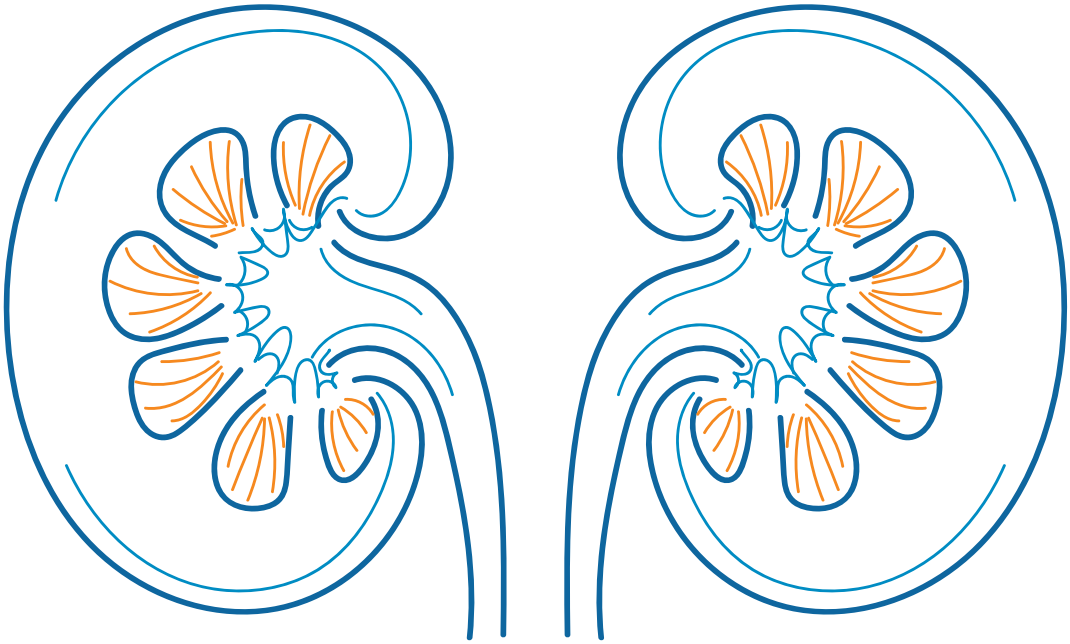
Optimal cardiovascular treatment strategies in kidney disease: casual inference from observational data

Fu, E.L.



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.

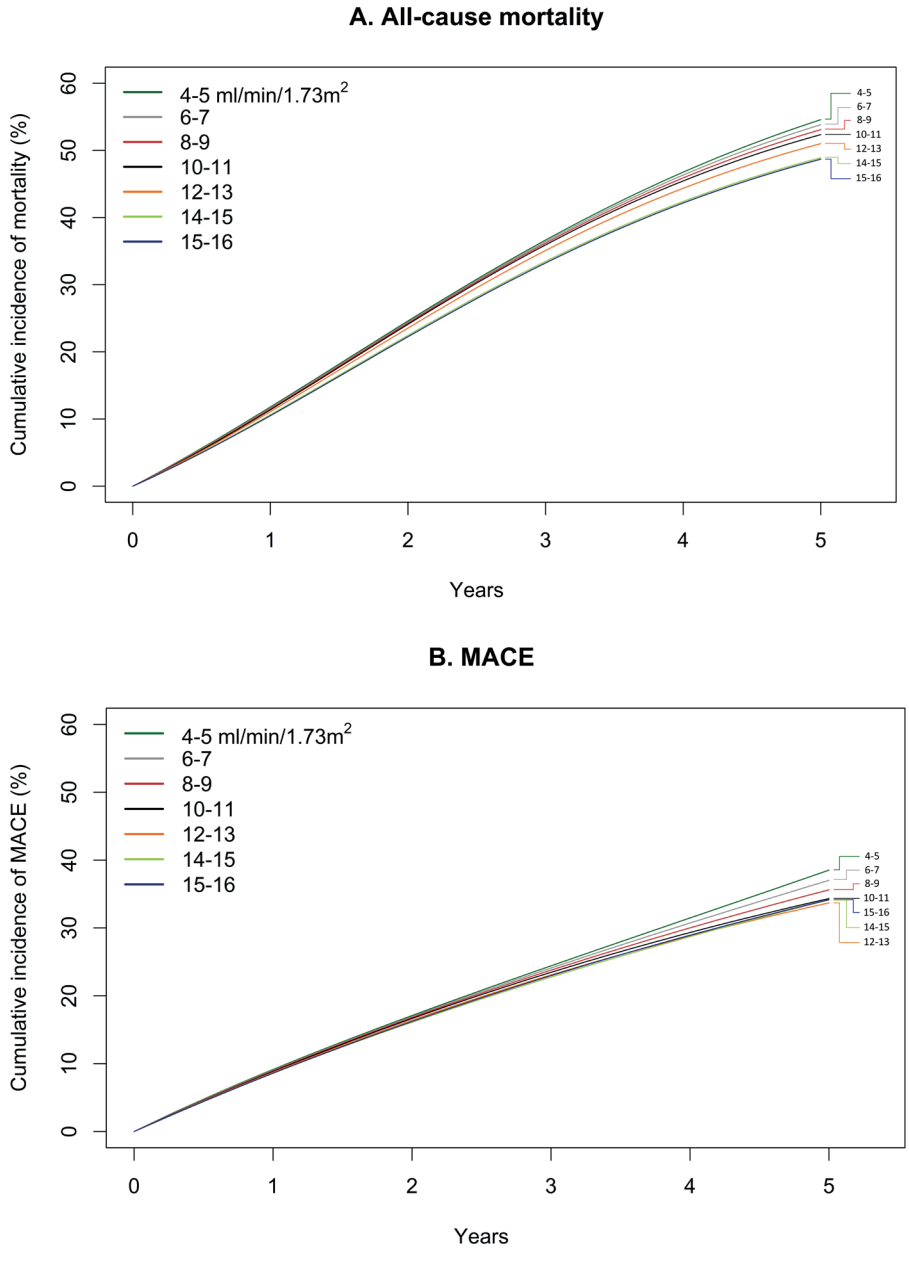


Table 2. 5-year absolute risks, risk differences, risk ratios and hazard ratios for all-cause mortality associated with initiating dialysis with eGFR values between 4 and 19 ml/min/1.73m² in increments of 1 ml/min/1.73m² with 6-7 ml/min/1.73m² as reference.

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.

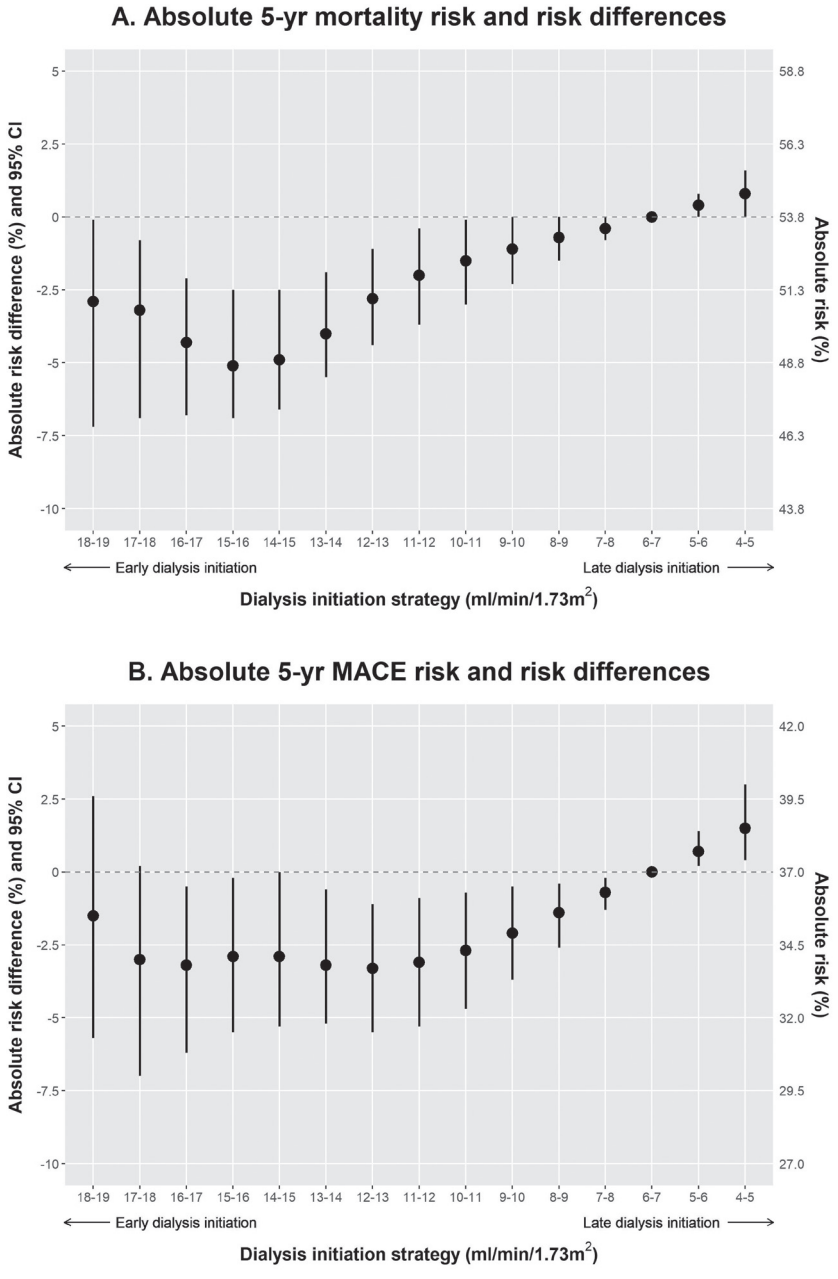
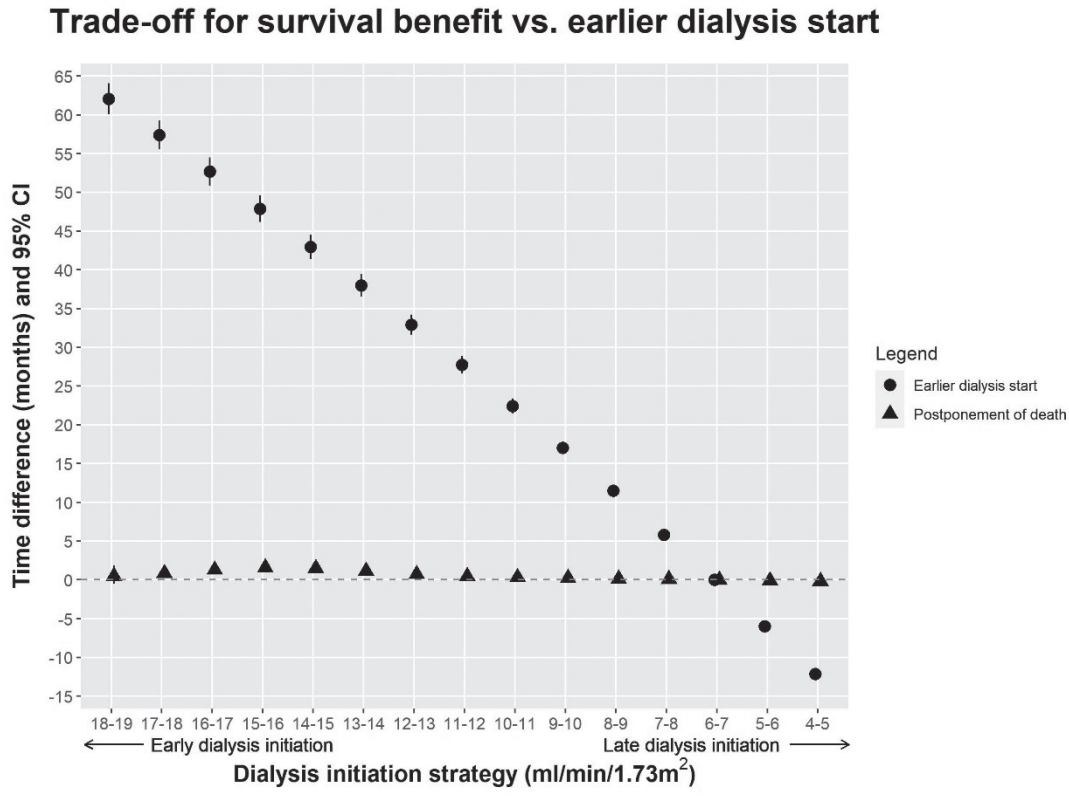


Figure 3. Trade-off between additional survival time (5-year RMST difference) and time that dialysis has to be initiated earlier, for dialysis initiation strategies with eGFR values between 4 and 19 ml/min/1.73m² in increments of 1 ml/min/1.73m², with 6-7 ml/min/1.73m² as reference. Note that a positive 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**).

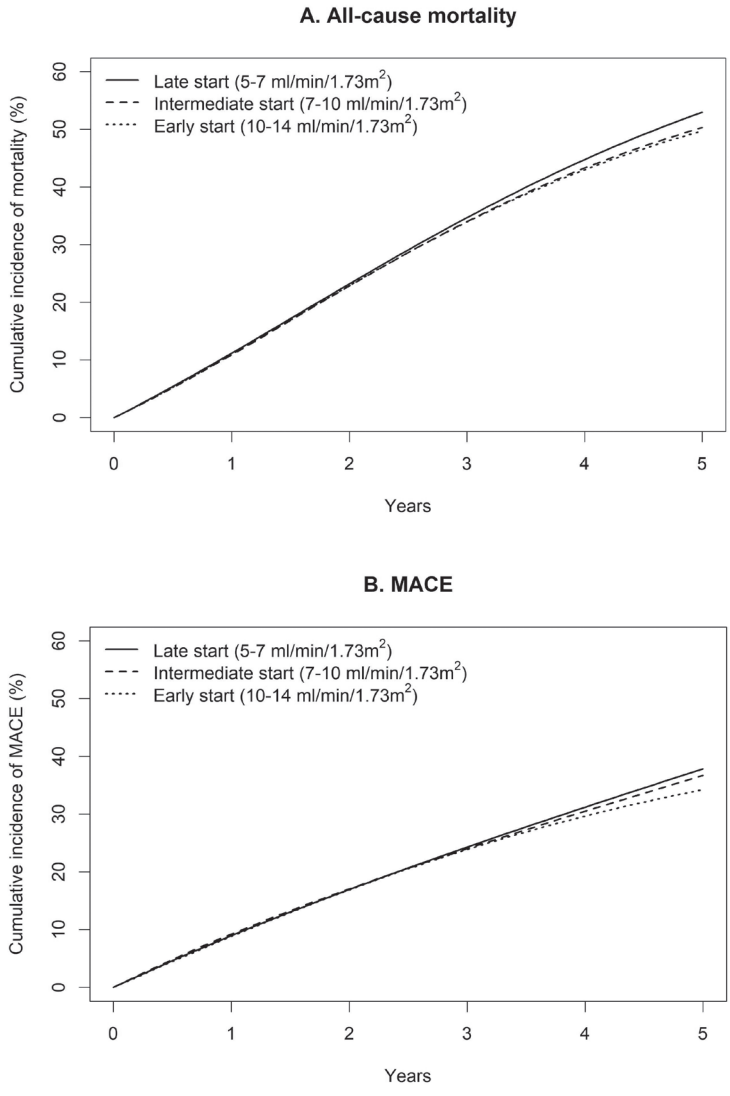
Supporting and sensitivity analyses

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.