Supplemental Table of Contents

Supplementary methods

Study design and data sources

We conducted a population-based cohort study, using linked administrative and laboratory provincial data from Alberta, Canada (Alberta Health database).[1] The Alberta Health database contains information on demographic data, emigration dates, vital statistics, and diagnostic and procedural information for inpatient and outpatient physician services. Over 99% of Alberta residents are registered with Alberta Health and have universal access to hospital care, laboratory testing and physician services. This study was approved by the Conjoint Health Research Ethics Board of the University of Calgary, with a waiver of patient consent.

Statistical analysis details

Cumulative incidence functions. To study the absolute risk of kidney failure accounting for the competing risk of death and the absolute risk of death without kidney failure, we estimated the crude (Aalen–Johansen estimator)[2] and adjusted cumulative incidence functions of these two events and their 95% confidence intervals (CI).[3,4] The cumulative incidence function (CIF) is the probability of failing from a specific cause k at time t. We obtained adjusted cumulative incidence functions for each event *indirectly* from a cause-specific hazard model of both kidney failure and death.[4-7] We checked consistency of results by estimating CIFs *directly* from a log-cumulative sub-distribution hazard of both events simultaneously.[8,9] We used plots and tables to summarize the crude and adjusted cumulative incidence functions and reported the same data on the corresponding crude (Kaplan–Meier) and adjusted naïve incidence functions (NIFs) obtained from the same cause-specific hazard model using conventional methods that censor for the competing event.[4,5,7]

Flexible parametric models use restricted cubic spline functions of time to model survival data and can be formulated on a hazard or other scales. Flexible parametric models enable proportional hazards to be fit but can be extended to model time-dependent effects on the hazard scale (or other scales). This approach has many advantages over the semi-parametric Cox model including the ease with which smooth predictions can be made, the modeling of complex time-dependent effects and the investigation of absolute as well as relative effects. As compared to other parametric models that have specific distributional requirements, flexible parametric models through the use of splines can adapt to a baseline hazard of any shape. We used flexible parametric regression to estimate the two cause-specific hazards of kidney failure and death simultaneously and derive cumulative incidence functions indirectly.[5]

Model building and verification. We used flexible parametric models specifying up to 7 degrees of freedom to model time from index date to the first event that occurred. We adjusted all models for sex, eGFR, albuminuria category, diabetes and presence of cardiovascular disease and tested first-order interactions between age and eGFR, albuminuria and eGFR, diabetes and albuminuria, cardiovascular disease and albuminuria, age and sex, and age and cardiovascular disease. We used fractional polynomial analyses and martingale residuals analyses to assess the form of the relationship between continuous covariates (age and eGFR) and outcome. We used semi-parametric modeling[7,9,10] and residual analyses to test proportionality of each time-invariant effect and model goodness-of-fit by overlaying the fitted cumulative cause-specific hazard functions on the Nelson-Aalen curves. Since the chosen flexible parametric model was a proportional hazards model for all covariates except event type, during model building we checked that the hazard ratios associated with all covariates except event type were the same as those from the corresponding Cox model stratified by event type up to the second decimal place on the hazard scale.[7] We included event type as a covariate with time-dependent effects in flexible parametric modeling. We used likelihood ratio tests to compare nested models and information criteria for non-nested models. During model building we checked that results were consistent across study time.

Examples of STATA codes for semi-parametric and flexible parametric models.

For cause-specific hazard regression, the risk set needs to be arranged in as many records per person as there are competing events (both for Cox and for flexible parametric modeling), with the failure variable coded as 0 or 1, and an event type variable specifying the type of failure.

Since time is measured from study entry until either kidney failure or death (without delayed entry), the preliminary stset command does not need specification of both origin (start) date and end date.

stcox X, strata(event) efron	where X = list of covariates and interactions; estimation of the coefficient of the event type is possible, but can be ignored
stpm2 X event , df(7) scale(hazard)	this specification assumes PH for event type
stpm2 X, df(7) scale(hazard) tvc(event) dftvc(5)	this specification allows time-dependent effects

For sub-distribution hazard regression, the risk set needs to be arranged in one record per person and the failure variable has as many levels as there are competing events, plus the censoring code (usually 0).

```
stset stop_date, failure(status==1) scale(365.25)
stcrreg X, compete(status == 2)
```

where X = list of covariates and interactions; the value of the competing event needs to be specified (death=2)

```
stset stop_date, failure(status==2) scale(365.25)
stcrreg X, compete(status == 1)
```

where X = list of covariates and interactions; the value of the competing event needs to be specified (kidney failure =1)

```
stpm2cr [kidney_failure: X, scale(hazard) df(7)]
        [death: X, scale(hazard) df(7)],
        events(status) cause(1 2) cens(0)
```

Using the full likelihood, the stpm2cr function fits flexible parametric regression models for the cause-specific CIF. Rather than fitting a model to each cause-specific CIF separately (as stcrreg does), stpm2cr fits all cause-specific CIFs simultaneously. The above specification assumes proportional hazards for all causes. Time-dependent effects can be specified using the tvc option as in stpm2 (we did not have any in this study).

sTable 1: The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	Within the Title Page, page 1 and Methods section of the Abstract page 3	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	The Methods section of the Abstract page 3 The Methods section of the Abstract page 3 The Methods section of the Abstract page 3
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Pages 4		
Objectives	3	State specific objectives, including any prespecified hypotheses	Page 4		
Methods					
Study Design	4	Present key elements of study design early in the paper	Page 5		

Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Pages 5-7		
Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case	Pages 6-7 **N/A** **N/A** **N/A** **N/A**	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	Pages 5-7 Pages 5-7; Supplemental material Figure 1
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	Pages 7-8 & Appendix Methods	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be	Pages 7-8 & Appendix Methods

				reported, an explanation should be provided.	
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Pages 5-6		
Bias	9	Describe any efforts to address potential sources of bias	Pages 5-7		
Study size	10	Explain how the study size was arrived at	Figure 1		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	Pages 6-7		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical	Pages 5-7 Pages 5-7 5-7 5-7 N/A N/A		

		methods taking account of sampling strategy (e) Describe any sensitivity analyses	Pages 6-7		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	Page 12 Pages 5-7
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Pages 5-7
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	Figure 1 Figure 1 Figure 1	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	Figure 1
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*,	Page 7 & Table 1		

		demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)	 Table 1 Page 7		
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study* - Report numbers of outcome events or summary measures	Page 7 N/A N/A		
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Page 8, Figure 2 Table 1 N/A		

Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	Appendix		
Discussion					
Key results	18	Summarise key results with reference to study objectives	Page 9		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Page 11	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Page 11
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Pages 10-11		
Generalisability	21	Discuss the generalisability (external validity) of the study results	Pages 11-12		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Page 12		

Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Pages 5-7

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

sTable 2: Codes for identifying dialysis or transplantation

1) **Physician claims**: Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures codes

Codes	Code description
For dialysis	
13.99A	Hemodialysis treatment, unstable patient
13.99B	Hemodialysis treatment, stable patient
13.99C	Assessment and management of an unstable patient with acute/chronic renal failure treated by peritoneal dialysis
13.99D	Assessment and management of a stable patient with chronic renal failure treated by peritoneal dialysis
13.99O	Management of dialysis patients on home dialysis or receiving treatment in a remote hemodialysis unit (per week)
13.99OA	Management of patient on hemodialysis or peritoneal dialysis (per week)
13.99AB	Dialysis therapy, any modality, in the intensive care unit
For transplantation	
67.5	Transplant of kidney
67.59	Other kidney transplantation
67.59A	Renal transplantation (homo, hetero, auto)

2) **Hospitalizations**: Canadian Classification of Health Intervention codes

Codes	Code description
For transplantation	
1.PC.85.^^	Transplant, kidney
1.PC.85.LA-XX-J	Using living donor (allogenic or syngeneic) kidney
1.PC.85.LA-XX-K	Using deceased donor kidney
1.OK.85.XU-XX-K	Transplant, pancreas with duodenum and kidney with exocrine drainage via bladder [e.g. donor duodenum is grafted to bladder: duodenocystostomy]
1.OK.85.XV-XX-K	Transplant, pancreas with duodenum and kidney with exocrine drainage via intestine with homograft [e.g. donor duodenum is grafted to bowel]

sTable 3: Cause-specific hazard parametric model of kidney failure and death

			Kidney failure			Death		
	Group	Comparison	HR	L95	U95	HR	L95	U95
No CV	Female, no DM	A0 vs A1	6.49	2.89	14.57	1.34	0.88	2.02
		A2 vs A1	1.55	0.79	3.05	0.97	0.69	1.36
		A3 vs A1	2.72	1.51	4.88	0.91	0.63	1.32
	Male, no DM	A0 vs A1	6.09	2.60	14.27	0.74	0.45	1.23
		A2 vs A1	1.46	0.70	3.02	0.54	0.34	0.84
		A3 vs A1	2.55	1.34	4.86	0.51	0.31	0.81
	Female, DM	A0 vs A1	12.55	5.67	27.81	1.93	1.27	2.91
		A2 vs A1	1.61	0.81	3.19	1.01	0.71	1.43
		A3 vs A1	4.49	2.50	8.06	1.03	0.71	1.49
	Male, DM	A0 vs A1	11.79	5.09	27.28	1.07	0.65	1.77
		A2 vs A1	1.51	0.72	3.16	0.56	0.36	0.88
		A3 vs A1	4.21	2.21	8.04	0.57	0.36	0.92
CV	Female, no DM	A0 vs A1	11.13	4.46	27.78	9.02	5.36	15.20
		A2 vs A1	3.21	1.49	6.92	6.16	3.88	9.79
		A3 vs A1	5.27	2.68	10.38	5.56	3.43	8.99
	Male, no DM	A0 vs A1	10.44	4.02	27.12	5.02	2.79	9.05
		A2 vs A1	3.01	1.33	6.81	3.43	2.00	5.89
		A3 vs A1	4.95	2.38	10.29	3.09	1.77	5.41
	Female, DM	A0 vs A1	21.52	8.84	52.41	12.99	7.73	21.85
		A2 vs A1	3.32	1.54	7.19	6.43	4.04	10.23
		A3 vs A1	8.70	4.43	17.10	6.29	3.89	10.15
	Male, DM	A0 vs A1	20.20	7.96	51.26	7.24	4.02	13.03
		A2 vs A1	3.12	1.37	7.08	3.58	2.08	6.16
		A3 vs A1	8.17	3.93	16.96	3.50	2.01	6.11
Age 70	A0	eGFR*	0.36	0.31	0.41	0.84	0.77	0.90
	A1	eGFR*	0.48	0.44	0.53	0.85	0.80	0.89
	A2	eGFR*	0.50	0.46	0.54	0.89	0.84	0.94
	A3	eGFR*	0.51	0.48	0.54	0.91	0.86	0.97
Age 80	A0	eGFR*	0.35	0.30	0.40	0.82	0.76	0.88
	A1	eGFR*	0.47	0.43	0.52	0.83	0.80	0.86
	A2	eGFR*	0.48	0.44	0.53	0.87	0.83	0.91
	A3	eGFR*	0.49	0.46	0.53	0.90	0.85	0.95
No CV	Female, eGFR 25	Age 70 vs 60	0.65	0.62	0.67	1.75	1.67	1.82
		Age 80 vs 70	0.58	0.55	0.61	1.91	1.85	1.98
	Male, eGFR 25	Age 70 vs 60	0.68	0.66	0.71	1.93	1.85	2.01
		Age 80 vs 70	0.61	0.58	0.64	2.11	2.04	2.19
	Female, eGFR 20	Age 70 vs 60	0.67	0.64	0.70	1.78	1.69	1.88
		Age 80 vs 70	0.60	0.57	0.63	1.96	1.86	2.05
	Male, eGFR 20	Age 70 vs 60	0.70	0.67	0.74	1.97	1.87	2.08
		Age 80 vs 70	0.63	0.60	0.67	2.16	2.05	2.27
CV	Female, eGFR 25	Age 70 vs 60	0.59	0.56	0.61	1.50	1.44	1.56
		Age 80 vs 70	0.53	0.50	0.55	1.64	1.59	1.69
	Male, eGFR 25	Age 70 vs 60	0.62	0.60	0.65	1.66	1.59	1.73
		Age 80 vs 70	0.56	0.53	0.59	1.81	1.76	1.87
	Female, eGFR 20	Age 70 vs 60	0.61	0.58	0.64	1.53	1.45	1.62
		Age 80 vs 70	0.55	0.52	0.58	1.68	1.60	1.76
	Male, eGFR 20	Age 70 vs 60	0.64	0.61	0.67	1.69	1.60	1.78
		Age 80 vs 70	0.58	0.54	0.61	1.85	1.77	1.94

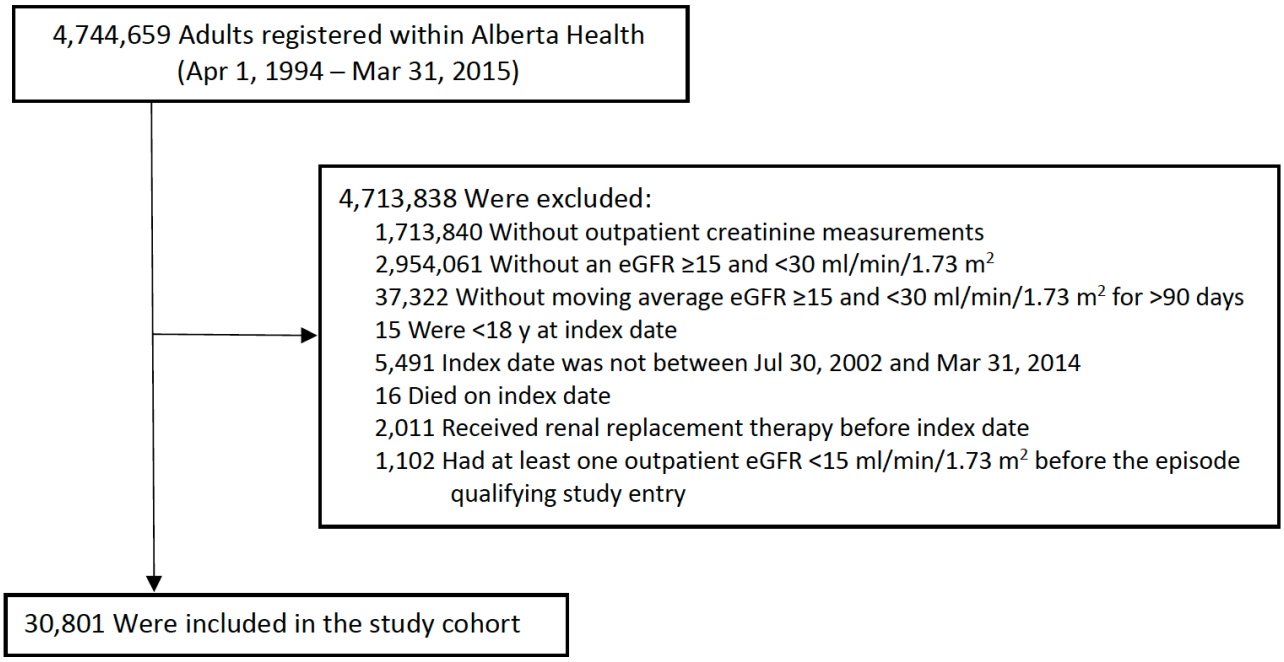
Legend: Flexible parametric model of kidney failure (N=5,511) and death (16,285). HR indicates hazard ratio; L95 and U95 indicate 95% confidence limits; (*) eGFR indicates estimated glomerular-filtration rate (in ml/min/1.73 m^2); A0 (missing), A1, A2, and A3 are categories of albuminuria – see Table 1 for details; CV indicates cardiovascular disease. Age is expressed in years. The model is a proportional hazard model for all included covariates except event type (kidney failure or death), whose effects are time-dependent (the corresponding Cox proportional hazard model is stratified by event type). The model has 54 DF, with event-specific effects specified for all terms, 7 splines for time and 3 for time-dependent effects, allowing different baseline hazard by event type. In addition to the main terms (age, eGFR, albuminuria category, male sex, diabetes and cardiovascular disease) the model includes a squared term (for age) and the following interactions: ageXeGFR, albuminuriaXeGFR, DMXalbuminuria, CVXalbuminuria, ageXmale and ageXCV.

sTable 4: Cause-specific hazard semi-parametric model of kidney failure and death

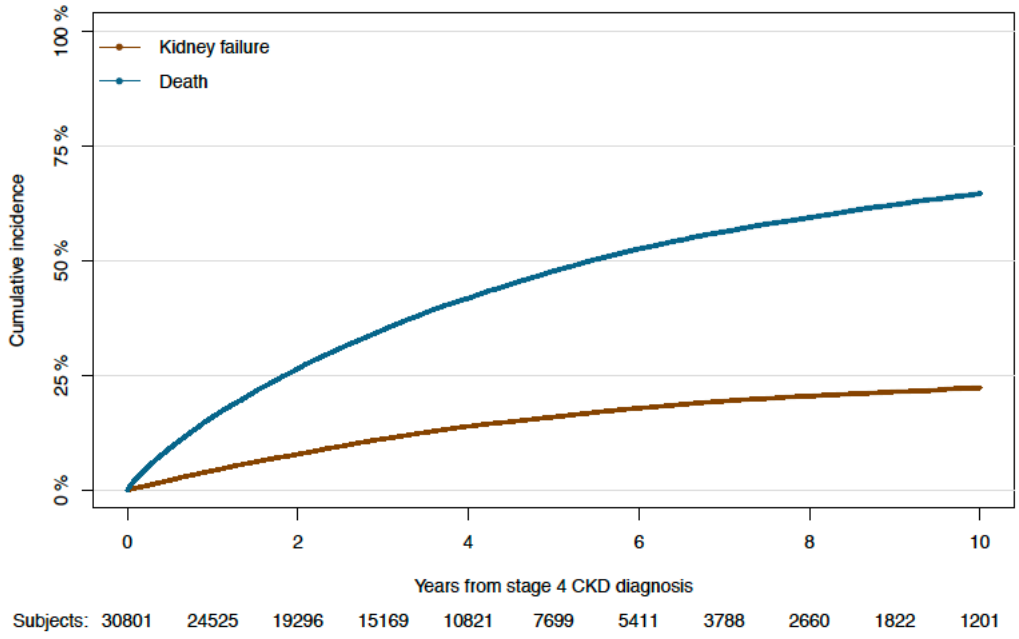
			Kidney failure			Death		
Group		Comparison	HR	L95	U95	HR	L95	U95
No CV	Female, no DM	A0 vs A1	6.48	2.89	14.52	1.34	0.88	2.02
		A2 vs A1	1.55	0.79	3.05	0.97	0.68	1.36
		A3 vs A1	2.72	1.52	4.88	0.91	0.62	1.31
	Male, no DM	A0 vs A1	6.07	2.59	14.19	0.75	0.45	1.23
		A2 vs A1	1.46	0.70	3.02	0.54	0.34	0.84
		A3 vs A1	2.55	1.34	4.85	0.51	0.31	0.81
	Female, DM	A0 vs A1	12.51	5.65	27.69	1.92	1.27	2.91
		A2 vs A1	1.61	0.81	3.19	1.01	0.71	1.43
		A3 vs A1	4.48	2.50	8.05	1.03	0.70	1.49
	Male, DM	A0 vs A1	11.72	5.07	27.11	1.07	0.65	1.78
		A2 vs A1	1.51	0.72	3.15	0.56	0.36	0.88
		A3 vs A1	4.20	2.20	8.01	0.57	0.36	0.92
CV	Female, no DM	A0 vs A1	11.00	4.41	27.45	9.07	5.38	15.28
		A2 vs A1	3.19	1.48	6.89	6.19	3.90	9.83
		A3 vs A1	5.26	2.67	10.35	5.57	3.44	9.01
	Male, no DM	A0 vs A1	10.30	3.97	26.74	5.06	2.81	9.12
		A2 vs A1	2.99	1.33	6.76	3.45	2.01	5.93
		A3 vs A1	4.92	2.37	10.24	3.11	1.78	5.44
	Female, DM	A0 vs A1	21.25	8.73	51.74	13.06	7.77	21.97
		A2 vs A1	3.31	1.53	7.15	6.46	4.06	10.28
		A3 vs A1	8.67	4.41	17.02	6.30	3.90	10.18
	Male, DM	A0 vs A1	19.91	7.85	50.48	7.29	4.05	13.13
		A2 vs A1	3.10	1.37	7.03	3.60	2.09	6.20
		A3 vs A1	8.12	3.91	16.85	3.52	2.01	6.14
Age 70	A0	eGFR*	0.36	0.32	0.41	0.84	0.77	0.90
	A1	eGFR*	0.48	0.44	0.53	0.85	0.80	0.89
	A2	eGFR*	0.50	0.46	0.54	0.89	0.84	0.94
	A3	eGFR*	0.51	0.48	0.54	0.92	0.86	0.97
Age 80	A0	eGFR*	0.35	0.30	0.40	0.82	0.76	0.88
	A1	eGFR*	0.47	0.43	0.52	0.83	0.80	0.86
	A2	eGFR*	0.48	0.44	0.53	0.87	0.83	0.91
	A3	eGFR*	0.49	0.46	0.53	0.90	0.85	0.95
No CV	Female, eGFR 25	Age 70 vs 60	0.65	0.62	0.67	1.75	1.68	1.82
		Age 80 vs 70	0.58	0.55	0.61	1.91	1.85	1.98
	Male, eGFR 25	Age 70 vs 60	0.68	0.66	0.71	1.93	1.85	2.01
		Age 80 vs 70	0.61	0.53	0.64	2.11	2.04	2.19
	Female, eGFR 20	Age 70 vs 60	0.67	0.64	0.70	1.79	1.69	1.88
		Age 80 vs 70	0.60	0.57	0.63	1.96	1.86	2.05
	Male, eGFR 20	Age 70 vs 60	0.70	0.67	0.74	1.97	1.87	2.08
		Age 80 vs 70	0.63	0.60	0.67	2.16	2.06	2.27
CV	Female, eGFR 25	Age 70 vs 60	0.59	0.56	0.61	1.50	1.44	1.57
		Age 80 vs 70	0.53	0.50	0.55	1.64	1.59	1.69
	Male, eGFR 25	Age 70 vs 60	0.62	0.60	0.65	1.66	1.59	1.73
		Age 80 vs 70	0.56	0.53	0.59	1.81	1.76	1.87
	Female, eGFR 20	Age 70 vs 60	0.61	0.58	0.64	1.53	1.45	1.62
		Age 80 vs 70	0.55	0.52	0.58	1.68	1.60	1.76
	Male, eGFR 20	Age 70 vs 60	0.64	0.61	0.67	1.69	1.60	1.78
		Age 80 vs 70	0.58	0.55	0.61	1.85	1.77	1.94

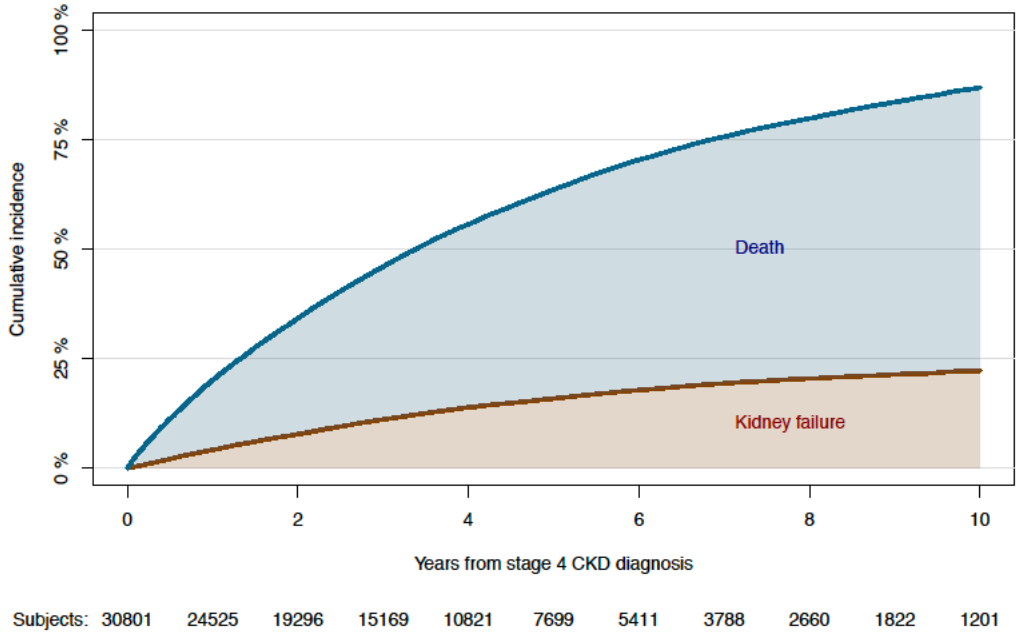
Legend: Cox model of kidney failure (N=5,511) and death (16,285). HR indicates hazard ratio; L95 and U95 indicate 95% confidence limits; (*) eGFR indicates estimated glomerular-filtration rate (in ml/min/1.73 m^2); A0 (missing), A1, A2, and A3 are categories of albuminuria – see Table 1 for details; CV indicates cardiovascular disease. Age is expressed in years. The model is a proportional hazard model for all included covariates, stratified by event type (kidney failure or death). In addition to the main terms (age, eGFR, albuminuria category, male sex, diabetes and cardiovascular disease) the model includes a squared term (for age) and the following interactions: ageXeGFR, albuminuriaXeGFR, DMXalbuminuria, CVXalbuminuria, ageXmale and ageXCV.

sFigure 1: Derivation of study cohort



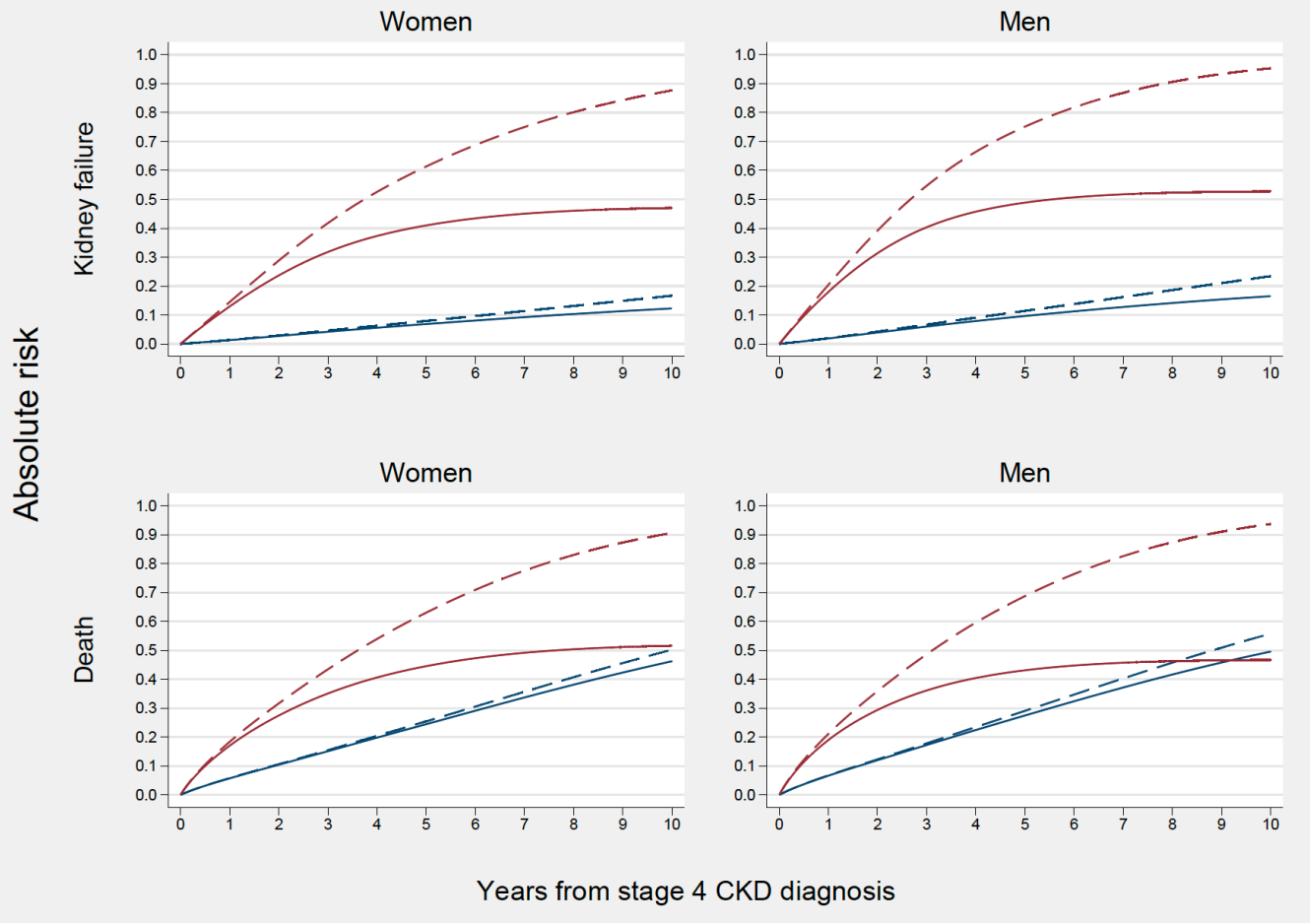
sFigure 2: Overlaid and stacked cumulative incidence functions





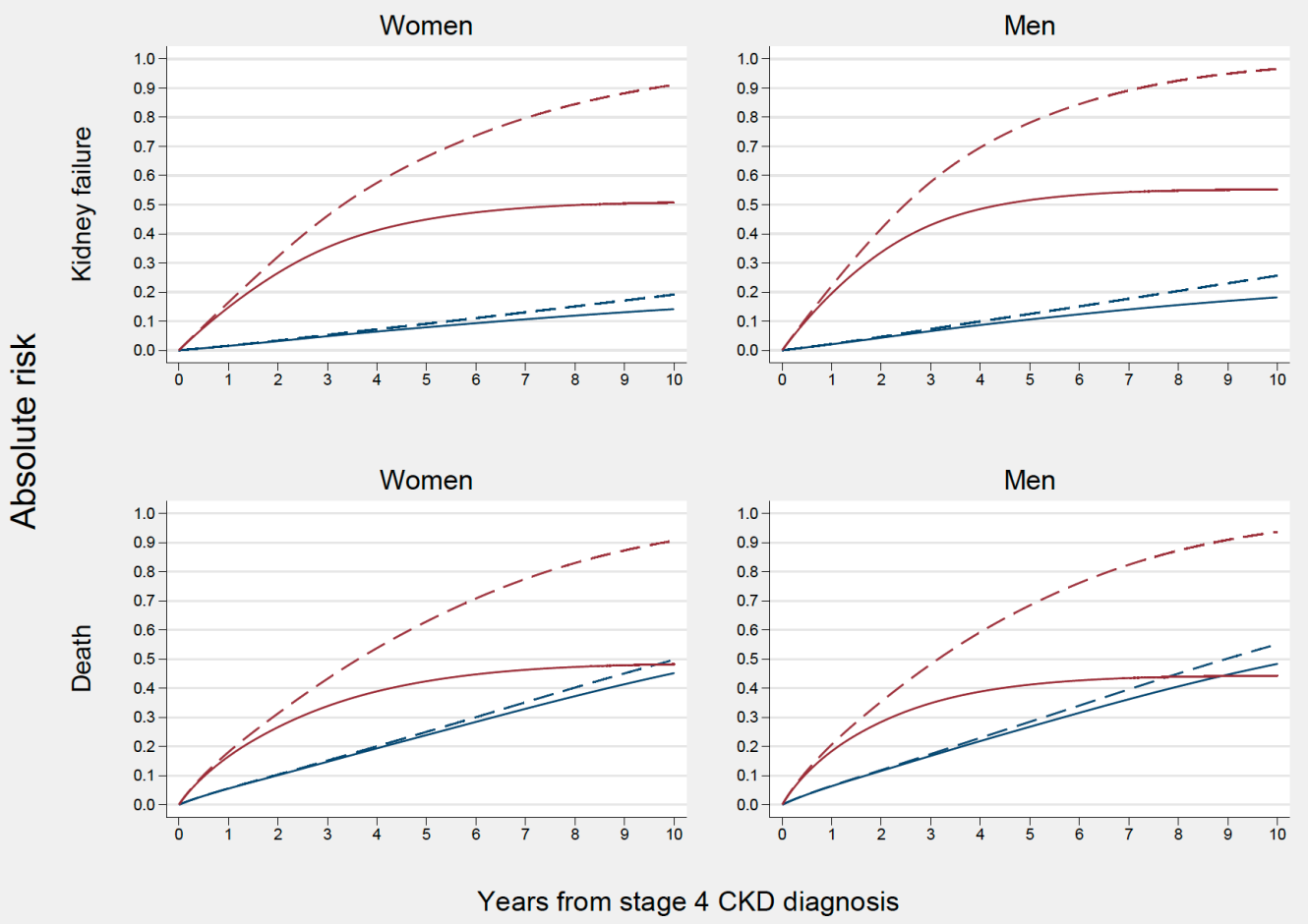
Legend: Crude cumulative incidence functions (Aalen–Johansen estimator).

sFigure 3: Overlaid model-based cumulative and naïve incidence functions (sensitivity analysis 1)



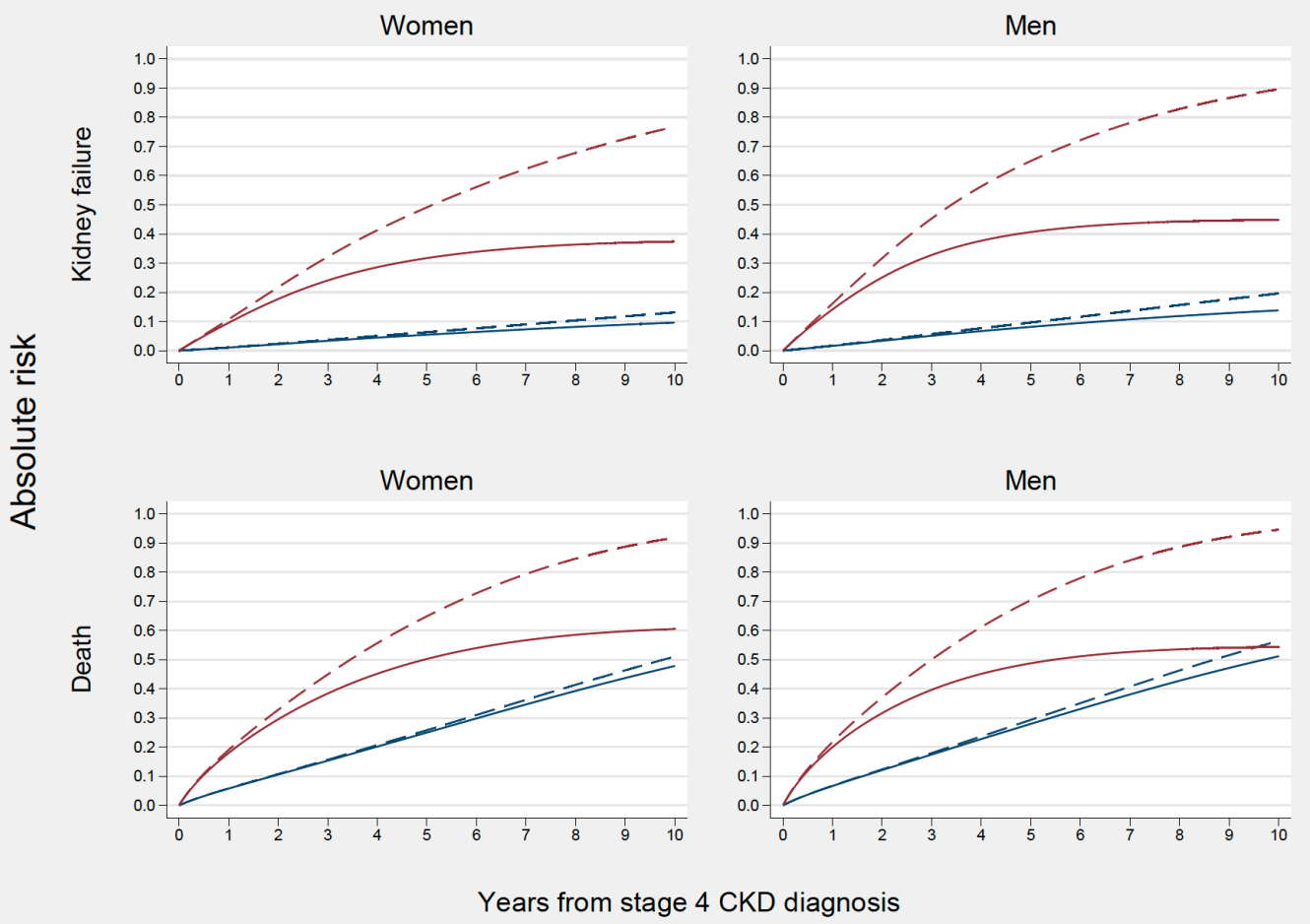
Legend: Cumulative incidence functions (solid lines) and naïve incidence functions (dashed lines) estimated from the final cause-specific hazard model of kidney failure or death (sTable 3). Maroon lines: high-risk person (person with diabetes, cardiovascular disease, albuminuria >300 mg/day and with an eGFR of 20 ml/min/m^2); navy lines: low-risk person (person without diabetes or cardiovascular disease, and with an eGFR of 25 ml/min/m^2 and albuminuria <30 mg/day). In all cases age is 75 years. In this sensitivity analysis, kidney failure was defined as initiation of renal replacement therapy or sustained eGFR<10 ml/min/1.73 m^2, instead of initiation of renal replacement therapy or moving average eGFR <10 ml/min/1.73 m^2 (5,253 kidney failure events and 16,472 death events; person-year at risk: 107,080).

sFigure 4: Overlaid model-based cumulative and naïve incidence functions (sensitivity analysis 2)



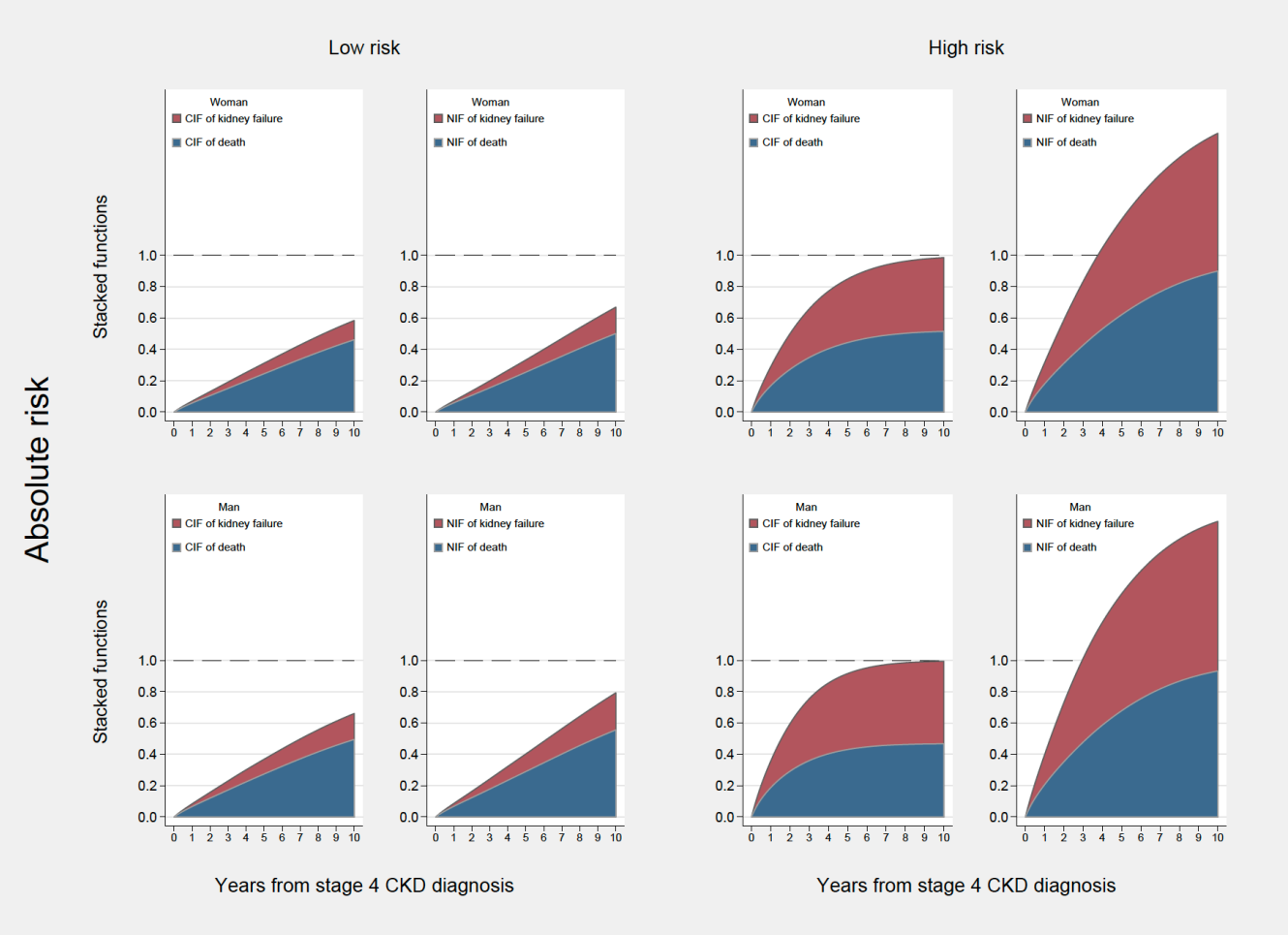
Legend: Cumulative incidence functions (solid lines) and naïve incidence functions (dashed lines) estimated from the final cause-specific hazard model of kidney failure or death (sTable 3). Maroon lines: high-risk person (person with diabetes, cardiovascular disease, albuminuria >300 mg/day and with an eGFR of 20 ml/min/m^2); navy lines: low-risk person (person without diabetes or cardiovascular disease, and with an eGFR of 25 ml/min/m^2 and albuminuria <30 mg/day). In all cases age is 75 years. In this sensitivity analysis, we excluded participants with incomplete data on albuminuria (analysis restricted to complete cases, N=27,823; 5,181 kidney failure events and 14,247 death events; person-year at risk: 97,731).

sFigure 5: Overlaid model-based cumulative and naïve incidence functions (sensitivity analysis 3)



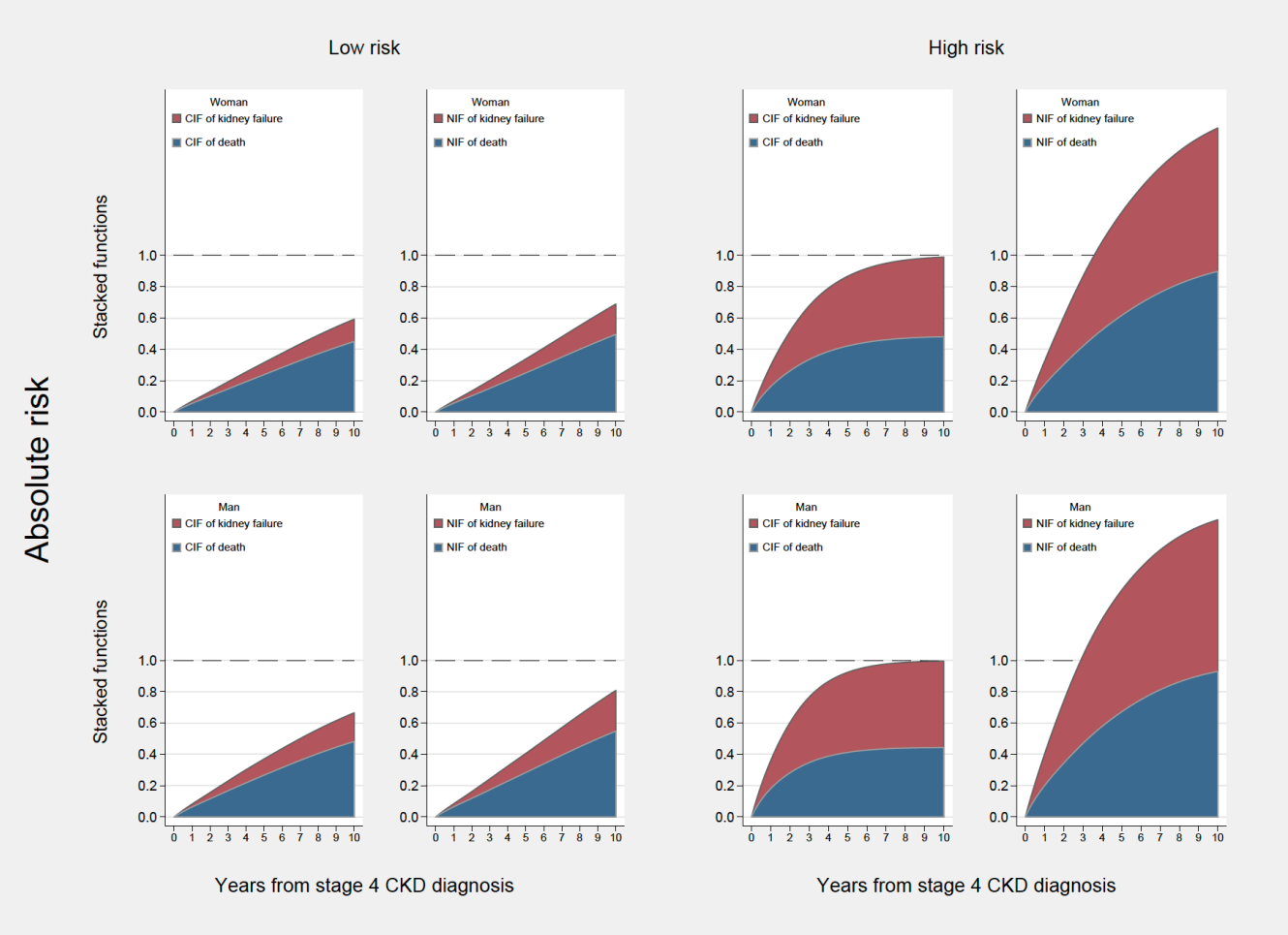
Legend: Cumulative incidence functions (solid lines) and naïve incidence functions (dashed lines) estimated from the final cause-specific hazard model of kidney failure or death (sTable 3). Maroon lines: high-risk person (person with diabetes, cardiovascular disease, albuminuria >300 mg/day and with an eGFR of 20 ml/min/m^2); navy lines: low-risk person (person without diabetes or cardiovascular disease, and with an eGFR of 25 ml/min/m^2 and albuminuria <30 mg/day). In all cases age is 75 years. In this sensitivity analysis, kidney failure was defined as initiation of renal replacement therapy, instead of initiation of renal replacement therapy or moving average eGFR <10 ml/min/1.73 m^2 (4,758 kidney failure events and 16,826 death events; person-year at risk: 107,979).

sFigure 6: Stacked model-based cumulative and naïve incidence functions (sensitivity analysis 1)



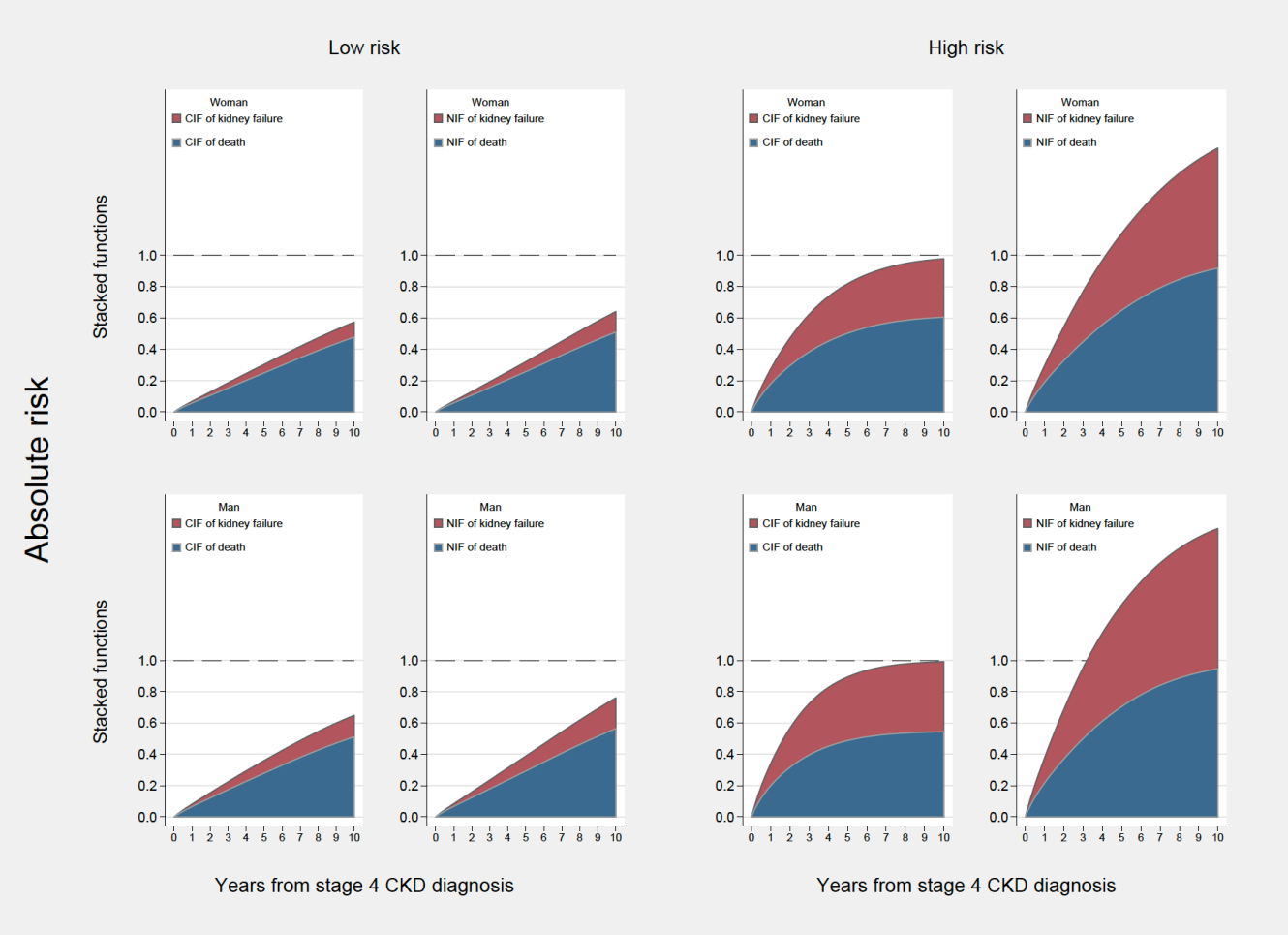
Legend: Cumulative incidence functions and naïve incidence functions estimated from the final cause-specific hazard model of kidney failure or death (sTable 3). The reference line indicates maximum possible risk. Top panels: female sex; bottom panels: male sex. Left panels: low-risk person (person without diabetes or cardiovascular disease, and with an eGFR of 25 ml/min/m^2 and albuminuria <30 mg/day); right panels: high-risk person (person with diabetes and cardiovascular disease, with an eGFR of 20 ml/min/m^2 and albuminuria >300 mg/day). In all cases age is 75 years. In this sensitivity analysis, kidney failure was defined as initiation of renal replacement therapy or sustained eGFR<10 ml/min/1.73 m^2, instead of initiation of renal replacement therapy or moving average eGFR <10 ml/min/1.73 m^2 (5,253 kidney failure events and 16,472 death events; person-year at risk: 107,080).

sFigure 7: Stacked model-based cumulative and naïve incidence functions (sensitivity analysis 2)



Legend: Cumulative incidence functions and naïve incidence functions estimated from the final cause-specific hazard model of kidney failure or death (sTable 3). The reference line indicates maximum possible risk. Top panels: female sex; bottom panels: male sex. Left panels: low-risk person (person without diabetes or cardiovascular disease, and with an eGFR of 25 ml/min/m^2 and albuminuria <30 mg/day); right panels: high-risk person (person with diabetes and cardiovascular disease, with an eGFR of 20 ml/min/m^2 and albuminuria >300 mg/day). In all cases age is 75 years. In this sensitivity analysis, we excluded participants with incomplete data on albuminuria (analysis restricted to complete cases, N=27,823; 5,181 kidney failure events and 14,247 death events; person-year at risk: 97,731).

sFigure 8: Stacked model-based cumulative and naïve incidence functions (sensitivity analysis 3)



Legend: Cumulative incidence functions and naïve incidence functions estimated from the final cause-specific hazard model of kidney failure or death (sTable 3). The reference line indicates maximum possible risk. Top panels: female sex; bottom panels: male sex. Left panels: low-risk person (person without diabetes or cardiovascular disease, and with an eGFR of 25 ml/min/m^2 and albuminuria <30 mg/day); right panels: high-risk person (person with diabetes and cardiovascular disease, with an eGFR of 20 ml/min/m^2 and albuminuria >300 mg/day). In all cases age is 75 years. In this sensitivity analysis, kidney failure was defined as initiation of renal replacement therapy, instead of initiation of renal replacement therapy or moving average eGFR <10 ml/min/1.73 m^2 (4,758 kidney failure events and 16,826 death events; person-year at risk: 107,979).

References

1. Hemmelgarn BR, Clement F, Manns BJ, et al. Overview of the Alberta Kidney Disease Network. *BMC Nephrol.* 2009;10:30.
2. Coviello V, Boggess M. Cumulative incidence estimation in the presence of competing risks. *Stata Journal.* 2004;4(2):103-112.
3. Lambert PC, Royston P. Further development of flexible parametric models for survival analysis. *Stata J.* 2009;9(2):265-290.
4. Lambert PC, Wilkes SR, Crowther MJ. Flexible parametric modelling of the cause-specific cumulative incidence function. *Stat Med.* 2017;36(9):1429-1446.
5. Hinchliffe SR, Lambert PC. Extending the flexible parametric survival model for competing risks. *Stata J.* 2013;13(2):344-355.
6. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine.* 2007;26(11):2389-2430.
7. Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics.* 1995;51(2):524-532.
8. Mozumder SI, Rutherford MJ, Lambert PC. stpm2cr: A flexible parametric competing risks model using a direct likelihood approach for the cause-specific cumulative incidence function. *Stata J.* 2017;17(2):462-489.
9. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association.* 1999;94(446):496-509.
10. Lambert PC. The estimation and modelling of cause-specific cumulative incidence functions using time-dependent weights. *Stata J.* 2017;17(1):181-207.