Challenges of studying and predicting chronic kidney disease progression and its complications using routinely collected electronic healthcare records

FAYE CLEARY

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy
of the
University of London

NOVEMBER 2024

Department of Non-Communicable Disease Epidemiology

Faculty of Epidemiology and Population Health

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by Medical Research Council London Intercollegiate Doctoral Training Partnership

Declaration

I, Faye Cleary, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Chronic kidney disease (CKD) affects over 10% of the population worldwide. Complications include cardiovascular morbidity, death, and progression to kidney failure requiring dialysis. The number of CKD cases is growing, with implications for patients and healthcare services. Action is needed to support earlier diagnosis and better targeted care with the potential to delay disease progression and reduce associated complications.

Increasing availability of electronic healthcare records (EHRs) can support study of CKD and its progression in the general population. Regular kidney function testing which is recommended in clinical practice has the potential to capture patients with CKD and subsequent progression of disease. Large sample sizes and long duration of follow-up can support study of rare outcomes such as kidney failure and decline in kidney function which may progress slowly over many years in some patients.

However, there is variation in recognition of CKD in clinical practice and there are challenges in detecting CKD due to its asymptomatic nature in the early stages, relying on blood tests to detect. Availability of kidney function test results is likely to depend on patient risk factors, healthcare seeking behaviours and healthcare provider factors ("informative testing"). This may lead to selection bias and impact reliability of research findings.

This thesis aimed to explore and highlight the challenges resulting from issues of data quality and completeness inherent to EHRs when used to study the epidemiology of progression of CKD, and to present approaches to overcome these challenges.

Firstly, a systematic review showed substantial risks of selection bias in previous research, due to selection procedures for study inclusion and completeness of data captured during follow-up for outcomes. Generally, large proportions of patients were excluded from analysis due to missing data, with little reflection on the implications of bias in study results and unrepresentative samples. Statistical methodology varied widely, with varying capability of handling missing data.

Secondly, a feasibility study investigated data quality and completeness for kidney function tests conducted in UK primary care. Testing was uncommon in adults overall, but there was high frequency of repeat testing in patients with risk factors for CKD, with the potential to capture most patients with CKD. However, reasons for missing data weren't clear, and data

may be disproportionately missing due to certain risk factors or management in secondary care. Data quality issues due to historical laboratory reporting problems led to under-estimation of decline in kidney function but impact was small in most patients.

Thirdly, we studied the association between GP practice completeness of diagnostic coding for CKD and patient-level hospitalisation outcomes in patients with CKD in England. The use of a practice-level exposure aimed to reduce risks of unmeasured confounding. Being registered at a higher coding practice was associated with lower rates of hospitalisations for CV events, after adjustment for other practice factors.

Finally, we developed new risk prediction equations for kidney failure requiring dialysis, using EHRs capturing the entire healthcare system in Stockholm, Sweden. Previously validated equations require data that is not routinely collected in most patients, but our analysis included 98% of patients identified with CKD, by including predictor variables that are routinely available. New models were precisely estimated and achieved high discrimination.

EHRs hold huge value to study progression of CKD due to large sample size and long duration of regularly collected kidney function tests. However, issues of informative missingness and sampling bias have not been appropriately acknowledged and addressed in previous research. Future research should ensure that research questions can be answered with available data using appropriate statistical techniques, and with improved transparency of potential for selection bias. Research in this thesis has strengthened evidence for the importance of diagnostic coding for CKD in clinical practice in reducing risk of complications of CKD, by enabling improved patient care. New risk models have the potential to improve equality of healthcare, enabling risk prediction in all patients with CKD, but require validation in the UK.

Acknowledgements

I would like to express thanks to my supervisors, Dorothea Nitsch and David Prieto-Merino, who have supported me throughout my research studies, and have provided expert guidance to enhance the quality of my research. It has been a privilege to learn from them both.

I would like to thank the Medical Research Council for funding my research and the LSHTM Epidemiology and Population Health Equality and Diversity Committee for further financial support which enabled me to manage working with a young baby and support with childcare costs.

I would like to thank LSHTM support staff including Lauren Dalton, Jenny Fleming, Lara Crawford and student support services.

I thank Richard Silverwood and Sally Hull for providing feedback on my research plans through my upgrading seminar, and guiding me to re-direct my research plans appropriately.

I am thankful to Juan-Jesus Carrero and colleagues who invited me to the Karolinska Institute and shared their data with me, which enhanced the quality and findings of my research.

Thanks also to all the co-authors I have had the pleasure of working with and learning from over the last few years.

Finally, thank you to my husband Jay for his constant unwavering support, as we have seen through the COVID pandemic, getting married, a pregnancy, and our first baby, through the course of my studies, and who has acted as chief breadwinner to support me and our family, enabling me to pursue my research studies. He is the absolute best! Plus, of course, thanks to my delightful 2-year-old Lenny, for bringing a lot of joy to my world and helping me to keep perspective when PhD life has occasionally gotten rocky and overwhelming… Thanks also to my mum for extra free childcare!

Acronyms

ACEis	angiotensin converting enzyme inhibitors
(u)ACR	(urinary) albumin:creatinine ratio
AF	atrial fibrillation
AKI	acute kidney injury
ARBs	angiotensin-receptor blockers
BP	blood pressure
CHD	coronary heart disease
CI	confidence interval
CKD	chronic kidney disease
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration (equation)
CKD-PC	Chronic Kidney Disease Prognosis Consortium
CNIs	calcineurin inhibitors
COPD	chronic obstructive pulmonary disease
COVID-19	coronavirus 19
CPRD	Clinical Practice Research Datalink (database)
C-statistic	concordance statistic
CV(D)	cardiovascular (disease)
DBP	diastolic blood pressure
ESKD	end-stage kidney disease (same meaning as end-stage renal disease)
ESRD	end-stage renal disease (same meaning as end-stage kidney disease)
eGFR	estimated glomerular filtration rate
EHRs	electronic healthcare records
FDA	(United States) Food and Drug Administration
GP	general practitioner
GFR	glomerular filtration rate
HES	Hospital Episodes Statistics
HF	heart failure
HR	hazard ratio
ICD-10	International Classification of Diseases version 10
IDI	integrated discrimination improvement
IDMS	isotope dilution mass spectrometry

IMD	Index of Multiple Deprivation
IQR	inter-quartile range
KDIGO	Kidney Disease Improving Global Outcomes
KDOQI	(United States) Kidney Disease Outcomes Quality Initiative (group)
KFRE	kidney failure risk equation
KM	Kaplan-Meier
KRT	kidney replacement therapy (same meaning as renal replacement therapy)
LOCF	last observation carried forward
LSHTM	London School of Hygiene & Tropical Medicine
MAR	missing at random
MCAR	missing completely at random
MDRD	Modification of Diet in Renal Disease (study equation)
MHRA	Medicines and Healthcare products Regulatory Agency
MNAR	missing not at random
NCKDA	National Chronic Kidney Disease Audit
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NOMESCO	Nordic Medico-Statistical Committee (procedure codes)
NRI	net reclassification index
NSAIDS	non-steroidal anti-inflammatory drugs
ONS	Office for National Statistics
PAD	peripheral artery disease
PH	proportional hazards
PPI	patient and public involvement
QOF	Quality and Outcomes Framework
(u)PCR	(urinary) protein:creatinine ratio
RCT	randomised controlled trial
RECORD	REporting of studies Conducted using Observational Routinely-collected Data (guidelines)
RRT	renal replacement therapy (same meaning as kidney replacement therapy)
SBP	systolic blood pressure
SCREAM	Stockholm CREAtinine Measurements (database)
SD	standard deviation
SGLT2(i)	Sodium-glucose co-transporter-2 (inhibitors)

SNOMED CT	Systematized Nomenclature of Medicine Clinical Terms
SRR	Swedish Renal Registry
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology (guidelines)
UKRR	UK Renal Registry

Table of Contents

List of tables

Additional tables are included within research papers and are separately referenced within-in those papers in the relevant research paper chapters. Supplementary tables attached to research articles are included in appendices, referenced within the research papers, with appendix location listed in the contents page.

List of figures

Additional figures are included within research papers and are separately referenced within-in those papers in the relevant research paper chapters. Supplementary figures attached to research articles are included in appendices, referenced within the research papers, with appendix location listed in the contents page.

Chapter 1

Background

1.1 Chapter summary

This chapter lays out the background of the importance of improving knowledge in the epidemiology of chronic kidney disease (CKD) and its progression, how the use of electronic healthcare records (EHRs) can support this aim, and key analytical challenges that are faced. Aims and objectives of the thesis are presented in *Chapter 2*.

1.2 Overview of epidemiology of CKD

Disease burden

CKD (stages 1-5) affects more than 10% of the population in both developed and developing countries [1-2] and studies have shown that CKD is one of the leading causes of death globally [3-4]. In 2016, the Health Survey for England (using nationally representative data) estimated prevalence of CKD stages 1-5 at 12.7% and prevalence of CKD stages 3-5 at 5.1% among adults aged 16 and over [5]. These estimates are now out of date, and current prevalence is uncertain. The total burden of CKD is growing in the UK and globally due to a combination of factors, including increasing prevalence of risk factors, population growth and an ageing population [6]. While the number of cases of CKD is increasing, so to is the proportion of patients experiencing later-stage disease (mostly due to higher prevalence of CKD in the elderly), which is accompanied by greater risk to health [7]. CKD is projected to become the fifth highest cause of years of life lost globally by 2040 [8].

Causes and mechanism of disease

CKD involves gradual loss of kidney function over time, sustained due to various causes and with various mechanisms involved [9]. Most common causes include diabetes, hypertension, cardiovascular disease, and obesity, demonstrating the important role played by lifestyle factors in the onset and progression of CKD. Other less common causes include genetic kidney diseases and use of nephrotoxic drugs, such as calcineurin inhibitors (CNIs) and

lithium. Depending on the cause, different mechanisms of damage may occur. Diabetes and hypertension are the leading causes of CKD globally, and account for approximately two-thirds of all cases [1]. High blood sugar resulting from diabetes can damage blood vessels in the kidneys, resulting in damage to filters and leakage of substances such as protein [10]. High blood pressure puts excessive pressure on already damaged kidneys, further damaging and weakening the blood vessels [11].

Complications of CKD

Complications of CKD include increased cardiovascular (CV) risk, mortality, acute kidney injury (AKI) and increased susceptibility to infection [12-13]. In rare cases, CKD may progress to kidney failure [14], requiring routine dialysis or kidney transplantation in order to sustain life (although some patients may undergo conservative care). Risks of adverse outcomes associated with CKD are closely related to disease severity, with worse outcomes resulting from further progressed disease [12]. Quality of life is also significantly reduced with later stage disease, particularly in the case of kidney failure, due to requirements to spend significant periods of daily life in clinical facilities connected to dialysis machines, and substantial symptom burden, which may include considerable fatigue, drowsiness and headaches [7]. Further details on prognostic factors for complications of CKD are detailed in *Section 1.3 (Definition and classification of CKD)*.

Economic burden

In addition to burden on patient health, CKD presents a high economic burden to healthcare services [7,15]. The bulk of this cost is due to kidney replacement therapy (KRT) which is required in a minority of patients experiencing kidney failure but is very expensive, with most patients requiring 4 hours of dialysis, 3 times per week [7]. In the UK, routine dialysis was estimated to cost the NHS approximately £34,000 per patient per year in 2023 [7]. Approximately 30,000 patients in the UK currently require dialysis due to kidney failure, and this figure is expected to rise substantially in the coming years, leading to increased costs [7,14]. CKD is currently estimated to account for 3.2% of total NHS costs, with further costs to the general economy resulting from CKD, such as reduced capacity to work [7].

Recognition of importance of CKD

Despite the growing burden and important consequences of CKD, recognition of the importance of CKD among healthcare professionals is variable, and generally quite poor,

with considerable opportunity for improvement [16]. Research spending is also relatively low, when considering the magnitude of impact of the disease to public health [7]. It is thought that early diagnosis and appropriate management may be able to prevent or slow down disease progression [17]. However, CKD is often asymptomatic in the early stages, leading to challenges in detection and diagnosis in routine care (*further detailed in Section 1.3*). Testing efforts in those at risk are required to confirm and diagnose CKD [18], requiring patient presentation to healthcare services. Consequently, challenges in detection disproportionately affect certain socioeconomic groups, leading to health inequalities [19].

More research is needed on the burden, causes, and consequences of CKD and its progression that reflects the population of patients with CKD, as well as investment in development of new treatments [7]. Such research will improve our ability to maintain appropriate management strategies capable of reducing the impact of CKD on public health (and associated cost implications). There is a particular need to improve efforts to personalise care, targeting monitoring and treatment efforts to those patients who will receive greatest benefit [9].

1.3 Definition and classification of CKD

History of CKD guideline development

International guidelines have been developed to support evaluation and management of CKD by the Kidney Disease: Improving Global Outcomes (KDIGO) organisation [9]. These guidelines provide definitions for CKD and disease staging, which have evolved through the course of guideline updating [9,20-21]. They require assessment of laboratory parameters which are indicative of disease severity. This includes estimated glomerular filtration rate (eGFR) from blood tests and detection of albuminuria (presence of albumin protein in the urine) through urine samples (preferably using urine albumin:creatinine ratio [uACR]) [9,18]. More details on these parameters and how they are used as prognostic markers for CKD are provided later in this section.

The US-based Kidney Disease Outcomes Quality Initiative [KDOQI] 2002 CKD guideline (also adopted by the international community) initially defined 5 categories of CKD severity based on eGFR (and a chronicity criterion), enabling improved focus on earlier stages of disease aimed at improving early CKD detection, where previous focus was on late stage

disease [22-23]. Increasing international interest led to development of the KDIGO 2012 CKD guideline, extending staging criteria to include categories of albuminuria, known to be associated with risks of key adverse outcomes [24]. Finally, the most recent KDIGO CKD guideline was published in 2024, particularly emphasising the importance of individual risk prediction [9]. Specific UK guidelines are also provided by the National Institute for Health and Care Excellence (NICE), most recently updated in 2021 [18].

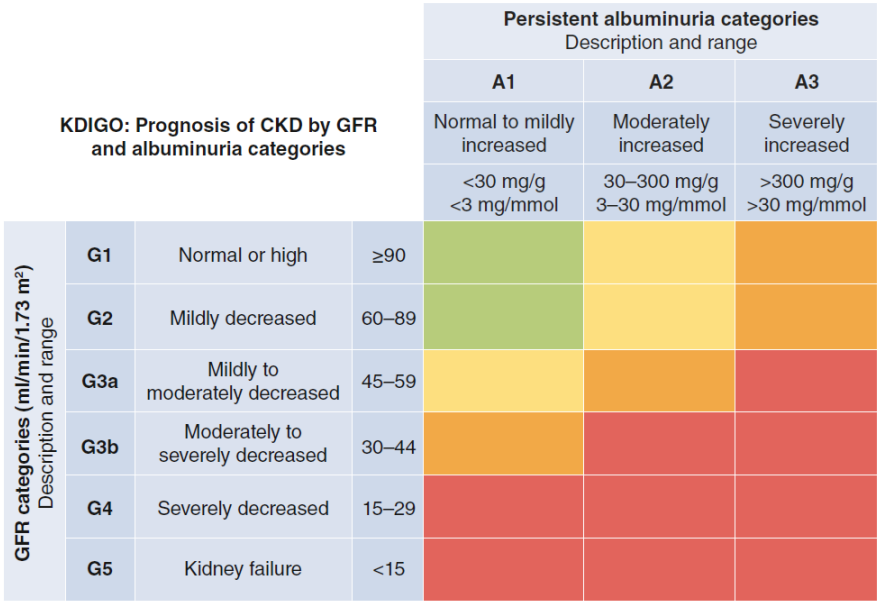
Standard definition and disease staging

CKD is defined as:

> "abnormalities of kidney structure or function, present for a minimum of 3 months, with implications for health. CKD is classified based on **C**ause, **G**lomerular filtration rate (GFR) category (G1–G5), and **A**lbuminuria category (A1–A3), abbreviated as **CGA**." *(KDIGO 2024)*

Both eGFR and albuminuria are independently associated with outcomes of people with CKD [9]. KDIGO disease staging criteria are displayed in **Figure 2**. The heat map shown broadly indicates risks of numerous important adverse outcomes (including CV risk, mortality, CKD progression, kidney failure, AKI), on an aggregate population level basis.

Figure 2. Categories of severity of CKD in terms of eGFR and uACR, with heat map broadly indicating risks of numerous adverse outcomes (CV risk, mortality, CKD progression, kidney failure, AKI)



Green: low risk (if no other markers of kidney disease, no CKD); Yellow: moderately increased risk; Orange: high risk; Red: very high risk. GFR, glomerular filtration rate.

This figure is reproduced from KIDGO 2024 clinical practice guidelines [9].

It is important to note that individual risks for any specific adverse event may vary substantially within any category, and the use of risk prediction equations is recommended to estimate individual risks to inform clinical decision-making, taking into account various patient factors influencing risks [9].

Laboratory parameters for CKD staging: eGFR and uACR

We have detailed how the use of laboratory parameters, eGFR and uACR, aid in definition, staging and subsequent prognostication of important outcomes associated with CKD. We now explain what these laboratory markers are and how they may be measured.

eGFR

As the term "glomerular filtration rate" suggests, GFR is the rate at which the kidneys filter the blood. ("Glomeruli" constitute the network of tiny blood vessels in the kidneys which are responsible for filtration.) This is not straightforward to measure clinically, and equations have been developed to estimate GFR ("eGFR") based on serum biomarkers, called GFR-estimating equations [25-27]. eGFR is reported in units of ml/min/1.73m^2 (volume filtered, per unit of time, per body surface area), and although not factually correct, it is often thought of as "percentage of kidney function" due to the scale of measurement, with normal results equating to approximately 100 ml/min/1.73m^2 [28]. This can be useful to facilitate a basic understanding of the concept of GFR.

eGFR <15 ml/min/1.73m^2 (GFR category G5; CKD stage 5) indicates kidney failure. Earlier stages are categorised as G1 to G4 (CKD stages 1-4) representing less severe losses in kidney function (**Figure 2**). Other (non-GFR) evidence of kidney damage is required to confirm CKD for stages 1-2.

This PhD primarily involved research focused on CKD stages 3-5 (defined by 2 eGFR measures <60 ml/min/1.73m^2 separated by a minimum of 90 days), where kidney function is at least mildly to moderately decreased. The chronicity criterion requiring reduced eGFR over at least 3 months aims to ensure capture of irreversible decline in kidney function, as opposed to acute changes resulting from short-term illness that may be reversible and can be a "red herring" when it comes to establishing kidney disease and severity of disease. Details of GFR-estimating equations and underlying serum biomarkers are provided in *Section 1.5 (Measuring prognostic markers required for evaluation of CKD and CKD progression)*.

uACR

Albuminuria is a marker of kidney damage indicating increased glomerular permeability that results in leakage of the albumin protein from the kidneys to the urine [24,29]. It is more common with certain causal mechanisms of kidney damage (e.g. resulting from diabetes) [24]. Albuminuria is captured by a urine test, with higher levels of urine albumin indicating worse kidney damage. The recommended measure is uACR.

Alternatives urinary measures are protein:creatinine ratio (PCR) and dipstick tests (which are less precise). Conversion equations have been developed to approximate uACR using PCR and dipstick results, with reasonable performance [30].

Detection and diagnosis challenges

There are challenges faced in detection and diagnosis of CKDdue to its asymptomatic nature in the early stages, and many patients may not seek care until later stages of disease when prognosis is worsened. CKD is therefore commonly undiagnosed, particularly in early stages, with increasing likelihood of diagnosis as disease progresses [31]. The 2016 National Chronic Kidney Disease Audit (NCKDA) in England and Wales showed significant variation in CKD diagnosis between primary care practices, measured by completeness of electronic diagnostic coding for CKD in patients with evidence of CKD stages 3-5 confirmed by laboratory parameters (2 x eGFR<60, over 90+ days) [16]. Approximately 30% of CKD patients were uncoded overall.

1.4 Recommended management of CKD

General principles of guidelines to support the management of CKD are around monitoring, treatment and referral to specialist care, with a particular emphasis on treatment efforts appropriate to underlying disease aetiology [9]. In the UK, the vast majority of patients with CKD are managed in primary care, with referral to specialist care in specific cases requiring more focussed care.

Monitoring

Due to the challenges of asymptomatic disease, testing for CKD is recommended for patients with risk factors for CKD [18]. Once CKD has been established based on repeat eGFR test results, regular monitoring of eGFR and albuminuria is advised with frequency depending on severity of disease and risk of disease progression [9,18]. NICE guidelines for recommended

numbers of annual tests are shown in **Table 1**. The UK Quality and Outcomes Framework (QOF) previously incentivised testing for CKD, resulting in improved recognition around testing for CKD and increases in frequency of eGFR tests requested [32].

Table 1. NICE minimum number of eGFR monitoring checks per year for adults, children and young people with or at risk of chronic kidney disease

	ACR category A1	ACR category A2	ACR category A3
GFR category G1	0 to 1	1	1 or more
GFR category G2	0 to 1	1	1 or more
GFR category G3a	1	1	2
GFR category G3b	1 to 2	2	2 or more
GFR category G4	2	2	3
GFR category G5	4	4 or more	4 or more

This table is reproduced from NICE clinical practice guidelines [18].

While albuminuria is a strong predictor of adverse outcomes in patients with CKD, routine monitoring is generally low, which impacts on disease staging and prognostication in clinical practice [16,33].

Treatments

Recommended treatments for CKD involve management of risk factors, and protection against CV risk and CKD progression [9,34-35]. This includes:

- Blood pressure management (angiotensin-converting enzyme [ACE] inhibitors and angiotensin II receptor blockers [ARBs])
  [reduces CV events, CKD progression and mortality]
  Recommended for use in 83.3% of the CKD population (diabetes, hypertension, albuminuria), but data shows that only 53.3% of those eligible are receiving this standard of care (UK) [7].
  ACEi/ARBs are recommended in patients with albuminuria regardless of need for blood pressure control [18].

- SGLT2 inhibitors
  [reduces CKD progression, mortality and CV events]
  Approved for type 2 diabetes since 2013, and approved in 2022 for use in CKD patients with or without diabetes [7]; indicated for CKD stage 3

- Statin therapy
  [reduces CV risk (and may slow CKD progression)]

- Patient education and recommended lifestyle modification, including dietary changes (low salt, low phosphorous, low-protein) to reduce kidney workload, physical activity and smoking cessation
- Initiation of KRT, involving routine dialysis or kidney transplantation
  Only used in the event of kidney failure; required for survival.
  Some (predominantly elderly) patients choose conservative care instead of KRT.

Avoidance of non-steroidal anti-inflammatory drugs (NSAIDs) and nephrotoxic drugs is also advised, where possible.

Referral

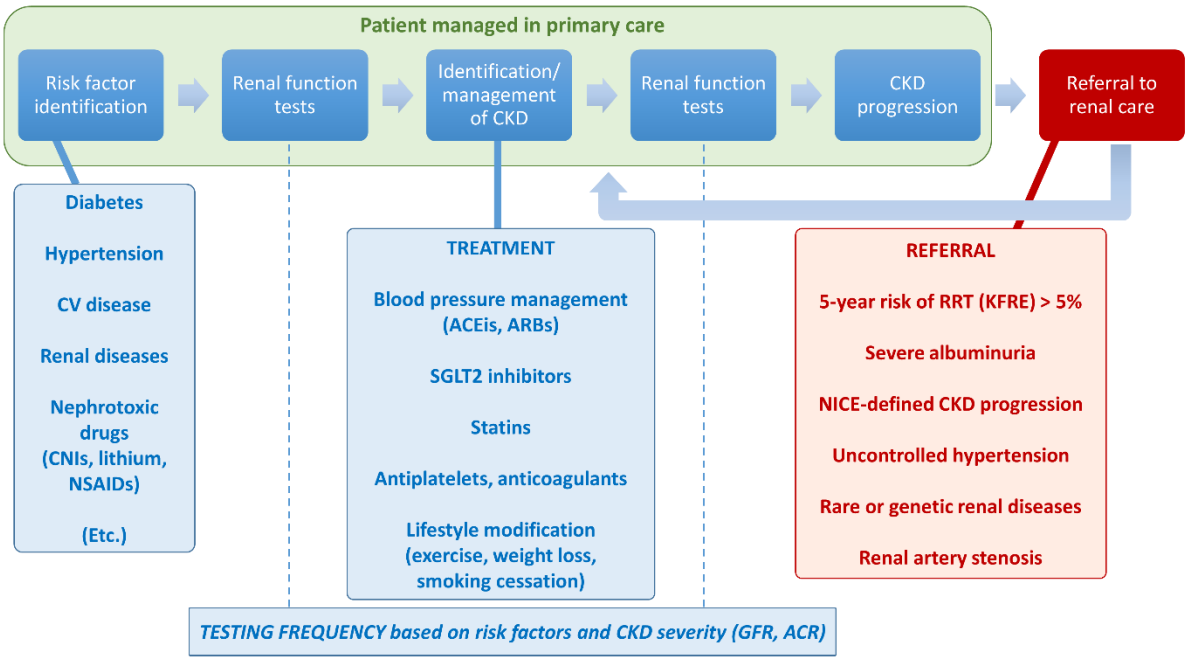
Referral to specialist nephrology care will be recommended in different circumstances in different healthcare systems, depending on resources available and healthcare priorities. NICE recommends referral to a specialist based on the following criteria [18]:

- 5-year risk of KRT > 5%
  (estimated using the kidney failure risk equation [KFRE])
- uACR ≥ 70 mg/mmol, unless managed
- uACR ≥ 30 mg/mmol & haematuria
- Reduction in eGFR ≥ 25% over 12 months, or change in eGFR category
- Rate of decline of eGFR > 15 ml/min/1.73m^2/year
- Uncontrolled hypertension
- Rare or genetic kidney diseases
- Renal artery stenosis

Guidelines also highlight the important of a multi-disciplinary approach to managing CKD, which may involve collaboration between nephrologists, primary care physicians, dieticians, pharmacists, and other healthcare professionals. There is a particular emphasis on the importance of early detection and personalised management.

Figure 1 demonstrates the care pathway for CKD in the UK healthcare system.

Figure 1. Care pathway for intended management of CKD in the UK



ACEis = angiotensin converting enzyme inhibitors; ARBs = angiotensin-receptor blockers; CV = cardiovascular; CNIs = calineurin inhibitors; KFRE = kidney failure risk equation; NSAIDs = non-steroidal anti-inflammatory drugs; RRT = renal replacement therapy; SGLT2 inhibitors = sodium-glucose co-transporter-2 inhibitors; NICE = National Institute for Health and Care Excellence

1.5 Measuring prognostic markers required for evaluation of CKD and its progression

We have introduced the key measures (eGFR and uACR) involved in diagnosis and staging of CKD, and the fact that GFR-estimating equations are used to approximate underlying GFR. Assessing rate of disease progression also requires estimation of GFR over time, but accurate estimation of GFR is not necessarily straightforward.

GFR-estimating equations

Serum creatinine is the most commonly used biomarker for estimation of GFR, due to testing being readily available and cheaper than other markers (such as cystatin c) [36]. Creatinine is a waste product of muscle metabolism, with approximately constant levels of production in the body over time for an individual, and which is filtered from the body solely by the kidneys [37]. Accumulation of creatinine indicates reduced kidney function, and the inverse association between serum creatinine levels and underlying GFR is exploited by GFR-estimating equations [25-27].

However, creatinine levels are also influenced by non-GFR-determinants, predominantly muscle mass and protein intake. As a result, GFR-estimating equations adjust for patient factors that are associated with muscle mass, including age and sex. Race variables have previously featured in equations, but have been removed due to concerns about patient labelling based on race, where race itself is not a determinant of filtration marker levels, but rather is associated with other biological determinants [27]. Despite age and sex adjustments, significant noise remains in the estimation of GFR using GFR-estimating equations, which complicate reliable identification and staging of CKD [26].

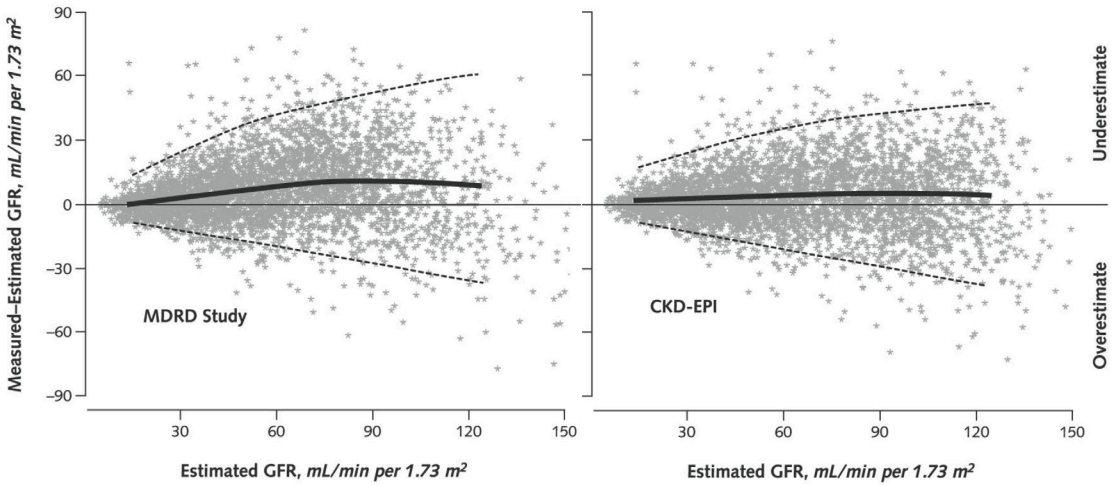
GFR-estimating equations have evolved over time, as improvements have been implemented, and different filtration markers have been used. The newest equations (CKD Epidemiology Collaboration [CKD-EPI] 2021 [27]) have been recommended for use in clinical practice (based on creatinine [eGFRcr] and both creatinine and cystatin C [eGFRcr-cys]) [9] but are not yet validated in the UK [38], where use of the CKD-EPI 2009 is still recommended [39]. This PhD used the Modification of Diet in Renal Disease (MDRD) equation and CKD-EPI 2009 equation, which were used in clinical practice at the time of data collection.

Accuracy of GFR-estimating equations

Figure 3 shows the variation in accuracy of GFR-estimating equations used in this thesis (MDRD; CKD-EPI 2009). These graphics demonstrate marked differences in estimated and underlying GFR (using a gold standard method for measuring GFR) in some patients. Performance is dependent on the true level of GFR, with better performance as kidney function declines. There is improved performance for CKD-EPI over MDRD, with less severe under-estimation on average and better precision in general.

Despite difficulties in accurately estimating GFR in individual patients, within-patient changes in creatinine (and eGFR) over time are likely to be reasonably stable, due to consistent weight in most individuals [9,40]. Factors that may inflate within-patient variability in creatinine, other than changes in kidney function, include very marked changes in protein consumption, or very marked changes in muscle mass [9]. Temporary (reversible) changes in kidney function may also result from acute illness and dehydration.

Figure 3. Accuracy of GFR-estimating equations used in thesis



This figure is reproduced from published research developing the CKD-EPI equation [26].

1.6 Overview of electronic healthcare records

Historically, medical records were recorded predominantly on paper. However, over recent decades, such records have been gradually transferred to digitalised computer systems [41-42]. Most patient interactions with the healthcare system are now recorded directly on computers and stored digitally, including medical diagnoses, prescriptions and test results. The extent of digitalisation of healthcare records depends on the healthcare system and areas within the healthcare system. Roll-out of computerised systems and regulations for digital recording varies between regions and between countries. While adoption of digital healthcare is now widespread among developed countries, there are delays in implementation in developing countries [43-44].

Digitally recorded healthcare data are commonly referred to as “electronic healthcare records”. In the UK, and globally, there are ongoing initiatives to digitalise healthcare records [45-46]. Scope is increasing and improving gradually over time [41]. Such initiatives are likely to benefit both individual patients and society as a whole.

Broadly, benefits of healthcare digitalisation [41-42] and some specific examples may include:

- To support direct patient care:
    - Ease of healthcare provider access to patient records
    - Use of standardised coding infrastructure commonly utilising internationally recognised terminologies to support specific and accurate recording of information such as disease diagnoses, medical procedures and medication prescribing
    - Electronic flagging of allergies and contra-indications to support safer prescribing
- Improved patient engagement:
    - Ease of patient access to medical records, via online access
- Facilitating healthcare management and resource allocation:
    - Financial management and budget planning
    - Public health monitoring of current disease and outcomes burden and projection of future disease burden, resources required and associated costs
    - To support identification of patients with specific risk factors, for use in screening programs or vaccine delivery programs
- Audit of quality of care:
    - To ensure that patient care is appropriately provided according to national guidelines
    - To support administration of pay-for-performance quality of care initiatives requiring electronically coded data
    - Monitoring of hospital admissions for adverse drug reactions
- For epidemiological and other healthcare research:
    - Population-based studies, to investigate determinants of disease and ascertain prevalence and burden of disease

- Post-market surveillance of drug effectiveness and safety using real-world data
- Clinical prediction modelling studies to support advancement of personalised medicine efforts
- Use of EHR data within pragmatic trials
- Supporting feasibility assessment and recruitment for clinical trials

Briefly, specific challenges relevant to the widespread implementation of EHRs and their secondary use (as outlined above) include: data security and protection of patient information, impact on clinician workload, availability of fit-for-purpose software, integration of data between different areas of the healthcare system, completeness and accuracy of data entry, variation in coding systems, and withdrawal of patient consent to share data for research in some cases.

A detailed overview of the UK (and Swedish) healthcare system, resulting availability of EHRs and its relevance to data used in this thesis are provided in *Chapter 4*.

1.7 Rationale for use of EHRs to study CKD and its progression

Unmet needs in research on CKD and its progression

We have previously stated the need for more research studying the burden, nature, causes and consequences of CKD and its progression in the general population, and have laid out some particular challenges related to detection, diagnosis and in establishing progression of CKD resulting from measurement error in eGFR.

Ideal data for studying CKD progression

RCTs. Randomised controlled trials (RCTs) are the gold standard study design for medical research investigating causal effects of exposures, due to fundamental principles of: randomisation; blinding (where possible and ethical); and prospective follow-up. Randomisation of exposures of interest ensures that patient characteristics are approximately balanced with respect to exposure status, and therefore that any differences in outcomes can be deemed attributable to the exposure. This study design reduces risks of confounding (i.e.

observed differences in outcomes occurring partly as a result of other factors which differ between exposure groups), which naturally occurs in the use of real-world data due to patient characteristics commonly being associated with one another. Blinding further protects against placebo effects, where possible and ethical. Considerable efforts are made in RCTs to follow up patients for outcomes over planned study time-frames, and prevent losses to follow-up where possible, ensuring completeness of data for analysis.

The ideal data to study CKD, health care intervention and subsequent CKD outcomes would be a clinical trial. However, such studies are expensive and time-consuming. They are more likely to be limited to studying higher risk populations with shorter term follow up or limited to outcomes with higher event rates due to analytical power considerations. This may limit their utility in the study of CKD progression, where important clinical outcomes such as kidney failure are rare, and may take a decade or more to accrue, in a large sample of patients. Longer-term follow-up may also lead to higher likelihood of non-compliance and drop-outs, and there may be ethical issues in continuing interventions if patients experience side effects, for example. Furthermore, many exposures cannot be studied in clinical trials due to ethical issues (e.g. smoking).

Evidence from clinical trials supports recommended therapeutic intervention strategies for CKD, but these studies reflect selective populations, and data on actual healthcare implementation is lacking [47]. A pragmatic trial without exclusion of patients would avoid some of these issues.

Prospective cohort studies. Prospective cohort studies possess the next highest level of evidence to RCTs, due to sampling on the basis of exposures identified and measured at study entry (which are likely to be accurately and completely measured) and prospective follow-up (allowing outcomes to be accurately and completely measured). In principle, exposures of interest and all relevant confounding factors can be accurately measured, with adjustment for relevant confounders in analyses.

However, like RCTs, these studies have limitations due to running costs and time-frames for follow-up, with potential challenges in loss to follow-up if follow-up duration required to develop outcomes of interest is long. There may therefore be limitations in the extent of research questions which can be answered by such studies. Like clinical trials, they may also be affected by volunteer bias.

Retrospective cohort studies. The next level of evidence would be a retrospective cohort study. Such studies require researchers to "look back" to ascertain initial exposures, relying on exposure information captured in medical records or patient recall, which may result in misclassification bias. While strength of evidence is poorer, benefits are reduced costs and a shorter time required to complete the study, given that the follow-up period has already occurred.

Opportunities presented by EHRs

General advantages of EHRs. While the initial purpose of data collection may vary in databases containing routinely collected EHRs (for example, to support routine management of patient care, for quality monitoring or for tracking public health outcomes), the increasingly vast scale of such databases offers great value for observational research [48]. Data may include detailed collection of a range of patient characteristics including basic demographics, longitudinal measures of test results, disease diagnoses, prescriptions, and patient outcomes. They may also cover large representative samples of the general population.

With improving healthcare digitalisation, we are likely to see further expansion and enhancement of EHR resources for research purposes in the future. Use of EHRs for epidemiological research can be carried out at low cost compared to other study designs with higher reliability of evidence.

Specific advantages for studying progression of CKD. Key opportunities in the study of CKD progression lie in the fact that there is an increasing body of longitudinal assessments of serum creatinine (and eGFR) held in EHRs. In the UK, availability of data on kidney function tests in typical EHR databases tends to be limited to those requested in primary care [48], but in other health systems such as in Stockholm, Sweden, data is available in all care settings [49]. Recommendations in clinical guidelines [18], previous financial initiatives for testing (QOF) [32] and general improvements in awareness around testing for CKD may enhance availability of kidney function tests in patient populations for which improved knowledge on CKD progression is important. Data may be captured over long periods of time, sufficient to study rare outcomes such as kidney failure, and sample sizes accumulated across populations have the potential to be very large. The use of standard coding terminology (such as Read

codes, SNOMED codes and ICD-10 codes [50]) support consistent recording of data, and aid in data extraction, aggregation and analysis.

1.8 Key analytical challenges in use of EHRs to study CKD progression

In using routinely collected data for the study of CKD progression, it is unlikely that all data required for analysis (including exposures, outcomes and confounders) will be available and accurately captured in all patients in the target population at specific time points that we would desire for analysis, and that said data can be considered a fair representation (as if it were taken randomly) for a given patient (e.g. representing steady-state health for an individual).

We first introduce some key principles and standard statistical terminology on missing data mechanisms, before presenting key analytical biases that may occur when studying CKD progression using EHRs.

Missing data mechanisms

Standard definitions for mechanisms of missing data are defined in **Table 2** [51]. The underlying mechanism of missing data in an analysis affects the likelihood of resultant biases occurring and hence the reliability of study results. While some checks can be performed on data to check for evidence of biases, the only way to confirm the missingness mechanism is to observe the missing data which is not usually possible. This means it is important to carefully consider the process by which the data are generated and what might influence presence of data when deciding on appropriate analytical techniques. Appropriate approaches to handle missing data in analysis depend on the missingness mechanism.

Some common approaches to handling missing data are:

- Complete case analysis
  Limit analysis to patients with complete data for all analysis variables

- Simple imputation techniques (e.g. last observation carried forward)
  Impute the missing data based on other data that may provide a "best guess" for the data value

- Multiple imputation

  Where a data point is missing for an individual, impute the data point by randomly drawing/predicting an observation from the distribution of data for similar patients for which the data point was observed (i.e. conditional on variables associated with the missing data variable)

- Linear mixed models

  If longitudinal data are missing, this can sometimes be handled by linear mixed models (or similar methods), where it is assumed that patient trajectories follow a distribution around a common mean trajectory. This is particularly useful for repeated measures that are unequally spaced with variable frequency of results, provided distributional assumptions are adhered to, and missing results follow the same distribution as non-missing results (which may sometimes be conditional on observed variables)

Analytical challenges relevant to EHRs

Irregular and incomplete capture of eGFR. In order to study CKD progression using EHRs, we require repeat tests of eGFR over time in a representative sample of the target population. We have established that laboratory results for kidney function tests will only be available for analysis if they have been requested and recorded by a health professional and are captured in a data source that we have access to. However, eGFR tests are unlikely to be captured regularly within specific time intervals, and frequencies of testing and time between tests are likely to vary between individuals due to various factors which may or may not be measured.

Missingness that occurs randomly can be accommodated by statistical methods such as linear mixed models, but data are more likely to be available or missing for reasons related to risks of CKD progression, where healthcare professionals will request tests based on perceived clinical need (informative testing).

Table 2. Standard definitions for mechanisms of missing data

Missingness mechanism	**Definition**	**Examples**	**Appropriate data handling**
Missing completely at random (MCAR)	Occurs when likelihood of data (required for analysis) being missing is completely random and does not depend on the value of the variable itself or on the value of any other variables relevant to the analysis (including exposure, outcome, covariates or confounders [measured or unmeasured])	eGFR was missing for a single GP practice on a particular day due to a failure to send blood samples to the laboratory; a cyber attack led to missing data for all patients at a particular hospital over a limited period	Complete case analysis (final results may be less precise, but are not systematically biased)
Missing at random (MAR)	Occurs when likelihood of data (required for analysis) being missing depends on observed variables only	eGFR was more commonly missing in those aged under 60, but missingness occurs randomly within age groups (and is unrelated to any other variables)	Multiple imputation (we can predict values of missing eGFR data using age data, based on patients with complete data)
Missing not at random (MNAR)	Occurs when likelihood of data (required for analysis) being missing depends on unobserved values, including possibly the missing value itself	eGFR is disproportionately missing in patients who are referred to secondary care during study follow-up; we cannot predict missing data based on data observed in our dataset, as patients who are referred are likely to have different eGFR trajectories to patients who remain in the dataset	Pattern mixture models; Joint longitudinal survival models

Inaccurate or incomplete measurement of study covariates. Another challenge may result from missing or inaccurate coding in related study variables such as confounders or covariates related to CKD and its progression. An example would be in measurement of comorbidities, where diagnostic codes are used to identify chronic conditions from EHRs. While there are standardised codes available for recording diagnoses in EHRs which support accurate recording, completeness of coding of co-morbidities may depend on patient factors [52-53]. It is likely that co-morbidity recording will be specific but may not be sensitive.

Accuracy of diagnostic coding may also depend on who is responsible for recording the data. For example, in UK secondary care, ICD10 codes are added to discharge letters, usually by a team of coders who examine clinical notes and discharge letters. In this case, accuracy depends on coders being adequately trained and clinical notes being clear and detailed [54].

Another particular issue is availability of uACR data. While its importance is clear due to its known strong association with CKD progression, it is not routinely collected in the majority of patients with CKD [16,33].

A glossary of general analytical biases which are encountered frequently in this thesis are presented in *Section 1.10*. Here, we summarise key analytical biases in the context of studies of CKD progression using EHRs (**Table 3**).

Table 3. Summary of key analytical biases in context of CKD progression studies using EHRs

Bias	Description / Source
Informative testing	Timing and frequency of eGFR tests is likely to be related to patient risk factors, disease severity and value of the measurement itself, patient health seeking behaviours, healthcare provider factors and financial incentivisation, which may lead to selection bias
Selection bias	Potentially occurring as a result of informative testing; may impact identification of representative sample of target population (exacerbated by chronicity criterion for detection of CKD, increasing threshold number of tests required for inclusion, but may be mitigated if care providers appropriately test for chronicity as recommended by management guidelines); informative testing may also impact completeness of follow-up for

	GFR-related outcomes (higher risk patients may be more likely to receive follow-up eGFR tests, or indeed the sickest patients or patients with particular illnesses may be lost to primary care follow-up, for example due to being monitored in secondary care or initiating KRT, at which point eGFR results become unstable)
Informative censoring	Informative testing during follow-up may lead to early truncation of follow-up for eGFR-based outcomes in some patients which occurs non-randomly, for example due to referral to specialist care; Occurrence of competing events may also truncate follow-up, e.g. initiation of KRT or death
Ascertainment bias	May occur if events are disproportionately identified depending on patient factors; May occur for eGFR-based outcomes as a result of informative testing during follow-up, and ascertainment may vary according to risk factors or the value of the outcome itself; Less likely to be a concern for "hard" clinical outcomes, such as hospitalisation for CV events or death which are likely to be completely captured, but possible for less severe outcomes such as AKI which may sometimes go undiagnosed, where identification may depend on patient characteristics, for example due to intensity of healthcare provider engagement
Survival bias	A specific type of selection bias; May occur in complete case analyses, where patients are sampled on the basis of complete follow-up (e.g. certain number of eGFR tests, captured over certain time-frame), where patients are only included in analysis if they survive for long enough to have complete follow-up data; Limits the analyses that can be carried out using cross-sectional data, which may look back at historical data, where patients are sampled on the basis of being alive at end of follow-up (patients who are alive are more likely to have a healthier history of eGFR decline than those who have experienced a rapid decline and have died)

Misclassification bias	May occur if any categorical study variables are inaccurately captured; AKI is a risk factor and consequence of CKD progression, but ICD-10 codes may not capture all cases of AKI, and capture may depend on patient characteristics or may be missing for unknown reasons; Incomplete capture of co-morbidities may occur if electronic diagnoses are not recorded in all patients consistently; Severity of CKD captured by baseline eGFR or for use in CKD stage subgroup comparisons may be inaccurately measured due to timing of the qualifying result which may not be at baseline or may be captured at different timepoints to other baseline variables
Measurement error	May occur if any continuous study variables are inaccurately measured, such as eGFR
Unobserved confounding	Relevant for causal inference analyses, where factors known to impact risk of CKD progression are not captured in certain databases, for example we may only have access to primary care data, but confounders may be captured in secondary care; Confounders may exist that we have not considered or are not possible to accurately measure, such as healthcare seeking behaviour.
Competing events	Events which may prevent measurement of eGFR data required to assess CKD progression or prevent capture of other relevant clinical outcomes include occurrence of competing outcomes which are not of primary interest, such as mortality, initiation of KRT (or possibly even referral to specialist care, if this leads to truncation of observations of eGFR)

1.9 Defining progression of CKD using EHRs

Why do we need to define CKD progression?

We stated earlier the importance of earlier identification of CKD and efforts to delay disease progression to improve public health *(Section 1.2)*. There are different rates and pathways of CKD progression between individuals due to interplay of various casual factors which may

vary over time, in a heterogeneous population of affected patients [55-56]. Known factors affecting the rate of progression of CKD include lower GFR, greater levels of albuminuria, underlying cause of CKD, as well as various patient demographics, comorbidities and lifestyle factors [24].

We may wish to identify CKD progression for different reasons, in both clinical practice and for research purposes. For example, identifying evidence of clinically important CKD progression during earlier stage disease in clinical practice may prompt improved treatment efforts and improved prognosis. Or, identifying clinically important CKD progression outcomes (which may precede kidney failure) for use in epidemiological research may allow causal inference analyses aiming to identify important causes of CKD progression. Defining CKD progression in a way that is clinically important and statistically robust is therefore important.

This thesis uses terms “CKD progression”, “progression of CKD”, and “progressive CKD” interchangeably to refer to progression of CKD, which may be defined or measured in a variety of ways.

Considerations on clinically important measures

The most traditionally and historically accepted measure of CKD progression has been progression to kidney failure requiring KRT, due to being clinically important and reliably measured. However, this outcome is rare, affecting a minority of CKD patients [14], and large sample sizes are required to study such outcomes. Many patients who are at risk of clinically important CKD progression may never expect to progress to kidney failure due to higher likelihood of a competing event (which may be associated with CKD progression) such as CV mortality [12]. Understanding causes and consequences of CKD progression in such patients remains important, in which case use of “earlier” CKD progression outcomes would aid the study of such patients. Opportunity to intervene at earlier stage of disease and better powering of clinical studies will benefit from earlier measures of disease progression which have a strong association with important clinical outcomes associated with CKD progression (KRT, mortality, CV events, etc) [57].

Considerations on defining measures suitable for EHR data

In order to evaluate CKD progression using EHRs, we must work with the type of data that we have available (i.e. eGFR measures which may be captured irregularly, with varying frequency, for varying reasons).

Some key considerations in defining CKD progression measures are as follows:

- A strong definition (and measurement procedure) for CKD progression should distinguish long-term clinically meaningful changes in kidney function in all patients we wish to analyse.
- Measures based on eGFR should aim to overcome any problems caused by known inaccuracies in estimation of eGFR between patients as well as within-patient fluctuations in eGFR. While within-patient changes in creatinine are likely to mostly reflect true change in underlying kidney function, KDIGO guidelines state that a change of less than 25% between 2 eGFR results may reflect physiological variation rather than true progression [24].
- Care is needed to avoid incorrect labelling of CKD progression, where acute reversible changes are in fact observed, which may be the case if kidney function tests are captured due to patients seeking care due to illness.
- Methods should take into account varying frequency and irregular spacing of kidney function tests over the time period of interest, which may vary between patients. Some patterns of missingness may inhibit or impact ability to evaluate changes in eGFR and may be related to underlying GFR decline.
- Accuracy of estimation of rate of decline is likely to improve with the number of serum creatinine measures and duration of follow-up, but there is a trade-off since being too strict with data completeness criteria will result in fewer patients being analysed and less representative samples.
- eGFR trajectories may not be linear in all patients and this will need to be considered when decided on appropriate measurement procedures for CKD progression. Extrapolation of eGFR slopes estimated over short follow-up periods to infer long-term rates of decline may be unsuitable in some cases where underlying GFR trajectories are non-linear.

Existing measures of CKD progression

Some measures of CKD progression have already been proposed, for example by KDIGO [24,58] for use in routine care and by the FDA [59-60] for use in clinical trials. Existing definitions are briefly presented in **Table 4** and **Table 5**. Results demonstrate a graded association between the size of percent changes between measures and subsequent outcomes, and between regression slope measures (decline of 5 units per year) and outcomes, which are clinically important.

A graphic of implied threshold trajectories for these measures (**Figure 4**) is shown to demonstrate variation in assumptions around the nature of possible eGFR trajectories which may partly underpin existing measures. This includes extrapolation over time and consideration of different baselines of eGFR, based on certain assumptions (linear changes for absolute changes; log-linear changes for percent changes).

Table 4. KDIGO 2012 definitions of CKD progression (evaluated using routine data in Canada [24,58])

Definition of CKD progression	**Measurement procedure**	**Available data for evaluation**	**Hazard ratio (+95% CI) for all-cause mortality (vs stable reference group)**	**Hazard ratio (+95% CI) for ESKD (vs stable reference group)**
CERTAIN DROP: Drop in GFR category (G1, G2, G3a, G3b, G4, G5) AND a ≥25% drop in eGFR from baseline	Percent change in 2 x eGFR measures over ≥ 6 months	N = 598,397 (median 2.4 yrs)	1.89 (1.83, 1.95)	5.11 (4.56, 5.71)
RAPID PROGRESSION: a sustained decline in eGFR of more than 5 ml/min/1.73 m^2/year	Slope regression analysis for patients with ≥ 3 eGFR measures over ≥ 4 years	N = 529,312 (median 2.4 yrs)	*not provided*	12.5 (10.0, 15.5)

* Analyses were adjusted for age, sex, hypertension, diabetes, proteinuria, Charlson co-morbidities and baseline (first) GFR. ESKD = end-stage kidney disease

Table 5. FDA definitions of CKD progression (evaluated using 1.7 million patients with 2 x serum creatinine in 35 cohorts from the Chronic Kidney Disease Prognosis Consortium [CKD-PC] [60])

Definition of CKD progression	**Measurement procedure**	**Hazard ratio for all-cause mortality (vs stable reference group)**	**Hazard ratio for ESKD (vs stable reference group)**
30% reduction in GFR from baseline over 2 years	Percent change in 2 x eGFR measures, over 1.5 to 2.5 years	GFR<60: 1.8 (1.6, 1.9) GFR≥60: 1.6 (1.4, 1.8)	GFR<60: 5.4 (4.5, 6.4) GFR≥60: 6.7 (3.9, 11.5)
40% reduction in GFR from baseline over 2 years		GFR<60: 2.3 (2.1, 2.5) GFR≥60: 2.4 (2.0, 2.9)	GFR<60: 10.2 (8.2, 12.7) GFR≥60: 15.3 (8.5, 27.2)
57% reduction in GFR from baseline over 2 years (prior gold standard)		GFR<60: 3.7 (3.2, 4.4) GFR≥60: 3.8 (2.8, 5.2)	GFR<60: 32.1 (22.3, 46.3) GFR≥60: 57.2 (21.9, 149.1)

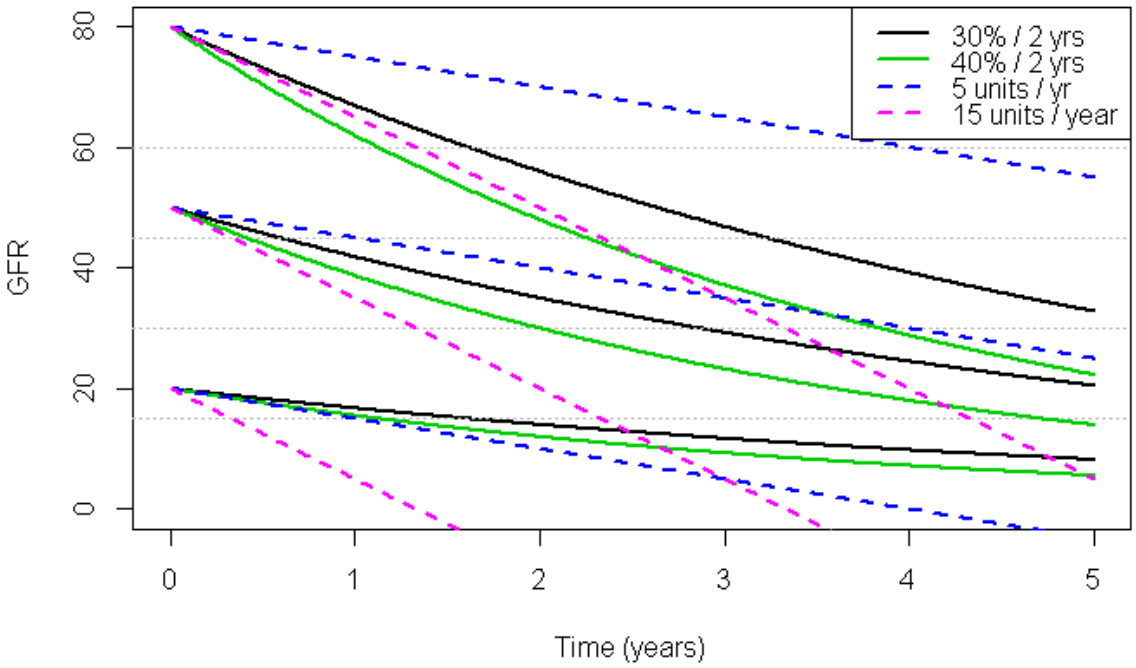
ESKD = end-stage kidney disease

For context, NICE guidelines recommend referral to specialist care based on evidence of any of the following definitions of "CKD progression" (among other specific factors) [18]:

- a 5-year risk of KRT > 5% (based on 4-variable KFRE), or
- a sustained decrease in eGFR of 25% or more and a change in eGFR category within 12 months, or
- a sustained decrease in eGFR of 15 ml/min/1.73 m2 or more per year

We therefore sought to include the latter additional progression measure (eGFR decline > 15 units per year) in **Figure 4**, due to its relevance to the UK care pathway for CKD.

Figure 4. Implied threshold trajectories for CKD progression definitions, extrapolated over time: FDA percent change 30% per 2 years; FDA percent change 40% per 2 years; KDIGO absolute change 5 ml/min/1.73 m2/year; NICE absolute change 5 ml/min/1.73 m2/year



Threshold trajectories for different CKD progression definitions vary considerably in both rate and shape of progression, dependent on the baseline or starting values, use of absolute vs percent change definitions, and duration of evaluation, highlighting some of the challenges facing treating clinicians and researchers in identifying meaningful CKD progression.

Some reflections on potential suitability of evaluation methods for CKD progression, when evaluated through EHR studies are:

- Percent change measures based on 2 eGFR results avoid the issue of making assumptions about the shape of trajectories. However, valuable information could be lost if trajectories do in fact follow a general pattern that could be modelled.
- Regression methods allow utilisation of numerous repeat measures over time, with the potential to exploit all available data, although this requires availability of repeated measures as well as assumptions about the shape of disease trajectories, and variation in eGFR may vary as disease progresses.
- Percent change (regression) outcomes seem the most mathematically convenient as they have the capacity to capture distinct changes in GFR for all patients, irrespective

of baseline GFR, while ruling out presence of infeasible changes. Also, by utilising thresholds of >25% change, the likelihood of falsely declaring small fluctuations with no clinical value as progressive CKD is reduced.

This thesis does not make further detailed attempts to define and evaluate different definitions of CKD progression but does comment further on methods used in other studies (systematic review) and strives to study useful eGFR-based outcomes of CKD progression with clinical utility (risk prediction modelling).

1.10 Glossary of statistical and epidemiological terminology used in thesis

Table 6 lists and defines key statistical and epidemiological terminology and concepts used repeatedly throughout this thesis. Missingness mechanisms terminology were defined separately in **Table 2**.

Table 6. Glossary of key statistical and epidemiological terminology used in thesis

Term	**Explanation**
Randomised controlled trial (RCT) [experimental study]	Gold standard study design for establishing causal relationships between exposures and outcomes, using fundamental principles of randomisation of exposures, blinding (where possible) and prospective follow-up for outcomes. Randomised control of exposures aims to ensure balance in other characteristics which may influence risks of outcomes, meaning that observed differences in outcomes can be attributed to exposure of interest.
Prospective cohort study [observational study]	A group of people ("cohort") are followed through a specified period of time in order to study outcomes of interest. Exposures of interest (and relevant confounders) are observed and defined at study entry. Patients are followed up regularly through the study period to ascertain data on outcomes.
Retrospective cohort study [observational study]	Similar to prospective cohort study, a group of people ("cohort") to be studied are identified based on their observed exposure status, but this is identified retrospectively using past records or patient history. Other relevant data required for analysis

	including subsequent outcomes and confounders are also identified retrospectively.
Confounding	Phenomenon occurring in observational research, whereby observed associations between an exposure and outcome of interest may be distorted due to associations with a common cause ("confounder"), and therefore observed magnitudes of association between the exposure and outcome of interest may not have a direct causal interpretation
Residual confounding	Occurs when analyses adjust for confounding variables but sources of confounding remain unaccounted for, either due to failure to measure confounding variables (that may be known or unknown) or insufficient measurement or detail of capture of existing confounders
Selection bias	Occurs when patients analysed in a study are not representative of the study population of interest, where those included in the study are systematically different to those excluded, or completeness of follow-up in the study population is dependent on exposure or outcome values, resulting in distorted (and systematically biased) associations of interest, which may not reflect the population of interest.
Misclassification bias	Occurs when study data are not accurately recorded, which may lead to diluted study associations (where misclassification occurs randomly) or systematically biased observed study associations (where probability of misclassification is related to study variables)
Ascertainment bias	Occurs when completeness of recording of outcomes data is related to patient factors which are relevant to the analysis, leading to systematically biased observed associations
Competing events	Occurs when occurrence of study outcomes may be prevented by other outcomes which are not of primary interest, leading to complications in analyses and interpretation
Survival bias	A specific type of selection bias, often resulting from inference being limited to patients who have survived for long enough to

	have sufficient data for analysis; may result in disproportionate exclusion or under-representation of patients with more severe disease or poorer prognosis, who may be more likely to become missing from data sources due to death or poor health
Informative missingness	Phenomenon which may be prominent in analysis of routinely collected electronic healthcare records, where data required for analysis are missing not at random, with systematic differences in data completeness resulting from patient factors relevant to analysis, which may bias analysis results
Informative testing	Occurs when using routinely collected test results in analysis of EHRs, where availability of test results is not random but is dictated by reasons for testing which depend on patient characteristics; test results are captured at varying time intervals and frequencies, associated with patient risk factors and disease progression profile, potentially leading to biased observed associations
Informative censoring	Occurs in analysis of routinely collected EHRs, when there is early truncation of follow-up occurring due to competing events or due to informative missingness of test results required for evaluation of outcomes
Internal validation	Relevant to risk prediction modelling; Evaluation of model performance in the development dataset used to develop the prediction model
External validation	Relevant to risk prediction modelling; Evaluation of model performance in a different (external) dataset than that used to develop the prediction model
Discrimination	Relevant to risk prediction modelling; quantifies ability of model to distinguish between patients with and without outcomes; ability of model to rank patients from high to low risk; the higher the discrimination, the more powerful the model is to support risk stratification and subsequent decision-making in prioritisation of care in clinical practice

Calibration	Relevant to risk prediction modelling; quantifies ability of model to accurately estimate risks; good calibration is important to accurately communicate risks to patients and/or policy makers and interpret risks in the wider context of other health concerns for an individual

Chapter 2

Aim and objectives

2.1 Chapter summary

This chapter describes the aim and objectives of the thesis and outlines the structure of the thesis document.

2.2 Aim

To explore and highlight the extent of challenges faced resulting from issues of data quality and completeness inherent to EHRs when used to study the epidemiology of progression of CKD, advising on potential implications for reliability of study results and to present approaches to deal with analytical biases and overcome these challenges

2.3 Objectives

This thesis addresses the following 4 research objectives:

1. To describe statistical methodology used in previous research studying progression of CKD using EHRs, and evaluate efforts to address data quality issues through appropriate study design, analysis and reporting *(systematic review)*;

2. To investigate availability of data in UK EHRs that would be required for studies evaluating progression of CKD, and explore the impact of key data quality issues on the ability to accurately estimate slopes of decline in kidney function *(feasibility analysis)*;

3. To investigate the impact of completeness of practice electronic diagnostic coding for CKD on individual risks of adverse outcomes known to be associated with progression of CKD (*coding analysis*);

4. To adapt existing risk prediction equations for kidney failure which require data that is not routinely available in the majority of CKD patients, to allow improved risk stratification across the entire CKD population (*risk prediction modelling*).

2.4 Thesis structure

This is a combination book and research paper style thesis. The bulk of research completed is presented in research paper chapters, with additional book-style chapters surrounding this work, to provide a coherent whole. A flow chart is presented detailing the flow of chapters of the thesis and brief details of content in each chapter (**Figure 5**).

An outline of the structure of the thesis is as follows.

Chapter 1: presents relevant **background** to the research conducted in this thesis including the importance of improving knowledge in the epidemiology of chronic kidney disease (CKD) and its progression, how the use of electronic healthcare records (EHRs) can support this aim, and key analytical challenges that are faced

Chapter 2: introduces the overarching **aim** and key **objectives** of the PhD

Chapter 3: brief summary of work done in **research paper 1** (*systematic review of statistical methodology used in studies of CKD progression using EHRs*), final published paper, and bullet point consolidation of key finding and implications

Chapter 4: rationale for use of **data sources**, background surrounding UK and Swedish healthcare systems, and overview of EHR databases used in research papers 2, 3 and 4

Chapter 5: summary of work done in **research paper 2** (*feasibility analysis for study of CKD progression using EHRs*), final published paper, and bullet point consolidation of key findings and implications

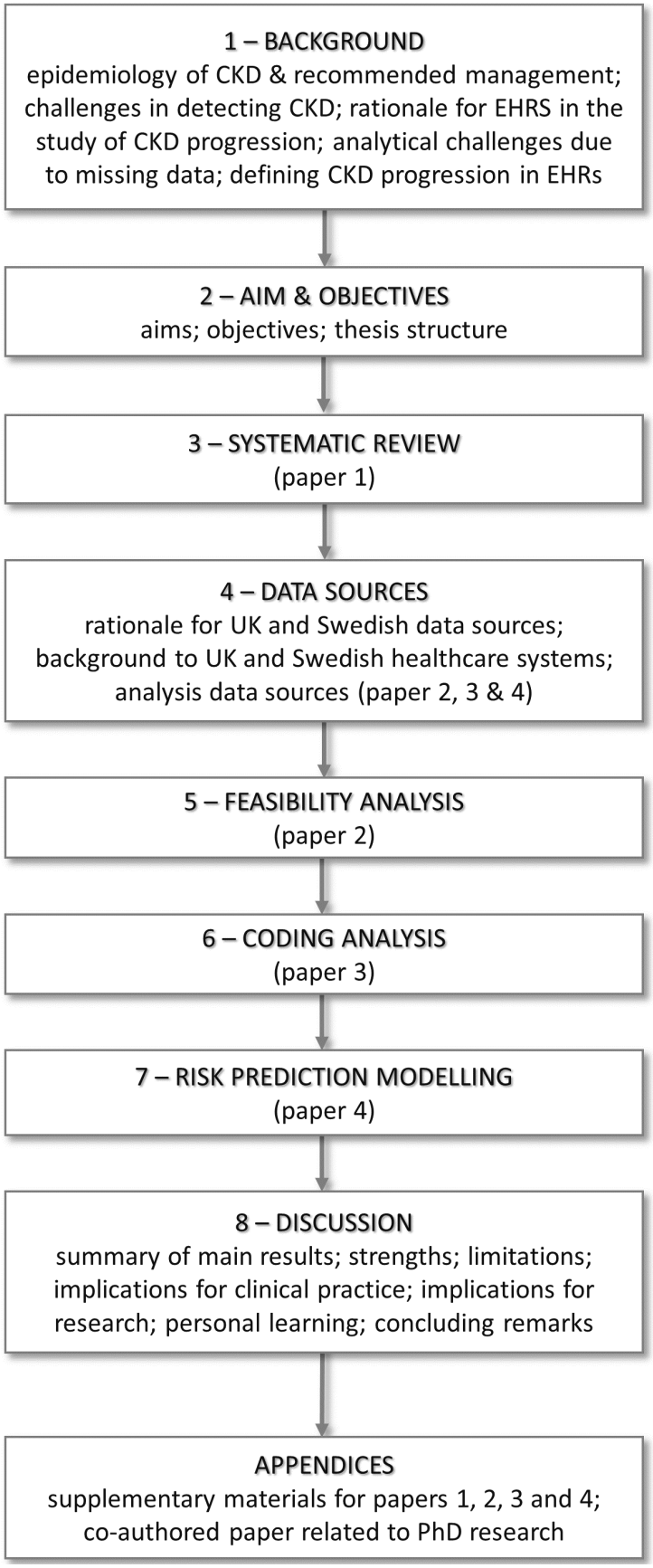
Chapter 6: summary of work done in **research paper 3** (*association between practice CKD coding and individual hospitalisation outcomes*), final published paper, and bullet point consolidation of key findings and implications

Chapter 7: summary of work done in **research paper 4** (kidney failure risk prediction modelling using routinely available data), final published paper, and bullet point consolidation of key findings and implications

Chapter 8: critical **overarching discussion** consolidating findings across PhD research studies, including summary of main results, strengths and limitations, implications for clinical practice, implications for research, personal learning and concluding remarks

Finally, appendices are provided at the end of the thesis document, including supplementary materials from each of the 4 research papers (which are referred to in the relevant research paper chapters) and co-authored work (which is referred to in Chapter 3).

Figure 5. Flow of chapters of the thesis and brief details of content in each chapter



Chapter 3

A systematic review of statistical methodology used to evaluate progression of chronic kidney disease using electronic healthcare records (paper 1)

3.1 Chapter summary

This chapter provide a summary of work done in research paper 1, presents the original published research paper, and lists key findings and implications in the context of the overall thesis.

3.2 Summary of work in the context of aims of PhD

Background

In *Chapter 1 (Background)*, we described the challenges of using EHRs to study progression of CKD, including informative testing of kidney function and potential differential ascertainment of outcomes by CKD risk factors. Such biases in data collection may impact reliability of study results if not handled appropriately in study design and methodology, with risks of overstated findings if biases are not discussed in the context of study results.

Methods

We sought to conduct a thorough review of the literature to date (final data extraction in August 2021) for studies assessing changes in kidney function over time using EHRs in patients with CKD. Particular items of interest were: how changes in kidney function were measured (and how progression of CKD was defined, if at all); statistical methodology used; and how issues with data completeness were handled and discussed.

Results

In 80 studies meeting study eligibility criteria, we identified considerable issues in transparency of reporting of analysis criteria, data completeness, and recognition of the

implications of missing data on the reliability of study conclusions. In studies with sufficient data to evaluate data completeness, it was common for large proportions of the target study populations to be excluded on the basis of incomplete data either at baseline or during follow-up for outcomes. Methods capable of handling missing longitudinal data and informative losses to follow up, such as joint longitudinal survival models, were used in a minority of studies. We also identified substantial heterogeneity in definitions of progression of CKD.

Conclusions

Many studies were likely to have overstated the reliability of findings and representativeness of study results to populations of interest. This study revealed a lack of consensus in the research community on clinically important and statistically robust measures in the study of CKD progression, which may lead to future difficulties in harmonising evidence to consolidate existing research findings.

3.3 Research paper 1

See next page for original published research paper, and appendix 1 for supplementary materials.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1703701	**Title**	Miss
First Name(s)	Faye		
Surname/Family Name	Cleary		
Thesis Title	Challenges of studying and predicting chronic kidney disease progression and its complications using routinely collected electronic healthcare records		
Primary Supervisor	Dorothea Nitsch		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Plos ONE		
When was the work published?	July 2022		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	FC was the lead reviewer; DN and DP were the supporting reviewers. FC was responsible for screening all articles for eligibility, involving scrutiny of abstracts and full-text review. DN and DP screened a sample of 50 articles each for eligibility to ensure clarity of inclusion/exclusion criteria and consistency in agreement for inclusion. FC was responsible for data extraction for all eligible research articles. In addition, key items that were the subject of this review were validated by DN and DP who independently extracted specific items as detailed in methods section. FC drafted the research paper. DN and DP contributed to draft revisions.

SECTION E

Student Signature	Faye Cleary
Date	23 May 2024

Supervisor Signature	Dorothea Nitsch
Date	10/06/2024

PLOS ONE

RESEARCH ARTICLE

A systematic review of statistical methodology used to evaluate progression of chronic kidney disease using electronic healthcare records

Faye Cleary*, David Prieto-Merino‡, Dorothea Nitsch‡

Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, United Kingdom

‡ DPM and DN also contributed equally to this work.
* faye.cleary@lshtm.ac.uk

OPEN ACCESS

Citation: Cleary F, Prieto-Merino D, Nitsch D (2022) A systematic review of statistical methodology used to evaluate progression of chronic kidney disease using electronic healthcare records. PLoS ONE 17(7): e0264167. https://doi.org/10.1371/journal.pone.0264167

Editor: Mabel Aoun, Faculty of Medicine, Saint-Joseph University, LEBANON

Received: February 12, 2021

Accepted: February 5, 2022

Published: July 29, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0264167

Data Availability Statement: This is a systematic review of previously published research, available in the public domain. All relevant data extracted

Abstract

Background

Electronic healthcare records (EHRs) are a useful resource to study chronic kidney disease (CKD) progression prior to starting dialysis, but pose methodological challenges as kidney function tests are not done on everybody, nor are tests evenly spaced. We sought to review previous research of CKD progression using renal function tests in EHRs, investigating methodology used and investigators' recognition of data quality issues.

Methods and findings

We searched for studies investigating CKD progression using EHRs in 4 databases (Medline, Embase, Global Health and Web of Science) available as of August 2021. Of 80 articles eligible for review, 59 (74%) were published in the last 5.5 years, mostly using EHRs from the UK, USA and East Asian countries. 33 articles (41%) studied rates of change in eGFR, 23 (29%) studied changes in eGFR from baseline and 15 (19%) studied progression to binary eGFR thresholds. Sample completeness data was available in 44 studies (55%) with analysis populations including less than 75% of the target population in 26 studies (33%). Losses to follow-up went unreported in 62 studies (78%) and 11 studies (14%) defined their cohort based on complete data during follow up. Methods capable of handling data quality issues and other methodological challenges were used in a minority of studies.

Conclusions

Studies based on renal function tests in EHRs may have overstated reliability of findings in the presence of informative missingness. Future renal research requires more explicit statements of data completeness and consideration of i) selection bias and representativeness of sample to the intended target population, ii) ascertainment bias where follow-up depends on risk, and iii) the impact of competing mortality. We recommend that renal progression

from reviewed articles are captured in the manuscript and its supporting Information files.

Funding: This work was supported by the Medical Research Council (MR/N013638/1), grant awarded to FC. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

studies should use statistical methods that take into account variability in renal function, informative censoring and population heterogeneity as appropriate to the study question.

Introduction

Chronic kidney disease (CKD) is a growing public health problem [1, 2]. Risks associated with CKD include cardiovascular morbidity, death, and in rare cases progression to end-stage renal disease (ESRD) requiring renal replacement therapy (RRT) [3]. Severity of disease, mechanism of renal damage and rate of progression of disease vary between patients, and the disease may change course over time in response to changing risk factors [4, 5]. While a minority of patients progress to ESRD, the cost of RRT presents a substantial economic burden to public health services and is likely to increase further over the coming years as prevalence of RRT rises alongside population growth and an ageing population [6, 7]. Increasing adoption of electronic healthcare records (EHRs) offers an opportunity to study progression of kidney disease in real-world care, that may enable improved decision-making in clinical practice. Whilst there is the promise of big sample sizes to be analysed, constraints on data availability of renal function test results may complicate reliable evaluation in EHRs. Frequency of monitoring of renal function is likely to vary in routine care according to differing individual patient risk profiles, local healthcare policy, physician-related factors, area of management within the healthcare system, social factors, or temporary illness. This may lead to some members of the target population being less likely to be followed up for renal function, potentially leading to selection and ascertainment biases in the study of CKD progression that may result in unreliable conclusions.

There are other methodological challenges in evaluation of CKD progression that are not specific to EHRs that should be considered by researchers. Deterioration in renal function over time is most commonly detected through changes of the estimated glomerular filtration rate (eGFR), usually derived from serum creatinine, sex, age, and ethnicity. Such creatinine-based GFR-estimating equations are imprecise, particularly at high levels of eGFR [8, 9]. Major changes in renal function in the context of acute illness are a sign of acute kidney injury (AKI). Although AKI is at least partially reversible in surviving patients, a history of AKI may accelerate subsequent loss in renal function. However, when researchers study eGFR decline over time, often statistical models are used that ignore the impact of acute drops in renal function on the subsequent trajectory. Population heterogeneity (caused by variation in risk factors both at baseline and evolving over time) may complicate analyses that assume a common mean linear trajectory of renal function loss over time, and it may be necessary to use more sophisticated methods if this assumption is violated that take this variability into account. Unmeasured confounding may also present issues, particularly if important confounders are not considered in the analysis. Competing events such as initiation of RRT or death complicate evaluation of progression outcomes. A previous systematic review by Boucquemont et al. in 2014 [10] reviewed statistical methods used to identify risk factors for progression of CKD, covering research on cohort studies published between 2002 and 2012. They summarised most used outcome measures and statistical models, critiquing handling of bias due to informative censoring, competing risks, correlation due to repeated measures, and non-normality of response, and proposed recommendations for best practice statistical methods and software packages.

We performed a systematic review of all longitudinal analyses of renal function tests investigating the nature, burden or consequences of CKD progression using EHRs. We aimed to establish how data issues inherent to EHRs and methodological challenges were handled, how CKD progression was defined, what statistical methods were used and whether data issues were acknowledged in the context of reliability of study conclusions.

Materials and methods

Protocol and registration

There is no published protocol available for this systematic review. Prior to completion of data extraction, this review was registered in the PROSPERO international prospective register of systematic reviews (registration number CRD42020182587).

Eligibility criteria

This is a review of statistical methodology covering all research studying the nature, burden or consequences of CKD progression using EHRs. Our intention was to focus on how researchers used renal function tests to study CKD progression. Initiation of dialysis is already a well-established clinically important outcome and as this was not the subject of the review, we excluded dialysis endpoints (as a measure of CKD progression) from review. Populations that had already initiated RRT at baseline or that were sampled on the basis of RRT initiation were excluded from review, since such populations are not appropriate for studying progression of CKD. (This criterion does not exclude patients that initiated RRT during follow-up.) Measures of CKD progression may constitute either exposures or outcomes of analysis. PICOS criteria are listed in the table below. There are no restrictions on sample size, population location or date of publication. Only studies reported in English language are included.

Participants	Include: Adults aged ≥18 with CKD stages 3–5; Studies that involve both CKD and non-CKD patients are also included, e.g. diabetes Exclude: Patients who have initiated RRT (dialysis or transplant), even if data is collected for renal function prior to RRT initiation; Patients with AKI (unless chronic changes are also studied); Non-human subjects; Children
Intervention/ Exposure	No restriction if CKD progression is measured as the outcome, rather than exposure. If CKD progression is analysed as an exposure, restrictions of this measure apply (see outcome definition).
Comparators/ Control	No restriction.
Outcome	No restriction on outcomes if CKD progression is measured as an exposure, rather than outcome. If the outcome is a measure of CKD progression: Include: Measures of chronic change in renal function based on multiple measures of eGFR or any other measure that may be used to infer eGFR (e.g. serum creatinine, cystatin-C, iohexol clearance), e.g. rate of change, change from baseline, regression slope, time to change or threshold eGFR Exclude: All other measures of renal function, e.g. proteinuria; Studies of acute AKI or short term follow up (<6 months) of renal function following a procedure; Single time-point analyses; Time to RRT as single outcome.
Study design	Include: Retrospective analysis of routinely collected electronic healthcare records which may include retrospective cohort studies, case-control studies and cross-sectional studies (if a measure of past progression is included) Exclude: Case reports, Clinical trials, prospective cohort studies or any other study design with pre-planned data collection strategy for research purposes.

Searches

We performed electronic searches of MEDLINE, EMBASE, Global Health and Web of Science databases through to 11th August 2021. A copy of the search strategy is provided in the supplementary materials S2 File.

Study selection

This study had one lead reviewer and two supporting reviewers. The lead reviewer was responsible for screening all articles for eligibility, which involved scrutiny of abstracts followed by full-text review. The two supporting reviewers independently screened a sample of 50 articles each for eligibility. Consistency of agreement and reasons for disagreement were discussed. Clarity of inclusion/exclusion criteria was updated following discussion and prior to completion of eligibility review by the lead reviewer.

Data collection process

The lead reviewer was responsible for data extraction for all eligible research articles. In addition, key items that were the subject of this review were validated by supporting reviewers who independently extracted the following items for all articles: (1) measure of change in renal function; (2) statistical methods used in analysis of changes in renal function; and (3) definitions of progression of CKD, if any. The lead reviewer developed a data extraction form in an Excel spreadsheet, which was reviewed and approved by supporting reviewers in the initial stages of data extraction.

Data items

Information extracted from eligible research articles included details of the study population, study methodology and how data quality issues and other methodological issues were handled. Extracted items are listed below.

Study population. Data collection timeframe; Country of residence; Mean age; Percent male; Primary morbidity under study / reason for inclusion; Data source / healthcare setting

Study methodology. Date of publication; Study design; Research aims; Sample size (before and after exclusions for reasons of data completeness [for details, see below explanation of data completeness inclusion criteria and calculations of percentage of target population analysed]); Measure of renal function; Measure of change in renal function over time; Definition of progression (if any); Whether change in renal function was exposure or outcome; Duration of follow up for changes in renal function; Data completeness inclusion criteria and the minimum number of renal function tests required for analysis; Statistical tools used; Statistical model used.

Some additional results were derived to quantify data completeness for analysis, including the percentage of the target population that were analysed after application of data completeness inclusion criteria and the percentage of patients that dropped out of analysis during the intended follow up period having met criteria for inclusion in analysis. Here, "data completeness inclusion criteria" refer to the study-specific inclusion criteria applied prior to main analyses being performed that aimed to retain only those patients with sufficient data completeness to be deemed suitable for analysis, with such criteria expected to vary between studies.

Percentage of target population analysed was defined as:

$$\frac{number\ of\ patients\ analysed\ (meeting\ population\ criteria\ after\ exclusions\ due\ to\ data\ completeness)}{number\ of\ patients\ meeting\ population\ criteria\ prior\ to\ exclusions\ due\ to\ data\ completeness} \times 100$$

This was computable in some but not all studies, as it requires data on the total number of patients included in analysis as well as the number of patients that met population criteria before data completeness exclusion criteria were applied. (In propensity score matched cohort studies, propensity score matching criteria are included in population criteria, and we only compute percentage of target population analysed in the propensity score matched cohort, where this is possible.)

Percentage of study population lost to follow up was defined as:

$$\frac{number\ of\ analysed\ patients\ lost\ to\ follow\ up\ during\ the\ intended\ follow\ up\ period}{number\ of\ patients\ analysed} \times 100$$

Again, this was computable in some but not all studies, as it requires data on the number of patients analysed and the number of those patients that dropped out during the intended follow up period, for example due to death, initiation of RRT or other lack of follow up in routine care which could be for many different reasons.

Handling of data quality issues and other methodological challenges. Of the items below, details extracted included whether items were mentioned, whether information was provided on data completeness [if relevant], whether implications were acknowledged, whether challenges were tackled methodologically and any statistical methods used to attempt to overcome challenges:

Handling of sample completeness / representativeness of the target population; Handling of informative drop-outs/censoring; Handling of missing longitudinal data; Handling of missing covariate data; Distributional checks/issues; Handling of within-patient correlation and variability of kidney function over time; Handling of population heterogeneity; Handling of confounding.

Risk of bias in individual studies

Assessment of bias in individual studies was one of the main aims of this systematic review. Key measures of bias evaluated in individual studies were the percentage of the sample target population that were analysed and the percentage of the analysed study population that were lost to follow up. Study-specific measures were reported and bar charts were produced for these measures to demonstrate the potential for bias in individual studies due to informatively missing data.

Synthesis of results

This review was descriptive with simple aggregation of collected data items only and no statistical analysis was performed. 4 separate summaries are provided to describe study population characteristics, study methodology used, acknowledgment and handling of data quality issues and other methodological challenges, and definitions of CKD progression. For studies exploring multiple outcomes or conducting multiple analyses of changes in renal function, the outcomes and analyses considered the primary focus regarding renal progression in each paper are summarised in the review.

Risk of bias across studies

There was no single effect size of interest in this study and no meta-analysis was performed, as the review focussed on methodology used and investigators' handling of data quality issues. Publication bias was therefore challenging to evaluate, as funnel plots and statistical tests could not be used. Efforts were made to maximise coverage of peer-reviewed literature in this field, including extraction of articles from 4 major databases. If research is missing from review due to publication

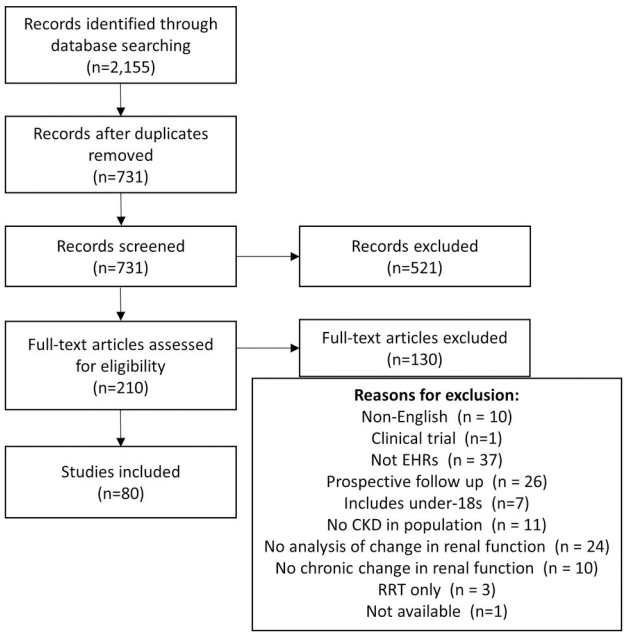


Fig 1. Flow chart of study selection.

https://doi.org/10.1371/journal.pone.0264167.g001

in non-English languages, then data quality issues in such missing studies are likely to be similar to those in English language studies that were included. There will be clinical audit studies that are not peer-reviewed; these studies are likely to be of a similar of worse quality than reviewed studies because peer-reviewed literature is expected to go through certain research quality checks. In any case, as peer-reviewed literature is more likely to be used to inform policy than other research, this is arguably the optimal collection of research to assess the aims of this review.

Results

731 unique articles were identified from database searching, of which 80 met study eligibility criteria (Fig 1). Primary reasons for exclusion were not using EHRs, pre-planned data collection for research purposes such as a prospective cohort study, and studies with a single renal function test rather than longitudinal analysis of repeated measures of renal function. Other reasons for exclusion were ineligible populations, such as studies including children, restricted to RRT populations or studies that did not include CKD patients, such as studies of the incidence of CKD. All included studies retrospectively analysed routinely collected healthcare data. It was not always clear whether electronic or paper records were used, and while efforts were taken to differentiate this, it is possible that some included studies may have involved manual data extraction from paper records. 70 studies (88%) clearly stated the use of EHRs. In the 10 studies that did not state this, the time-frame for data collection and location of research suggested that electronic healthcare systems were likely to have been used, but we could not verify this. These studies have been summarised separately in the supplementary materials. A full list of reviewed studies is also included in the supplementary materials S3 File.

Study population characteristics

Table 1 summarises characteristics of study populations analysed in reviewed articles. Research was most commonly conducted in the UK (25%) and USA (30%), followed by East Asian

Table 1. Summary of study populations studied (N = 80).

Study population characteristics	N (%)
Primary decade of follow up	
2010–2019	35 (43.8%)
2000–2009	36 (45.0%)
1990–1999	3 (3.8%)
Not available	6 (7.5%)
Country	
Europe	**28 (35.0%)**
UK	20 (25.0%)
Germany	2 (2.5%)
Italy	2 (2.5%)
Norway	2 (2.5%)
Multiple European countries	2 (2.5%)
North America	**25 (31.3%)**
USA	24 (30.0%)
Canada	1 (1.3%)
Asia	**25 (31.3%)**
South Korea	6 (7.5%)
China	5 (6.3%)
Taiwan	7 (8.8%)
Japan	6 (7.5%)
Thailand	1 (1.3%)
Oceania	**1 (1.3%)**
Australia	1 (1.3%)
South America	**1 (1.3%)**
Colombia	1 (1.3%)
Africa	**0**
Mean age[a]	
Median (IQR)	64 (56, 71)
30–49	7 (8.8%)
50–59	20 (25.0%)
60–69	29 (36.3%)
70–80	22 (27.5%)
Not stated	2 (2.5%)
Percent male	
Median (IQR)	52% (44%, 63%)
$\leq$ 34%	6 (7.5%)
35–44%	15 (18.8%)
45–54%	24 (30.0%)
55–64%	16 (20.0%)
$\geq$ 65%	19 (23.8%)
Main morbidity /reason for inclusion	
CKD	21 (26.3%)
Diabetes	16 (20.0%)
General population	8 (10.0%)
Diabetic nephropathy / kidney disease	5 (6.3%)
Atrial fibrillation	5 (6.3%)
Multiple CKD risk factors	2 (2.5%)

(*Continued*)

Table 1. (Continued)

Study population characteristics	N (%)
IgA nephropathy	2 (2.5%)
Infections (Hepatitis C, HIV)	3 (3.8%)
Transplant recipients (liver, heart)	3 (3.8%)
Autoimmune diseases (lupus, IgG4 related, vasculitis)	3 (3.8%)
Gout/hyperuricemia	2 (2.5%)
Other*	10 (12.5%)
Data source / clinical setting	
Multiple care settings	23 (28.8%)
Primary care	19 (23.8%)
Outpatient	17 (21.3%)
Diabetes clinic	6 (7.5%)
Renal clinic	3 (3.8%)
Diabetic-renal clinic	1 (1.3%)
Not specified	7 (8.8%)
Hospital	11 (13.8%)
Tertiary care	6 (7.5%)
Not stated	4 (5.0%)

[a]Other morbidities/reason for inclusion were urinary system disorders, hyperkalemia, obesity, osteoporosis, primary aldosteronism, abdominal aortic aneurysm, acute renal embolism, light chain deposition disease, lung cancer and renal cancer.

https://doi.org/10.1371/journal.pone.0264167.t002

countries, including South Korea (8%), China (6%), Taiwan (9%) and Japan (8%). Research in non-English-speaking countries may be missing from review. Typically (based on median), studied populations had a mean age of 64 and were 52% male, although there was substantial variation between studies in these characteristics. Most commonly studied morbidities were CKD (26%) and diabetes (20%) although research covered a range of different populations, including (non-renal) transplant recipients and specific renal diseases. 10% studied the general population, with a further 3% studying patients with general risk factors for CKD. Clinical settings of retrieved databases varied widely, including primary care (23%), un-specified hospital settings (14%), outpatient clinics (21%), and 29% of studies used linked data across multiple care settings.

Study methodology

Study methodology is summarised in Table 2 and a listing of key items by study is also provided in the supplementary materials S4 Table. Use of EHRs for observational research increased rapidly in recent years, with 74% of reviewed studies published in the last 5.5 years. The overwhelming majority of research was focussed on risk factor identification and causal inference (82%), with only a handful of studies attempting risk prediction (9%). Other aims included estimation of incidence or prevalence (4%) and descriptive characterisations of changes in renal function (4%). Sample size ranged drastically from 24 up to 1,597,629, with a median sample size of 1,114.

eGFR was the most commonly used measure of renal function (94%). Measures of change in renal function and methods of derivation were highly variable. Regression of absolute changes in eGFR was most common (26% of studies), although methods varied with many using mixed models but others using individual linear regression. Calculation of

Table 2. Study methodology (N = 80).

Study methodology features	N (%)
Date of publication	
2015–2021	59 (73.8%)
2010–2014	14 (17.5%)
2005–2009	6 (7.5%)
2000–2004	1 (1.3%)
Study design	
Retrospective cohort study	74 (92.5%)
Cross-sectional study	4 (5.0%)
Case-control study	2 (2.5%)
Research aims	
Risk factor identification / causal inference	65 (81.3%)
Risk prediction	7 (8.8%)
Estimation of incidence/prevalence	3 (3.8%)
Descriptive characterisation of changes in renal function	3 (3.8%)
Identification of sub-populations	1 (1.3%)
Audit of care provision	1 (1.3%)
Sample size	
Median (IQR)	1114 (209, 9876)
≤ 99	10 (12.5%)
100–499	18 (22.5%)
500–999	11 (13.8%)
1,000–9,999	22 (27.5%)
≥ 10,000	19 (23.8%)
Measure of renal function	
eGFR	**75 (93.8%)**
MDRD	33 (41.3%)
CKD-EPI	28 (35.0%)
MDRD, CKD-EPI combination	1 (1.3%)
Taiwan CKD-EPI	1 (1.3%)
Japanese formula	3 (3.8%)
Not specified	9 (11.3%)
Estimated creatinine clearance	**2 (2.5%)**
Cockcroft and Gault	2 (2.5%)
Serum creatinine	**2 (2.5%)**
Inverse serum creatinine	**1 (2.5%)**
Measure of change in renal function over time[a]	
eGFR	**75 (93.8%)**
Regression slope (absolute changes)	20 (25.0%)
Individual linear regression	8 (10.0%)
Linear mixed model	10 (12.5%)
Growth model	1 (1.3%)
Generalised estimating equations	1 (1.3%)
Regression slope (absolute and percent changes)	1 (1.3%)
Linear mixed model	1 (1.3%)
Rate of change between measures	5 (6.3%)
Rate of change, not clearly defined	4 (5.0%)
Rate of percentage change, not clearly defined	3 (3.8%)

(*Continued*)

Table 2. (Continued)

Study methodology features	N (%)
Raw absolute change from baseline	10 (12.5%)
Raw percent change from baseline	13 (16.3%)
Raw percent change between measures	1 (1.3%)
Binary progression to threshold eGFR	6 (7.5%)
Binary progression (changes/threshold combination)	3 (3.8%)
Transition between CKD stages	6 (7.5%)
Trajectory shape class (mixed model)	1 (1.3%)
Model predicted percent change per year	1 (1.3%)
Model predicted eGFR at multiple time points	1 (1.3%)
Estimated creatinine clearance	**2 (2.5%)**
Regression slope (absolute scale)	1 (1.3%)
Raw percent change from baseline	1 (1.3%)
Serum creatinine	**2 (2.5%)**
Raw absolute change from baseline	1 (1.3%)
Binary progression to threshold serum creatinine	1 (1.3%)
Inverse serum creatinine	**1 (1.3%)**
Regression slope (absolute changes)	1 (1.3%)
Change in renal function as outcome or exposure	
Outcome	74 (92.5%)
Exposure (if exposure, outcome listed below)	6 (7.5%)
Referral to renal care	1 (1.3%)
CV events	1 (1.3%)
Multiple outcomes (CV, hospitalisation, death)	1 (1.3%)
Advanced CKD (stage 4)	1 (1.3%)
Bleeding events	1 (1.3%)
Duration of follow up for renal function changes	
Median (IQR), years	3.0 (1.6, 4.4)
< 1 year	7 (8.8%)
1–4.9 years	48 (60.0%)
5–9.9 years	14 (17.5%)
≥ 10 years	1 (1.3%)
Not stated	10 (12.5%)
Minimum number of renal function measures for inclusion	
0	1 (1.3%)
1	7 (8.8%)
2	24 (30.0%)
3	15 (18.8%)
4	5 (6.3%)
5	1 (1.3%)
6	4 (5.0%)
Not stated	23 (28.8%)
Percentage of target population used in analysis	
<50%	17 (21.3%)
50% - 75%	9 (11.3%)
75% - 90%	5 (6.3%)
90% - 95%	5 (6.3%)
>95%	8 (10.0%)

(*Continued*)

Table 2. (Continued)

Study methodology features	N (%)
Not available	36 (45.0%)
Percentage of study population lost to follow up	
< 25%	2 (2.5%)
25% - 50%	3 (3.8%)
> 50%	1 (1.3%)
Not available	62 (77.5%)
Complete case analysis (only including records of people with follow-up data)	11 (13.8%)
Statistical tools used[b]	
Descriptive results only	5 (6.3%)
Simple statistical tests	9 (11.3%)
Linear regression models	8 (10.0%)
ANOVA/ANCOVA	2 (2.5%)
Kaplan-Meier estimation / life table analysis	3 (3.8%)
Generalised linear models (GLMs)	11 (13.8%)
Cox proportional hazards regression	18 (22.5%)
Competing risks survival models	3 (3.8%)
Mixed modelling methods	12 (15.0%)
Other latent variable methods	2 (2.5%)
Generalised estimating equations (GEEs)	2 (2.5%)
Joint longitudinal survival modelling	2 (2.5%)
Structural equation modelling	1 (1.3%)
Multiple imputation	5 (6.3%)
Machine learning methods	3 (3.8%)
Statistical model used[b]	
Risk factor identification / causal inference	**N = 65**
Difference in means t-test	2 (3.1%)
Mean difference paired t-test	4 (6.2%)
Simple non-parametric tests (Mann-Whitney U)	1 (1.5%)
Difference in proportions chi-squared test	2 (3.1%)
ANOVA	1 (1.5%)
ANCOVA	1 (1.5%)
Linear regression	7 (10.8%)
Logistic regression	10 (15.4%)
Kaplan Meier estimation /life table analysis	3 (4.6%)
Cox proportional hazards regression	16 (24.6%)
Competing risk survival models	3 (4.6%)
Linear mixed model	10 (15.4%)
Generalised estimating equations (GEEs)	2 (3.1%)
Joint longitudinal survival model	2 (3.1%)
Structural equation modelling	1 (3.1%)
Risk prediction	**N = 7**
Kalman filter (time series model)	1 (14.3%)
Naïve Bayes classifier	1 (14.3%)
Logistic regression	4 (57.1%)
Cox proportional hazards regression	1 (14.3%)
Random forest regression	2 (28.6%)
Linear mixed model	1 (14.9%)

(*Continued*)

Table 2. (Continued)

Study methodology features	N (%)
Estimation of incidence/prevalence	**N = 3**
Crude estimation	3 (100%)
Identification of sub-populations	**N = 1**
Trajectory clustering using latent variables	1 (100%)
Audit of care provision	**N = 1**
Linear mixed model	1 (100%)

[a]More specific details of measures of changes in renal function in individual studies assessing CKD progression and corresponding statistical analysis methods are shown in Table 4, including where time-to-event models were used in the presence of unequal follow up or censoring.

[b]Multiple items possible for a single study but focus only on main analysis of CKD progression.

https://doi.org/10.1371/journal.pone.0264167.t003

absolute changes and percent changes in eGFR were also common (14% and 17% respectively), but duration of follow up varied substantially between studies. Other less common measures were rates of change calculated between measures, regression slopes on the percent scale, and binary measures for progression to thresholds of eGFR or CKD stages. 7 studies (9%) analysed rates of change in eGFR that were not clearly defined as either regression slopes or rates of change between measures. Other renal function measures studied were Cockcroft and Gault estimated creatinine clearance (3%), serum creatinine (3%) and inverse serum creatinine (1%).

Most studies (93%) analysed changes in renal function as an outcome, with only 6 studying changes in renal function as an exposure. Typical (median) duration of follow up for renal function was 3 years, but ranged from 3 months to 14 years, and was not stated in 13% of studies. Duration of follow up also commonly varied significantly between patients within individual studies, mostly due to variation in data completeness with regards to availability and timing of serum creatinine test results on the health record. Inclusion criteria relating to availability of repeat eGFR measures varied and was commonly not stated (29%). The percentage of the target population analysed could not be calculated for 36 studies (45%) due to insufficient data (Fig 2A). The study population constituted less than 50% of patients in the target population for 17 studies (21%), and less than 75% of the target population in 26 studies (33%) (Fig 2B). Statistics on data completeness were rarely stated explicitly and were often difficult to ascertain. Rates of loss to follow up were even more difficult to ascertain, and many studies sampled patients on the basis of varying levels of completeness of follow up. In 11 studies (14%), quantifying the impact of loss to follow up was not possible due to sampling based on complete follow up, and in 62 studies (78%) no data was reported on losses to follow up. The supplementary listing of individual studies provides a more detailed breakdown of analysis criteria, percentage of target population analysed and rates of loss to follow up.

Statistical methods for analysing CKD progression depended on whether the renal function measure was continuous (e.g. rate of change in eGFR) or binary (e.g. >30% change in eGFR from baseline at repeat measurement), which varied between studies. Most commonly used statistical methods were linear mixed models, linear regression, logistic regression, and Cox proportional hazards regression. Many studies used simple statistical tests, despite the inability of these methods to adjust for confounders commonly present in observational data. More sophisticated methods taking into account differential drop-outs due to death were rare. 2 studies used joint longitudinal survival models and 3 studies used competing risks survival models.

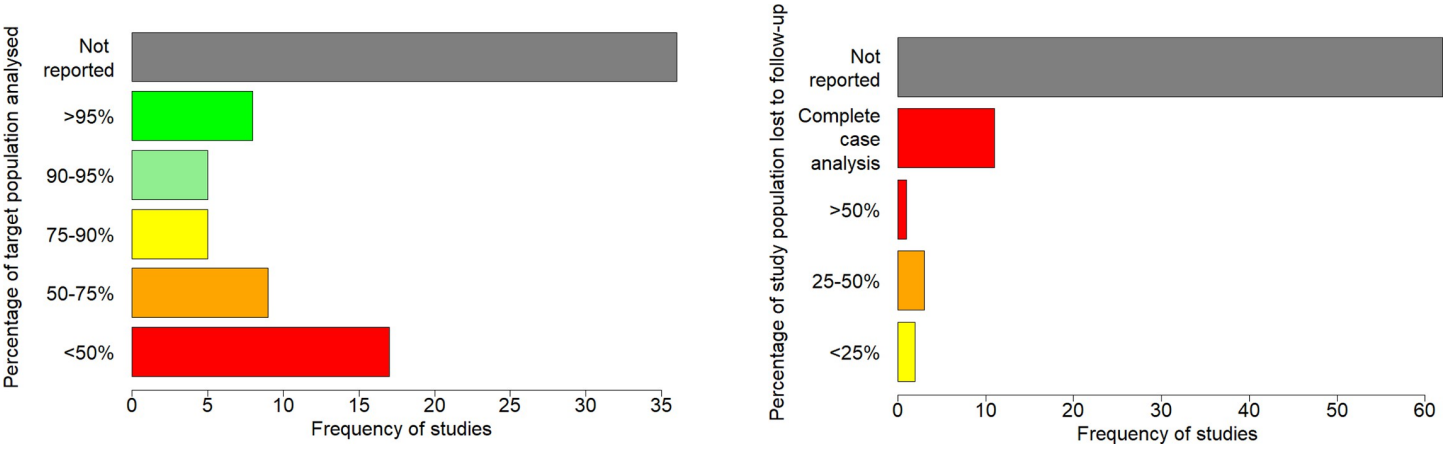


Fig 2. Risk of selection bias (A) and ascertainment bias (B) in individual studies.

https://doi.org/10.1371/journal.pone.0264167.g002

Handling of data quality issues and methodological challenges

Table 3 summarises how data quality issues and methodological challenges were dealt with in reviewed articles. EHR databases used for analysis rarely had good quality data on renal function, i.e. collected regularly over time and completely for all patients in the target population. A few studies attempted to improve sample completeness, for example by using imputation methods to avoid exclusions. Studies selected patients for analysis on the basis of varying levels of data completeness, relating to number of measures and duration of follow up, and many studies would have excluded patients from analysis completely on the basis of insufficient data over time. 64% of studies at least partially acknowledged this as introducing bias, 18% provided some data on sample completeness without acknowledging implications and 16% did not mention sample completeness or representativeness at all. Very few studies mentioned losses to follow up during the study period or potential reasons for loss to follow up and 61% of studies did not mention the issue of informative censoring at all. Only 6 studies (8%) tackled the issue methodologically, for example by accounting for the competing risk of death through joint longitudinal survival models and competing risks survival models.

Most studies (59%) did not mention (or tackle) the issue of missing longitudinal data on renal function tests over time. One in 6 studies did however use mixed modelling methods (16%) which may partially deal with the issue. 4 studies (5%) attempted to deal with missing longitudinal data through imputation methods. 40% of studies failed to mention missing covariate data despite covariate analysis, while 20% did not perform covariate adjustment. 25% at least partially acknowledged the issue and 16 studies (20%) made some attempt to handle missing covariate data through imputation methods, data linkage or other adjustment for missingness.

Distributional checks for renal function measures were rare, with only 5 studies (6%) mentioning distributional checks or considering alternative error distributions. Regarding the issue of variability in renal function over time and within-patient correlation, 25% did not mention (or tackle) such issues at all, 40% tackled the issue methodologically, 30% partially tackled or acknowledged the issue and a further 5% fully acknowledged such issues. 21% of studies used patient random effects to account for within-patient correlation, and 28% used outcomes which are likely to identify an important and real change.

Most studies acknowledged some aspects of population heterogeneity in analyses. At the most basic level, covariate adjusted analyses were used to account for baseline differences

Table 3. Critique of handling of data quality and methodological challenges (N = 80).

Handling of data quality and methodological challenges	N (%)
Representativeness of sample to target population	
Not mentioned	13 (16.3%)
Mentioned care pathway and inclusion criteria, but not sample completeness	2 (2.5%)
Mentioned sample completeness, but not implications	14 (17.5%)
Partially acknowledged implications of sample completeness	37 (46.3%)
Fully acknowledged implications of sample completeness	10 (12.5%)
Tackled methodologically	4 (5.0%)
Methods of handling[a]	
None	68 (85.0%)
Detailed/comprehensive database of EHRs used	5 (6.3%)
Multiple imputation (to avoid exclusions)	4 (5.0%)
Other imputation methods (to avoid exclusions)	3 (3.8%)
Handling of informative drop-outs/censoring	
Not mentioned	49 (61.3%)
Mentioned care pathway follow up, but not losses to follow up (inc. death)	2 (2.5%)
Mentioned losses to follow up, but not implications	7 (8.8%)
Partially acknowledged implications of losses to follow up	13 (16.3%)
Fully Acknowledged implications of losses to follow up	3 (3.8%)
Tackled methodologically	6 (7.5%)
Methods of handling[a]	
None	71 (88.8%)
Complete follow up	1 (1.3%)
Joint modelling of longitudinal changes and time to drop out (including death)	2 (2.5%)
Sensitivity analysis in drop-outs	1 (1.3%)
Competing risks survival models	4 (5.0%)
Sensitivity analysis adjusting for competing risks	1 (1.3%)
Handling of missing longitudinal data	
Not mentioned	47 (58.8%)
Mentioned care pathway follow up, but not data completeness	4 (5.0%)
Mentioned data completeness, but not implications	7 (8.8%)
Partially acknowledged implications of data completeness	13 (16.3%)
Fully acknowledged implications of data completeness	1 (1.3%)
Tackled methodologically	8 (10.0%)
Methods of handling[a]	
None	62 (77.5%)
LOCF	1 (1.3%)
Imputation with mean/median	2 (2.5%)
Mixed modelling	13 (16.3%)
Generalised estimating equations	1 (1.3%)
Multiple imputation	1 (1.3%)
Handling of missing covariate data	
Not relevant (no covariate analysis)	16 (20.0%)
Not mentioned (despite covariate analysis)	32 (40.0%)
Mentioned data completeness, but not implications	2 (2.5%)
Partially acknowledged implications of data completeness	17 (21.3%)
Fully acknowledged implications of data completeness	3 (3.8%)
Tackled methodologically	7 (8.8%)

(*Continued*)

Table 3. (Continued)

Handling of data quality and methodological challenges	N (%)
Methods of handling[a]	
None	64 (80.0%)
LOCF	2 (2.5%)
Imputation with mean	4 (5.0%)
Multiple imputation	5 (6.3%)
Complete data was available for all covariates	2 (2.5%)
Data linkage to improve data completeness	1 (1.3%)
Adjustment for missingness	2 (2.5%)
Distributional checks/issues	
Not mentioned	70 (87.5%)
Mentioned or partially addressed	5 (6.3%)
Fully Acknowledged	0
Tackled	5 (6.3%)
Methods of handling[a]	
None	75 (93.8%)
Distributional checks	4 (5.0%)
Consideration of alternative error distributions	1 (1.3%)
Handling of within-patient correlation / variability in kidney function over time	
Not mentioned	20 (25.0%)
Mentioned or partially addressed	24 (30.0%)
Fully Acknowledged	4 (5.0%)
Tackled	32 (40.0%)
Methods of handling[a]	
None	35 (43.8%)
Random effects / latent variables	17 (21.3%)
Generalised estimating equations	2 (2.5%)
Modelling of stochastic process	1 (1.3%)
Outcome likely to identify real change	22 (27.5%)
Measures capturing AKI explicitly excluded	1 (1.3%)
Paired t-test	3 (3.8%)
Handling of population heterogeneity	
Not mentioned	1 (1.3%)
Mentioned or partially addressed	36 (45.0%)
Fully Acknowledged	3 (3.8%)
Tackled	40 (50.0%)
Method of handling[a]	
None	8 (10.0%)
Adjustment for covariates	21 (26.3%)
Interaction terms	9 (11.3%)
Stratified or separate/subgroup analysis	34 (42.5%)
Latent classes	1 (1.3%)
Random effects	3 (3.8%)
ANOVA/ANCOVA	2 (1.5%)
Propensity score methods	1 (1.3%)
Features in machine learning classification	1 (1.3%)
Handling of confounding (risk factor / causal inference analyses only)	N = 65
Not mentioned	7 (10.8%)

(*Continued*)

Table 3. (Continued)

Handling of data quality and methodological challenges	N (%)
Mentioned or partially addressed	17 (26.2%)
Fully Acknowledged	3 (4.6%)
Tackled	38 (58.5%)
Methods of handling[a]	
None	12 (18.5%)
Adjustment for baseline confounders	46 (70.8%)
Propensity score methods	6 (9.2%)

[a]Methods/approaches for handling issues are listed, regardless of whether the corresponding issues were fully tackled in analysis.

https://doi.org/10.1371/journal.pone.0264167.t004

between patients (26%). Other methods included stratification or subgroup analyses to study distinct populations (43%), interaction terms allowing differing trajectories of renal function according to patient characteristics (11%) and random effects (4%). For studies performing causal analyses, 59% tackled the issue of confounding, mostly through baseline adjustment. A subset (11%) did not mention (or tackle) confounding at all, with some studies performing simple statistical tests such as t-tests and chi-squared tests despite the potential for confounding by indication.

Definitions of CKD progression

Table 4 provides a list of CKD progression measures used in individual studies, grouped by method of derivation. A listing is provided rather than aggregate summary due to the substantial variation in the way researchers defined CKD progression across the literature. Terms used included progression, rapid progression, fast progression, rapid decline, progressive decline, progressive renal impairment, renal function deterioration and worsening renal function, while some did not provide labels, simply stating the outcome as a threshold percent change in renal function for example. There is no consistency between studies in the way these terms apply to different outcomes.

Discussion

We performed a systematic review of peer-reviewed literature studying progression of CKD using routinely collected EHR data. Handling of data quality issues was generally poor, with unclear reporting of analysis criteria, data completeness and discussion of the implications of missing data on reliability of conclusions. For studies with sufficient data, representativeness of samples to target populations was likely to be poor with large numbers of patients excluded from analysis on the basis of poor data completeness at baseline and during follow-up thereby likely introducing selection bias. Methods capable of handling missing longitudinal data and informative losses to follow up, such as joint longitudinal survival models, were only used in a minority of studies and many studies are likely to have overstated the reliability of findings and applicability to populations of interest. Measures of change in renal function and definitions of progression varied substantially between studies, revealing a lack of consensus on clinically important and statistically robust measures in the study of CKD progression.

Unlike prospective cohort studies and clinical trials which prospectively identify patients for research and take efforts to follow up patients regularly and completely over time, retrospective analysis of routine healthcare data relies on data collected for the purposes of clinical

Table 4. Listing of CKD progression measures in reviewed articles (52 of 80 articles).

Methods	Rule[a]	Term	Author [ref][b]	Year	Avg follow up	Sample size	Other methods[a]
Individual linear regression	eGFR slope decline: > **3 ml/min/1.73m^2/year**	Progressors	Chase HS et al. [11]	2014	6 years	481	Naïve Bayes classifier; logistic regression
	eGFR slope decline: > **median (8.1) ml/min/ 1.73m^2/year**	Relatively rapid eGFR decline	Wang Y et al. [12]	2019	2 years	128	Logistic regression
	eGFR slope decline: > **mean (1.5) ml/min/1.73m^2/ year**	Faster decline	Abdelhafiz AH et al. [13]	2012	14 years	100	Logistic regression
Linear mixed model	eGFR slope decline: > **5 ml/min/1.73m^2/year**	Rapid progression	Eriksen BO et al. [14]	2006	3.7 years	3,047	Slope interactions
	eGFR slope decline: > **4 ml/min/1.73m^2/year**	Rapid progression	Jalal K et al. [15]	2019	> = 3 years	10,927	N/A
	eGFR slope decline: > **3 ml/min/1.73m^2/year**	eGFR slope decline	Cabrera CS et al. [16]	2020	4.3 years	30,222	Cox PH regression
	eGFR slope decline: > **0 ml/min/1.73m^2/year**	Progressors (vs non-progressors)	Eriksen et al. [17]	2010	4 years	1,224	2-level model
	eGFR slope decline: > **0 ml/min/1.73m^2/year**	eGFR decline	Annor FB et al. [18]	2015	4 years	575	Structural equation modelling
	eGFR predicted percent rate of decline: > **5% per year**	Progression	Diggle PJ et al. [19]	2015	4.5 years	22,910	Piecewise linear mixed model
Absolute change between measures	eGFR drop at any time: > **10 ml/min/1.73m^2**	Progression	Butt AA et al. [20]	2018	3 months	17,624	Difference in proportions chi-squared test
Percent change between measures	eGFR percent drop: >**10%; >20%**	Progression	Singh A et al. [21]	2015	1 year	6,435	Logistic regression
	eGFR percent drop: >**15%**	Progressive renal impairment	Evans RDR et al. [22]	2018	5 years	24	Descriptive result only
	eGFR percent drop: >**20%**	Transient or persistent renal function decline	Jackevicius CA et al. [23]	2021	Approx. 1.4 years	49,458	Cox PH regression
	eGFR percent drop: >**25%**	Progression	Lai YJ et al. [24]	2019	1 year	1,620	Cox PH regression
	eGFR percent drop: >**25%** (**AND** increase in CKD stage)	Progression	Vejakama P et al. [25]	2015	4.5 years	32,106	Competing risks survival models
	eGFR percent drop: >**30%**	"30% decline in eGFR"	Posch F et al. [26]	2019	1.4 years	14,432	Cox PH regression
	eGFR percent drop: >**30%**	Renal function decline	Hsu TW et al. [27]	2019	5 years	5,046	Cox PH regression
	eGFR percent drop: >**30%**	Rapid eGFR decline	Inaguma D et al. [28]	2020	2 years	9,911	Logistic regression; Random forest regression
	eGFR percent drop: >**30%**	eGFR decline	Peng YL et al. [29]	2020	1.5 years	1,050	Cox PH regression
	eGFR percent drop: >**30%**	(no label)	Yao X et al. [30]	2017	11 months	9,796	Cox PH regression
	eGFR percent drop: >**30%**	"Loss of eGFR >30%"	Lamacchia O et al. [31]	2018	4 years	582	Logistic regression
	eGFR percent drop: >**30%**	eGFR loss	Viazzi F et al. [32]	2018	4 years	535	Logistic regression
	eGFR percent drop: >**30%**	Clinically important decline	Rej S et al. [33]	2020	3.1 years	6,226	Cox PH regression
	eGFR percent drop: >**30%; 30–50%; and 50%**	Progression	Yoo H et al. [34]	2019	5.7 years	478	Kaplan meier with log-rank test
	eGFR percent drop: >**40%** (or RRT initiation)	RRT40	Tangri N et al. [35]	2021	3.9 years	32,007	Cox PH regression
	eGFR percent drop: >**50%**	Renal survival endpoint	Lv L et al. [36]	2017	3.1 years	208	Cox PH regression
	Serum creatinine percent increase: >**50%**	Worsening renal function	Li XM et al. [37]	2016	1.8 years	44	Descriptive results only
	Estimate creatinine clearance percent drop: >**0%**	Decline in creatinine clearance	Gallant JE et al. [38]	2005	1 year	658	Descriptive results only

(*Continued*)

Table 4. (Continued)

Methods	Rule[a]	Term	Author [ref][b]	Year	Avg follow up	Sample size	Other methods[a]
Rate of change between measures	eGFR drop per time elapsed (assumed):	Progressive GFR decline	Herget-Rosenthal S et al. [39]	2013	3 years	803	Logistic regression
	> 2.5 ml/min/1.73m^2/year						
	eGFR drop per time elapsed: **> 3 ml/min/1.73m^2/year**	Rapid progression	Morales-Alvarez MC et al. [40]	2019	Not stated	594	Descriptive comparisons
	eGFR drop per time elapsed: **> 5 ml/min/1.73m^2/year**	eGFR decline	Nderitu P et al. [41]	2014	9 months	4,145	Logistic regression
	eGFR drop per time elapsed: **> 5 ml/min/1.73m^2/year**	Fast progression	Koraishy FM et al. [42]	2017	Not stated	2,170	Logistic regression
	eGFR drop per time elapsed (assumed): **> 5 ml/min/1.73m^2/year**	Progressive CKD	Johnson F et al. [43]	2015	Not stated	200	Difference in proportions chi-squared test
	eGFR drop per time elapsed: **> 5 ml/min/1.73m^2/year**	Rapid decline	Chakera A et al. [44]	2015	7 years	147	Logistic regression
	eGFR percent drop per time elapsed (assumed): **>5% per year**	Rapid kidney function decline	Chen H et al. [45]	2014	3 years	365	Logistic regression
Change in CKD stage, based on measures	Population: incident CKD stage 3 (2 x eGFR < 60 over > 3 months);	CKD progression from stage 3 to 4	Perotte A et al. [46]	2015	Not stated	2,908	Cox proportional hazards regression
	Outcome: **2 x eGFR <30 over >3 months**						
	Increase in CKD stage: **By one or more stages**	Worsening in CKD stage	Cummings DM et al. [47]	2011	7.6 years	791	Logistic regression
	Increase in CKD stage: **By one or more stages (eGFR values or diagnostic codes)**	Declining kidney function	Horne L et al. [48]	2019	Not stated	195,178	Crude estimation of incidence rate
	Increase in CKD stage: **By one or more stages (eGFR values or coded RRT)**	CKD stage worsening	Robinson DE et al. [49]	2021	Approx. 3.7 years	19,324	Competing risks survival models
	Increase in CKD stage: **By one stage**	Progression of kidney dysfunction to next CKD stage	Nicolos GA et al. [50]	2020	5 years	Approx 37,000	Life-table analysis
	Increase in CKD stage / risk category: **To very high risk category (eGFR <30 and proteinuria (-); eGFR <45 and proteinuria (±); eGFR < 60 and proteinuria (+))**	Diabetic kidney disease progression	Yanagawa T et al. [51]	2021	6.2 years	681	Cox PH regression
	Change in CKD stage: **From and to any stage, summarised by initial and final stage**	Transition between CKD stages	Vesga JI et al. [52]	2021	6-month intervals	1,783	Crude estimation
Binary progression to threshold value	Threshold eGFR: **median eGFR < 30, for at least 3 consecutive months**	Nephrotoxicity	Oetjens M et al. [53]	2014	8.8 years	115	Cox PH regression
	Threshold eGFR: **2 x eGFR<30 over ≥90 days with no intermediate eGFR>30**	Advanced CKD	Neuen BL et al. [54]	2021	2.9 years	91,319	Cox PH regression
	Threshold eGFR: **2 x eGFR<30 over ≥90 days with no intermediate eGFR>30 (or a stage 4–5 code)**	Incident CKD stages 4–5	Weldegiorgis M et al. [55]	2019	7.5 years	1,397,573	Cox PH regression
	Threshold eGFR: **< 45 ml/min/1.73m^2**	Progression to CKD stage 3b	Niu SF et al. [56]	2021	3.0 years	3,114	Cox PH regression
	Threshold eGFR: **< 15 ml/min/1.73m^2**	Renal survival endpoint	O'Riordan A et al. [57]	2009	3.2 years	54	Kaplan meier estimation; log-rank test
	Threshold eGFR: **ESRD (eGFR<15 or dialysis)**	Progression to ESRD	Tsai CW et al. [58]	2017	4.2 years	739	Cox PH regression
Binary progression (changes/threshold combination)	eGFR percent drop: **>50%**	Renal event	Leither MD et al. [59]	2019	5.3 years	196,209	Cox PH regression
	AND						
	Threshold eGFR: **2 x eGFR <30**						
	eGFR percent drop: **>50%**	"ESRD or an irreversible reduction in eGFR"	Liu D et al. [60]	2019	3.7 years	455	Cox PH regression
	OR						
	Threshold eGFR: **ESRD**						
	eGFR percent drop: **>50%**	CKD progression	Rincon-Choles H et al. [61]	2017	2.8 years	1,676	Competing risks survival models
	OR						
	Threshold eGFR: **ESRD**						

(*Continued*)

Table 4. (Continued)

Methods	Rule[a]	Term	Author [ref][b]	Year	Avg follow up	Sample size	Other methods[a]
Latent class non-linear mixed models	Prediction of latent eGFR **trajectory class**, 6 categories	Trajectory category*	VanWagner LB et al. [62]	2018	1 year	671	Logistic regression, conditional on class

[a]In time-to-event analyses (e.g. Cox PH regression, competing risks survival models), the rule for progression can be met at any time during data collection, utilising repeated test results over time. In binary analyses (e.g. logistic regression), the rule is applied once per patient, likely at a specific time which may vary between studies.
[b]For consistency, article reference numbers [ref] also match those provided in the supplementary S3 File listing of reviewed studies.

https://doi.org/10.1371/journal.pone.0264167.t005

care. While monitoring guidelines may be in place in healthcare systems that aim to ensure regular follow up of patients at risk of CKD progression, such guidelines may be followed at the discretion of healthcare providers, and frequency of testing and time between tests is likely to be influenced by patient risk. If patients are sampled for analysis on the basis of threshold levels of data completeness over time, there is a risk of disproportionately including patients in analysis that are followed up more regularly as a result of their evolving risk profile (selection bias) and that remain both alive and free of RRT long enough to meet the follow up criteria (survival bias). In addition, if data is collected in a single care setting but patients are managed in different care settings based on their risk, data may be informatively missing where patients move between care settings (ascertainment bias). It is highly likely that studies using EHRs that exclude patients from analysis due to poor data completeness or fail to follow up patients equally among different risk groups will have unreliable results, and results may reflect an unknown subgroup of the target population. The use of such studies to inform clinical decision-making may therefore fail to benefit the community as hoped.

There are a number of methodological challenges in longitudinal analysis of renal function that are not necessarily specific to EHRs but that are important considerations for researchers, discussed in more detail in [10, 63] and introduced earlier. In the absence of acute kidney injury, mixed effects models with patient random effects may improve estimation of changes over time compared to individual linear regressions which may lead to more extreme slope estimations. Such models allow sharing of information between patients, assuming a common mean trajectory, and they allow patients to be included in analysis with variable levels of data completeness to avoid excluding patients from analysis unnecessarily. Other benefits are the ability to perform the entire analysis (comparing exposures and outcomes) in a single model, without the loss of information and under-estimation of standard errors that may result from a 2-step model that estimates individual changes prior to further modelling. CKD is a heterogeneous disease, with various possible contributing causes and pathways of progression. Linear mixed models typically assume a common mean trajectory but other methods are available if this assumption is too strong. While random slope models allow individual trajectories to vary around a common mean slope, more sophisticated models such as latent class mixed models allow modelling of trajectory groups which may be linear or non-linear and correspond to sub-populations of patients. Another challenge is competing risk of mortality and how to handle the initiation of RRT in the analyses of repeated renal function tests, where such events are likely to be associated with rate of decline. An analysis that does not account for informative censoring may lead to biased results. Joint longitudinal survival models and competing risks survival models can be used to account for competing risks if data is available (this may require data linkage to external databases to obtain information on competing event dates).

A major finding of this review was the extreme variation in definitions of CKD progression used, and the clinical importance of each definition was unclear. More work has been done in

the last decade to identify clinically important measures of progression of CKD. In 2012, the United States Food and Drug Administration (FDA) commissioned research to identify new endpoints of CKD progression for use in clinical trials [64, 65]. Definitions were developed using data from the Chronic Kidney Disease Prognosis Consortium (CKD-PC) that showed strong association with important clinical outcomes of progression to ESRD and all-cause mortality, including thresholds of reduction in eGFR between measures of 30% and 40% over approximately 2 years, stratified by baseline eGFR. Further research that aims to define new outcomes of smaller clinically meaningful changes in renal function would be useful, as this may enable earlier identification of progression of CKD that would be useful in clinical practice, and future EHR studies could adopt such outcomes for research.

Strengths of this review include the large number of databases utilised and studies reviewed and detailed data extraction efforts, allowing a comprehensive evaluation of how well data quality issues were handled and acknowledged. The review was however limited to peer-reviewed articles and those that clarified in their abstract that repeated renal function tests were used in analysis. Limitations include the limitation to articles written in English, lack of inclusion of grey literature and issues with ascertaining whether EHRs were used as opposed to other methods of extraction from paper records. Despite this, the majority of data issues present will be the same regardless of whether electronic or paper records were used. Retrospective studies using traditional paper records will suffer from the same problems as those using electronic health records: incomplete records, variation in logging practices, addressing AKI when modeling CKD progression, loss to follow-up and competing risks.

Conclusions

Many studies using EHRs to study progression of CKD do not fully acknowledge the biases that result from poor data quality inherent in EHRs and reporting was poor. While some studies have defined CKD progression measures similar to those validated by FDA in 2012 [64, 65] showing an understanding of identifying clinically important changes in renal function, recommendations following the systematic review by Boucquemont et al. review in 2014 [10] have not been implemented on a broader scale. Observational studies using EHRs should follow the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) [66, 67] and REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) [68] guidelines, which aim to improve transparency and clarity in reporting of research. Research publications should clearly state the care pathway and intended follow up framework, data completeness eligibility criteria, the percentage of the target population excluded based on those criteria, whether there were differences in characteristics of those included vs. excluded and according to important risk factors, as well as rates of loss to follow up. Where possible, researchers should attempt to ascertain reasons for loss to follow up, which may involve linkage to external data. Researchers should consider using existing validated outcomes of CKD progression and we hope that heterogeneity in definitions of CKD progression will improve over time. Focussing research questions on populations for which regular data collection is performed as part of routine care may offer a route to better quality data on changes of renal function over time and important changes in renal function will be easier to identify accurately in patients with reduced renal function at baseline, such as those with established CKD where GFR-estimating equations perform better.

Supporting information

S1 File. PRISMA checklist.
(DOC)

S2 File. MEDLINE database search strategy.
(DOCX)

S3 File. List of reviewed studies.
(DOCX)

S1 Table. Summary of study populations, where unclear if EHRs used.
(DOCX)

S2 Table. Study methodology, where unclear if EHRs used.
(DOCX)

S3 Table. Critique of handling of data quality and methodological challenges, where unclear if EHRs used.
(DOCX)

S4 Table. Listing of key features of all included studies, sorted by year of publication.
(DOCX)

S1 Data. Data extraction spreadsheet.
(XLSX)

Acknowledgments

Only the listed authors contributed to the work reported in this manuscript.

Author Contributions

Conceptualization: Faye Cleary, David Prieto-Merino, Dorothea Nitsch.

Data curation: Faye Cleary.

Funding acquisition: Faye Cleary.

Investigation: Faye Cleary.

Methodology: Faye Cleary, David Prieto-Merino, Dorothea Nitsch.

Project administration: Faye Cleary.

Supervision: David Prieto-Merino, Dorothea Nitsch.

Validation: David Prieto-Merino, Dorothea Nitsch.

Writing – original draft: Faye Cleary.

Writing – review & editing: David Prieto-Merino, Dorothea Nitsch.

References

1. Eckardt KU, Coresh J, Devuyst O, Johnson RJ, Kottgen A, Levey AS, et al. Evolving importance of kidney disease: from subspecialty to global health burden. Lancet. 2013; 382(9887):158–69. https://doi.org/10.1016/S0140-6736(13)60439-0 PMID: 23727165
2. Hill NR, Fatoba ST, Oke JL, Hirst JA, O'Callaghan CA, Lasserson DS, et al. Global Prevalence of Chronic Kidney Disease—A Systematic Review and Meta-Analysis. PLoS One. 2016; 11(7):e0158765. https://doi.org/10.1371/journal.pone.0158765 PMID: 27383068
3. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med. 2004; 351(13):1296–305. https://doi.org/10.1056/NEJMoa041031 PMID: 15385656

4. Li L, Astor BC, Lewis J, Hu B, Appel LJ, Lipkowitz MS, et al. Longitudinal progression trajectory of GFR among patients with CKD. Am J Kidney Dis. 2012; 59(4):504–12. https://doi.org/10.1053/j.ajkd.2011.12.009 PMID: 22284441

5. Caravaca-Fontán F, Azevedo L, Luna E, Caravaca F. Patterns of progression of chronic kidney disease at later stages. Clin Kidney J. 2018; 11(2):246–53. https://doi.org/10.1093/ckj/sfx083 PMID: 29644066

6. UK Renal Registry. 20th Annual Report of the Renal Association. NEPHRON 2018; 139 (suppl1). Available: renal.org/audit-research/annual-report

7. UK Renal Registry. UK Renal Registry 22nd Annual Report–data to 31/12/2018, Bristol, UK. Available: renal.org/audit-research/annual-report

8. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009; 150(9):604–12. https://doi.org/10.7326/0003-4819-150-9-200905050-00006 PMID: 19414839

9. The Renal Association. About eGFR. [cited 8 January 2021] Available: https://renal.org/health-professionals/information-resources/uk-eckd-guide/about-egfr

10. Boucquemont J, Heinze G, Jager KJ, Oberbauer R, Leffondre K. Regression methods for investigating risk factors of chronic kidney disease outcomes: the state of the art. BMC Nephrology. 2014; 15(1):45. https://doi.org/10.1186/1471-2369-15-45 PMID: 24628838

11. Chase HS, Hirsch JS, Mohan S, Rao MK, Radhakrishnan J. Presence of early CKD-related metabolic complications predict progression of stage 3 CKD: a case-controlled study. BMC Nephrology. 2014; 15:187. https://doi.org/10.1186/1471-2369-15-187 PMID: 25431293

12. Wang Y, Zhao L, Zhang J, Wu Y, Zhang R, Li H, et al. Implications of a Family History of Diabetes and Rapid eGFR Decline in Patients With Type 2 Diabetes and Biopsy-Proven Diabetic Kidney Disease. Frontiers in Endocrinology. 2019; 10 (no pagination). https://doi.org/10.3389/fendo.2019.00855 PMID: 31920969

13. Abdelhafiz AH, Tan E, Levett C, Minchin J, Nahas ME. Natural history and predictors of faster glomerular filtration rate decline in a referred population of older patients with type 2 diabetes mellitus. Hospital practice (1995) Hospital practice. 2012; 40(4):49–55.

14. Eriksen BO, Ingebretsen OC. The progression of chronic kidney disease: a 10-year population-based study of the effects of gender and age. Kidney International. 2006; 69(2):375–82. https://doi.org/10.1038/sj.ki.5000058 PMID: 16408129

15. Jalal K, Anand EJ, Venuto R, Eberle J, Arora P. Can billing codes accurately identify rapidly progressing stage 3 and stage 4 chronic kidney disease patients: a diagnostic test study. Bmc Nephrology. 2019;20.

16. Cabrera CS, Lee AS, Olsson M, Schnecke V, Westman K, Lind M, et al. Impact of CKD Progression on Cardiovascular Disease Risk in a Contemporary UK Cohort of Individuals With Diabetes. Kidney International Reports. 2020; 5(10):1651–60. https://doi.org/10.1016/j.ekir.2020.07.029 PMID: 33102957

17. Eriksen BO, Tomtum J, Ingebretsen OC. Predictors of declining glomerular filtration rate in a population-based chronic kidney disease cohort. Nephron. 2010; 115(1):c41–50. https://doi.org/10.1159/000286349 PMID: 20173349

18. Annor FB, Masyn KE, Okosun IS, Roblin DW, Goodman M. Psychosocial stress and changes in estimated glomerular filtration rate among adults with diabetes mellitus. Kidney Research and Clinical Practice. 2015; 34(3):146–53. https://doi.org/10.1016/j.krcp.2015.07.002 PMID: 26484039

19. Diggle PJ, Sousa I, Asar O. Real-time monitoring of progression towards renal failure in primary care patients. Biostatistics. 2015; 16(3):522–36. https://doi.org/10.1093/biostatistics/kxu053 PMID: 25519432

20. Butt AA, Ren Y, Puenpatom A, Arduino JM, Kumar R, Abou-Samra AB. Effectiveness, treatment completion and safety of sofosbuvir/ledipasvir and paritaprevir/ritonavir/ombitasvir + dasabuvir in patients with chronic kidney disease: an ERCHIVES study. Alimentary Pharmacology & Therapeutics. 2018; 48(1):35–43. https://doi.org/10.1111/apt.14799 PMID: 29797514

21. Singh A, Nadkarni G, Gottesman O, Ellis SB, Bottinger EP, Guttag JV. Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration. Journal of Biomedical Informatics. 2015; 53:220–8. https://doi.org/10.1016/j.jbi.2014.11.005 PMID: 25460205

22. Evans RDR, Cargill T, Goodchild G, Oliveira B, Rodriguez-Justo M, Pepper R, et al. Clinical Manifestations and Long-term Outcomes of IgG4-Related Kidney and Retroperitoneal Involvement in a United Kingdom IgG4-Related Disease Cohort. Kidney International Reports. 2019; 4(1):48–58. https://doi.org/10.1016/j.ekir.2018.08.011 PMID: 30596168

23. Jackevicius CA, Lu LY, Ghaznavi Z, Warner AL. Bleeding Risk of Direct Oral Anticoagulants in Patients With Heart Failure And Atrial Fibrillation. Circulation-Cardiovascular Quality and Outcomes. 2021; 14(2):155–68. https://doi.org/10.1161/CIRCOUTCOMES.120.007230 PMID: 33541109

24. Lai YJ, Lin YC, Peng CC, Chen KC, Chuang MT, Wu MS, et al. Effect of weight loss on the estimated glomerular filtration rates of obese patients at risk of chronic kidney disease: the RIGOR-TMU study.

Journal of Cachexia, Sarcopenia and Muscle. 2019; 10(4):756–66. https://doi.org/10.1002/jcsm.12423 PMID: 30938491

25. Vejakama P, Ingsathit A, Attia J, Thakkinstian A. Epidemiological Study of Chronic Kidney Disease Progression: A Large-Scale Population-Based Cohort Study. Medicine. 2015; 94(4). https://doi.org/10.1097/MD.0000000000000475 PMID: 25634196

26. Posch F, Ay C, Stoger H, Kreutz R, Beyer-Westendorf J. Exposure to vitamin k antagonists and kidney function decline in patients with atrial fibrillation and chronic kidney disease. Research and Practice in Thrombosis and Haemostasis. 2019; 3(2):207–16. https://doi.org/10.1002/rth2.12189 PMID: 31011705

27. Hsu TW, Hsu CN, Wang SW, Huang CC, Li LC. Comparison of the effects of denosumab and alendronate on cardiovascular and renal outcomes in osteoporotic patients. Journal of Clinical Medicine. 2019; 8(7). https://doi.org/10.3390/jcm8070932 PMID: 31261703

28. Inaguma D, Kitagawa A, Yanagiya R, Koseki A, Iwamori T, Kudo M, et al. Increasing tendency of urine protein is a risk factor for rapid eGFR decline in patients with CKD: A machine learning-based prediction model by using a big database. PLoS ONE [Electronic Resource]. 2020; 15(9):e0239262. https://doi.org/10.1371/journal.pone.0239262 PMID: 32941535

29. Peng YL, Tain YL, Lee CT, Yang YH, Huang YB, Wen YH, et al. Comparison of uric acid reduction and renal outcomes of febuxostat vs allopurinol in patients with chronic kidney disease. Scientific Reports. 2020; 10(1):10734. https://doi.org/10.1038/s41598-020-67026-1 PMID: 32612180

30. Yao X, Tangri N, Gersh BJ, Sangaralingham LR, Shah ND, Nath KA, et al. Renal Outcomes in Anticoagulated Patients With Atrial Fibrillation. Journal of the American College of Cardiology. 2017; 70 (21):2621–32. https://doi.org/10.1016/j.jacc.2017.09.1087 PMID: 29169468

31. Lamacchia O, Viazzi F, Fioretto P, Mirijello A, Giorda C, Ceriello A, et al. Normoalbuminuric kidney impairment in patients with T1DM: Insights from annals initiative. Diabetology and Metabolic Syndrome. 2018; 10(1). https://doi.org/10.1186/s13098-018-0361-2 PMID: 30083251

32. Viazzi F, Greco E, Ceriello A, Fioretto P, Giorda C, Guida P, et al. Apparent treatment resistant hypertension, blood pressure control and the progression of chronic kidney disease in patients with type 2 diabetes. Kidney and Blood Pressure Research. 2018; 43(2):422–38. https://doi.org/10.1159/000488255 PMID: 29698970

33. Rej S, Herrmann N, Gruneir A, McArthur E, Jeyakumar N, Muanda FT, et al. Association of Lithium Use and a Higher Serum Concentration of Lithium With the Risk of Declining Renal Function in Older Adults: A Population-Based Cohort Study. The Journal of clinical psychiatry. 2020; 81(5). https://doi.org/10.4088/JCP.19m13045 PMID: 32841553

34. Yoo H, Park I, Kim DJ, Lee S. Effects of sarpogrelate on microvascular complications with type 2 diabetes. International Journal of Clinical Pharmacy. 2019. https://doi.org/10.1007/s11096-019-00794-7 PMID: 30864083

35. Tangri N, Reaven NL, Funk SE, Ferguson TW, Collister D, Mathur V. Metabolic acidosis is associated with increased risk of adverse kidney outcomes and mortality in patients with non-dialysis dependent chronic kidney disease: an observational cohort study. BMC Nephrology. 2021; 22(1). https://doi.org/10.1186/s12882-021-02385-z PMID: 34011303

36. Lv L, Chang DY, Li ZY, Chen M, Hu Z, Zhao MH. Persistent hematuria in patients with antineutrophil cytoplasmic antibody-associated vasculitis during clinical remission: chronic glomerular lesion or low-grade active renal vasculitis? BMC Nephrology. 2017; 18(1):354. https://doi.org/10.1186/s12882-017-0763-7 PMID: 29207950

37. Li XM, Rui HC, Liang DD, Xu F, Liang SS, Zhu XD, et al. Clinicopathological characteristics and outcomes of light chain deposition disease: an analysis of 48 patients in a single Chinese center. Annals of Hematology. 2016; 95(6):901–9. https://doi.org/10.1007/s00277-016-2659-1 PMID: 27056200

38. Gallant JE, Parish MA, Keruly JC, Moore RD. Changes in renal function associated with tenofovir disoproxil fumarate treatment, compared with nucleoside reverse-transcriptase inhibitor treatment. Clinical Infectious Diseases. 2005; 40(8):1194–8. https://doi.org/10.1086/428840 PMID: 15791522

39. Herget-Rosenthal S, Dehnen D, Kribben A, Quellmann T. Progressive chronic kidney disease in primary care: modifiable risk factors and predictive model. Preventive Medicine. 2013; 57(4):357–62. https://doi.org/10.1016/j.ypmed.2013.06.010 PMID: 23783072

40. Morales-Alvarez MC, Garcia-Dolagaray G, Millan-Fierro A, Rosas SE. Renal Function Decline in Latinos With Type 2 Diabetes. Kidney International Reports. 2019; 4(9):1230–4. https://doi.org/10.1016/j.ekir.2019.05.012 PMID: 31517142

41. Nderitu P, Doos L, Strauss VY, Lambie M, Davies SJ, Kadam UT. Analgesia dose prescribing and estimated glomerular filtration rate decline: a general practice database linkage cohort study. BMJ Open. 2014; 4(8):e005581. https://doi.org/10.1136/bmjopen-2014-005581 PMID: 25138808

42. Koraishy FM, Hooks-Anderson D, Salas J, Scherrer JF. Rate of renal function decline, race and referral to nephrology in a large cohort of primary care patients. Family Practice. 2017; 34(4):416–22. https://doi.org/10.1093/fampra/cmx012 PMID: 28334754

43. Johnson F, Phillips D, Talabani B, Wonnacott A, Meran S, Phillips AO. The impact of acute kidney injury in diabetes mellitus. Nephrology. 2016; 21(6):506–11. https://doi.org/10.1111/nep.12649 PMID: 26452246

44. Chakera A, MacEwen C, Bellur SS, Chompuk LO, Lunn D, Roberts ISD. Prognostic value of endocapillary hypercellularity in IgA nephropathy patients with no immunosuppression. Journal of Nephrology. 2016; 29(3):367–75. https://doi.org/10.1007/s40620-015-0227-8 PMID: 26318019

45. Chen H, Liu C, Fu C, Zhang H, Yang H, Wang P, et al. Combined application of eGFR and albuminuria for the precise diagnosis of stage 2 and 3a CKD in the elderly. Journal of Nephrology. 2014; 27(3):289–97. https://doi.org/10.1007/s40620-013-0011-6 PMID: 24609886

46. Perotte A, Ranganath R, Hirsch JS, Blei D, Elhadad N. Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis. Journal of the American Medical Informatics Association. 2015; 22(4):872–80. https://doi.org/10.1093/jamia/ocv024 PMID: 25896647

47. Cummings DM, Larsen LC, Doherty L, Lea CS, Holbert D. Glycemic Control Patterns and Kidney Disease Progression among Primary Care Patients with Diabetes Mellitus. Journal of the American Board of Family Medicine. 2011; 24(4):391–8. https://doi.org/10.3122/jabfm.2011.04.100186 PMID: 21737763

48. Horne L, Ashfaq A, MacLachlan S, Sinsakul M, Qin L, LoCasale R, et al. Epidemiology and health outcomes associated with hyperkalemia in a primary care setting in England. BMC Nephrology. 2019; 20 (1):85. https://doi.org/10.1186/s12882-019-1250-0 PMID: 30841854

49. Robinson DE, Ali MS, Pallares N, Tebe C, Elhussein L, Abrahamsen B, et al. Safety of Oral Bisphosphonates in Moderate-to-Severe Chronic Kidney Disease: A Binational Cohort Analysis. Journal of Bone and Mineral Research. 2021; 36(5):820–32. https://doi.org/10.1002/jbmr.4235 PMID: 33373491

50. Nichols GA, Deruaz-Luyet A, Brodovicz KG, Kimes TM, Rosales AG, Hauske SJ. Kidney disease progression and all-cause mortality across estimated glomerular filtration rate and albuminuria categories among patients with vs. Without type 2 diabetes. BMC Nephrology. 2020; 21(1).

51. Yanagawa T, Koyano K, Azuma K. Retrospective study of factors associated with progression and remission/regression of diabetic kidney disease-hypomagnesemia was associated with progression and elevated serum alanine aminotransferase levels were associated with remission or regression. Diabetology International. 2021; 12(3):268–76. https://doi.org/10.1007/s13340-020-00483-1 PMID: 34150435

52. Vesga JI, Cepeda E, Pardo CE, Paez S, Sanchez R, Sanabria RM. Chronic kidney disease progression and transition probabilities in a large preventive cohort in colombia. International Journal of Nephrology. 2021; 2021 (no pagination).

53. Oetjens M, Bush WS, Birdwell KA, Dilks HH, Bowton EA, Denny JC, et al. Utilization of an EMR-biorepository to identify the genetic predictors of calcineurin-inhibitor toxicity in heart transplant recipients. Pacific Symposium on Biocomputing. 2014:253–64. PMID: 24297552

54. Neuen BL, Weldegiorgis M, Herrington WG, Ohkuma T, Smith M, Woodward M. Changes in GFR and Albuminuria in Routine Clinical Practice and the Risk of Kidney Disease Progression. American Journal of Kidney Diseases. 2021. https://doi.org/10.1053/j.ajkd.2021.02.335 PMID: 33895181

55. Weldegiorgis M, Smith M, Herrington WG, Bankhead C, Woodward M. Socioeconomic disadvantage and the risk of advanced chronic kidney disease: results from a cohort study with 1.4 million participants. Nephrology Dialysis Transplantation. 2020; 35(9):1562–70. https://doi.org/10.1093/ndt/gfz059 PMID: 31329936

56. Niu SF, Wu CK, Chuang NC, Yang YB, Chang TH. Early Chronic Kidney Disease Care Programme delays kidney function deterioration in patients with stage I-IIIa chronic kidney disease: an observational cohort study in Taiwan. Bmj Open. 2021; 11(1). https://doi.org/10.1136/bmjopen-2020-041210 PMID: 33468527

57. O'Riordan A, Dutt N, Cairns H, Rela M, O'Grady JG, Heaton N, et al. Renal biopsy in liver transplant recipients. Nephrology Dialysis Transplantation. 2009; 24(7):2276–82. https://doi.org/10.1093/ndt/gfp112 PMID: 19293134

58. Tsai CW, Lin SY, Kuo CC, Huang CC. Serum Uric Acid and Progression of Kidney Disease: A Longitudinal Analysis and Mini-Review. PLoS ONE [Electronic Resource]. 2017; 12(1):e0170393. https://doi.org/10.1371/journal.pone.0170393 PMID: 28107415

59. Leither MD, Murphy DP, Bicknese L, Reule S, Vock DM, Ishani A, et al. The impact of outpatient acute kidney injury on mortality and chronic kidney disease: a retrospective cohort study. Nephrology Dialysis Transplantation. 2019; 34(3):493–501. https://doi.org/10.1093/ndt/gfy036 PMID: 29579290

60. Liu D, You J, Liu Y, Tang X, Tan X, Xia M, et al. Serum immunoglobulin G provides early risk prediction in immunoglobulin A nephropathy. International Immunopharmacology. 2019; 66:13–8. https://doi.org/10.1016/j.intimp.2018.10.044 PMID: 30415190

61. Rincon-Choles H, Jolly SE, Arrigain S, Konig V, Schold JD, Nakhoul G, et al. Impact of Uric Acid Levels on Kidney Disease Progression. American Journal of Nephrology. 2017; 46(4):315–22. https://doi.org/10.1159/000481460 PMID: 29032376

62. VanWagner LB, Montag S, Zhao L, Allen NB, Lloyd-Jones DM, Das A, et al. Cardiovascular Disease Outcomes Related to Early Stage Renal Impairment After Liver Transplantation. Transplantation. 2018; 102(7):1096–107. https://doi.org/10.1097/TP.0000000000002175 PMID: 29557907

63. Shou H, Hsu JY, Xie D, Yang W, Roy J, Anderson AH, et al. Analytic Considerations for Repeated Measures of eGFR in Cohort Studies of CKD. Clin J Am Soc Nephrol. 2017; 12(8):1357–65. https://doi.org/10.2215/CJN.11311116 PMID: 28751576

64. Levin A, Agarwal R, Herrington WG, Heerspink HL, Mann JFE, Shahinfar S, et al. International consensus definitions of clinical trial outcomes for kidney failure: 2020. Kidney International. 2020; 98(4):849–59. https://doi.org/10.1016/j.kint.2020.07.013 PMID: 32998816

65. Levey AS, Inker LA, Matsushita K, Greene T, Willis K, Lewis E, et al. GFR decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration. Am J Kidney Dis. 2014; 64(6):821–35. https://doi.org/10.1053/j.ajkd.2014.07.030 PMID: 25441437

66. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Int J Surg. 2014; 12(12):1500–24. https://doi.org/10.1016/j.ijsu.2014.07.014 PMID: 25046751

67. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008; 61(4):344–9. https://doi.org/10.1016/j.jclinepi.2007.11.008 PMID: 18313558

68. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015; 12(10):e1001885. https://doi.org/10.1371/journal.pmed.1001885 PMID: 26440803

3.4 Key findings and implications for PhD

- **Research that uses EHRs to study progression of CKD has increased rapidly** over recent decades, predominantly for the purposes of casual inference.
- Research overwhelmingly studied **changes in eGFR**, with substantial **variation in methods** for evaluating such changes, which included regression slopes, absolute or percent changes between measures, and progression to eGFR thresholds (among others).
- Most commonly used statistical methods were linear mixed models, linear regression, logistic regression, and Cox regression. More sophisticated methods taking into account drop-outs due to death or for other reasons were rare.
- There was **substantial variation in the way researchers defined CKD progression** including measures used and terms applied, revealing a lack of consensus on clinically important and statistically robust measures.
- Significant **failures in transparency of data completeness** and its implications on study results and conclusions were common.
- In existing studies, there is commonly a **high risk of selection bias and ascertainment bias**, occurring due to methods used to select patients for analysis based on data availability criteria, and limited use of statistical methods that account for missing data, biases that are not usually well discussed.
- To improve transparency, future research should clearly state the care pathway and intended follow-up framework, data completeness eligibility criteria, percentage of the target population excluded based on those criteria, and whether there were differences in characteristics of those included vs. excluded.

3.5 Related co-authored research

After finalising the systematic review, I was co-author on an invited review article entitled: "Defining measures of kidney function in observational studies using routine health care data: methodological and reporting considerations". My contribution included sharing my knowledge on insights in heterogeneity in measures of CKD progression identified in the literature and reviewing and suggesting edits on the research article text. A copy of this article is included in *Appendix 2*.

3.6 Recommendations for future research

Following completion of the systematic review, I consolidated my learnings to propose advice on considerations regarding study methodology and transparency of reporting in future studies. Recommendations are listed in *Appendix 3*.

Chapter 4

Data sources

4.1 Chapter summary

This chapter states the rationale for UK and Swedish data sources used in this thesis, providing background on UK and Swedish healthcare systems, availability of EHRs for research, and details of databases used for analyses

4.2 Rationale for UK data sources

At the outset of the PhD, our focus was towards improving knowledge on CKD progression in the UK. Research was enabled by excellent EHR data resources being available in the UK to support observational research. UK data sources were used for research papers 2 and 3.

4.3 Background to UK databases

In the UK, universal healthcare is provided to all citizens through the National Health Service (NHS). The first point of access is generally through primary care general practitioners (GPs), who are the gatekeepers for referral to secondary care services, with over 98% of the population registered with a GP [61].

There are continuing efforts to computerise EHRs in the UK [62], as well as to extract some of this data for research purposes [63]. This has been very successful for primary care. Healthcare recording in primary care services has been computerised since the early 1990s [64] with direct patient care being coded in drop-down menus, and electronic prescriptions being issued. However, many secondary and tertiary care services still use paper in some aspects of their care delivery, and/or just recently have switched over to electronic patient records.

Electronically recorded healthcare records in primary care have enabled data from routine care to be digitally extracted and transformed into databases that can be used for

epidemiological research [63]. In contrast, secondary care services used an approach of retrospectively coding paper records that were generated for clinical care. The most prominent example is for example the Hospital Episode Statistics (HES) which codes up discharge letters issued by specialists for hospitalised patients in secondary care after the event [54,65].

Collation of existing healthcare record data for research may be disjointed due to differences in the way data from the clinical consultation are recorded and/or are extracted in different settings. Variation in data completeness may occur due to patient healthcare seeking behaviours, healthcare practitioner behaviours, and as a result of financial incentivisation of disease recording and management occurring for example due to the QOF framework in primary care [32]. There are also pay-for-performance targets for secondary care which used to be in action, called Cquins [66].

To fully understand the health experience of patients, different databases are required to be linked together to improve data completeness across the full care pathway.

Primary care data

UK primary care databases arise from a number of limited software suppliers [48]. If made available for research, these provide valuable resources owing to extensive population coverage. Available data includes clinical diagnoses and laboratory test results ordered in primary care. There is potential for data collection over long periods of time provided patients do not change their address/GP. In addition, some information about secondary care interactions may be recorded, following discharge letters sent from hospitals to the GP, but may be incomplete [54]. Numerous large comprehensive research data resources have been established containing primary care EHRs (including for example the Clinical Practice Research Datalink [CPRD][48], The Health Improvement Network [THIN][67] and QResearch [68]), with pseudonymisation removing patient identifiable information. Such resources are generally improving and expanding over time.

Secondary care data

In England, secondary care data are available for all NHS hospitals through the HES database, which includes both emergency and non-emergency admissions and outpatient data [69]. Admissions data are recorded by “spell” of care, which is a defined period of care under

a particular treating consultant, and there may be multiple spells during a single admission. Data collected in each spell include patient demographics and clinical diagnoses, but no laboratory assessments are recorded in HES.

Disease registries

Disease registries may provide more accurate and complete data on disease diagnoses and related data than other routine healthcare databases. An example is the UK Renal Registry (UKRR) database although this was not available for this PhD research. The UK renal registry extracts secondary care laboratory results from all English renal databases in secondary care, and gets sent secondary care AKI laboratory alerts from all laboratories in England [70]. This requires overcoming practical issues with frequently changing secondary care software suppliers which are more diverse than those for primary care.

Other administrative data

The Office for National Statistics (ONS) maintains a death registry, which records date and cause of death, for all citizens in England and Wales [71].

Linking routinely collected data without individual consent

There are significant challenges in creating data resources that map the entirety of the patient journey, as all these data are collected in routine consultation without individual patient consent for these data to be used for later research. Hence, there are safeguards implemented under the Data Protection Act which need to be met for data to be linked for the purposes of specific research. Data linkages are done in England mostly by what was called NHS Digital, and requests for such linkages required permissions/data access requests which may take years to obtain. For my PhD I used already available and linked data with permissions in place to carry out the studies presented in this work. I detail these data resources below.

4.4 National Chronic Kidney Disease Audit (NCKDA) database

In 2016, a large national audit of kidney care was carried out in England and Wales, covering approximately 14% of the population of England and Wales, which was representative of the overall population in terms of age and sex [16]. The resultant National Chronic Kidney

Disease Audit (NCKDA) database held all creatinine test results and laboratory reported eGFR results reported between 2008 up until data extraction (mostly occurring in 2016) in adults registered at recruited primary care practices, as well as demographic characteristics and coded risk factors for CKD evaluated at data extraction.

Rationale for utilisation of this database for this PhD was the substantial coverage of the database and large sample size of identified patients with both electronically coded CKD and biochemical evidence confirming presence of CKD. Limitations were the cross-sectional nature of the database, sampling only those patients who were alive at data extraction and had at least one of the following: any creatinine test since 2008; any coded risk factor for CKD; or any renal code.

This database was used for research paper 2 *(feasibility analysis for evaluation of the natural history of kidney disease in the general CKD population)* and research paper 3 *(association between practice CKD coding and subsequent hospitalisations and death [England only])*, with linkage to HES and ONS mortality (**Figure 6**). Details of study cohorts are presented in the respective research paper chapters.

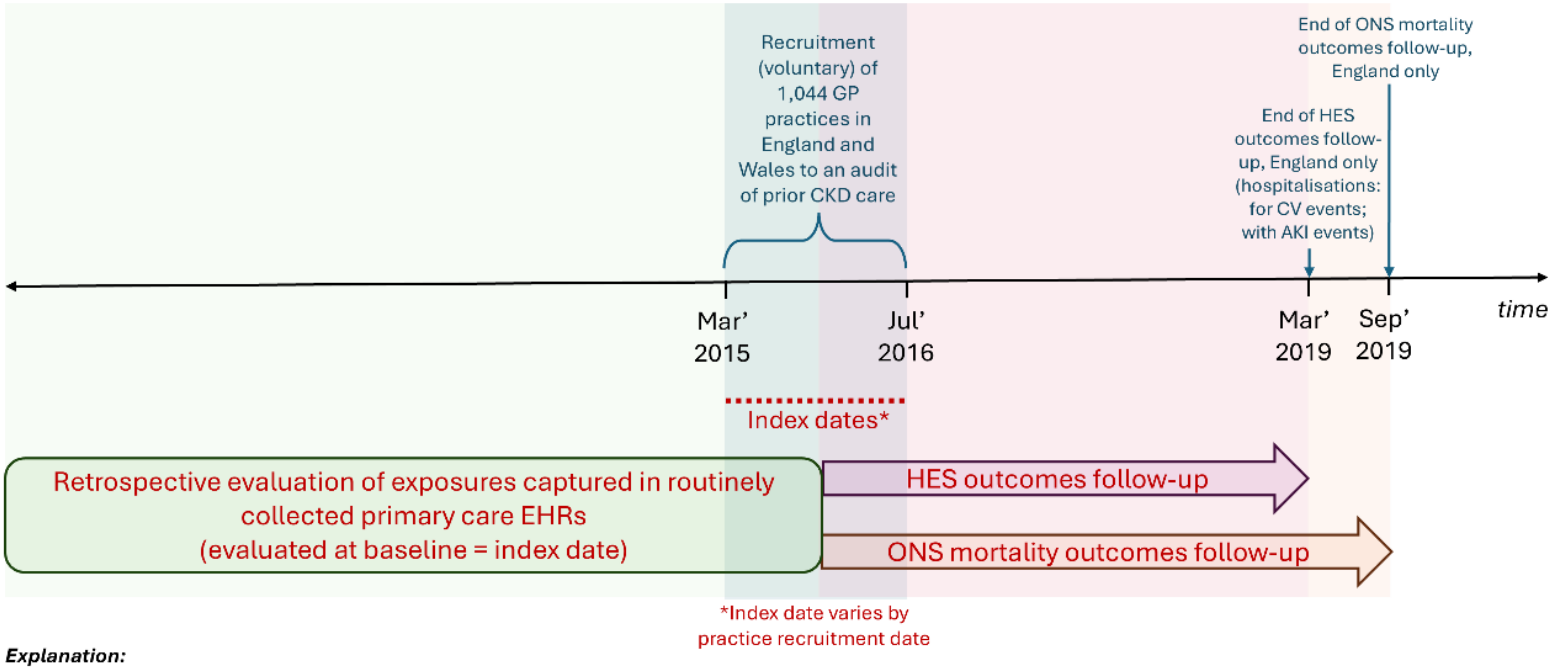
Linkage to HES Admitted Patient Care (APC) facilitated identification of the following outcomes: hospitalisation for CV events; hospitalisation for heart failure; and hospitalisation with AKI. Details of definitions of outcomes are presented in the relevant research paper chapter.

Linkage to ONS mortality was used to identify death outcomes.

4.5 Problems with UK data sources and rationale for Swedish data

The NCKDA database is a cross-sectional database, which limits the types of epidemiological studies that can be conducted in the study of CKD progression. While it captured all serum creatinine test results between 2008-2016, it was limited to patients who were alive at data extraction (2016). This means it could not be used, for example, to identify patients with incident CKD and follow those patients up for eGFR-based progression outcomes, due to survival bias. For example, patients who had progressed rapidly and died would have been unfairly excluded from analysis.

Figure 6. Timeframes for data capture using the NCKDA database and linked HES and ONS mortality data for analyses in research papers 2 and 3, including index dates, retrospective evaluation of primary care exposure data and prospective capture of hospital and mortality outcomes data



Explanation:

1,044 GP practices in England and Wales were recruited to an audit of CKD care between March 2015 and July 2016, with extraction of historical records for individuals who were alive at the time of GP data extraction and had at least one of the following: any creatinine test between 2008 and data extraction; any coded NICE-defined risk factor for CKD; or any renal code. Biochemical evidence of CKD (2 x eGFR, over 90+ days) was evaluated (confirmed) in the 2 years prior to data extraction, using the most recent measures available. History of co-morbidities and patient demographics were evaluated at the index date (time of GP data extraction).

In research paper 2, frequency and completeness of historical eGFR testing was evaluated by risk factor status at baseline (using denominators capturing all patients in analysed GP practices, if not all patients in the relevant risk group were extracted, where appropriate).

In research paper 3, the cohort of patients with evidence of biochemical CKD at the index date (GP data extraction) in 695 GP practices in England were linked to HES and ONS mortality, and were subsequently followed for risks of hospitalisations and death, exploring the association between GP practice coding for CKD at baseline and individual patient outcomes occurring in the next ~4 years.

Another problem is that it only included results of tests requested in primary care, with secondary care results reported elsewhere and not available for research. However, the predominant issue was that we were unable to link to the UK Renal Registry in the course of this PhD, that is required to capture outcomes of kidney failure requiring KRT.

As a result, we reached out to colleagues at the Karolinska Institute who have built the Stockholm CREAtinine Measurements (SCREAM) database [49]. It captures routinely collected healthcare data in the region of Stockholm, Sweden, and is maintained for research purposes. It is an integrated healthcare database providing a detailed view of the end-to-end patient healthcare experience and outcomes for those assessed for kidney function in routine care (described in more detail below).

Longitudinal collection of creatinine results over a long period of time and linkage to renal registry data make it an excellent data resource for studying progression of CKD. Although our initial research aims were to study CKD progression in the UK, the Swedish population are expected to be similar in many ways, and we intend to carry out (or encourage) further validation of work carried out using the SCREAM database in the UK, in due course.

The SCREAM database was used for research paper 4.

4.6 Stockholm CREAtinine Measurements (SCREAM) database

As in the UK, Sweden benefits from universal healthcare coverage, and extracted healthcare data is therefore likely to capture a (reasonably) representative sample of the population of interest, subject to data completeness. Like in the UK, initial care is provided by primary care services, with referrals to specialist care recommended in specific circumstances requiring further care [72]. Patients may also seek care privately.

The SCREAM project was initiated in 2010, with the primary aims to estimate the burden and consequences of chronic kidney disease (CKD) and to identify inappropriate drug use [49,73]. It involves maintenance of a large integrated healthcare database, which continues to be updated over time, and currently holds all laboratory tests of patients residing or accessing healthcare in the region of Stockholm, who underwent creatinine assessments between 2006-2018. In Sweden, healthcare is managed by regional providers, with the region of Stockholm constituting 20-25% of the total population of Sweden [49]. It includes linkage of various

healthcare data sources, with key components of the database used for this PhD project described below.

Laboratory data

Inclusion criteria for SCREAM requires being a Stockholm resident between 2006-2018 with at least one measurement of serum creatinine or albuminuria collected in any care setting. All kidney function tests carried out during this time period are collected, as well as various other relevant laboratory results. We used this data in research paper 4 *(kidney failure risk prediction modelling)* to identify a CKD cohort for analysis, to evaluate availability of albuminuria, to estimate historical slopes of eGFR decline, and to identify eGFR-based kidney failure outcomes (with further details in the next chapter).

Regional healthcare utilisation data

Regional healthcare utilisation data are split into primary care data, outpatient consultations and in-hospital data, and include ICD-10 coded diagnoses and NOMESCO procedure codes recorded in each care setting, including both public and private healthcare interactions [49]. Demographics data are also included. We used this data to define analysis covariates, including demographic characteristics, co-morbidities and recent experience of adverse events.

Disease registries

The Swedish Renal Registry (SRR) includes data for Swedish patients referred to a nephrologist and diagnosed with CKD [49]. In particular, of relevance to this PhD, it holds data on initiation of KRT, allowing accurate and complete identification of such outcomes. We used this data to accurately ascertain dates of initiation of KRT, the primary outcome for research paper 4.

Drug registry

The Swedish Prescribed Drug Registry is a nationwide registry established in 2005 which holds information on all dispensed prescription drugs at Swedish pharmacies [74,49]. This register had complete coverage (>99.7%) of all dispensed drugs dispensed between 2006 and 2019. We used this data to identify prescriptions of anti-hypertensive drugs.

Other administrative data

The Swedish Population Registry holds complete data for deaths, including date and cause of death, for all Swedish citizens [49], which was used in statistical modelling with censoring for death.

Chapter 5

Feasibility of evaluation of the natural history of kidney disease in the general population using electronic healthcare records (paper 2)

5.1 Chapter summary

This chapter provide a summary of work done in research paper 2, presents the original published research paper, and lists key findings and implications in the context of the overall thesis.

5.2 Summary of work in the context of aims of PhD

Background

To explore further the opportunities and challenges in using EHRs to study progression of CKD, we investigated whether data is captured in routinely collected EHRs which would be needed to address research questions of interest. This includes data to assess changes in kidney function over time (i.e. repeat tests of eGFR) collected consistently in a representative sample of patients with CKD.

Methods

This study used national audit data extracted from 1044 GP practices in England and Wales to identify completeness of testing of creatinine (required to estimate GFR) in all adult patients, in those with CKD risk factors, and in those identified with CKD, in UK primary care. We assessed availability of repeat creatinine tests and described changes in renal function estimated using individual linear regressions, comparing risk factor subgroups. We also explored the impact of data quality issues that result from changes in creatinine calibration practices over time, differences in laboratory reporting practices, and differing availability of test results between patients, reporting how such factors impact accuracy of estimation of eGFR slopes.

Results

Data completeness was poor in the general adult population but was very good in patients with risk factors for CKD and established CKD. While we found considerable issues with accuracy of estimation of eGFR slopes due to data quality issues, magnitude of over-estimation of slopes (under-estimation of decline) was modest in the majority of patients of interest.

Conclusions

We expect data issues to be less prominent in future data as creatinine calibration issues are improving since 2012, and issues of accuracy of estimation are less common in patients with established and later stage CKD. It was difficult to establish whether eGFR results may have been disproportionately missing in any important patient groups, such as those with later stage CKD who may be managed in secondary care.

5.3 Research paper 2

See next page for original published research paper, and appendix 4 for supplementary materials.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1703701	**Title**	Miss
First Name(s)	Faye		
Surname/Family Name	Cleary		
Thesis Title	Challenges of studying and predicting chronic kidney disease progression and its complications using routinely collected electronic healthcare records		
Primary Supervisor	Dorothea Nitsch		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Clinical Kidney Journal (CKJ)		
When was the work published?	October 2020		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	FC, DN and DP developed the ideas for content of this article. FC carried out data management and analysis and drafted the article. All authors reviewed several draft versions of the manuscript, suggested updates to analysis and interpretations and approved the final manuscript.

SECTION E

Student Signature	Faye Cleary
Date	23 May 2024

Supervisor Signature	Dorothea Nitsch
Date	10/06/2024

Clinical Kidney Journal, 2021, vol. 14, no. 6, 1603–1609

doi: 10.1093/ckj/sfaa175
Advance Access Publication Date: 22 October 2020
Original Article

ORIGINAL ARTICLE

Feasibility of evaluation of the natural history of kidney disease in the general population using electronic healthcare records

Faye Cleary[1], David Prieto-Merino[1], Sally Hull[2], Ben Caplin [3] and Dorothea Nitsch[1]

[1]Department of Non-Communicable Disease Epidemiology, Faculty of Epidemiology and Population Heath, London School of Hygiene and Tropical Medicine, London, UK, [2]Clinical Effectiveness Group, Centre for Primary Care and Public Health, Queen Mary University of London, London, UK and [3]Department of Renal Medicine, University College London Medical School, London, UK

Correspondence to: Faye Cleary; E-mail: faye.cleary@lshtm.ac.uk

ABSTRACT

Background. Knowledge about the nature of long-term changes in kidney function in the general population is sparse. We aim to identify whether primary care electronic healthcare records capture sufficient information to study the natural history of kidney disease.

Methods. The National Chronic Kidney Disease Audit database covers ∼14% of the population of England and Wales. Availability of repeat serum creatinine tests was evaluated by risk factors for chronic kidney disease (CKD) and individual changes over time in estimated glomerular filtration rate (eGFR) were estimated using linear regression. Sensitivity of estimation to method of evaluation of eGFR compared laboratory-reported eGFR and recalculated eGFR (using laboratory-reported creatinine), to uncover any impact of historical creatinine calibration issues on slope estimation.

Results. Twenty-five per cent of all adults, 92% of diabetics and 96% of those with confirmed CKD had at least three creatinine tests, spanning a median of 5.7 years, 6.2 years and 6.1 years, respectively. Median changes in laboratory-reported eGFR (mL/min/1.73 m^2/year) were −1.32 (CKD) and −0.60 (diabetes). Median changes in recalculated eGFR were −0.98 (CKD) and −0.11 (diabetes), underestimating decline. Magnitude of underestimation (and between-patient variation in magnitude) decreased with deteriorating eGFR. For CKD Stages 3, 4 and 5 (at latest eGFR), median slopes were −1.27, −2.49 and -3.87 for laboratory-reported eGFR and −0.89, −2.26 and −3.75 for recalculated eGFR.

Conclusions. Evaluation of long-term changes in renal function will be possible in those at greatest risk if methods are identified to overcome creatinine calibration problems. Bias will be reduced by focussing on patients with confirmed CKD.

Keywords: CKD, CKD progression, creatinine, electronic healthcare records, primary care

Received: 3.4.2020; Editorial decision: 6.7.2020

INTRODUCTION

Chronic kidney disease (CKD) is an irreversible reduction in kidney function that may progress over prolonged time without symptoms. In rare cases, the disease can progress to end-stage renal disease (ESRD) requiring renal replacement therapy (RRT) [1–3]. More common complications preceding ESRD include increased cardiovascular risk, acute kidney injury (AKI), hospital admission and mortality, with increasing risks associated with lower levels of kidney function [4]. Slowing of progression of kidney disease is therefore of great importance to reduce morbidity and burden on healthcare services. Due to its asymptomatic nature, the characteristics of kidney disease progression in the general population from onset to the requirement of dialysis are not well-understood. Improvements in knowledge may lead to better decision-making with potential to delay progression and improve patient outcomes.

In the UK, >99% of the population are registered with a general practitioner (GP), with GPs acting as the gatekeeper to non-emergency specialized care. The National Institute for Health and Care Excellence (NICE) offers evidence-based guidance on managing patients with CKD in primary care and advice on criteria for referral to secondary care. Referral is recommended in a minority of patients including those with CKD Stages 4–5, proteinuria, rapidly declining glomerular filtration rate (GFR), uncontrolled hypertension and genetic renal diseases [3]. In 2004, the Quality and Outcomes Framework was introduced to incentivize long-term condition management in primary care [5]. Performance measures introduced included creatinine testing in patients at high risk of CKD and maintenance of a register of all adults with CKD Stages 3–5 [6, 7]. Recognition of CKD and testing for renal function in primary care has since increased [8].

This article presents the results of a feasibility study investigating whether it is possible to study the natural history of kidney disease using data from primary care electronic healthcare records (EHRs). Using a large database of EHR data extracted in England and Wales in 2015–16, it explores availability of repeat creatinine tests and attempts to describe changes in renal function, within risk factor subgroups. Issues surrounding reliability of estimation of changes in renal function are evaluated, including testing frequency, changes in creatinine calibration practices and gaps in primary care monitoring.

MATERIALS AND METHODS

Database

The National Chronic Kidney Disease Audit database was used for analysis. The audit was a cross-sectional study set up to investigate CKD identification and management in primary care in England and Wales in 2014–16 [9, 10]. It evaluated performance of renal function testing in patients at risk of CKD and coding of CKD for patients with established biochemical CKD Stages 3–5, identified by two estimated GFR (eGFR) measures <60 mL/min/1.73m^2 a minimum of 90 days apart. Data were extracted from 1044 GP practices, for all adult patients alive and registered at the GP practice at data extraction with coded NICE-defined CKD risk factors or at least one creatinine test result recorded between 2008 and data extraction. Data collected included basic demographic characteristics, CKD risk factor codes and all serum creatinine and reported eGFR results recorded between 2008 and data extraction. Age–sex stratified practice list size data and practice ethnicity breakdown were also collected.

Variables

At the time of the audit, the majority of laboratories reported eGFR using the Modification of Diet in Renal Disease (MDRD) study equation, although this is unlikely to be adjusted for ethnicity which is not typically available to the laboratory. We recalculated eGFR using the MDRD study equation, adjusted for age, sex and ethnicity. CKD Stages 3–5 were identified by two recalculated eGFR measures <60 mL/min/1.73m^2 a minimum of 90 days apart, with at least one measure recorded in the last 2 years prior to data extraction. CKD stage was identified using recalculated eGFR for the most recently recorded creatinine test. Throughout this article, the term CKD will refer to patients with biochemically confirmed CKD Stages 3–5 unless otherwise stated. Coded CKD Stages 3–5, as defined by a Read code, were also explored in some analyses. Urinary albumin:creatinine ratio (ACR) may also be used to evaluate severity of CKD, although uptake of repeat ACR testing in primary care is low [9], and change in eGFR is more commonly used to identify progression of renal disease.

Risk factors explored were diabetes, hypertension, cardiovascular disease (CVD) and CKD stage. Co-morbidities were defined by the presence of relevant Read codes recorded at any time prior to data extraction. Analyses of hypertension excluded patients with a diabetes code to reduce likelihood of effects being driven by co-occurring diabetes.

Frequency of repeat creatinine tests was defined as the number of creatinine test results recorded for each patient between 2008 and data extraction. Duration of follow-up was difficult to ascertain due to lack of data on time of registration at a GP practice. Duration of coverage of tests was defined as the time between the first and last creatinine test. Loss to follow-up was arbitrarily defined as having no creatinine test in the last 3 years prior to data extraction but with at least three creatinine tests recorded prior. Read codes were used to identify initiation of RRT, and all creatinine test results captured post-initiation of RRT were excluded from analysis.

Statistical analysis

Availability of repeat creatinine tests. The percentage of adults with at least three creatinine tests was summarized by risk factor to evaluate data completeness for estimation of slopes of change in eGFR. Denominators for underlying health condition risk factor groups were determined by summing the number of patients with coded risk factors in the database. Denominators for the entire adult population and age, sex and ethnicity groups were determined using practice list size data. Missing list size for 56 Welsh practices was imputed using the average list size in Wales. In patients with at least three creatinine tests, the frequency of tests, duration of coverage of tests, average time between tests, percentage lost to follow-up and percentage initiating RRT in those lost to follow-up were summarized by risk factor. Potential reasons for gaps in primary care monitoring might be low priority for testing due to good health or management in secondary care due to advancement of disease. Summaries were repeated in patients with diabetes only, with additional stratification by age and sex.

Slope estimation. *Laboratory reporting practices: creatinine calibration and eGFR reporting rules.* When serum creatinine blood

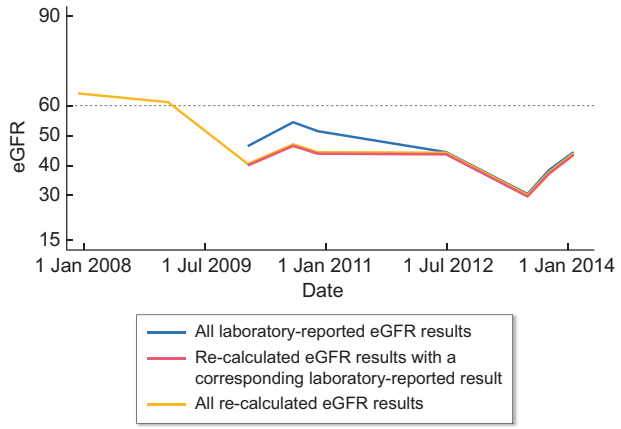


FIGURE 1: Example profile of laboratory-reported and recalculated eGFR results available for an individual patient in the EHR.

tests are ordered in UK primary care, laboratories are required to report eGFR if laboratory-specific criteria are met (usually eGFR < 60 or eGFR < 90) corresponding to thresholds of accuracy for GFR-estimating equations. Prior to estimating GFR, creatinine concentrations must be calibrated to an international reference standard (isotope dilution mass spectrometry). In recent years, laboratories have reported correctly calibrated creatinine results, although historically un-calibrated results may have been reported [12, 13]. Creatinine results in the EHR may, therefore, not be comparable within or between patients over time and it may not be straightforward to identify which results were calibrated from the EHR. Coded data extracted from the EHR does not include information on laboratory creatinine calibration practices or eGFR reporting rules and both may vary by laboratory and over time.

Figure 1 shows an example profile of both laboratory-reported and recalculated eGFR measures over time in an individual patient. In this example, while recently reported and recalculated results are a close match, older recalculated results appear to be underestimated. This is likely due to failure to calibrate historical creatinine results to international standards. Slope analysis using recalculated eGFR may lead to overestimation of slopes (underestimation of rate of decline) if recent measures are more accurate but older measures are underestimated. Also, some recalculated eGFR results do not have a corresponding reported eGFR result due to laboratory reporting conventions and, therefore, slope estimation using reported eGFR may be biased by selective inclusion of test results. Descriptive checks were performed to assess the frequency of reported and recalculated eGFR test results agreeing by +/−1 and by +/−3 mL/min/1.73 m^2, for all available eGFR results in the database, stratified by calendar year and CKD stage.

Slope of change in eGFR was estimated using linear regression for all patients with at least three valid test results, with separate regression models for each patient. This approach is similar to that used by GPs in routine care to estimate individual changes in kidney function and may be subject to measurement error. Mixed modelling was not used in this analysis, due to concerns about the model assumptions imposed. Analysis was carried out separately for laboratory-reported eGFR and recalculated eGFR to evaluate sensitivity of estimation of slopes to

Table 1. eGFR slope regression analysis criteria

Analysis	Test results included
Reported GFR	All laboratory reported eGFR results
MDRD (1)	Recalculated eGFR results for all creatinine test results with a corresponding reported GFR result
MDRD (2)	Recalculated eGFR results for all creatinine test results

method of evaluation of eGFR and analysis of recalculated eGFR was repeated using only those test results with a corresponding laboratory-reported eGFR to evaluate sensitivity of estimation to laboratory-imposed selective inclusion of test results (Table 1). Values of eGFR outside of the valid range (15–150 mL/min/1.73 m^2) and an excess of reported eGFR values of 60 and 90 (likely coded in GP records as >60 or >90 but appearing inaccurately in the database simply as 60 and 90) were excluded from analysis.

Comparisons of slope of eGFR. Boxplots of slope of eGFR were stratified by slope estimation method and by risk factor, CKD stage and testing frequency. Only patients with at least three reported eGFR results were included to restrict comparison to the same population. To reduce impact of outliers, whiskers represent 5% and 95% percentiles. Distribution of slopes in all patients with at least three valid recalculated eGFR test results was tabulated for reference, constituting a different population of likely healthier patients.

Individual differences in slope estimates. Difference between slope estimates was computed for each patient. Discrepancy between reported and recalculated eGFR slopes using corresponding test results only [Reported GFR—MDRD (1)] shows the effect on slope estimation of creatinine calibration issues. Discrepancy between reported eGFR and recalculated eGFR slopes using all creatinine test results [Reported GFR – MDRD (2)] shows the effect on slope estimation of creatinine calibration issues and laboratory reporting restrictions combined.

Boxplots of the distribution of individual differences in slopes were produced by risk factor and CKD stage. Repeat sensitivity boxplots were stratified by ethnicity (coded black or not) to rule out differences being driven by failures to correct for ethnicity in laboratory-reported results. Descriptive paired *t*-tests were used to identify any statistically significant mean difference in slopes for each comparison by risk factor.

RESULTS

Study population

The audit database covered a population of ∼6.5 million adults and was representative of the general population in terms of age and sex. Of the underlying adult population, ∼6% of patients had a diabetes code, 18% had a hypertension code, 6% had a CVD code, 4% had a CKD code and 4% had confirmed CKD.

Availability of repeat tests

About 2.2 million patients (34%) had at least one creatinine test, 1.6 million (25%) had at least three creatinine tests and 1.1 million (17%) had at least three valid laboratory-reported eGFR results. Approximately 5000 patients (<0.1%) had a code for RRT initiation at any time, with around half of those codes

Table 2. Availability of repeat creatinine tests in primary care in all adults and by risk factor

Risk factor	Number of patients	Patients with ≥3 tests (N, %)	Test frequency[a] (median + IQR)	Duration of test coverage, years[a] (median + IQR)	Time (months) between tests[a] (median + IQR)	No test in last 3 years[a] (N, %)	Coded RRT if no test in last 3 years (N, %)
All adults	6 513 000	1 597 629 (24.5%)	7 (5, 10)	5.7 (4.2, 6.4)	8.4 (6.1, 11.2)	39 091 (2.4%)	2.0%
Age							
18–39	2 301 700[b]	59 187 (2.6%)	4 (3, 6)	4.0 (2.5, 5.5)	9.3 (6.2, 13.3)	3 338 (5.6%)	2.6%
40–59	2 214 100[b]	419 144 (18.9%)	5 (4, 8)	5.1 (3.4, 6.1)	9.2 (6.6, 12.5)	14 732 (3.5%)	1.7%
60–79	1 578 600[b]	824 468 (52.2%)	7 (5, 10)	5.9 (4.5, 6.5)	8.4 (6.2, 11.0)	15 960 (1.9%)	2.2%
80+	418 600[b]	294 830 (70.4%)	9 (6, 13)	6.1 (5.0, 6.6)	7.4 (5.4, 9.8)	5 061 (1.7%)	1.9%
Sex							
Male	3 200 400[b]	765 907 (23.9%)	7 (5, 10)	5.7 (4.2, 6.4)	8.4 (6.1, 11.0)	17 634 (2.3%)	2.8%
Female	3 312 600[b]	831 715 (25.1%)	7 (4, 10)	5.7 (4.2, 6.4)	8.5 (6.1, 11.4)	21 457 (2.6%)	1.4%
Ethnicity							
Black	111 300[b]	17 917 (16.1%)	6 (4, 9)	5.2 (3.4, 6.3)	8.5 (6.1, 11.6)	492 (2.7%)	3.7%
Non-black	6 401 700[b]	1 579 712 (24.7%)	7 (5, 10)	5.7 (4.2, 6.4)	8.4 (6.1, 11.2)	38 599 (2.4%)	2.0%
Diabetes	394 568	364 565 (92.4%)	10 (7, 14)	6.2 (5.1, 6.7)	6.6 (5.0, 8.5)	2 053 (0.6%)	14.3%
Hypertension	1 102 781	959 922 (87.0%)	8 (5, 11)	5.9 (4.7, 6.5)	8.2 (6.0, 10.6)	16 000 (1.7%)	3.9%
CVD	390 506	351 273 (90.0%)	9 (6, 13)	6.1 (5.0, 6.6)	7.4 (5.4, 9.6)	3 362 (1.0%)	7.6%
CKD code	266 358	251 792 (94.5%)	11 (7, 15)	6.2 (5.2, 6.7)	6.3 (4.6, 8.5)	3 495 (1.4%)	20.0%
Confirmed CKD	256 568	247 352 (96.4%)	10 (7, 15)	6.2 (5.2, 6.7)	6.4 (4.6, 8.7)	N/A[c]	N/A[c]
CKD stage[d] (last GFR)							
1 (90+)	456 902	319 127 (69.8%)	6 (4, 9)	5.5 (3.9, 6.4)	8.7 (6.4, 11.5)	7 657 (2.4%)	0.03%
2 (60–90)	1 342 474	937 219 (69.8%)	7 (4, 9)	5.6 (4.1, 6.4)	8.9 (6.6, 11.7)	24 636 (2.6%)	0.1%
3 (30–60)	371 893	318 931 (85.8%)	9 (6, 13)	6.0 (4.7, 6.6)	7.0 (5.0, 9.4)	5 831 (1.8%)	1.0%
4 (15–30)	19 016	18 137 (95.4%)	15 (9, 21)	6.3 (5.3, 6.8)	4.6 (3.2, 6.4)	334 (1.8%)	51.2%
5 (<15)	4 293	3 743 (87.2%)	13 (8, 21)	5.5 (3.3, 6.6)	4.0 (2.7, 6.1)	605 (16.2%)	88.6%

[a]In patients with ≥3 tests.
[b]Population age, sex and ethnicity breakdown are estimated based on aggregate data provided at the practice level.
[c]Loss to follow-up not evaluable in confirmed CKD since group definition requires creatinine measurement in last 2 years.
[d]CKD stage evaluated in all patients with at least one creatinine test result (34% of all adults).

dated prior to 2008 when creatinine data collection began. Of those with a RRT code post-2008, 1583 (60.5%) had at least three GFR tests prior to Read-coded RRT initiation.

Table 2 presents the availability of repeat creatinine tests in all adults and by risk factor. In patients with at least three creatinine tests, the median number of tests was 7, spanning a median of 5.7 years. About 2.4% of these patients had no test performed in the last 3 years of follow-up. Availability of repeat tests and testing frequency was considerably higher and loss to follow-up was lower in high-risk groups, particularly diabetes. Patients lost to follow-up commonly had a RRT code, particularly those with diabetes, coded CKD and eGFR indicating late-stage CKD. For repeat results in diabetes patients, see Supplementary data.

Comparison of slopes of eGFR

Descriptive checks showed better agreement between reported and recalculated eGFR for more recent measures and for later stages of CKD. Slope of recalculated eGFR was estimated in 1.6 million patients and slope of reported eGFR was estimated in 1.1 million patients. The duration of coverage of tests for slope analyses and percentage agreement statistics are provided in the Supplementary data.

Figure 2A shows the distribution of slopes of change in eGFR by risk factor and slope estimation method, among all patients with at least three reported GFR results. The median slope varies by risk factor and estimation method and is consistently higher for analyses using recalculated eGFR than for analysis of reported eGFR. The median slope of recalculated eGFR using all creatinine tests [MDRD (2)] was consistently higher in the population of patients with at least three recalculated eGFR results (a more complete population, not shown) than in those with at least three reported GFR results. (For numerical figures and for population breakdown by age, sex and ethnicity, see Supplementary data.)

Figure 2B shows the distribution of individual differences in slope estimates. Descriptive paired t-tests showed strong statistical significance for a non-zero mean difference in slopes for all slope comparisons by risk factor, $P < 0.001$. Positive differences show systematic overestimation of slope of change in eGFR (underestimation of decline) when using recalculated eGFR compared with reported GFR results, with a median overestimation of ~0.2 mL/min/1.73 m^2/year across subgroups for comparison using the same test results [MDRD (1)—Reported], increasing to ~0.3 mL/min/1.73 m^2/year, for comparison not restricted to the same test results [MDRD (2)—Reported]. Discrepancies are lower in CKD than in other risk groups. Ninety-five per cent of differences between 'MDRD (1)' and 'Reported GFR' slope estimates in CKD patients lie between −0.25 and 1.6 mL/min/1.73 m^2/year, which may not be clinically important.

About 1.1% of patients with at least three reported eGFR test results had coded black ethnicity. Sensitivity analysis excluding patients with coded black ethnicity (i.e. for which laboratory

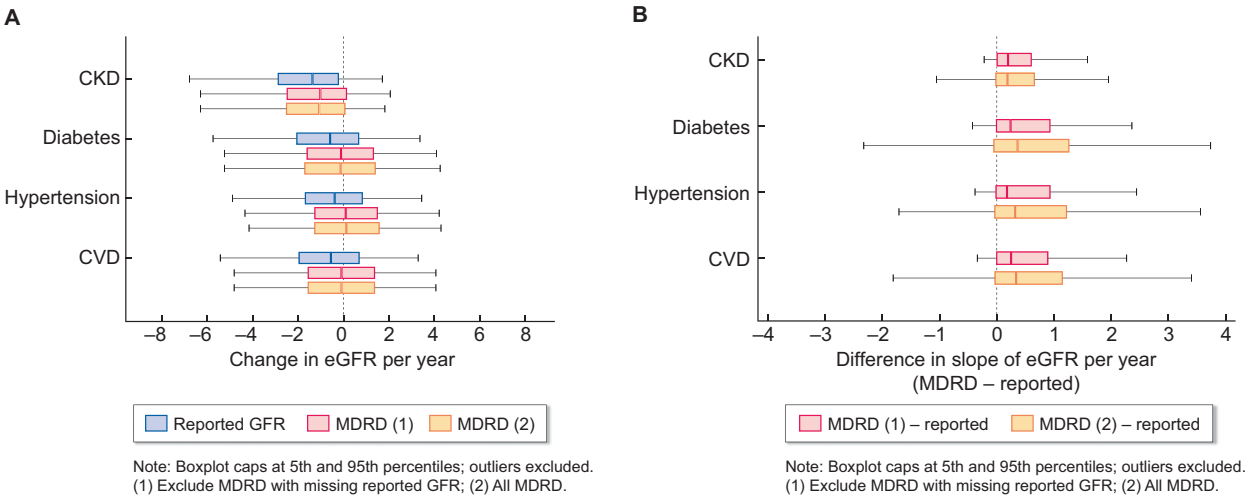


FIGURE 2: Distribution of slopes of change in eGFR (A) and distribution of differences between recalculated and reported GFR slopes (B) in patients with at least three reported eGFR results, by risk factor and method of estimation of slope of eGFR.

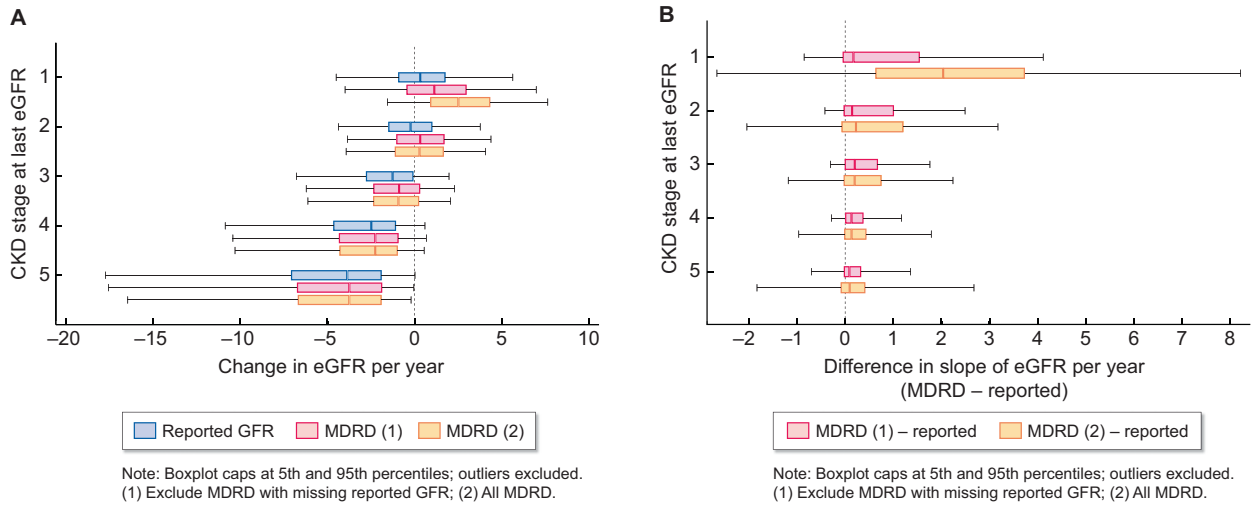


FIGURE 3: Distribution of slope of change in eGFR (A) and distribution of differences between recalculated and reported GFR slopes (B) in patients with at least three reported eGFR results, by CKD stage (1–5) at most recent measure and method of estimation of slope of eGFR.

ethnicity correction is not required) had no effect on observed overestimation. (For boxplots, see Supplementary data.)

Figure 3A shows the distribution of slopes by CKD stage. As expected, slope of decline is steeper in patients reaching later stages of disease. There is greater variation in slopes at later stages of disease, with some patients appearing to decline much more rapidly than others. Figure 3B shows discrepancy in slopes between estimation methods by CKD stage. Discrepancy diminishes considerably as kidney function worsens and slope estimation is highly sensitive to the estimation method for patients with latest eGFR in the normal range.

Comparison of slopes by frequency of tests (see Supplementary data) showed markedly reduced variability for patients with five or more tests, likely due to increased precision of estimation for increasing number of tests, although plausibly driven by patients with worse kidney function having more tests.

DISCUSSION

The aim of this study was to identify whether it may be feasible to study the natural history of kidney disease using EHRs held in UK primary care. The database used was large and representative of the UK population. While testing frequency was low in the general population, high-risk groups were tested regularly, sufficient to study long-term longitudinal changes in renal function. It is possible that we may not capture a representative sample of CKD patients if the sickest patients are managed solely in secondary care throughout creatinine data collection. Informative loss to follow-up from primary care may also be a concern, although rates of loss to follow-up were low, particularly in diabetes.

A major issue that may compromise evaluation of longitudinal changes in renal function using primary care EHRs is lack of creatinine calibration to international standards in historical

results reported by laboratories [11, 12]. It is very challenging to identify when calibration practices may have changed from the EHR but failure to correct historical measures would mean that measures are not comparable within or between patients, and estimates of change in eGFR will be overestimated in many patients (underestimating decline). One solution may be to use laboratory-reported eGFR but lower values are more likely to be reported than those in (or closer to) the normal range, leading to selective inclusion of test results. It may be possible to develop statistical methods capable of identifying the time point(s) in individuals at which creatinine calibration practices have changed and apply appropriate correction factors to uncalibrated results. Some authors have attempted to do this [13]. Another approach would be to restrict analysis to test results reported post-2012, when calibration issues are less common, but this would impact duration of follow-up. Restricting analysis to patients with CKD would also reduce overestimation to levels that may not be of clinical importance.

This study did not consider the possibility of temporary losses in renal function that may occur due to an acute event. Although scheduled annual review tests in primary care are likely to be carried out on a relatively stable population, tests may also be carried out on patients who present to their GP due to ill health and it is not known how many AKI events may be captured in primary care eGFR data. Longer-term drops in renal function following an acute event are also possible and changes over time may be non-linear in some patients [4], which may require a more complex modelling approach. Slope distributions reported in this article may, therefore, not be clinically reliable.

Primary care EHRs in the UK are an excellent source of data on changes in renal function over a long duration of follow-up in patients at the highest risk of CKD and CKD progression. Future studies aiming to study longitudinal changes in renal function should take care to handle data quality issues present in EHRs. In particular, failure to account for creatinine calibration problems may lead to underestimation of decline in renal function over time. The study population should be selected taking into account data availability and reliability of analytical methods. Future studies may also need to account for informative loss to follow-up. We recommend the use of joint modelling of longitudinal changes and the drop-out process [14] with linkage to external databases to help establish reasons for loss to follow-up.

SUPPLEMENTARY DATA

Supplementary data are available at *ckj* online.

ACKNOWLEDGEMENTS

This research was made possible by free access to data from the National Chronic Kidney Disease Audit. Authors F.C., D.N., B.C. and S.H. were previously involved in design, analysis and reporting of the audit. Experience with the database and audit findings led to many of the ideas presented in this article.

FUNDING

This work was supported by the Medical Research Council [grant number MR/N013638/1].

AUTHORS' CONTRIBUTIONS

F.C., D.N. and D.P. developed the ideas for content of this article. F.C. carried out data management and analysis and drafted the article. All authors reviewed several draft versions of the manuscript, suggested updates to analysis and interpretations and approved the final manuscript.

CONFLICT OF INTEREST STATEMENT

D.N. reports grants related to the submitted work from the Health Quality Improvement Partnership, funded by NHS (National Health Service) England and Wales and subcontracted to the London School of Hygiene and Tropical Medicine by BMJ Informatica as well as grants outside of this work from GlaxoSmithKline. B.C. reports grants outside of this work from AstraZeneca.

REFERENCES

1. UK Renal Registry (2019). *UK Renal Registry 21st Annual Report – Data to 31/12/2017, Bristol, UK*. https://www.renalreg.org/publications-reports (17 February 2020, date last accessed)
2. Kerr M. *Chronic Kidney Disease in England: The Human and Financial Cost. Insight Health Economics* 2012. https://www.england.nhs.uk/improvement-hub/wp-content/uploads/sites/44/ 2017/11/Chronic-Kidney-Disease-in-England-The-Human-and-Financial-Cost.pdf (17 February 2020, date last accessed)
3. National Institute for Health and Care Excellence (2014). *Chronic Kidney Disease in Adults: Assessment and Management (NICE Guideline CG182)*. https://www.nice.org.uk/guidance/cg182 (17 February 2020, date last accessed)
4. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Inter Suppl* 2013; 3: 1–150
5. Health and Social Care Information Centre (2005). *National Quality and Outcomes Framework Statistics for England 2004/05*. https://files.digital.nhs.uk/publicationimport/pub01xxx/pub01946/qof-eng-04-05-rep.pdf (17 February 2020, date last accessed)
6. Health and Social Care Information Centre (2005). *National Quality and Outcomes Framework Statistics for England 2004/05: Annex A – Quality Indicators*. https://files.digital.nhs.uk/publicationimport/pub01xxx/pub01946/qof-eng-04-05-defi-indi-anx.pdf (17 February 2020, date last accessed)
7. Usher-Smith J, Young A, Chatfield S *et al*. Chronic kidney disease in primary care. *Br J Cardiol* 2007; 14: 221–228
8. Feakins B, Oke J, McFadden E *et al*. Trends in kidney function testing in UK primary care since the introduction of the quality and outcomes framework: a retrospective cohort study using CPRD. *BMJ Open* 2019; 9: e028062
9. Nitsch D, Caplin B, Hull S, on behalf of the National CKD Audit and Quality Improvement Programme in Primary Care *et al*. *First National CKD Audit Report* 2017. https://www.lshtm.ac.uk/files/ckd_audit_report.pdf (17 February 2020, date last accessed)
10. Mudie K, Cleary F, Caplin B, on behalf of the National CKD Audit and Quality Improvement Programme in Primary Care *et al*. *Second National CKD Audit Report* 2017. https://www.lshtm.ac.uk/media/9951 (17 February 2020, date last accessed)

11. Miller W, Myers G, Eckfeldt J. Reporting calculated GFR from serum creatinine. 2016. https://acutecaretesting.org/-/media/acutecaretesting/files/pdf/reporting-calculated-gfr-from-serum-creatinine.pdf (17 February 2020, date last accessed)
12. Bargnoux A, Kuster N, Cavalier E *et al.* Serum creatinine: advantages and pitfalls. *J Lab Precis Med* 2018; 3: 71–71.
13. Poh N, McGovern A, de Lusignan S. Improving the measurement of longitudinal change in renal function: automated detection of changes in laboratory creatinine assay. *J Innov Health Inform* 2015; 22: 293–301
14. Boucquemont J, Heinze G, Jager K *et al.* Regression methods for investigating risk factors of chronic kidney disease outcomes: the state of the art. *BMC Nephrol* 2014; 15: 45–45

5.4 Key findings and implications for PhD

- Among the **entire adult population** covered by the nation audit data (cross-sectional study), 34% of patients had at least one creatinine test, and **25% had at least 3 creatinine tests**, with a median coverage of 5.7 years. Percentage of patients with at least 3 prior creatinine tests was **92% in those with coded diabetes**, **95% in those with coded CKD** and **96% in those with biochemical evidence confirming CKD** (2 x eGFR results <60, over at least 90 days), with high median numbers of tests in these groups (10, 11 and 10 respectively). There was also sufficient data to identify 3.9% of the adult population as having confirmed CKD at the time of data extraction.

- The impact of differences in creatinine calibration and laboratory reporting practices which were not appropriately identified in the data led to considerable issues in estimation of values of eGFR and resultant eGFR slope estimates, leading to **systematic over-estimation of slopes (underestimation of decline)**, with a median over-estimation of approximately 0.2-0.3 mL/min/1.73m^2/year, which was similar in magnitude across subgroups. Degree of over-estimation was higher in some patients than others, possibly due to timing of test results, where recent test results suffered generally less imprecision in estimation than historical results.

- This study was limited due to the **cross-sectional nature of the data**, only being able to assess patients who were alive at time of data extraction (**survival bias**), and evaluating completeness of test results retrospectively based on risk factors identified at data extraction. However, **strengths** of the study were the **size and representativeness of the database**, covering 14% of the population of England and Wales, leading to important **insights which are likely to represent the general population**.

- **Primary care EHRs** in the UK are an **excellent source of data on changes in renal function** over a long duration of follow-up in patients at the highest risk of CKD and CKD progression.

- Future studies aiming to study longitudinal changes in renal function should take into account **data quality issues** present in EHRs, including creatinine calibration issues and laboratory reporting practices, to avoid biased estimation of slopes of change in

eGFR. These **issues are likely to diminish over time** if new data are studied as results reported beyond 2012 are less likely to be affected by miscalibration. Slope estimation **biases are also less marked** in the study of **patients with established and later stage CKD**.

Chapter 6

Association between practice coding of chronic kidney disease in primary care and subsequent hospitalisations and death: a cohort analysis using national audit data (paper 3)

6.1 Chapter summary

This chapter provide a summary of work done in research paper 3, presents the original published research paper, and lists key findings and implications in the context of the overall thesis.

6.2 Summary of work in the context of aims of PhD

Background

Electronic diagnostic coding for CKD stages 3-5 is recommended as part of pay-for-performance quality metrics in English primary care (QOF) [75]. Despite this, the 2016 National CKD Audit found that only 70% of CKD cases were coded, with high variability in completeness of coding between GP practices [16].

CKD coding behaviour may be influenced by healthcare practitioner beliefs on the importance of CKD coding and patient risk factors for CKD. In addition, CKD can only be detected when repeat tests for kidney function have been carried out, separated by at least 3 months (see *Chapter 1 Section 1.3*), which depends on patient healthcare seeking behaviours and whether GPs order blood tests.

We hypothesised that coding for CKD is likely to improve patient outcomes, due to a knock-on effect of improved healthcare practices in those with coded CKD such as prescription of anti-hypertensives, testing of proteinuria, and prescription of statins. We therefore sought to

investigate whether there is a causal relationship between CKD coding and adverse outcomes in individuals with CKD.

Methods

Designing an appropriate epidemiological study is challenging due to likely confounding by risk factors, healthcare seeking behaviours and GP factors, which may be complex and difficult to measure using available data. Confounders such as diabetes and hypertension rely on diagnostic coding, which may be inaccurately recorded, and some known confounders weren't captured in our database. Residual or unmeasured confounding may lead to biased study results. Furthermore, informative testing may lead to unrepresentative samples and inaccurate measurement of disease severity (a key analysis confounder).

We analysed the association between practice completeness of CKD coding (percentage of confirmed CKD cases which were appropriately coded) and patient-level hospital outcomes. We believed that practice CKD coding performance would be unrelated to practice case-mix, thereby cutting any associations between practice-level study exposure and patient-level confounders. This approach may be likened to a natural experiment, where characteristics of the population are, on average, similar between GP practices (even after accounting for varying CKD coding levels), and the process of exposure assignment (practice CKD coding) to individuals may be considered approximately random. We additionally adjusted for practice-level confounders which may be associated with practice CKD coding, including variables related to practice risk profile, practice testing behaviours and practice disease severity in CKD patients detected for analysis.

Results

Study population characteristics proved to be well-balanced according to quantiles of practice CKD coding performance, except at the extremes of coding performance. There was evidence of a lower risk of hospitalisation outcomes (such as CV events) for higher levels of practice CKD coding in fully adjusted models, which is likely to be clinically important.

Conclusions

Results point towards evidence of a causal relationship between CKD coding and CKD-related adverse events, where we have taken efforts to account for unobserved confounding and adjusted for practice-level confounders. There are risks of unmeasured practice confounders that we may have failed to identify and capture. However, our findings are

compatible with other research highlighting the benefits of particular interventions (which we found to be mildly associated with practice CKD coding) in reducing risks of adverse events, including use of anti-hypertensives and statins [76-78].

6.3 Research paper 3

See next page for original published research paper, and appendix 5 for supplementary materials.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1703701	**Title**	Miss
First Name(s)	Faye		
Surname/Family Name	Cleary		
Thesis Title	Challenges of studying and predicting chronic kidney disease progression and its complications using routinely collected electronic healthcare records		
Primary Supervisor	Dorothea Nitsch		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	BMJ Open		
When was the work published?	September 2022		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	DN, FC, BC, SH, LK and DP-M planned the analysis. FC conducted analysis, interpreted results and reported the work. DN, LK, DP, DW, SH, RS, RF, DA, SD, KG, FL and BC contributed insights in the interpretation of results, provided updates to manuscript text, and read and approved the final manuscript. FC is the guarantor and accepts full responsibility for the finished work and conduct of the study, had access to the data, and controlled the decision to publish.

SECTION E

Student Signature	Faye Cleary
Date	23 May 2024

Supervisor Signature	Dorothea Nitsch
Date	10/06/2024

BMJ Open

Association between practice coding of chronic kidney disease (CKD) in primary care and subsequent hospitalisations and death: a cohort analysis using national audit data

Faye Cleary ,[1] Lois Kim,[2] David Prieto-Merino,[1] David Wheeler,[3] Retha Steenkamp,[4] Richard Fluck,[5] David Adlam,[6] Spiros Denaxas,[7,8] Kathryn Griffith,[9] Fiona Loud,[10] Sally Hull ,[11] Ben Caplin ,[3] Dorothea Nitsch[1]

To cite: Cleary F, Kim L, Prieto-Merino D, *et al*. Association between practice coding of chronic kidney disease (CKD) in primary care and subsequent hospitalisations and death: a cohort analysis using national audit data. *BMJ Open* 2022;**12**:e064513. doi:10.1136/bmjopen-2022-064513

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-064513).

Received 06 May 2022
Accepted 07 September 2022

For numbered affiliations see end of article.

Correspondence to
Faye Cleary;
faye.cleary@lshtm.ac.uk

ABSTRACT

Objective To examine the association between practice percentage coding of chronic kidney disease (CKD) in primary care with risk of subsequent hospitalisations and death.

Design Retrospective cohort study using linked electronic healthcare records.

Setting 637 general practitioner (GP) practices in England.

Participants 167 208 patients with CKD stages 3–5 identified by 2 measures of estimated glomerular filtration rate <60 mL/min/1.73 m^2, separated by at least 90 days, excluding those with coded initiation of renal replacement therapy.

Main outcome measures Hospitalisations with cardiovascular (CV) events, heart failure (HF), acute kidney injury (AKI) and all-cause mortality

Results Participants were followed for (median) 3.8 years for hospital outcomes and 4.3 years for deaths. Rates of hospitalisations with CV events and HF were lower in practices with higher percentage CKD coding. Trends of a small reduction in AKI but no substantial change in rate of deaths were also observed as CKD coding increased. Compared with patients in the median performing practice (74% coded), patients in practices coding 55% of CKD cases had a higher rate of CV hospitalisations (HR 1.061 (95% CI 1.015 to 1.109)) and HF hospitalisations (HR 1.097 (95% CI 1.013 to 1.187)) and patients in practices coding 88% of CKD cases had a reduced rate of CV hospitalisations (HR 0.957 (95% CI 0.920 to 0.996)) and HF hospitalisations (HR 0.918 (95% CI 0.855 to 0.985)). We estimate that 9.0% of CV hospitalisations and 16.0% of HF hospitalisations could be prevented by improving practice CKD coding from 55% to 88%. Prescription of antihypertensives was the most dominant predictor of a reduction in hospitalisation rates for patients with CKD, followed by albuminuria testing and use of statins.

Conclusions Higher levels of CKD coding by GP practices were associated with lower rates of CV and HF events, which may be driven by increased use of antihypertensives and regular albuminuria testing, although residual confounding cannot be ruled out.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ A large database of 167 208 patients with biochemical evidence of chronic kidney disease (CKD) was used for analysis, covering 637 general practitioner practices in England who volunteered to participate in an audit of care and which were representative of the general population in terms of age and sex.

⇒ Risk of confounding due to patient characteristics is reduced by studying the association between practice level (rather than patient level) CKD coding and patient-level outcomes, where practice casemix is not expected to differ with practice coding rates.

⇒ Practice behaviours associated with CKD coding performance that are not believed to occur as a consequence of CKD coding may confound associations but were adjusted for as far as possible.

⇒ Average duration of follow-up between assessment of practice coding performance and end of data collection for outcomes was limited to approximately 4 years; longer term effects of CKD coding may, therefore, not be captured in this study.

INTRODUCTION

Chronic kidney disease (CKD) is a growing public health problem.[1–3] Consequences of CKD include cardiovascular (CV) morbidity, acute kidney injury (AKI) and premature mortality, with increasing risks as disease progresses.[4] The burden of CKD and associated healthcare costs are increasing,[5 6] yet recognition of the disease in routine practice is often poor and varies between healthcare providers.[7 8] This may lead to delayed intervention and worsen prognosis in many patients with CKD.

In the UK, computer systems used by general practitioner (GP) practices allow electronic coding of patient clinical information, enabling consistent and specific recording

and ease of access to coded data.[9] The National Institute for Health and Care Excellence provides recommendations for regular renal function testing and CKD management in primary care,[10] and the Quality and Outcomes Framework provides financial incentives for GPs to maintain a practice register of patients with CKD stages 3–5.[11]

The National Chronic Kidney Disease Audit (NCKDA) conducted in England and Wales in 2015–2016 found that approximately 30% of biochemically confirmed CKD stages 3–5 cases were not given an appropriate CKD code in primary care electronic healthcare records (EHRs).[8] Among patients with biochemical evidence of CKD stages 3–5, those registered with a CKD code had significantly lower rates of CV hospitalisations, AKI and mortality than those without CKD codes.[12] However, analyses only adjusted for age, sex and coded diabetes, hypertension and CV disease (CVD), due to limited data availability. Therefore, the audit report cautioned that a causal association cannot be established for the reported benefits of CKD coding, as results will be affected by confounding by patient health seeking behaviours and unmeasured morbidities. Some have cautioned against overdiagnosis of CKD that may fail to benefit the overall health of the population,[13] and more research is needed to study the benefits of CKD coding.

In attempt to overcome the issue of unmeasured confounding experienced in original analyses of the NCKDA,[12] we set out to examine the association between completeness of CKD coding of the GP practice at which patients were registered and individual adverse outcomes known to be associated with CKD. We hypothesised that practice CKD coding performance would not be associated with individual patient characteristics within practices (casemix), thereby removing some potential confounding, and, if appropriately adjusted for practice behaviours, analysis would provide stronger evidence for a causal effect of coding of CKD on outcomes. Additional aims were to explore the role of practice behaviours that may occur as a result of CKD coding in reducing risk of adverse outcomes. Evidence that higher levels of practice CKD coding improved patient outcomes would have important ramifications for GPs and may influence policy to improve recognition of CKD in primary care.

METHODS

Study design

We carried out a retrospective cohort study using routinely collected EHRs.

Data sources

The NCKDA database holds selected data from the EHRs of 695 GP practices in England and was used to identify a cohort of patients with CKD for analysis and to define exposure variables. Data extraction ranged from March 2015 to July 2016 as practices were gradually recruited into the audit, with the majority of practices recruited by July 2015. In brief, the audit is a snapshot of care at the time point of audit data extraction. Details of the audit and data collection strategy are specified elsewhere.[8] The NCKDA database was linked to Hospital Episode Statistics (HES) holding information on all hospital admissions in England, and Office for National Statistics mortality data, to followup patients for adverse outcomes. Linkage was carried out by National Health Service (NHS) Digital using NHS number, and hospital record information with pseudoanonymised linkage IDs were provided for analysis.

Study population

The analysis cohort included all adult patients in the NCKDA database extracted from eligible GP practices in England and with biochemical evidence of CKD (hereafter referred to simply as 'CKD' and/or 'confirmed CKD'), defined as at least two records of estimated glomerular filtration rate (eGFR) <60 mL/min/m^2 separated by at least 90 days, with the most recent measure recorded within the last 2 years prior to data extraction and excluding any patients with coded initiation of renal replacement therapy.

Primary exposure

At the time of data extractions (2015–2016), Read codes were used to electronically record patient findings in GP computer systems.[14] Variables defined based on eGFR use the isotopedilution mass spectrometry calibrated Modification of Diet in Renal Disease study equation, the standard GFR-estimating equation in use during the period of data collection.

Practice CKD coding performance was characterised as the percentage of patients with CKD in a practice with a CKD stages 3–5 Read code, hereafter referred to as practice CKD coding (performance) or percent coded CKD. Practice CKD coding performance was defined at practice data extraction, which marked the index date for commencement of follow-up for outcomes.

Practices with fewer than 50 total CKD cases were excluded from analysis due to anticipated excess noise in measurement of the primary exposure.

Outcomes

Four outcomes of interest were studied: (1) hospitalisation for CV events, (2) hospitalisation for heart failure (HF), (3) hospitalisation with AKI and (4) all-cause mortality, defined in online supplemental table 1. Follow-up began at the time of practice data extraction and was capped at 1 March 2019 for hospital outcomes and 1 September 2019 for deaths.

Practice features

Features of a practice that may confound the association between practice CKD coding and patient adverse outcomes were defined and categorised as: those reflecting overall practice risk profile; practice testing behaviours for CKD; and characteristics of the identified practice CKD population. Practice features that may improve CKD outcomes, some of which may lie on the

causal pathway between practice CKD coding and patient adverse outcomes, were also defined. Practice percentage variables were defined by summing the number of patients meeting relevant risk factor criteria in each practice and dividing by relevant practice denominators. Practice list size data were used to determine size of the adult population in each practice.

Patient-level risk factors identified in NCKDA data (and used to calculate practice percentages) were defined based on presence of any relevant Read code in the EHR prior to data extraction and included:

- Diabetes—any diabetes code not superseded by a diabetes resolved code
- Hypertension—any hypertension code
- CVD—any CVD code
- CKD stages 3b-5—confirmed CKD and latest eGFR <45
- Statin use—any statin prescription or contraindication code
- Antihypertensive use—any prescription or contraindication code for angiotensin-converting enzyme inhibitors (ACEi) or angiotensin-II receptor blockers (ARBs)
- Blood pressure (BP) targets met—last BP within target range within last year before data extraction: systolic blood pressure (SBP) <130 mm Hg and diastolic blood pressure (DBP) <80 mm Hg in those with diabetes or last urine albumin to creatinine ratio (ACR) ≥70 mg/mmol or last protein to creatinine ratio (PCR) ≥100 mg/mmol; SBP <140 mm Hg and DBP <90 mm Hg in all other patients
- Influenza vaccination—any influenza vaccination code in the last year
- Pneumococcal vaccination—any pneumococcal vaccination code in the last 5 years

Additional patient-level risk factors not available in NCKDA data were defined using HES data and included:

- Recent chronic obstructive pulmonary disease (COPD) admission—admission in last 3 years prior to (NCKDA) practice data extraction with a COPD ICD-10 code (J44) as primary diagnosis
- Recent cancer admission—admission in last 3 years prior to (NCKDA) practice data extraction with a cancer ICD-10 code (C00-C97, excluding non-melanoma skin cancers C44) as primary diagnosis

Practice characteristics reflecting overall practice risk profile

GPs' awareness on how to identify CKD may depend on the overall burden of conditions associated with CKD in their practice. Practice prevalence of diabetes, hypertension and CVD were determined by summing the number of adult patients meeting patient-level comorbidity definitions, with adult population size as denominator. Practice list size data stratified by age and sex were used to determine mean practice age and percent of adults that were male. Practice deprivation was summarised using the median rank of the Index of Multiple Deprivation (IMD) score among all patients extracted from the GP practice (which was limited to patients with CKD risk factors, creatinine assessments or renal codes).

Practice testing behaviours

Testing behaviours which may impact the patient vintage (ie, the underlying duration of CKD at detection by the GP) and the types of patients with CKD selected for analysis were also defined within practices, including percentage of diabetes patients with a GFR test in the last year, percentage of patients with CKD with a GFR test in the last year and percentage of the practice adult population with confirmed CKD.

Practice characteristics of the detected CKD population

Underlying practice morbidity and testing behaviours may impact the types of patients with CKD detected and therefore included in analysis. Percent CKD stages 3b-5, percent recent COPD admission and percent recent cancer admission were defined using patient-level risk factor definitions, with number of total CKD cases as denominator.

Practice behaviours that may improve CKD outcomes

Practice behaviours expected to be related to CKD coding, some of which may be on the causal pathway from CKD to improved outcomes were defined as: percent usage of ACEi/ARBs in hypertension, percent usage of statins in diabetes, percent usage of statins in CVD, percent meeting BP target in last year in CKD, percent ACR/PCR test in last year in CKD, percent influenza vaccination in last year in CKD and percent pneumococcus vaccination in past 5 years in CKD stages 4–5. Practice behaviour variables were dichotomised at the median value.

Statistical methods

Baseline characteristics of the study population were summarised by sextile of practice percent coded CKD to determine balance in patient characteristics according to primary exposure. Practice characteristics were also summarised by sextile of practice CKD coding to identify any associations with other practice characteristics.

Main Cox regression analyses

Cox proportional hazards regression (time to first event) was used to evaluate the association between practice CKD coding and each of the four patient outcomes. Hospitalisation outcomes were censored for death. A 5 knot spline was used for the primary exposure, providing flexibility to demonstrate the nature of association between practice coding and outcomes across the spectrum of practice CKD coding performance, without overfitting. The following adjustments for practice characteristics (included as continuous covariates) were carried out sequentially:

Model 1: adjusted for practice characteristics reflecting overall practice risk profile (primary analysis, planned a priori)

Model 2: adjusted for practice characteristics reflecting overall practice risk profile, practice testing behaviours

and practice characteristics of the detected CKD population (secondary analysis, data driven).

Adjusted HR curves for outcomes with 95% CIs were plotted across the spectrum of practice CKD coding, compared with average (median) practice CKD coding. Attributable fractions for the number of (first) events preventable by the median follow-up time among patients with CKD in practices at lower coding levels (17th percentile, bottom of sextile 2) if such practices instead coded at higher coding levels (83rd percentile, top of sextile 5) were estimated under assumption of causality following adjusted Cox regression (model 2), detailed in online supplemental information 1.

Additional analyses with a single linear continuous covariate for percent coded CKD were carried out after visual inspection of an approximately linear relationship for some model 2HR curves, allowing a more convenient clinical interpretation. These models were restricted to sextiles 2–5 of practice coding only (representing the 67% of most averagely performing practices) where linear trends were most apparent. Descriptive likelihood ratio tests were used to assess improvement in model fit using spline terms vs a single linear covariate.

Subgroup analyses

Subgroup analyses were carried out by diabetes status and by CKD severity (stage 3a, stage 3b–5). Practice percent coded CKD was recalculated within each subgroup for analysis, since coding behaviour differed substantially between subgroups.

Analyses of practice behaviours that may improve CKD outcomes

Further Cox regression analyses aimed to identify practice behaviours associated with improvements in all four patient outcomes, with adjustment for all (model 2) confounders as well as practice behaviours that may improve CKD outcomes. Practice CKD coding covariates were excluded from analysis to identify practice factors most predictive of outcomes, regardless of CKD coding performance and not conditional on CKD coding.

Patient and public involvement

Kidney Care UK supported the research questions, grant applications and related record linkage applications of the NCKDA. After NCKDA discontinuation, Kidney Care UK helped with ethics and section 251 permissions to maintain database access for research purposes. A patient representative (Fiona Loud) was involved in the NCKDA from inception, is a co-author and critically reviewed content of this paper.

RESULTS

Data completeness

Of 695 practices in England captured in the NCKDA database, 637 practices (92%) met criteria for analysis (at least 50 CKD cases), covering 99% of all patients with CKD from the original database (n=167 208) (figure 1). CKD coding rates did not differ after excluding ineligible practices, overall or by subgroup, but sample sizes were smaller in some subgroups after excluding practices with fewer than 50 CKD cases (online supplemental table 2).

Patient characteristics

Study population characteristics were generally well balanced between sextiles of practice CKD coding (table 1). There were trends of slightly higher rates of diabetes, hypertension and CVD coding and lower eGFR in patients in the highest coding practices (sextile 6) and slightly lower rates of comorbidities and higher eGFR in the lowest coding practices (sextile 1), indicating potential differences in either true underlying morbidity or risk factor coding in patients in practices performing at the extremes. Median month of data extraction and resulting follow-up duration were well balanced between sextiles, suggesting good balance in seasonal coverage.

Practice characteristics

Median practice percent coded CKD was 73.9% in the overall CKD population. It was higher in CKD stages 3b-5 (87.9%) than CKD stage 3a (64.8%), and higher in

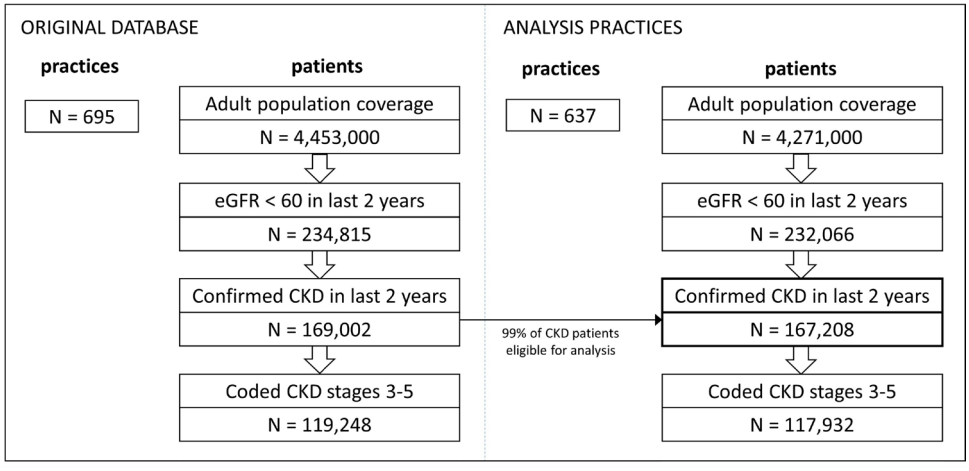


Figure 1 Flow chart of selection of study population (confirmed CKD in last 2 years). CKD, chronic kidney disease.

Table 1 Study population characteristics by practice coding sextile

Practice coding sextile (S)	S1	S2	S3	S4	S5	S6
Practice percent coded CKD	<54.8%	54.8%–65.5%	65.5%–73.9%	73.9%–80.7%	80.7%–87.5%	≥87.5%
No of patients with CKD	N=29 389	N=30 360	N=28 248	N=28 824	N=26 731	N=23 656
Coded CKD	13 105 (44.6%)	18 505 (61.0%)	19 744 (69.9%)	22 326 (77.5%)	22 524 (84.3%)	21 728 (91.8%)
Age, mean (SD)	76.7 (11.2)	77.1 (11.0)	77.0 (10.9)	77.2 (10.8)	77.1 (10.7)	77.5 (10.7)
Male	11 686 (39.8%)	12 240 (40.3%)	11 464 (40.6%)	11 709 (40.6%)	10 664 (39.9%)	9493 (40.1%)
Diabetes	7013 (23.9%)	7657 (25.2%)	6887 (24.4%)	7288 (25.3%)	6544 (24.5%)	6148 (26.0%)
Hypertension	19 475 (66.3%)	20 749 (68.3%)	19 300 (68.3%)	19 798 (68.7%)	18 860 (70.6%)	17 207 (72.7%)
CVD	9324 (31.7%)	9945 (32.8%)	9087 (32.2%)	9642 (33.5%)	8819 (33.0%)	8086 (34.2%)
Biochemical CKD stage						
3a	18 927 (64.4%)	19 201 (63.2%)	17 938 (63.5%)	18 151 (63.0%)	16 858 (63.1%)	14 774 (62.5%)
3b	8384 (28.5%)	8938 (29.4%)	8254 (29.2%)	8576 (29.8%)	7847 (29.4%)	7089 (30.0%)
4	1838 (6.3%)	1937 (6.4%)	1812 (6.4%)	1831 (6.4%)	1796 (6.7%)	1572 (6.6%)
5	240 (0.8%)	284 (0.9%)	244 (0.9%)	266 (0.9%)	230 (0.9%)	221 (0.9%)
Time since last GFR measure (months), median (IQR)	5.0 (2.0–9.4)	4.5 (1.8–8.7)	4.6 (1.9–8.6)	4.4 (1.8–8.5)	4.3 (1.8–8.3)	4.5 (1.9–8.3)
Prior COPD admission in last 3 years	346 (1.2%)	352 (1.2%)	365 (1.3%)	366 (1.3%)	323 (1.2%)	285 (1.2%)
Prior cancer admission in last 3 years	1278 (4.3%)	1332 (4.4%)	1266 (4.5%)	1212 (4.2%)	1122 (4.2%)	1044 (4.4%)
Follow-up duration (months) for HES outcomes (months), median (IQR)	47 (36–47)	47 (36–47)	45 (36–47)	46 (33–47)	45 (35–47)	45 (33–47)

CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; GFR, glomerular filtration rate; HES, Hospital Episode Statistics.

Table 2 Practice characteristics by practice coding sextile

Practice coding sextile	S1	S2	S3	S4	S5	S6
Practice percent coded CKD	<54.8%	54.8%–65.5%	65.5%–73.9%	73.9%–80.7%	80.7%–87.5%	≥87.5%
No of practices	n=106	n=106	n=106	n=107	n=106	n=106
Practice percent coded CKD, mean (SD)	44.2% (9.8%)	61.2% (3.5%)	70.0% (2.4%)	77.5% (2.0%)	84.1% (2.2%)	91.7% (3.1%)
Practice characteristics reflecting practice risk profile						
Mean adult age, mean (SD)[M1,M2,CP]	48.9 (4.1)	49.5 (4.7)	49.1 (3.2)	49.2 (3.9)	49.5 (3.8)	49.4 (3.4)
Percent male, mean (SD) [M1,M2,CP]	49.3% (1.9%)	48.6% (2.6%)	49.3% (1.9%)	49.1% (2.0%)	49.3% (2.0%)	49.1% (2.5%)
Practice median rank of IMD, median (IQR) [M1,M2,CP*]	18236 (9,574, 23,480)	17887 (11,835, 22,921)	17363 (9,987, 24,460)	17114 (10,392, 21,711)	15793 (10,946, 21,345)	16555 (12,742, 22,962)
Diabetes prevalence, mean (SD)† [M1,M2,CP]	5.6% (1.3%)	5.9% (1.4%)	6.0% (1.5%)	6.0% (1.5%)	6.0% (1.3%)	6.3% (1.5%)
Hypertension prevalence, mean (SD)† [M1,M2,CP]	16.1% (4.1%)	16.8% (3.9%)	16.6% (3.7%)	17.1% (4.1%)	17.6% (3.7%)	18.7% (4.2%)
CVD prevalence, mean (SD)† [M1,M2,CP]	5.5% (1.8%)	5.9% (2.4%)	5.8% (1.6%)	6.0% (1.8%)	6.2% (1.6%)	6.4% (1.8%)
Practice testing behaviours						
Percent GFR test in last year in diabetes, mean (SD) [M2,CP]	86.9% (12.8%)	89.8% (5.3%)	89.7% (4.7%)	89.7% (5.2%)	89.7% (10.4%)	90.3% (5.0%)
Median months since last GFR test in diabetes, median (IQR)	4.6 (3.7, 5.4)	4.2 (3.5, 5.0)	4.2 (3.7, 5.1)	4.4 (3.7, 5.1)	3.8 (3.3, 4.7)	4.4 (3.8, 5.1)
Percent GFR test in last year in CKD, mean (SD) [M2,CP]	83.3% (12.7%)	86.6% (4.9%)	87.0% (4.7%)	87.2% (5.6%)	86.6% (10.6%)	88.3% (5.5%)
Median months since last GFR test in CKD, median (IQR)	4.8 (4.1, 5.7)	4.6 (4.1, 5.2)	4.7 (4.0, 5.2)	4.5 (3.8, 5.3)	4.2 (3.6, 5.0)	4.4 (3.9, 5.2)
Biochemical CKD prevalence, mean (SD) [M2,CP]	4.2% (1.9%)	4.0% (1.5%)	3.9% (1.4%)	3.8% (1.5%)	4.0% (1.4%)	4.0% (1.5%)
Practice characteristics of the detected CKD population						
Percent CKD stages 3b-5, mean (SD) [M2,CP]	37.0% (5.4%)	37.5% (4.9%)	37.5% (4.8%)	38.3% (5.2%)	38.0% (4.3%)	38.0% (5.0%)
Percent recent COPD admission, mean (SD) [M2,CP]	1.2% (1.0%)	1.2% (0.7%)	1.4% (0.9%)	1.3% (1.0%)	1.3% (1.1%)	1.2% (1.0%)
Percent recent cancer admission, mean (SD) [M2,CP]	4.3% (1.7%)	4.4% (1.9%)	4.4% (1.8%)	4.2% (1.8%)	4.1% (1.8%)	4.4% (2.0%)
Practice behaviours that may improve CKD outcomes						
Percent usage of ACE/ARBs in hypertension, mean (SD)[CP]	76.0% (6.2%)	75.5% (6.1%)	77.1% (5.3%)	75.9% (5.9%)	76.5% (4.6%)	78.2% (5.5%)
Percent usage of statins in diabetes, mean (SD)[CP]	82.2% (5.8%)	83.2% (5.8%)	82.9% (5.5%)	84.2% (5.8%)	84.8% (5.3%)	85.0% (5.5%)
Percent usage of statins in CVD, mean (SD)[CP]	91.4% (3.4%)	92.2% (3.5%)	92.8% (2.9%)	93.0% (2.6%)	93.1% (2.6%)	93.0% (3.5%)
Percent meeting blood pressure target in last year in CKD, mean (SD)[CP]	56.8% (7.4%)	57.0% (6.2%)	57.9% (7.0%)	56.8% (7.4%)	58.3% (7.4%)	60.2% (6.7%)
Percent ACR/PCR test in last year in CKD, mean (SD)[CP]	40.1% (13.3%)	49.1% (13.4%)	55.9% (11.9%)	61.3% (13.3%)	62.5% (17.0%)	68.9% (15.7%)
Percent influenza vaccination in last year in CKD, mean (SD)[CP]	73.3% (15.4%)	77.0% (7.4%)	77.6% (5.7%)	78.8% (5.6%)	79.3% (9.5%)	80.5% (6.3%)
Percent pneumococcus vaccination in past 5 years in CKD stages 4–5, mean (SD)[CP]	15.3% (14.9%)	18.2% (16.8%)	15.2% (14.5%)	19.1% (17.0%)	17.1% (15.7%)	18.2% (18.6%)

M1 indicates variables included in statistical analysis model 1 (primary analysis).
M2 indicates variables included in statistical analysis model 2 (secondary analysis).
CP indicates variables included in statistical analysis of practice behaviour variables.
*IMD is ranked for all lower super output areas in England (1=worst, 32 844=best); summary statistics are based on IMD rankings of patients extracted from practices only and should be interpreted carefully.
†Prevalence statistics use adult population as denominator.
ACEi, angiotensin-converting enzyme inhibitors; ACR, albumin to creatinine ratio; ARB, angiotensin-II receptor blockers; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; GFR, glomerular filtration rate; IMD, Index of Multiple Deprivation; PCR, protein to creatinine ratio.

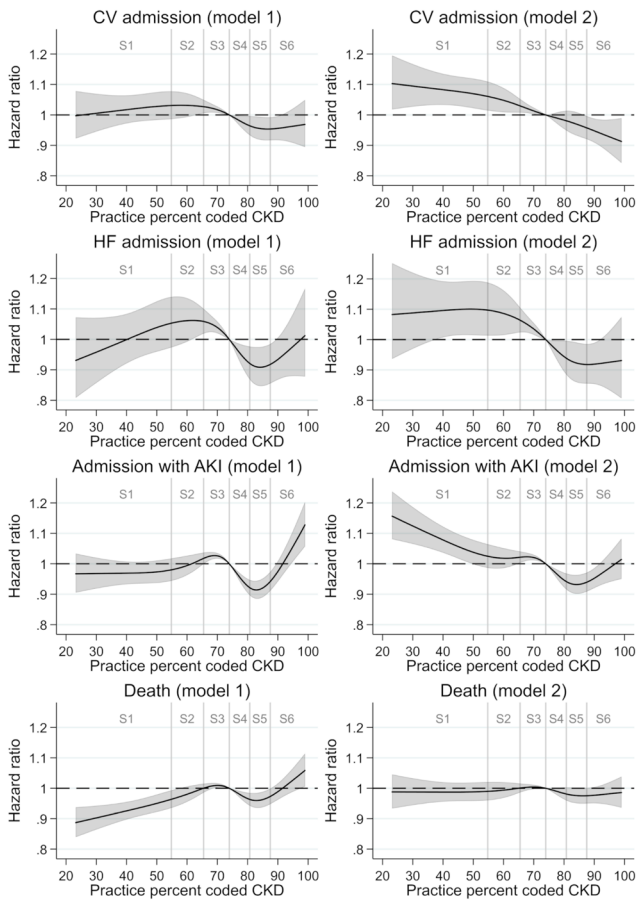


Figure 2 HR curves for all outcomes according to practice percent coded CKD. Analysis includes all patients with confirmed CKD. Primary exposure is (continuous) practice percent coded CKD, with median practice coding (74% coded) as the reference group. Model 1 adjusts for practice risk profile characteristics: mean age, percent male, median rank of index of multiple deprivation, diabetes prevalence, hypertension prevalence, CVD prevalence. Model 2 adjusts for model 1 variables and additionally for practice characteristics of CKD population and testing behaviours: percent of CKD cases at stages 3b–5, percent of patients with CKD admitted for COPD in last 3 years, percent of patients with CKD admitted for cancer in last 3 years, percent GFR test in last year in diabetes, percent GFR test in last year in CKD, percent of adult population with CKD. Labels S1–S6 descriptively indicate sextiles of practice percent coded CKD, with each sextile representing one sixth of all practices. GFR, glomerular filtration rate; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease.

those with diabetes (78.6%) than those without diabetes (71.4%). Coding performance was more variable in early stage CKD and non-diabetic CKD (online supplemental figure 1).

Practice characteristics were generally well balanced in coding sextiles 2–5 (table 2). At the extremes (sextiles 1 and 6), higher coding practices had on average higher prevalence of coded comorbidities, were more deprived, and performed more regular and complete GFR testing in those at risk. Practice behaviours that may improve CKD outcomes showed trends of improved performance in higher coding practices, across the spectrum of CKD coding. In particular, ACR/PCR testing was substantially higher in higher coding practices, and influenza vaccination rates were also markedly higher.

Outcomes

Of 167 208 patients with CKD identified from the NCKDA database, after national linkage we found that 563 deaths had occurred but had not been reported on primary care systems at date of data extraction, leaving 166 645 eligible for outcomes analysis. Median follow-up duration was 3.8 years (range 1 day to 3.9 years) for HES outcomes and 4.3 years (range 1 day to 4.4 years) for mortality, with no meaningful differences in follow-up between sextiles. Crude event rates by sextile are shown in online supplemental table 3,4 and online supplemental figure 2,3.

Main adjusted Cox regression analyses

Figure 2 demonstrates how individual patient risks of the four studied outcomes differ according to practice CKD coding, compared with a patient in an averagely performing practice. In model 1 analyses, inverted S-shaped HR curves suggest that confounding at the extremes of practice coding may distort the association between practice CKD coding and adverse outcomes. After further adjustments (model 2), curves become flatter (approaching linearity). Wide CIs at lower levels of CKD coding reflect poor precision in HR estimates due to sparse data. (Crude analyses and additional sequential adjustments in the CKD population and subgroups (detailed in online supplemental information 2) are shown in online supplemental figures 4–23).

There are strong trends, particularly among sextiles 2–5 (55%–88% coded), of reduced rates of outcomes with improved practice CKD coding in fully adjusted analyses (model 2). Compared with patients in the averagely performing practice (74% coded), patients in practices coding only 55% of CKD cases had a significantly higher rate of CV hospitalisations (HR 1.061 (95% CI 1.015 to 1.109)) and HF hospitalisations (HR 1.097 (95% CI 1.013 to 1.187)). Patients in practices coding 88% of CKD cases had a significantly reduced rate of CV hospitalisations (HR 0.957 (95% CI 0.920 to 0.996)) and HF hospitalisations (HR 0.918 (95% CI 0.855 to 0.985)), compared with the averagely performing practice. The percentage of preventable events over a period of 3.8 years (median follow-up time) for an improvement in practice coding from 55% to 88% of CKD cases (attributable fraction) was 9.0% for first CV hospitalisations and 16.0% for first HF hospitalisations, under assumption of causality after Cox modelling (model 2). Trends of a small reduction in AKI but no substantial change in rate of deaths were also observed as CKD coding increased. (Additional results of analysis of a single linear practice coding term (where appropriate) are shown in online supplemental table 5 and online supplemental figure 24).

Table 3 Adjusted HRs for the association between practice behaviour variables and CV events, sorted by point estimate

Practice behaviour	HR (95% CI)
Percent usage of ACEi/ARBs in hypertension (>76.6%)	0.956 (0.929 to 0.983)*
Percent ACR/PCR test in last year in CKD (>58.7%)	0.968 (0.939 to 0.998)*
Percent usage of statins in CVD (>93.0%)	0.972 (0.942 to 1.003)
Percent pneumococcus vaccination in past 5 years in CKD stages 4–5 (>12.5%)	0.982 (0.955 to 1.010)
Percent meeting blood pressure target in last year in CKD (>57.8%)	0.992 (0.963 to 1.028)
Percent usage of statins in diabetes (>84.1%)	0.995 (0.965 to 1.026)
Percent influenza vaccination in last year in CKD (>78.8%)	0.998 (0.968 to 1.028)

Analysis adjusted for practice characteristics: mean age, percent male, median rank of IMD, diabetes prevalence, hypertension prevalence, CVD prevalence, percent of CKD cases at stages 3b-5, percent of patients with CKD admitted for COPD in last 3 years, percent of patients with CKD admitted for cancer in last 3 years, percent GFR test in last year in diabetes, percent GFR test in last year in CKD, percent of adult population with CKD.
*95% confidence interval excludes 1
ACEi, ACE inhibitors; ACR, albumin to creatinine ratio; ARB, angiotensin-II receptor blocker; CKD, chronic kidney disease; CV, cardiovascular; CVD, cardiovascular disease; PCR, protein to creatinine ratio.

Subgroup analyses showed a steeper reduction in CV hospitalisations and HF hospitalisations with increasing practice coding among CKD stages 3b–5 than in CKD stage 3 both in spline regression analyses and those assuming linear effects of practice CKD coding (online supplemental figures 5,6,10,11,24).

Practice behaviours analyses

Analysis of the association between practice behaviours and CV hospitalisations adjusted for all confounders showed a significant reduction in rate of CV hospitalisations for patients with CKD belonging to practices with greater than average usage (median 76.6%) of ACEi/ARBs in hypertension compared with practices with lower than average usage (HR 0.956 (95% CI 0.929 to 0.983)) (table 3). Practice ACR/PCR testing in CKD was also associated with a reduction in rates of CV hospitalisations in CKD (HR 0.968 (95% CI 0.939 to 0.998)). Results for analyses of AKI, HF and deaths are available in online supplemental tables 6–8. In brief, usage of ACEi/ARBs was the most consistently dominant predictor, being strongly associated with a reduction in events across all outcomes. This was followed by ACR/PCR testing, which was associated with a reduction in all outcomes except deaths, and usage of statins in CVD, which was associated with a reduction in rate of CV and AKI events.

DISCUSSION

Higher levels of practice CKD coding were associated with lower rates of hospitalisation for CV and HF events among patients with confirmed CKD, after adjusting for practice characteristics. Reductions in hospitalisation rates were strongest in CKD stages 3b-5, although greatest opportunities for improvement are in CKD stage 3a where practice variation in coding was much wider. There was no difference in death rates according to practice CKD coding (although the relationship was less clear among CKD stages 3b-5). Findings were limited by duration of follow-up with longer-term benefits of CKD coding not yet apparent. Practice behaviours associated with CKD coding including usage of ACEi/ARB therapy and ACR/PCR testing were independent predictors of reduction in hospitalisation rates.

There are very limited studies looking at the impact of recognition and diagnostic coding of CKD in primary care as many health systems use coded disease as opposed to laboratory records to identify patients. It is possible that some patients with 2 eGFR measures <60 mL/min/1.73 m^2 more than 3 months apart are not recognised by GPs as having CKD because of concerns of overdiagnosis, for example, in elderly patients without hypertension or in patients recovering from AKI, despite these patients meeting the accepted definition of CKD. Furthermore, numerous studies have demonstrated disparities in CKD coding efforts with younger patients, those from deprived backgrounds, and ethnic minorities being less commonly coded than their counterparts,[7 15] leading to concerns around equity of care. Recent studies have shown an association between CKD coding and interventions known to reduce CV risk such as prescription of statins and antihypertensive agents,[15 16] and CKD coding may play a role in triggering further long-term treatment efforts with potential to reduce patient risks. Our study identified a reduced burden of CV and HF hospitalisations for practices coding more CKD, and a reduced burden of hospitalisations for practices providing more interventions (associated with CKD coding) that are likely to improve CKD outcomes.

A key strength of this study is the large sample size, including data from 167 208 patients with CKD. Data were extracted from GP practices in England with a similar age–sex distribution to the whole population, so findings are likely to be generalisable to the wider population. By studying the association between practice-level CKD coding and patient-level outcomes, we were able to eliminate a lot of confounding due to individual patient characteristics that would be present in a conventional study design using patient-level exposure with unmeasured confounders. This was demonstrated by the balanced risk profile in patient characteristics observed across sextiles (which is likely to extend also to unmeasured characteristics). This is a major benefit over original analyses of the NCKDA[12] (online supplemental figure 25), which

showed a very strong association between individual patient CKD coding and risk of outcomes (CV events, AKI, death) but with a high risk of confounding due to coding efforts being associated with perceived patient risk, and some potentially important risk factors missing from the database.

A potential weakness is that included practices had volunteered to participate in an audit of care. CKD coding may have been higher in recruited practices than the general population which may have impacted on estimated strengths of associations; benefits of coding in the wider practice population may be larger than estimated. Risk factor evaluation mostly relied on comorbidity coding and it is not clear whether small differences in risk factor prevalence reflected true morbidity or GP behaviour. Assessment of eligibility of patients for analysis also relied on availability of repeat creatinine tests over time, which may depend on patient risk, and earlier stage CKD cases or more severe cases managed solely in secondary care may be disproportionately missing. Nevertheless, this identified CKD population may stand to benefit most imminently from improvements in primary care, assuming GPs target further coding efforts to patients already identified as at risk and with creatinine test results compatible with CKD. While there was a small signal of more frequent creatinine testing with increasing practice CKD coding, this was only at the extremes, and distribution of CKD severity appeared generally very well balanced across practice coding sextiles. Practice characteristics were analysed differently depending on whether they were likely to confound analyses or lie on the causal pathway, however, we could not verify if our assumptions were reasonable, and misspecification could affect reliability of conclusions. For example, practice management of hypertension with ACEi/ARB therapy may plausibly confound analyses (if hypertension management and CKD coding share a common cause, such as practice funding or clinical expertise) or lie on the causal pathway (if management of hypertension occurs as a consequence of CKD coding). Our findings for AKI are likely affected by outcome misclassification as hospital codes were used to detect AKI events, which may have led to underestimation of the number of events and lack of power to detect an association. We did not have enough dialysis events to allow evaluation of the impact of practice coding on outcomes. These data precede the use of SLGT2-inhibitor drug treatment in HF and albuminuric kidney disease in UK primary care.

Conclusions

Rates of CV and HF events were lower for patients belonging to practices coding more CKD, supporting the argument that CKD coding in primary care may contribute to improvement in patient outcomes. While the presence of unmeasured confounding cannot be ruled out, this is in agreement with other studies conducted in this setting.[15–17] High-quality evidence supporting our findings is available from clinical trials and systematic reviews which underline the benefits of use of interventions in early-stage CKD, including ACEi/ARB therapy to control hypertension and statin therapy to reduce CV risk.[18–20] This study suggests that reductions in key adverse events for patients with CKD could be made by improvements to GP practice identification and coding of CKD as these are associated with subsequent care efforts that are known to prevent poor outcomes.

Author affiliations
[1]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
[2]Cardiovascular Epidemiology Unit, University of Cambridge, Cambridge, UK
[3]Department of Renal Medicine, University College London, London, UK
[4]UK Renal Registry, UK Kidney Association, Bristol, UK
[5]Department of Renal Medicine, Royal Derby Hospital, Derby, UK
[6]Department of Cardiovascular Sciences and NIHR Leicester Biomedical Research Centre, University of Leicester, Leicester, UK
[7]Institute of Health Informatics, University College London, London, UK
[8]British Heart Foundation Data Science Centre, London, UK
[9]No affiliation, retired, York, UK
[10]Director of Policy, Kidney Care UK, Alton, UK
[11]Wolfson Institute of Population Health, Queen Mary University of London, London, UK

Acknowledgements We thank all the patients who have provided their data to the audit analyses, and patient organisations, especially KidneyCare UK and the National Kidney Foundation, and those involved in the set up and conduct of the National Chronic Kidney Disease Audit, which provided a large database of patients with CKD, without which this research would not have been possible. This includes (apart from the listed authors): Matthew Harker, Yvonne Silove, Tasneem Hoosain, Nick Wilson, Ronnie Moodley, Chris Gush, Liam Smeeth, Ron Cullen, Fergus Caskey, Andy Syme, Richard Gunn, Paul Wright, Hugh Gallagher, Sion Edwards, Nick Palmer, Anita Sharma, Kate Cheema, Kathleen Mudie.

Contributors DN, FC, BC, SH, LK and DP-M planned the analysis. FC conducted analysis, interpreted results and reported the work. DN, LK, DP-M, DW, SH, RS, RF, DA, SD, KG, FL and BC contributed insights in the interpretation of results, provided updates to manuscript text, and read and approved the final manuscript. FC is the guarantor and accepts full responsibility for the finished work and conduct of the study, had access to the data, and controlled the decision to publish.

Funding This research was funded by the UK MRC, grant reference MR/N013638/1. The funding body had no role in the design of the study, collection, analysis, or interpretation of data and no role in writing of the manuscript.

Competing interests David C Wheeler has an ongoing consultancy contract with AstraZeneca and has received honoraria, consultancy fees or speaker fees from Amgen, Astellas, Bayer, Boehringer Ingelheim, GlaxoSmithKline, Gilead, Janssen, Napp/Mundipharma, Merck Sharp and Dohme, Tricida, Vifor and Zydus. David Adlam has received research funding from Abbott vascular to support a clinical research fellow; he has also received funding from AstraZeneca inc. for unrelated research and has undertaken consultancy for General Electric inc. to support research funds. Dorothea Nitsch reports grants unrelated to this work from the National Institute for Health Research, Medical Research Council (MRC), the Health Foundation and GlaxoSmithKline.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval The NCKDA research database has national ethical approval (REC reference: 17/LO/0049), section 251 approval (CAG reference number: 17CAG0058). NCKDA steering group approval and LSHTM ethics approval (reference: 21689) were granted for this analysis.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. The data that support the findings of this study are stored at University College London (UCL) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

ORCID iDs
Faye Cleary http://orcid.org/0000-0001-8401-2305
Sally Hull http://orcid.org/0000-0002-8691-7519
Ben Caplin http://orcid.org/0000-0001-9544-164X

REFERENCES

1 Bikbov B, Purcell CA, Levey AS, *et al*. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet* 2020;395:709–33.
2 Jager KJ, Fraser SDS. The ascending rank of chronic kidney disease in the global burden of disease study. *Nephrol Dial Transplant* 2017;32:ii121–8.
3 Schoolwerth AC, Engelgau MM, Hostetter TH, *et al*. Chronic kidney disease: a public health problem that needs a public health action plan. *Prev Chronic Dis* 2006;3:A57.
4 Kidney disease: improving global outcomes (KDIGO) CKD work group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease 2013;3:1–150.
5 Kerr M, Bray B, Medcalf J, *et al*. Estimating the financial cost of chronic kidney disease to the NHS in England. *Nephrol Dial Transplant* 2012;27 Suppl 3:iii73–80.
6 Renal Registry UK. Uk renal registry 22nd annual report – data to 31/12/2018, Bristol, UK, 2020. Available: https://ukkidney.org/audit-research/annual-report/22nd-annual-report-data-31122018 [Accessed 27 Jan 2022].
7 Kim LG, Cleary F, Wheeler DC, *et al*. How do primary care doctors in England and Wales code and manage people with chronic kidney disease? results from the National chronic kidney disease audit. *Nephrol Dial Transplant* 2018;33:1373–9.
8 Nitsch D, Caplin B, Hull S. On behalf of the National CKD audit and quality improvement programme in primary care. first national CKD audit report, 2017. Available: https://www.lshtm.ac.uk/files/ckd_audit_report.pdf [Accessed 28 Jul 2021].
9 Digital NHS, SNOMED CT, 2021. Available: https://digital.nhs.uk/services/terminology-and-classifications/snomed-ct [Accessed 27 Jan 2022].
10 National Institute for Health and Care Excellence. Chronic kidney disease in adults: assessment and management (NICE guideline CG182), 2014. Available: https://www.nice.org.uk/guidance/cg182 [Accessed 28 Jul 2021].
11 National Institute for Health and Care Excellence. Nice quality and outcomes framework indicator. Available: https://www.nice.org.uk/standards-and-indicators/qofindicators [Accessed 27 Jan 2022].
12 Mudie K, Cleary F, Caplin B. On behalf of the National CKD audit and quality improvement programme in primary care. second National CKD audit report, 2017. Available: https://www.lshtm.ac.uk/media/9951 [Accessed 28 Jul 2021].
13 Capurro D, Coghlan S, Pires DEV. Preventing digital overdiagnosis. *JAMA* 2022;327:525–6.
14 Digital NHS, Codes R, 2020. Available: https://digital.nhs.uk/services/terminology-and-classifications/read-codes [Accessed 27 Jan 2022].
15 Molokhia M, Okoli GN, Redmond P, *et al*. Uncoded chronic kidney disease in primary care: a cross-sectional study of inequalities and cardiovascular disease risk management. *Br J Gen Pract* 2020;70:e785–92.
16 Hull SA, Rajabzadeh V, Thomas N, *et al*. Improving coding and primary care management for patients with chronic kidney disease: an observational controlled study in East London. *Br J Gen Pract* 2019;69:e454–61.
17 Olaitan A, Ashman N, Homer K, *et al*. Predictors of late presentation to renal dialysis: a cohort study of linked primary and secondary care records in East London. *BMJ Open* 2019;9:e028431.
18 Xie X, Liu Y, Perkovic V, *et al*. Renin-Angiotensin system inhibitors and kidney and cardiovascular outcomes in patients with CKD: a Bayesian network meta-analysis of randomized clinical trials. *Am J Kidney Dis* 2016;67:728–41.
19 , Herrington W, Emberson J, *et al*, Cholesterol Treatment Trialists' (CTT) Collaboration. Impact of renal function on the effects of LDL cholesterol lowering with statin-based regimens: a meta-analysis of individual participant data from 28 randomised trials. *Lancet Diabetes Endocrinol* 2016;4:829–39.
20 , Wright JT, Williamson JD, *et al*, SPRINT Research Group. A randomized trial of intensive versus standard blood-pressure control. *N Engl J Med* 2015;373:2103–16.

6.4 Key findings and implications for PhD

- Where our study investigated the **association between practice CKD coding performance and patient-level outcomes** (such as risk of CV events), **there was little evidence of association** between practice CKD coding performance and patient characteristics (i.e. **between primary exposure and individual confounders**), except for slight differences at the extremes of CKD coding performance, heavily reducing risks of confounding due to patient characteristics.

- Practice CKD coding performance may be related to other practice characteristics that may impact outcome rates in CKD patients. We attempted to adjust for key features of practices which may affect outcomes, such as practice risk profile characteristics and differences in practice disease severity among those identified with CKD. However, there is still a **risk of residual confounding due to practice characteristics that we did not identify** as confounders.

- Cox regression analyses (with splines for practice CKD coding) that adjusted for practice risk profile characteristics alone resulted in inverted **S-shaped hazard ratio curves**, suggesting the presence of **residual confounding** at the extremes. **Additional adjustment** for practice testing behaviours and characteristics of identified practice CKD populations appeared to resolve much of this confounding (although this cannot be verified), leading to curves approaching a more **linear association**, and a **more intuitive basis for a possible causal relationship**.

- Practice characteristics that are likely to improve CKD outcomes (CKD care practices) were verified to be associated with CKD coding performance, suggesting a logical mechanism for a causal relationship between practice CKD coding and patient adverse outcomes. Furthermore, significant **reductions in adverse outcomes were associated with better CKD care practices**, supporting this logic.

- This study used a **novel approach** to deal with substantial (measured and unmeasured) **confounding** that would have been present at the patient-level, and additionally sought to capture and **adjust for testing biases** that resulted in different identified CKD populations between practices. We believe this study adds further evidence pointing towards a causal relationship indicating that **completeness of**

electronic diagnostic CKD coding and subsequent care efforts may be **protective against complications** that are associated with progression of CKD, in patients with CKD.

- We quantified the **percentage of hospitalisation events that could be prevented** (using attributable fractions) resulting from an improvement in practice CKD coding from 55% of CKD cases being coded to 88% of CKD cases being coded (corresponding to 17th and 83rd percentiles [bottom of sextile 2 to top of sextile 5] of the distribution of practice CKD coding), **assuming** that our fully adjusted models captured a **causal association** between practice CKD coding and patient outcomes. We estimated that **9.0% of CV hospitalisations and 16% of HF hospitalisations could be prevented by improving practice CKD coding from 55% to 88%**, over the median follow-up time of 3.8 years, which is likely to be clinically important.

Chapter 7

Developing a new albuminuria-free risk prediction equation for kidney failure in patients with chronic kidney disease: retrospective cohort study (paper 4)

7.1 Chapter summary

This chapter provide a summary of work done in research paper 4, presents the original published research paper, presents additional work investigating clinical utility of the developed model, and lists key findings and implications in the context of the overall thesis.

7.2 Summary of work in the context of aims of PhD

Background

Key aims of this PhD included exploring the magnitude of problems resulting from issues of data accuracy and completeness inherent in EHRs when used in the study of CKD progression, and finding new ways to adapt research methods to overcome these challenges.

One important area of research is in risk prediction modelling for CKD progression. Clinical prediction modelling is increasingly recognised as an important part of epidemiological research, as researchers and clinicians alike aim to better personalise healthcare to individual needs, recognising that there is no "one size fits all" treatment approach for all patients afflicted with a given health condition, but rather there are likely to be a number of factors which play into patient risks and appropriate treatments. This is certainly true for CKD which is a heterogeneous disease, with an array of different causes, severities and subsequent consequences.

Equations have previously been developed to predict risks of kidney failure, of which the most commonly recognised and best validated to date is the 4-variable kidney failure risk equation (KFRE) [79-80]. This equation was developed using EHR data for patients referred

to a renal clinic in Canada, and was initially validated using a similar dataset comprising a CKD registry in a different region [79]. It has since undergone multi-national validation in a more diverse range of CKD cohorts, showing consistently high discrimination (ability to rank patients according to underlying risk) and reasonable accuracy of predicted risks [80].

A key goal of the original KFRE study was to use data that is routinely available, allowing for implementation in clinical practice, which may have been successful in initial cohorts of interest, but does not extend well to broader CKD populations who may be managed in different care settings, and with differing care practices between populations and over time, in particular due to the requirement of urine albuminuria measurements for the risk prediction.

While uACR is known to be an important predictor of CKD progression and outcomes (and is included in key staging criteria for CKD, as shown in *Chapter 1*), testing for uACR is actually rather poor in general practice, particularly in earlier stages of CKD and in the absence of specific risk factors such as diabetes [16,81].

This leads us to an important issue, that despite recommendations (for example, in NICE guidelines) to use KFRE to predict risks of kidney failure both to educate patients about their risks and in decision-making about referral to specialist care (at least in the UK), this can only be done for a selected group of patients who have received appropriate testing. Testing rates for urinary ACRs have not improved over time and appear to be entrenched. This in turn may lead to failures to identify certain subgroups of patients of unknown characteristics and is a potential issue for equity of healthcare.

We therefore aimed to develop new equations for kidney failure, suitable for all patients with confirmed CKD, regardless of availability of uACR data.

Methods

This work used a large integrated healthcare database in Sweden using renal function test results captured in both primary and secondary outpatient care.

Results

Resultant developed equations achieved high discrimination for the primary outcome (kidney failure, defined as initiation of KRT), which was comparable to that of KFRE, evaluated both in the development cohort and validation cohort (using temporal split-sample validation).

Conclusions

We are excited about the prospects for this equation to be used in clinical practice, enabling patients at earlier stages of CKD (who are less likely to have uACR data) to be evaluated for risks of CKD progression. The equation could be used immediately in routine care in Stockholm, Sweden, and is likely to generalise well to the wider Sweden population. Further external validation is needed before this equation could be expanded for use in other populations, in particular where predictor variables may be measured or collected in different ways and with varying data completeness in other settings and populations, which may affect model performance.

7.3 Research paper 4

See next page for draft research paper, which is currently being prepared for submission, and appendix 6 for supplementary materials.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1703701	**Title**	Miss
First Name(s)	Faye		
Surname/Family Name	Cleary		
Thesis Title	Challenges of studying and predicting chronic kidney disease progression and its complications using routinely collected electronic healthcare records		
Primary Supervisor	Dorothea Nitsch		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	BMJ Medicine
Please list the paper's authors in the intended authorship order:	Faye Cleary, David Prieto-Merino, Rupert Major, Juan-Jesus Carrero, Dorothea Nitsch

Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The study was conceptualised by DN, FC, DP and JJ. Details of study design were led by FC, with substantial contributions from DN, JJ and DP. Funding was acquired by FC. Data, software and resources were provided by JJ for analysis. FC was responsible for analysis and reporting of the study. All authors contributed to the interpretation of the data, draft revisions and approved the final version to be published.

SECTION E

Student Signature	Faye Cleary
Date	31 May 2024

Supervisor Signature	Dorothea Nitsch
Date	10/06/2024

Title: Developing a new albuminuria-free risk prediction equation for kidney failure in patients with chronic kidney disease: retrospective cohort study

ABSTRACT

Objectives: The commonly adopted kidney failure risk equation (KFRE) requires urine albumin:creatinine ratio (uACR) data which is not routinely available in the majority of patients. We aimed to develop new risk prediction equations for kidney failure in patients with chronic kidney disease (CKD), that do not require uACR data.

Design: Retrospective cohort study

Setting: Stockholm CREAtinine Measurements (SCREAM) database, containing routinely collected electronic healthcare records in both primary and outpatient care from the region of Stockholm, Sweden

Participants: 116,158 adults with CKD stages 3-4, defined by 2 estimated glomerular filtration rate (eGFR) results <60 to ≥15 ml/min/1.73m², at least 90 days apart, with no intermediate eGFR ≥60 ml/min/1.73m², between 1 January 2010 and 31 December 2018

Main outcome measure: Kidney failure, defined by first initiation of kidney replacement therapy (KRT), recorded within 5 years following the index date

Results: Using temporal split-sample validation, development and validation cohorts included 85,012 patients (736 KRT events) and 28,338 patients (114 KRT events), respectively. Following Cox regression using automated backwards selection, the final model included 10 predictors (in order of statistical significance): eGFR, age, diabetes, sex, atrial fibrillation, anti-hypertensive drugs, peripheral artery disease, eGFR decline, acute kidney injury, and hypertension. Model discrimination was excellent in both the development cohort (C statistic 0.941 (95% CI 0.932 to 0.951)) and validation cohort (C statistic 0.944 (95% CI 0.923 to 0.965)). In 26,229 patients with uACR data, the 4-variable KFRE showed marginal improvement in discrimination over our new equation (KFRE: C statistic 0.950 (95% CI 0.942 to 0.958); new equation: C statistic 0.926 (0.915 to 0.936)). However, the KFRE under-estimated risk in our cohort, with an observed/expected event probability ratio of 2.11, suggesting re-calibration is required.

Conclusions: It is possible to predict risk of kidney failure in a general CKD population with high accuracy using data that is routinely available, without requiring uACR results.

KEY MESSAGES

What is already known on this topic

- The 4-variable Kidney Failure Risk Equation (KFRE) was developed in 2011 with multi-national validation in 2016, showing high discrimination for risk prediction of kidney failure (initiation of kidney replacement therapy) in patients with chronic kidney disease (CKD).
- KFRE is commonly used in routine care and is recommended by the National Institute for Health and Care Excellence (UK) to identify patients at high risk of kidney failure, who would benefit from increased monitoring, referral to nephrology care and/or tailored interventions to reduce risk of disease progression.
- The KFRE requires data on age, sex, estimated glomerular filtration rate (eGFR) and urine albumin: creatinine ratio (uACR) to predict risk of kidney failure in the next 2 or 5 years; however uACR is not routinely available for most patients with CKD.

What this study adds

- This study highlights that a minority of patients with CKD can be evaluated using the existing KFRE due to low routine testing rates for uACR.
- A new equation is presented using data that is likely to be routinely available in the general CKD population, with comparable discrimination performance to KFRE in a large general CKD population cohort in Stockholm, Sweden.

How this study might affect research, practice or policy

- Following suitable external validation in populations of intended use, our new risk prediction equation will allow risk prediction for kidney failure in the next 5 years for patients without uACR data.
- Use of our new equation to risk stratify patients with CKD may improve equity of care where patients currently identified using KFRE represent a selective population of those tested for uACR.

INTRODUCTION

Chronic kidney disease (CKD) is emerging as one of the most important public health challenges of recent decades[1-3]. Affecting approximately 10% of the population worldwide (CKD stages 1-5), the number of cases of CKD has been increasing[4-6], due to a perfect storm of factors, including increasing prevalence of risk factors for CKD such as obesity and diabetes, population growth and an ageing population. CKD is characterised by presence of long-term kidney damage which may progress over time, but is often silent in early stages of disease, requiring regular routine testing to detect. It is diagnosed through identification of reduction in estimated glomerular filtration rate (eGFR) requiring measurement of blood biomarkers, and/or presence of albumin in the urine (albuminuria), sustained over time[7]. It is a heterogeneous disease, varying in aetiology and likelihood and pathways of progression[8-9]. Later stage disease is associated with worse outcomes, including cardiovascular events, mortality and in rare cases progression to end-stage kidney disease (ESKD) requiring kidney replacement therapy (KRT), but presenting a substantial burden to healthcare services[10]. Earlier identification of patients at highest risk and better targeted care efforts have the potential to delay disease progression and improve patient outcomes.

Clinical prediction modelling is a rapidly growing field, as the need for informed clinical decision-making based on individualised risks is increasingly recognised, and large datasets of clinical information capable of supporting development of such models are increasingly available[11-15]. In 2011, equations were developed using data from a Canadian nephrology clinic to estimate the risk of kidney failure, with the goal of improving targeting of care to those patients most likely to progress. The resultant 4-variable kidney failure risk equation (KFRE) showed high accuracy of risk prediction in original analyses, and in multinational validation analyses reported in 2016 (with addition of a calibration factor recommended in some populations)[16-17]. The equation has since been recommended by the National Institute for Health and Care Excellence (NICE) to support prioritisation of care[18]. However, risk prediction requires data on age, sex, eGFR and urine albumin:creatinine ratio (uACR); and while age, sex and eGFR are likely to be readily available in patients at risk of CKD progression, uACR is much less frequently collected as part of routine care. In a UK audit of kidney care in 2016, 54% of patients with diabetes had an uACR (or protein:creatinine ratio (PCR)) test in the last year and 30% with hypertension had a test in the last 5 years[19]; a recent meta-analysis including multiple international cohorts from the CKD Prognosis Consortium had similar findings with only 35% of patients with diabetes and 4% of non-diabetic patients with hypertension receiving uACR tests over a pre-defined 2-year period[20]. The KFRE can not be used to estimate risks in patients without uACR data, and it is not known how well the model will perform in

the subset of patients who do not typically have uACR tests, as it has not been validated in such a population.

This study aimed to develop new (prognostic) equations for kidney failure, suitable for use in all patients with CKD stages 3-5, using data that is routinely available (and which does not require uACR data) to be used in clinical practice to support healthcare professionals with prioritisation of care. The main benefits would be to CKD patients without uACR results, enabling earlier detection of those with high risks for kidney failure to be prioritised for more focussed care.

MATERIALS AND METHODS

Data sources

We used data from the Stockholm CREAtinine Measurements (SCREAM) project, an integrated healthcare database including all kidney function tests conducted in routine care across all citizens in the region of Stockholm, Sweden [21-22]. Laboratory data are linked to Swedish Renal Registry data for dates of initiation of KRT, regional healthcare databases for demographics and clinical diagnoses recorded across all care settings, the Swedish Prescribed Drug Registry including complete collection of prescription drugs dispensed at Swedish pharmacies, and population registry for dates of death.

Participants

Adult (18+ years old) patients with CKD stages 3-4 were selected for analysis based on eGFR results, defined by 2 eGFR results <60 ml/min/1.73m² and ≥15 ml/min/1.73m², at least 90 days apart, and with no intermediate eGFR ≥60 ml/min/1.73m². The 2009 CKD-EPI formula was used to estimate GFR in all analyses, assuming Caucasian, and Swedish laboratories have used standardised creatinine reporting throughout the study period. Only primary care and outpatient eGFR measures were used for cohort identification and analysis, as inpatient results may be affected by acute changes in kidney function which are not reflective of long-term chronic changes in kidney function. The index date was defined as the date of first evidence of CKD stages 3-4, at the 2nd qualifying eGFR result. Different cohort entry periods were explored to identify a suitable cohort for analysis with good availability of historical eGFR results prior to the index date (for estimation of eGFR slopes), taking into account the number of outcome events available for analysis and changes in availability of uACR over time for comparison with existing equations (**Supplementary Table 1**). The final cohort selected used creatinine results collected between 1 January 2010 and 31 December 2018. Follow up for outcomes was available from the index date to 31 December 2021.

Outcome

There are different possible approaches to identify 'kidney failure'. Our primary outcome was first initiation of KRT, consistent with previous studies [16-17]. This outcome benefits from being clinically important, clearly defined, with excellent data capture in our database, and is a strong proxy for "CKD progression". Also of interest is biochemical evidence of CKD stage 5 (eGFR<15 ml/min/1.73m²), where decisions to start KRT may vary between patients for reasons unrelated to observed eGFR. A sensitivity outcome of kidney failure was also studied, termed for clarity "non-rebounding" eGFR<15 or KRT, defined as the first instance of eGFR <15 ml/min/1.73m² which was not subsequently followed by eGFR ≥15 ml/min/1.73m² at any later date (or initiation of KRT).

Predictors

All candidate predictors were proposed based on clinical reasoning and were defined at the index date unless otherwise specified. Continuous variables were included in prediction models as linear covariates. Candidate predictors were defined as follows:

- Demographics: age, sex
- Medical history (defined by ICD-10 codes [**Supplementary Table 2**] any time prior to index date): diabetes mellitus, hypertension, heart failure (HF), coronary heart disease (CHD), atrial fibrillation (AF), stroke, peripheral artery disease (PAD), chronic obstructive pulmonary disorder (COPD)
- Renal history: baseline eGFR (at index date), prior eGFR slope of decline (estimated by simple linear regression models in individual patients with at least 3 historical eGFR results up to and including index date, where positive slopes indicate decline in kidney function), recent AKI event (hospitalisation due to AKI [ICD-10] in last year)
- Medications: use of angiotensin-converting enzyme (ACE) inhibitors and angiotensin II receptor blockers (ARBs) (defined by ATC codes, prescribed at any time within -12 months to + 3 months of index date)

New models were compared to KFRE which requires uACR data. Model comparisons were carried out in the subset of patients with an uACR test result available within -12 months to +3 months of the index date, in which case the nearest uACR result to the index date was used.

Sample size

Sample size was arrived at based on feasibility and consideration of event numbers available for different cohort entry criteria (**Supplementary Table 1**).

Missing data

The only data required for primary analyses that had missing values was eGFR slope, which occurred in a minority of patients. We opted to perform a complete case analysis, due to the low number of missing slope estimates being unlikely to affect generalisability, and since we deemed it to be a reasonable expectation to require some historical eGFR data for predicting risks in patients with CKD, especially if such data have predictive power.

Follow-up for the sensitivity outcome (non-rebounding eGFR<15 or KRT) is affected by creatinine testing patterns, which differ by patient and over time. Suitability of the outcome for analysis is impacted by the potential for ascertainment bias, if there is informative testing (likelihood of testing

being associated with patient risk factors). We evaluated testing patterns over time and by key risk factors, to identify any evidence of ascertainment bias, and hence suitability of the outcome to evaluate CKD progression.

Descriptive statistics

Population characteristics of the CKD cohort were summarised, overall and according to availability of uACR data (within -12 months to +3 months of the index date), to demonstrate the sub-populations for which risks can and cannot be predicted using the commonly adopted 4-variable KFRE.

Development vs. validation

We carried out a temporal split-sample validation, dividing the analysis cohort in a 3:1 ratio ordered by index date to arrive at a development and validation cohort, respectively. Index dates ranged from 6 April 2010 to 7 January 2016 in the development cohort and from 7 January 2016 to 27 December 2018 in the validation cohort. Although this approach reduced the size of the development dataset, we justified this due to the large dataset available for analysis which was well-powered to develop a precisely estimated model, and the desire to evaluate model calibration in different patients to those used for model performance, in particular to reveal any change in model performance over time. We chose a 3:1 split to make full use of the available data for model fitting (75% of full dataset), while retaining sufficient data (25% of full dataset) to precisely estimate performance in the validation cohort. We did not externally validate our model in another population or setting.

Statistical analysis

Cox proportional hazards regression was used to develop a prognostic model for risk of kidney failure, with censoring for death. Where the end goal was to develop an equation for the risk of kidney failure within the next 5 years, we opted to cap outcomes follow up at 5 years post-index date. Although inclusion of events beyond 5 years would increase analytical power, there is a risk of biased estimation if hazard ratios are not constant over the longer term. We carried out backward selection of predictor variables using automated variable selection, starting with the saturated model and removing variables with p-value <0.1 based on likelihood ratio tests, and repeated the process manually to verify the results. We used bootstrap re-sampling for internal validation in the development cohort using 200 bootstrap samples, and summarised model discrimination using the optimism-corrected Harrell's C-index. We separately evaluated model discrimination in the validation cohort, using Harrell's C-index. Model development was carried out using Stata command `stcox`,

with postestimation using `estat concordance`, and confidence intervals computed via the `somersd` package [23]. All estimation methods account for censoring.

We assessed calibration in the validation cohort, by comparing Kaplan-Meier estimates of 5-year predicted risks (censored for deaths) by quintiles of predicted risks, displayed graphically. (Methods for deriving 5-year predicted risks are presented in **Supplementary Information 1**). Magnitude of potential over-estimation of number of kidney failure events due to competing risk of death was explored.

Subgroup analyses

To verify whether model performance differs between different patient populations, prediction models were validated within subgroups, including by diabetes status, CKD stage, and availability of uACR data: model discrimination was evaluated by subgroup using Harrell's C index; calibration plots were also produced by subgroup. Subgroup analyses were carried out in the entire CKD cohort rather than the validation cohort, to increase precision of estimation, justified by the interest being in heterogeneity in model performance rather than precise evaluation of model performance.

Model comparisons

We compared our new 5-year risk prediction equation for kidney failure with the 4-variable KFRE using re-calibrated coefficients for non-North American populations [17] (**Supplementary Information 2**), in the subset of patients with uACR data, assessing discrimination using the C statistic, and calibration graphically. Histograms of linear predictions were produced by outcome status for each model, to further visualise discrimination performance. We assessed correlation in linear predictions using Pearson's correlation coefficient and scatter plots, to demonstrate how well predictions (or ranking thereof) would agree between models. We computed the percent shift in rank, which we defined as the difference in rank of linear predictions between models divided by the maximum rank, x100, where the resulting distribution demonstrates how differently the models rank patients in order of risk. We also assessed the distribution of difference in predicted risks between models. To assess calibration of KFRE in our analysis cohort, we computed the observed/expected event probability ratio, accounting for censoring [14].

Clinical utility

Measures have been developed to capture improvements in model performance due to inclusion of additional predictor variables, namely the net reclassification index (NRI) and integrated discrimination improvement (IDI) [24-25]. We decided against computation of such measures due to

reliance on predicted risks, where we identified mis-calibration issues with the KFRE which would unfairly favour our new equation with potentially misleading results.

Patient and public involvement

This study involved secondary use of electronic healthcare records data collected as part of routine clinical practice (the SCREAM database). The SCREAM project was initiated in 2010 and continues to the present, involving maintenance of a large integrated healthcare database, designed for research purposes, with the primary aims to estimate the burden and consequences of chronic kidney disease (CKD) in healthcare, and to identify inappropriate drug use, with end goals to benefit patients through improvements in clinical practice. The current study did not involve direct engagement of patients or members of the public, but utilises the laboratory tests and health trajectories of thousands of patients and we believe it is patient-focused. The main purpose of this study was to improve identification of patients at high risk of kidney failure, with a particular focus on recognition of those without ACR data who are in danger of being disregarded using currently adopted risk prediction equations, a potential issue for healthcare equity. We are liaising with patient representatives Miranda Scanlon (Lay Advisor Group Lead at Kidney Care UK) and Susan Lyon (Chair of the UK Kidney Association Patient Council) on plans for dissemination of results to patients, which may involve, for example, preparing an article for a patient charity magazine. Both patient representatives have kindly reviewed this research article for content relevant to patients; they have proposed some minor changes which we have incorporated, including statement of the expected benefits to patients of this research; they have also encouraged us to ensure appropriate dissemination of results to patients.

Protocol and registration

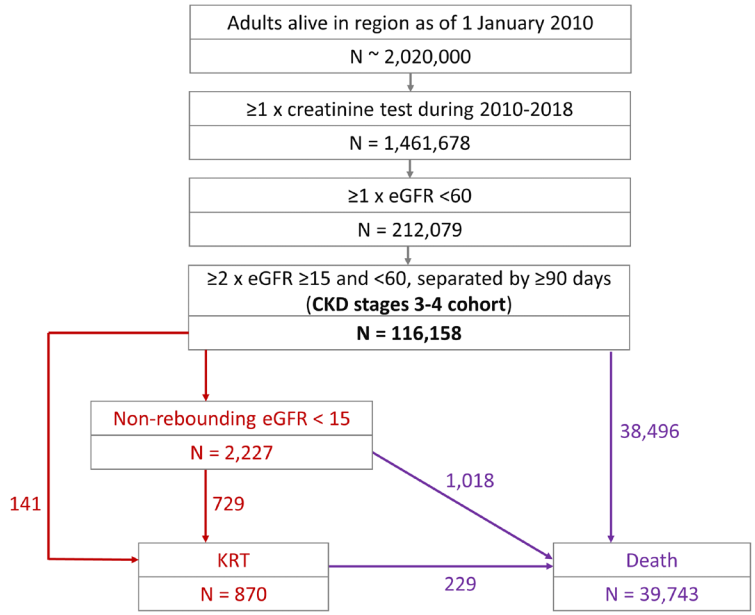
No protocol was prepared for this study and the study was not registered.

RESULTS

Participants

116,158 patients met cohort entry criteria for CKD stages 3-4 (**Figure 1, Supplementary Figure 1**). For comparison, 2.02 million adults lived in Stockholm region as of 1 January 2010. Median time between qualifying eGFR tests was 7.0 months (**Supplementary Figure 2**) and varied by CKD stage (stage 3a: 7.2 months, stage 3b: 6.6 months, stage 4: 5.4 months). It is not clear if differences reflect more frequent testing for later stage CKD or simply increased likelihood of 2 consecutive results below 60.

Figure 1: Flow chart of analysis cohort identification and subsequent outcome events.
Outcome events and deaths include those recorded within the next 5 years post index date. Of those estimated to be living in the region at the beginning of the data collection period, approximately 72% had a creatinine test, 10% had at least one eGFR result < 60 ml/min/1.73m², and 6% met CKD stages 3-4 cohort selection criteria (which excludes those who had already initiated KRT). (These percentages are approximate, as some patients become adults during the data collection period, and some adults die during the data collection period.) Possible pathways for patient outcomes include: (a) Non-rebounding eGFR only [N=480]; (b) Non-rebounding eGFR and KRT only [N=729]; (c) Non-rebounding eGFR, KRT and death [N=167]; (d) Non-rebounding eGFR and death only [N=1,018]; (e) KRT only [N=79]; (f) KRT and death only [N=62]; (g) Death only [N=38,496]. Adding combinations of this breakdown, one may deduce total events for each outcome as shown in the flow chart.



Descriptive results

26,658 (23.0%) of CKD patients had an uACR test within -12 months to +3 months of the index date. Patients with concurrent uACR data were more likely to be younger and male, were more likely to have later stage CKD, diabetes, and hypertension, were more likely to have had a recent hospitalisation for AKI, were more likely to be prescribed ACEi/ARB medications, and are likely to represent a more morbid population (**Table 1**).

Table 1: Baseline characteristics of all patients with CKD stages 3-4, overall and stratified by availability of uACR data (within -12 to +3 months of index date). eGFR = estimated glomerular filtration rate; uACR = urinary albumin:creatinine ratio

Patient characteristics	Chronic kidney disease stages 3-4 (2 x eGFR<60, over 90+ days)		
	Complete cohort	**Cohort without uACR**	**Cohort with uACR**
N	**N = 116,158**	**N = 89,500**	**N = 26,658**
Age, median (IQR) *Age missing, n (%)*	79 (71, 85) *0 (0%)*	80 (73, 86) *0 (0%)*	74 (67, 81) *0 (0%)*
Sex female, n (%) *Sex missing, n (%)*	66,429 (57.2%) *0 (0%)*	54,158 (60.5%) *0 (0%)*	12,271 (46.0%) *0 (0%)*
Diabetes, n (%)	27,953 (24.1%)	14,411 (16.1%)	13,542 (50.8%)
Hypertension, n (%)	87,125 (75.0%)	64,811 (72.4%)	22,314 (83.7%)
Coronary heart disease, n (%)	12,287 (10.6%)	9,266 (10.4%)	3,021 (11.3%)
Heart failure, n (%)	27,418 (23.6%)	27,418 (24.3%)	5,691 (21.4%)
Atrial fibrillation, n (%)	28,208 (24.3%)	22,706 (25.4%)	5,502 (20.6%)
Stroke, n (%)	14,196 (12.2%)	11,345 (12.7%)	2,851 (10.7%)
Peripheral Arterial disease, n (%)	11,448 (9.9%)	8,546 (9.6%)	2,902 (10.9%)
Chronic obstructive pulmonary disease, n (%)	21,499 (18.5%)	16,626 (18.6%)	4,873 (18.3%)
eGFR, median (IQR)	51 (44, 56)	52 (44, 56)	51 (42, 56)
Chronic kidney disease stage, n (%) 3a 3b 4	 84,065 (72.4%) 25,873 (22.3%) 6,220 (5.4%)	 65,730 (73.4%) 19,859 (22.2%) 3,911 (4.4%)	 18,335 (68.8%) 6,014 (22.6%) 2,309 (8.6%)
eGFR frequency[a] median (IQR) 2 measures 3-5 measures ≥ 6 measures eGFR coverage, yrs, median (IQR)	 10 (7, 16) 2,808 (2.4%) 17,186 (14.8%) 96,164 (82.8%) 5.6 (4.2, 8.2)	 10 (6, 15) 2,378 (2.7%) 14,829(16.6%) 72,293 (80.7%) 5.5 (4.2, 7.9)	 13 (8, 20) 430 (1.6%) 2,357 (8.8%) 23,871 (89.6%) 6.4 (4.3, 9.3)
Prior decline in eGFR slope (units per year)[a] Missing (not computed), n (%)	2.51 (1.20, 4.27) 2,808 (2.4%)	2.41 (1.12, 4.11) 2,361 (2.7%)	2.86 (1.50, 4.77) 447 (1.6%)
Recent acute kidney injury, n (%)	1,795 (1.6%)	875 (1.0%)	920 (3.5%)
uACR (mg/g) median (IQR) <30 30-299 ≥300	See right column for those with data	N/A	 1.7 (0.6, 7.8) 23,429 (87.9%) 2,866 (10.7%) 362 (1.4%)
Use of anti-hypertensives	71,184 (61.3%)	50,427 (56.3%)	20,757 (77.9%)

[a]Includes all measures between 2006-2018 prior to (and including) the index date, excluding inpatient measures

Outcomes

96,427 patients (83.0%) had 5 years of follow-up (time between index date and end of outcomes data collection, capped at 5 years); the remaining 19,731 patients (17.0%) had at least 3 years of follow-up, and median follow up of 4 years 2 months (**Supplementary Figure 3**). There were 870 KRT events (primary outcome), 2,227 non-rebounding eGFR<15 events, 2,368 composite non-rebounding eGFR<15 or KRT events (sensitivity outcome) and 39,743 deaths recorded between the index date and end of outcomes follow-up. Crude event rates were 1.9 events per 1000 patient years for KRT and 5.2 events per 1000 patient years for non-rebounding eGFR<15 or KRT. Event times are shown in **Supplementary Figure 4**.

Data completeness for outcome assessment

The number of outcome events, deaths, and number of patients with a valid eGFR test in each year following the index date are shown in **Supplementary Table 3.** The median number of pre-KRT eGFR records within the next 5 years after the index date was 7 (IQR 4-13). Patients with CKD were not all routinely tested each year, with approximately 70-80% of CKD patients who were alive and had not started KRT at the beginning of each year being tested in the next calendar year. (Estimates are conservative as some patients may die or start KRT early in the year [approximately 8% of patients at risk in each year]). There were small but important differences in eGFR testing by CKD risk factors. Kaplan Meier failure curves demonstrate higher event rates in 2010 than subsequent years, which occurred as a consequence of identifying more prevalent CKD cases in the first year of follow-up, as opposed to more incident CKD cases identified in subsequent years (**Supplementary Figures 5-7).**

Agreement between outcome definitions

Of 3,723 patients with a record of eGFR <15 ml/min/1.73m², 759 (20.4%) initiated KRT. Many patients with eGFR < 15 did not start KRT, and eGFR < 15 may not be sustained over time, with rebounds >15 at a later date. Of 2,227 patients with non-rebounding eGFR < 15 within 5 years of index date, 729 (32.7%) started KRT within 5 years of index date. When looking at the 870 patients initiating KRT, 141 (16%) did not have previous non-rebounding eGFR <15 ml/min/1.73m² as outpatient on file. In 729 patients experiencing both events, median time between non-rebounding eGFR<15 and subsequent KRT was 9.3 months (but time between events depends on remaining follow up time). Those that went on to start KRT mostly survived (25% died by end of follow up; median time to death since eGFR<15 of 8 days). Those with CKD5 who did not start KRT mostly died (68% died by end of follow up; median time to death since eGFR<15 of 25 days).] A venn diagram demonstrating overlap in event occurrence in the CKD cohort for non-rebounding eGFR < 15, KRT and deaths is available in **Supplementary Figure 8**.

Model development

Of 116,158 patients in the CKD cohort, 2,808 patients were excluded from analysis due to missing eGFR slope data, leaving 850 KRT events and 2,318 non-rebounding eGFR<15 or KRT events for analysis. After splitting in a 3:1 ratio, the development and validation cohorts included 85,012 patients (experiencing 736 KRT events) and 28,338 patients (experiencing 114 KRT events), respectively (**Supplementary Table 4**).

Primary analysis (KRT)

Predictors selected for the KRT prediction model (in order of statistical significance, based on p-value) which *increased* the risk of KRT were: lower baseline eGFR, lower baseline age, history of diabetes, male sex, no history of AF, concurrent prescription of ACEi/ARBs, history of PAD, steeper eGFR decline, no recent AKI, history of hypertension (**Table 2;** the final prediction equation is presented in **Supplementary Information 1**).

Table 2: Hazard ratios (and 95% confidence intervals) for new risk models for KRT developed in the development cohort, with discrimination statistics evaluated in both development cohort and validation cohort. eGFR = estimated glomerular filtration rate; KRT = kidney replacement therapy.

	Hazard ratios (and 95% confidence intervals)	
	KRT outcome	**eGFR outcome**
eGFR at baseline, per 5 mL/min/1.73m^2	0.53 (0.51 to 0.54)	0.50 (0.49 to 0.51)
eGFR slope decline, per 5 mL/min/1.73m^2/yr	1.03 (1.01 to 1.06)	-
Acute kidney injury in last year	0.57 (0.36 to 0.89)	0.72 (0.56 to 0.91)
Age, per 10 yr	0.49 (0.47 to 0.51)	0.69 (0.67 to 0.71)
Female sex	0.54 (0.46 to 0.63)	0.56 (0.51 to 0.61)
Diabetes	2.15 (1.85 to 2.50)	1.45 (1.32 to 1.59)
Hypertension	1.26 (1.03 to 1.55)	-
Heart failure	-	1.14 (1.02 to 1.27)
Coronary heart disease	-	0.84 (0.73 to 0.97)
Atrial fibrillation	0.56 (0.43 to 0.74)	0.78 (0.69 to 0.89)
Peripheral Arterial disease	1.50 (1.20 to 1.89)	1.26 (1.10 to 1.43)
Chronic obstructive pulmonary disease	-	-
Use of anti-hypertensives	1.62 (1.29 to 2.04)	-
C statistic[a] (development cohort)	0.941 (0.932 to 0.951)	0.883 (0.875 to 0.892)
C statistic (validation cohort)	0.944 (0.923 to 0.965)	0.837 (0.808 to 0.866)

[a]optimism-corrected Harrell's C statistic, computed via bootstrap resampling

Sensitivity analysis (non-rebounding eGFR <15 or KRT)

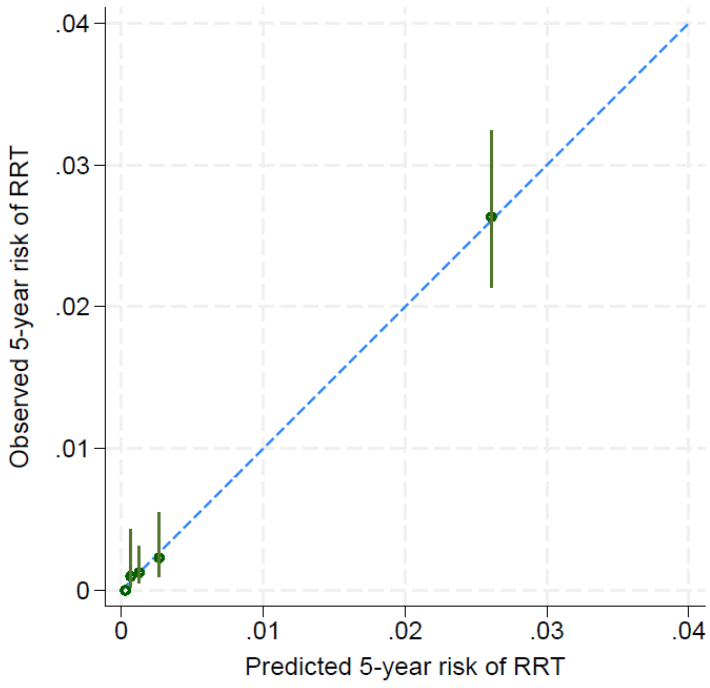
Similarly, predictors selected for the sensitivity outcome model which *increased* the risk of outcome were: lower baseline eGFR, lower baseline age, male sex, history of diabetes, no history of AF, history of PAD, no recent AKI, no CHD, history of HF (**Table 2**).

Model performance

Primary analysis (KRT)

Model discrimination according to the C statistic was high in both the development and validation cohorts (**Table 2**), with no apparent reduction in performance over time. Linear predictions showed good separation according to outcome status in distribution plots (**Supplementary Figure 9)**. Calibration plots showed no evidence of systematic under- or over-estimation in the validation cohort (**Figure 2**). Accuracy of predicted risks appeared reasonably good within subgroups, with some small deviations from linearity (agreement between observed and expected risks) which may not be clinically meaningful (**Supplementary Figure 10).** There was a trend of under-estimation of risks in patients with uACR data and over-estimation in patients without uACR data, suggesting that model performance could be further improved by incorporating predictors of uACR testing (or indeed uACR itself). Discrimination statistics by subgroup are presented in **Supplementary Table 5**.

Figure 2: Observed vs predicted probability of KRT at 5 years, by quintile of predicted risk, in validation cohort. Quintiles of predicted risk: 0% - 0.049%; 0.049% - 0.097%; 0.097% - 0.188%; 0.188% - 0.447%; 0.447% - 100%.



Sensitivity analysis (non-rebounding eGFR <15 or KRT)

Model discrimination was reasonably good for the sensitivity outcome (**Table 2**) but was markedly inferior to that seen in the primary analysis. Discrimination was lower in the validation cohort than the development cohort, which may be partly a result of decreasing follow-up duration over time due to censoring at end of data collection and possibly related to changes in testing rates over time, where the outcome occurrence depends closely on follow-up of creatinine over time. There was modest over-estimation of risks in the highest quintile of predict risks (range 1.357% - 100%) in the validation cohort (**Supplementary Figure 11**).

Predicted risks in those who died

Predicted risks from our model are to be interpreted as the risk of kidney failure that would be expected over a 5-year period, if death could be avoided. In the subset of patients with 5 years of follow-up (time between index date and end of data collection) who died during this period, the number of observed KRT events was lower than the number expected according to risk predictions (manually computed Observed/Expected event ratio of 0.30). This is a result of such patients being assumed to have a continuing risk of KRT after censoring for death. (This is not to be confused with the observed/expected event probability ratio which accounts for censoring [computed for KFRE]).

Comparison with 4-variable KFRE

Among 26,229 patients with uACR data analysed in our CKD cohort, the KFRE under-estimates risk of kidney failure on average, with an observed/expected event probability ratio of 2.11. This miscalibration appears to be mostly isolated to patients at the higher end of the spectrum of predicted risks (**Supplementary Figure 12**). Despite this, discrimination is excellent in this population (C index of 0.950 (95% CI 0.942 to 0.958)), a marginal improvement upon our new risk equation (C index of 0.926 (95% CI 0.915 to 0.936)). (For prediction distribution by outcome status, see **Supplementary Figures 13 and 14.)**

Correlation between linear prediction scores comparing the KFRE with our equation was 0.79 (scatter plot shown in **Supplementary Figure 15**). Distribution of differences in predicted risks and percent change in rank between equations are shown in **Supplementary Figures 16 and 17**. There is satisfactory agreement in rank of prediction scores between equations, with 50% of ranks of prediction scores within +/- 12% of one another (reflected by Q1, Q3). There are non-negligible differences in the way these 2 equations predict risk in individual patients with uACR data (**Supplementary Table 6**).

DISCUSSION

Statement of principal findings

Availability of uACR in patients with CKD is a substantial issue for risk prediction, with only 23% of the CKD cohort identified between 2010 to 2018 having an uACR result within the previous 12 months or following 3 months of the index date at which CKD was confirmed. Our new equation predicted risk of kidney failure (KRT) with very high discrimination, showing promise as a valuable tool for risk prediction, particularly for patients without uACR data in whom KFRE cannot be used. Calibration in the validation cohort was good, with small deviations in accuracy of predictions within subgroups which are unlikely to be clinically important. In patients with uACR data, the existing 4-variable KFRE showed marginal improvement in discrimination over our new equation, however re-calibration of KFRE is needed in the Swedish general CKD population due to observed miscalibration. Trends of under-estimation in patients with uACR data and over-estimation in patients without uACR data using our new equation suggest that there are important unmeasured factors still at play and risk prediction may be further improved by including additional predictors of uACR testing (where uACR itself is not routinely available). We identified substantial issues defining ESKD outcomes using eGFR data collected in routine practice due to incomplete eGFR data collection which varied over time and by risk factors (ascertainment bias) and with issues establishing chronicity which requires prolonged eGFR follow up.

Strengths and weaknesses of the study

This study had many important strengths, chiefly that the SCREAM database used for analysis covered a vast sample of the CKD population, including complete data collection for kidney function tests conducted across primary care and outpatient care in the region of Stockholm. We selected variables for analysis based on likelihood of availability in routine care, resulting in risk prediction equations with clinical utility in the wider CKD population. Very few patients were excluded due to missing data (for eGFR slope derivation) and we carried out a complete case analysis in 98% of the identified CKD cohort. Results are therefore highly likely to be generalisable to the Sweden and other White European heritage CKD populations. Despite KRT being a rare outcome and the relatively low risk CKD population studied, the large sample size of our database provided 736 outcome events for model development and 114 for validation, exploring 14 predictor variables, with a final model including 10 variables, leading us to achieve precise estimation of predictor coefficients in developed equations. KRT is a strong outcome, due to being clinically important, well-defined and reliably measured using renal registry data. Use of the Swedish Renal Registry to identify KRT is a particular strength of this study, as previous validation studies conducted in other settings have highlighted a

severe mismatch in chronic dialysis start dates if only hospital records are used to define dialysis start [26]. The large sample size also reduced the risk of overfitting, demonstrated by negligible difference in naïve C statistic (apparent performance, 0.941) and optimism-adjusted C statistic (0.941) computed as part of internal validation with bootstrapping procedure, showing very low 'optimism' of modelling strategy; there was also negligible difference between C statistics in the development and validation cohorts (0.941 and 0.944, respectively). The high discrimination observed suggests we have developed a powerful tool to rank patients according to risks. The validation cohort represented a lower risk population than the development cohort, due to including relatively fewer prevalent CKD cases and more incident CKD cases, and care practices may have changed slightly over time, yet model calibration remained strong when assessed using this more recent data, within the population and setting studied.

A key weakness of this study was that we only carried out model validation internally. It is not a requirement of model development to carry out external validation, particularly if the model development dataset is large, representative of the target population and reflective of data that will be used for risk prediction in clinical practice. However, model performance may vary in a different population or setting, for example due to different predictor effects, population case-mix, event rate, timing or methodology used for measurement of predictors, or outcome definition. Therefore, external validation will be required in alternative populations and settings intended for use (assessing transportability), with re-calibration if necessary. Additional validation in a population and setting similar to that used in our analysis would add further reassurance of reproducibility. Calibration plots were restricted to a subset of the original dataset (validation cohort), so it is not surprising to see good calibration in this population, although it is reassuring that good calibration was observed in different patients to those used in model development, assessed over a different time period. Observed trends in afore-mentioned reduced accuracy of risk prediction by uACR availability are not overly surprising, as those tested for uACR are likely a sicker group, and not all contributing factors will be captured by model predictors. While this does not impact the utility of our equation and differences in observed and predicted risks may not be clinically important, there is a signal here that there is potential to further improve our equation to more accurately predict risks. Computation of historical slopes of decline in eGFR required by default a minimum coverage of 90 days (due to CKD cohort identification criteria) and we required a minimum of 3 eGFR tests to compute historical slopes. eGFR is a very noisy measure and precision of estimation of slope of decline is likely to be low for patients with short duration of coverage and/or a low number of tests, hence we may fail to fully harness the information held in the underlying slope of decline in kidney function in our prediction model. There is limited scope to address this as increasing analysis constraints would reduce data

availability. Despite this, we found slope of decline to be a statistically significant predictor of kidney failure in our model.

Strengths and weaknesses in relation to other studies, discussing important differences in results

Original development of the 4-variable KFRE in 2011 used data from a Canadian nephrology clinic with external validation in the British Columbia CKD Registry including patients referred to nephrology clinics (patient identification 2001-2008) [16]. Multinational validation of KFRE in 2016 involved a meta-analysis of 31 cohorts participating in the Chronic Kidney Disease Prognosis Consortium (CKD-PC), selected based on data availability, with kidney failure risks ranging from 1.2 to 168.3 events per 1000 patient years [17]. In contrast, our analysis cohort had a crude event rate of 1.9 events per 1000 patient years, representing a much broader CKD population than previous studies. Original KFRE validation had a C statistic of 0.91 and multinational validation had a pooled C-statistic of 0.88, where our equation had a C statistic of 0.94 (but was limited to our original data source). The main strengths of our study compared to previous studies are the inclusion of a broader CKD population which is not selective based on receiving uACR tests or being referred to renal clinics, the fact that variables routinely available in the broader CKD population (not including uACR) are able to predict risk of kidney failure very well, and that our analysis used data collected in more recent years which is likely to better reflect more recent care practices and population health. However, as our validation was limited to a single data source, we do not yet have information on how well our equation performs in similar cohorts, higher risk CKD cohorts or in different geographical locations. Also, where existing equations use variables which are likely to have a strong causal association with the outcome and are unlikely to change in predictive nature over time (age, sex, eGFR, uACR), our equation includes variables which may be subject to measurement variation in different settings and includes variables depicting care practices which may vary between settings and over time. This may impact transportability and longevity of our equation.

Two important issues in developing risk prediction models for kidney failure are competing mortality (where death before the event of interest prevents occurrence of the event) and informative censoring (where risk profiles differ in those who are censored from analysis compared to those not censored). We used Cox modelling with censoring for death to develop our prediction model (the same approach used to develop KFRE[16-17]), which (among other things) assumes that conditional on covariate values, patients who die would have gone on to have had the same risk of kidney failure as those who do not die (had they not died), with an assumed continuing risk of KRT after censoring for death. In reality, patients who die before kidney failure may have a different risk profile to those that progress to kidney failure, and patients who die cannot go on to experience kidney failure. This

inevitably leads to fewer observed events in practice than is predicted by models over a 5-year period [27-28], which we identified to be the case. Other studies have used a competing risks modelling approach [29-31] which instead assumes a zero probability of kidney failure (after time of censoring for death) in those who die but may be problematic if it fails to identity patients at high risk of CKD progression who die before reaching kidney failure. This is a particular issue for CKD where one of those most common consequences of CKD progression is CKD-related death, which is difficult to define or study as an outcome of CKD progression. The competing risks approach may be useful in the setting of practical planning for dialysis initiation (where kidney failure is the only event of interest), but may be less useful in identifying patients who require more specialist care to reduce risk of disease progression (where kidney failure is an important outcome in itself, but is also a proxy for CKD progression in general with various important consequences) and there is a danger here of perpetuating clinical and health system biases against disadvantaged patients. Patients with high predicted risks of kidney failure from standard Cox models (which censor for death) such as our new equation and KFRE have a clinically important risk which warrants further care, regardless of whether their risk of death before kidney failure is higher, and surviving patients continue to be at high risk of kidney failure. A recent study used a machine learning "super learner" modelling approach (KDpredict), including cause-specific Cox models for predicting kidney failure and standard Cox models for predicting death, for the purpose of simultaneously predicting risks of both kidney failure and death, and to avoid "over-estimation" of kidney failure risks in those more likely to die [32]. While this study taps into an important need to provide clinicians with tools to put into context multiple important risks faced by individuals with CKD, there is a risk of mis-use if patients identified with lower risks of kidney failure (due to high competing risk of death) are incorrectly identified as low risk overall, where such patients are likely to be at increased risks of other important CKD-related outcomes, such as cardiovascular events and death. While accuracy of predictions were reportedly excellent for KDpredict, interpretation of the mechanism of action is less clear, making clinical scrutiny more challenging than for our new equation or KFRE. Strengths of the KDpredict model over KFRE may also have been overstated, where the KFRE is not correctly calibrated to the population under study, and would benefit from re-calibration in the analysis cohort prior to model comparisons.

Meaning of the study: possible explanations and implications for clinicians and policymakers

Our risk prediction equation could be used immediately in clinical practice in the general CKD population in Sweden to help treating physicians identify patients which may benefit from further care, which may include prioritisation of uACR testing, drug prescribing that is tailored to specific causality of CKD, or referral to renal clinics. We do not recommend any specific threshold risks for

further treatment, which depends on resources available, but note that other healthcare systems (UK) use a 5% risk threshold for identifying patients for referral[18]. Model co-efficients (or hazard ratios) for our model when interpreted individually may not necessarily represent causal effects (as this is not a requirement of prediction models). In our analysis, hazard ratios appeared to suggest that history of AF and history of AKI are protective against kidney failure; one possible explanation for this is a result of increased monitoring efforts which serve to improve outlook in these groups. In addition, by default, risk predictions apply to AKI survivors with baseline risk data. Similarly, prescription of ACEi/ARBs are associated with increased risks of outcome which is unsurprising given that these drugs are given to high-risk patients (ACEi/ARBs variable may act as a partial proxy of presence of albuminuria). Consistency of high model discrimination in both the development and validation cohorts which cover different time periods, as well as accurately predicted risks in the validation cohort, suggest that healthcare provider behaviours which contribute to functioning of the model (such as drug prescribing and increased monitoring of some patients) are likely to be entrenched over time, which is a sign of likely consistency of model performance in the future. Despite recommendations for uACR testing, uptake of uACR testing does not appear to be improving very quickly (see descriptive data [19, 33-34]) and in some settings has worsened since the COVID pandemic, so we cannot rely on availability of uACR to help predict kidney failure; this is why it is so important to use other methods to help practitioners identify high risk patients, especially because now there are drugs available to prevent decline of kidney function and dialysis onset, such as SLGT2 inhibitors. Failure to take up other methods of identification will lead to some high risk patients slipping through the net, experiencing unnecessary harm, and a potential issue for equity of healthcare if particular patient subgroups fail to be identified.

Unanswered questions and future research

We recommend our model to be validated every few years on more recent data to ensure predictions are still valid, in case of changes in care practices over time which may affect model performance. If care practices improve over time, then predicted risks in future patients may be over-estimated; this may be adequately addressed by updating the baseline risk in the prediction model equation to reflect the population of intended use (temporal re-calibration), but model performance should be tested in case re-calibration is required. External validation will be required in any other populations or settings of intended use. It would be useful to validate our equation in other healthcare systems with different healthcare practitioner behaviours and data collection practices. Any new intervention introduced that aims to use new risk prediction equations to guide clinical practice should be assessed for clinical utility.

REFERENCES

1. Kovesdy CP. Epidemiology of chronic kidney disease: an update 2022. Kidney Int Suppl 2022;12(1):7-11.
2. Eckardt K-UP, Coresh JP, Devuyst OP, et al. Evolving importance of kidney disease: from subspecialty to global health burden. Lancet 2013;382(9887):158-69.
3. Schoolwerth AC, Engelgau MM, Hostetter TH, et al. Chronic kidney disease: a public health problem that needs a public health action plan. Preventing chronic disease 2006;3(2):A57-A57.
4. Hill NR, Fatoba ST, Oke JL, et al. Global Prevalence of Chronic Kidney Disease - A Systematic Review and Meta-Analysis. PloS one 2016;11(7):e0158765 doi: 10.1371/journal.pone.0158765
5. Purcell CA, Abebe M, Adebayo OM, et al. 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.
6. Jager KJ, Fraser SDS. The ascending rank of chronic kidney disease in the global burden of disease study. Nephrology, dialysis, transplantation 2017;32(suppl_2):ii121-ii28.
7. Kidney disease: improving global outcomes (KDIGO) CKD work group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease 2013;3:1–150.
8. Li L, Astor BC, Lewis J, et al. Longitudinal progression trajectory of GFR among patients with CKD. Am J Kidney Dis 2012;59(4):504-12.
9. Caravaca-Fontán F, Azevedo L, Luna E, et al. Patterns of progression of chronic kidney disease at later stages. Clin Kidney J 2018;11(2):246-53.
10. Go AS, Chertow GM, Fan D, et al. Chronic Kidney Disease and the Risks of Death, Cardiovascular Events, and Hospitalization. The New England journal of medicine 2004;351(13):1296-305.
11. Steyerberg EW. Clinical prediction models: a practical approach to development, validation and updating. Second edition. Springer Nature Switzerland AG: Springer International Publishing 2019.
12. Ramspek CL, Jager KJ, Dekker FW, et al. External validation of prognostic models: what, why, how, when and where? Clinical Kidney Journal 2021;14(1):49-58.
13. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ 2024;384:e074819. doi: 10.1136/bmj-2023-074819
14. Riley RD, Archer L, Snell KIE, et al. Evaluation of clinical prediction models (part 2): how to undertake an external validation study. BMJ 2024;384:e074820. doi: 10.1136/bmj-2023-074820
15. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Medicine 2015;13(1):1.

16. Tangri N, Stevens LA, Griffith J, et al. A Predictive Model for Progression of Chronic Kidney Disease to Kidney Failure. JAMA : the journal of the American Medical Association 2011;305(15):1553-59.

17. Tangri N, Grams ME, Levey AS, et al. Multinational Assessment of Accuracy of Equations for Predicting Risk of Kidney Failure: A Meta-analysis. JAMA : the journal of the American Medical Association 2016;315(2):164-74.

18. National Institute for Health and Care Excellence. Chronic kidney disease: assessment and management (2021) NICE guideline NG203. Available from https://www.nice.org.uk/guidance/ng203/resources/chronic-kidney-disease-assessment-and-management-pdf-66143713055173 [Accessed 1 May 2024]

19. Nitsch D, Caplin B, Hull S, Wheeler D on behalf of the National CKD Audit and Quality Improvement Programme in Primary Care, First National CKD Audit Report 2017. Available from https://www.lshtm.ac.uk/files/ckd_audit_report.pdf [Accessed 1 May 2024]

20. Shin JI, Chang AR, Grams ME, et al. Albuminuria Testing in Hypertension and Diabetes: An Individual-Participant Data Meta-Analysis in a Global Consortium. Hypertension 2021;78(4):1042-52.

21. Runesson B, Gasparini A, RashidQureshi A, et al. The Stockholm CREAtinine Measurements (SCREAM) project: Protocol overview and regional representativeness. Clinical kidney journal 2016;9(1):119-27.

22. Gasparini A, Evans M, Coresh J, et al. Prevalence and recognition of chronic kidney disease in Stockholm healthcare. Nephrology, dialysis, transplantation 2016;31(12):2086-94.

23. Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. The Stata journal 2010;10(3):339-58.

24. Pencina MJ, D' Agostino Sr RB, D' Agostino Jr RB, et al. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Statistics in medicine 2008;27(2):157-72.

25. Kerr KF, Wang Z, Janes H, et al. Net Reclassification Indices for Evaluating Risk Prediction Instruments: A Critical Review. Epidemiology (Cambridge, Mass) 2014;25(1):114-21.

26. Santhakumaran S, Fisher L, Zheng B, et al. Identification of patients undergoing chronic kidney replacement therapy in primary and secondary care data: validation study based on OpenSAFELY and UK Renal Registry. BMJ Medicine 2024;3(1):e000807. doi: 10.1136/bmjmed-2023-000807

27. Coemans M, Verbeke G, Döhler B, et al. Bias by censoring for competing events in survival analysis. BMJ 2022;378:e071349. doi: 10.1136/bmj-2022-071349

28. Ramspek CL, Teece L, Snell KIE, et al. Lessons learnt when accounting for competing events in the external validation of time-to-event prognostic models. Int J Epidemiol 2022;51(2):615-25.

29. Ramspek CL, Evans M, Wanner C, et al. Kidney Failure Prediction Models: A Comprehensive External Validation Study in Patients with Advanced CKD. J Am Soc Nephrol 2021;32(5):1174-86.

30. Maher F, Teece L, Major RW, et al. Using the kidney failure risk equation to predict end-stage kidney disease in CKD patients of South Asian ethnicity: an external validation study. Diagn Progn Res 2023;7(1):22.

31. Grams ME, Sang Y, Ballew SH, et al. Predicting timing of clinical outcomes in patients with chronic kidney disease and severely decreased glomerular filtration rate. Kidney Int 2018;93(6):1442-51.

32. Liu P, Sawhney S, Heide-Jørgensen U, et al. Predicting the risks of kidney failure and death in adults with moderate to severe chronic kidney disease: multinational, longitudinal, population based, cohort study. BMJ 2024;385:e078063. doi: 10.1136/bmj-2023-078063

33. Office for Health Improvement & Disparities (OHID) and NHS Benchmarking Network. CVDPREVENT Second Annual Audit Report 2022 (for the audit period up to March 2021). Available from https://s3.eu-west-2.amazonaws.com/nhsbn-static/CVDPREVENT/2022/338%20CVDPREVENT%20Second%20Annual%20Audit%20Report%20FINAL.pdf [Accessed 1 May 2024]

34. Sullivan MK, Jani BD, Rutherford E, et al. Potential impact of NICE guidelines on referrals from primary care to nephrology: a primary care database and prospective research study. British Journal of General Practice 2023;73(727):e141-e47.

Acknowledgments

We thank all the patients who have provided their data to the SCREAM database, as well as all of those involved in the set up and maintenance of the SCREAM database, without which this research would not have been possible. We also thank patient representatives Miranda Scanlon and Susan Lyon for reviewing our work for relevance to patients, and for supporting us with plans for dissemination of our work to patient communities.

Contributors

The study was conceptualised by DN, FC, DP and JJ. Details of study design were led by FC, with substantial contributions from DN, JJ and DP. Funding was acquired by FC. Data, software and resources were provided by JJ for analysis. FC was responsible for analysis and reporting of the study. FC is the guarantor of the work. All authors contributed to the interpretation of the data, draft revisions and approved the final version to be published. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. Transparency: The lead author (the guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Funding

This research was funded by the UK MRC, grant reference MR/N013638/1, The Swedish Research Council and Region of Stockholm (ALF Medicine funding). The funding bodies had no role in the design of the study, collection, analysis, or interpretation of data and no role in writing of the manuscript.

Competing interests

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/. We declare support from the UK Medical Research Council, Swedish Research Council and Region of Stockholm (ALF Medicine funding) for the submitted work. RP is the owner of Kidney Failure Risk Ltd, which owns the domain names kidneyfailurerisk.co.uk and kfre.co.uk. The website was originally funded by AstraZeneca UK. University Hospitals of Leicester NHS Trust has a consulting agreement with Roche Diagnostics regarding risk algorithm implementation, but RP has not received any personal fees for this. RP has received personal fees for speaking from AstraZeneca UK, Bayer and Boehringer Ingelheim. We have no further financial relationships to declare in the previous 3 years that might have an interest in the submitted work.

We have no further relationships or activities to declare that could appear to have influenced this work.

Patient consent

Not deemed necessary by the Ethics committee because researchers never had access to identified data. Pseudoanonimization was performed by the Swedish National Board of Health and Welfare.

Ethical approval

Ethical approval for this research was approved by the Regional Ethics Review Board in Stockholm (reference: 2017/793-31). LSHTM ethics approval (ref: 25759) was also secured for this analysis.

Data availability

No data are publicly available due to GDPR regulations and the involvement of sensitive personal data held in electronic healthcare records. Academic collaborations are possible through exchanges at Karolinska Institutet. If interested, please send your proposal to juan.jesus.carrero@ki.se.

Provenance and peer review

Not commissioned; externally peer reviewed.

7.4 Comparison of clinical utility of new equation vs KFRE

While traditional markers used in the evaluation of clinical prediction models offer important insights into model performance, they do not offer specific insights in to how such models could be warranted for use in clinical practice [82-83]. Here, we compare the potential clinical utility of our new risk prediction equation to KFRE. We consider the difference in net benefit between two models as the measure of interest.

Use of "net benefit" to quantify clinical utility of prediction models

The "net benefit" quantifies the clinical utility of a decision strategy, or put simply, determines whether making clinical decisions based on model predictions (or alternatively, based on markers or tests) would do more harm than good [83]. This requires first making clinical judgments of the relative importance of true positives (the model predicts event occurrence, and event occurs) against false positives (the model predicts event occurrence, but event does not occur).

In our case, we must weigh up the relative importance of correctly identifying patients as "high risk" who go on to develop kidney failure against the importance of avoiding identifying patients as "high risk" who do not develop kidney failure. This relates to a probability threshold or cut-off, T, at which a clinical decision would be made to offer further care if model predicted risk is greater than T.

Defining clinically relevant risk thresholds for further care

In current UK clinical guidelines [18], referral to specialist renal care is recommended if risk of kidney failure in the next 5 years exceeds 5%, based on KFRE. Referral to renal care is therefore one example of a clinical decision that is likely to be supported by our new equation (and existing KFRE), and which we focus on here. A threshold risk of 5% suggests that the healthcare system is willing to offer further intervention for 20 patients in order to capture one patient with the event, or equivalently that the benefit of identifying a true positive is 19 times more important than the cost of identifying a false negative.

A decision strategy can be considered "beneficial" (net benefit > 0) if the number of correctly predicted outcome events (true positives) exceeds the appropriately weighted (x w) number of falsely predicted outcome events (false positives) [82]. The weighting is directly informed by

the risk threshold, such that w = odds(T). So for a risk threshold of 5%, net benefit > 0 if the number of true positives exceeds 1/19 times the number of false positives (or the number of false positives does not exceed 19 times the number of true positives). The optimal strategy for decision making about further treatment will be that which maximises the net benefit, i.e. that in which the number of true positives exceeds the weighted number of false positives to the largest degree.

Decision curve analysis

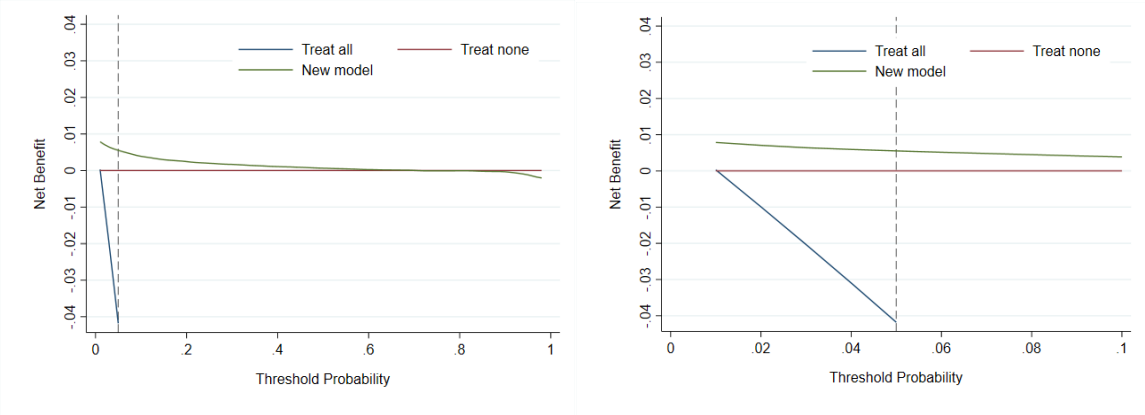
Decision curve analysis (DCA) can be used to compare the net benefit of different decision strategies to guide further treatment, against different potential clinical risk thresholds. We compute the net benefit and present decision curves covering a range of different plausible risk thresholds (using STATA package stdca), which may be used to guide decision-making in patient care. We first present decision curves for our new model in all patients, and then present decision curves comparing different models in all patients with ACR data (**Figure 7**). For completeness, we present curves for all possible risk thresholds as well as for those which are likely to be clinically important (e.g. up to 10%), but evaluation of model utility should be restricted to the range of potentially clinically relevant risk thresholds only.

Results and interpretation

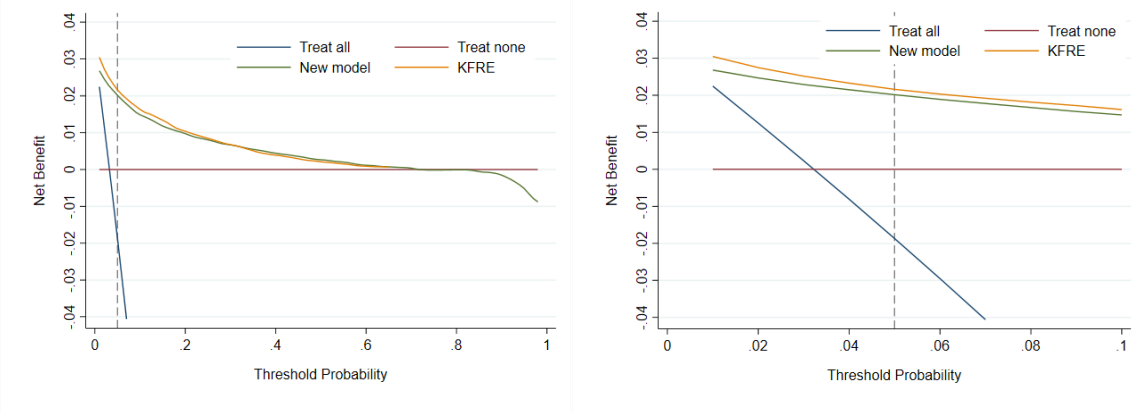
Figures 7A and **7B** shows that the "treat all" (i.e. always predict presence of outcome) strategy has negative benefit for all possible risk thresholds, and should not be used. This is due to the very low prevalence of kidney failure in the entire CKD population. In contrast, in the population of those with ACR data (**Figures 7C** and **7D**), there exist risk thresholds for which a "treat all" strategy would be preferable over a "treat none" (i.e. always predict absence of outcome) strategy. This is owing to the higher prevalence of kidney failure in those who tend to be tested for ACR. Therefore, if the population prevalence is the only information available about a patient's risk and this is greater than the risk threshold, then the logical decision is to treat (offer further care). In the case of a risk threshold of 5% (e.g. refer to nephrology specialist if risk > 5%), "treat all" strategies are not warranted, due to the high cost of identifying too many patients for treatment who do not go on to experience kidney failure, which would take up significant healthcare resources with limited benefit to patients.

Figure 7. Net benefits derived from a decision curve analysis, summarising the excess of clinical benefits compared to expected costs of different decision strategies, namely: (1) treat all; (2) treat none; (3) new model; (4) KFRE. Plot (A) compares the new model to treat all/none strategies across all possible risk thresholds, where plot (B) is the same figure on a reduced scale of risk thresholds of greater clinical relevance (up to 10%). Plot (C) additionally compares the new model to KFRE, reduced to the population of patients with ACR data, where plot (D) is the same figure on a reduced scale of risk thresholds of greater clinical relevance (up to 10%).

(A) (B)



(C) (D)



On the contrary, developed models (new model and KFRE) offer net benefit (>0) at all plausible risk thresholds of likely clinical relevance (**Figures 7B** and **7D**), and should therefore be adopted as supporting resources to aid treatment decision making (in the absence of other potentially better decision strategies which we have not evaluated or compared to). Among patients with ACR data, the KFRE offers marginal benefits over our new equation,

shown by higher net benefit of KFRE for all relevant risk thresholds (**Figure 7D**). However, we don't know if this finding would be generalisable to patients not typically tested for ACR (i.e. the broader CKD population) who represent a slightly different population of CKD patients.

The difference in net benefit between KFRE and our new equation at a risk threshold of 5% in the population of patients with ACR data is 0.00124, which is interpreted as the increase in proportion of true positives without a change in false positives when using KFRE instead of our new equation. That is, use of the marginally inferior new equation instead of KFRE in these patients would lead to 1 fewer true positives (without change in number of false positives) per 806 patients evaluated (= 1/0.00124). This difference may be more profound if the KFRE were correctly calibrated to the studied population (which we have observed is not the case), and we cannot say if this difference would be similar in the population of CKD patients without ACR data measured.

An alternative interpretation is that use of our new equation instead of KFRE would result in a proportion of 0.02356 (=1/19*0.00124) more false positives without change in the number of true positives, i.e. 1 more false positive per 42 patients evaluated, at a risk threshold of 5%.

Summary of results

In summary, the *cost* of using our new equation over KFRE (in patients typically tested for ACR) at a risk threshold of 5% could be described as:

(a) 1 **fewer** patients referred to renal clinics who **will** go on to develop kidney failure, per 806 patients evaluated, *or similarly,*
(b) 1 **additional** patient referred to renal clinics who **will not** go on to develop kidney failure, per 42 patients evaluated.

7.5 Key findings and implications for PhD

- In the identified CKD cohort, **only 23% of patients had concurrent uACR results** (recorded within -12 months to +3 months of the index date at which CKD was confirmed by laboratory test results), **clearly demonstrating that the existing KFRE equation is not usable for the majority of patients** with CKD in Stockholm. This is likely to extend to CKD cohorts in other countries, where multiple previous studies have shown low uACR testing rates in routine care.

- Unsurprisingly, **there were large differences in observed patient characteristics for those with and without concurrent uACR data** available in the patient healthcare record. Patients with uACR data were typically younger, more commonly male, with higher rates of coded diabetes and hypertension, later stage CKD and with higher usage of anti-hypertensive drugs.

- **New equations were developed (using data that is routinely available) with high discrimination** in both the development cohort (C statistic 0.941 (95% CI 0.932 to 0.951)) and validation cohort (C statistic 0.944 (95% CI 0.923 to 0.965)). Final model predictors (in order of statistical significance) were: baseline eGFR, age, history of diabetes, sex, history of atrial fibrillation, use of anti-hypertensive drugs, history of peripheral artery disease, estimated slope of eGFR decline, recent prior acute kidney injury, and history of hypertension.

- **Model discrimination was comparable to (but marginally lower than) KFRE** in those with uACR data ((KFRE: C statistic 0.950 (95% CI 0.942 to 0.958); new equation: C statistic 0.926 (0.915 to 0.936)). However, there was evidence of miscalibration of KFRE with severely under-estimated risks in evaluated patients (observed/expected event probability ratio of 2.11).

- Attempts to define eGFR-based kidney failure outcomes were not particularly successful in this study, due to differential frequency of testing observed between risk factor groups and over time (although model performance was in fact still quite good in our study). This was exacerbated in our study by different durations of follow-up between the development and validation cohorts, demonstrating that **more detailed thought is needed to carefully define such eGFR-based outcomes for use in EHR**

studies. These issues have been thoroughly discussed in recent co-authored work (*Appendix 2*).

- While **initiation of KRT is generally the better outcome for kidney failure** due to being clinically important and easy to ascertain using available EHR data, **downsides** are that this outcome is affected by health system funding of KRT, patient choice to decline KRT which may occur more so in the elderly, and **it may not reflect the actual population burden of most advanced kidney disease**. eGFR-based outcomes should not be disregarded, where interest for some research purposes may be in biochemical end-stage kidney disease, with the added benefit that higher sample sizes can be achieved.

- We **advocate for the use of standard Cox models with censoring for death** in developing and validating equations for kidney failure (rather than competing risks methodology, adjusting for competing risk of death), with an accompanying understanding of the interpretation of estimated risks as those that would be expected if death could be avoided. This seems the most appropriate to support clinical decision making to focus management efforts on those with high risks of CKD progression and associated poor outcomes, regardless of the risk of prior death preventing occurrence of eventual kidney failure in some patients.

- Our **new risk prediction equation is ready for implementation** in clinical practice in the general CKD population **in Stockholm**, Sweden, and should be re-validated on new data every few years to ensure predictions are still valid. It will be **most useful in patients without uACR data** for whom it is not possible to predict risks using KFRE.

- We are **hopeful that our new equation will be externally validated** in other populations and settings, which would hopefully provide further evidence of its usefulness in routine care. It may then be deployed in multiple healthcare systems if it is shown to be effective, with the **potential to improve equity of care and support efforts to identify patients at high risk of kidney failure at earlier stages of CKD**.

Chapter 8

Discussion

8.1 Chapter summary

This chapter summarises key findings of research studies conducted as part of this thesis, discusses key strengths and limitations of this work, advises on implications for clinical practice and suggests areas for future research. I reflect briefly on my personal learning during the research degree programme, and present some concluding remarks.

8.2 Summary of main results

Recap of aims and objectives

This thesis aimed to explore and highlight the challenges faced resulting from issues of data quality and completeness inherent to EHRs when used to study the epidemiology of progression of CKD, advising on potential implications for reliability of study results and to present approaches to deal with analytical biases and overcome these challenges.

This was achieved through:

(1) a systematic review describing study methodology used in previous research using EHRs to study progression of CKD;

(2) a feasibility analysis, exploring data completeness in EHRs from UK primary care, and identifying whether sufficient data is captured to study progression of CKD using such data sources;

(3) an analysis investigating the association between GP practice completeness of electronic coding for CKD and individual patient adverse outcomes of CKD (which are known to be associated with CKD progression); and

(4) development of risk prediction models for kidney failure (requiring KRT) which utilise routinely available data, as opposed to currently adopted models (KFRE) which require uACR data that is commonly missing in patients with CKD.

Key findings are presented below for each research paper.

Research paper 1: Systematic review

The increasing availability of EHR databases over recent decades has led to a marked explosion in volume of research that uses EHRs to study progression of CKD, predominantly for the purposes of risk factor identification and causal inference.

Choice of measures of change in eGFR varied widely, including regression slopes, absolute or percent changes between measures, and progression to eGFR thresholds (among others). Statistical methods used also varied widely, including linear regression, linear mixed models, Cox regression, logistic regression, and many more. 52 out of 80 (65%) reviewed articles defined threshold reductions in eGFR deemed of clinical interest, which we summarised as CKD progression measures.

Most studied did not appropriately account for potentially missing data in study design, analysis and reporting. In addition, due to selection procedures for study inclusion and/or due to completeness of data captured during follow-up for outcomes, eGFR tests are likely to be informatively missing (or missing not at random), and results of this systematic review emphasise the real and important possibility of selection bias in EHR studies.

The general lack of clarity and transparency of reporting on data completeness highlights a lack of awareness of the risks involved when studying CKD progression using EHRs.

Research paper 2: Feasibility analysis

Feasibility analysis using the NCKDA database retrospectively evaluated frequency of eGFR tests and slopes of eGFR, for adult patients registered at volunteering GP practices and alive at data extraction (2016), stratified by risk factors evaluated at data extraction. Among the entire adult population covered by the database, only 34% of patients had at least one creatinine test, and 25% had at least 3 creatinine tests, over a median period of 5.7 years. However, data capture was much higher in those with risk factors or established CKD. Percentage of patients with at least 3 creatinine tests was 92% in those with coded diabetes, 95% in those with coded CKD and 96% in those with CKD confirmed by laboratory tests, with high observed numbers of tests in most patients.

Uncalibrated creatinine results and variable laboratory reporting practices which were not appropriately identified in the data led to systematic over-estimation of eGFR slopes

(underestimation of decline), with a median over-estimation of approximately 0.2-0.3 mL/min/1.73m2/year, which was similar in magnitude across subgroups. There was, however, high variation in degree of over-estimation (up to 2+ mL/min/1.73m2/year in some patients).

Laboratory results reported after 2012 were less likely to be affected by uncalibrated creatinine, leading to improved accuracy in slope estimation. Slope estimation biases were also less common for patients with later stage CKD, due to improved accuracy of estimation at lower levels of eGFR.

Research paper 3: Coding analysis

The 2016 NCKDA database was linked to national data with (near) complete recording of hospital admissions (HES) to explore the association between GP practice completeness of coding for CKD and subsequent hospitalisations for adverse events, in patients with confirmed CKD. Primary analyses showed an "S-shaped" relationship between practice completeness of coding for CKD and individual risks of CV hospitalisations (assessed as hazard ratio curves, compared to median practice coding), with adjustment for practice "risk profile" characteristics such as prevalence of coded comorbidities and age and sex distribution. Observed trends appeared to indicate presence of residual confounding.

Further analyses adjusted for additional practice factors which we hypothesised to be associated with both practice CKD coding and likelihood of adverse outcomes. This included practice testing behaviours for detecting and monitoring CKD (testing biases), and practice disease severity among those identified with CKD. Fully adjusted analyses led to hazard ratio curves approximating a linear association between practice CKD coding and risks of hospitalisations for CV events, with statistically significant reduction in risks for increasing practice CKD coding performance. Results appeared to reflect a more intuitive basis for a causal relationship, strengthening evidence for the role of completeness of electronic diagnostic CKD coding in reducing the burden of complications of CKD.

Research paper 4: Risk prediction analysis

The KFRE is a well-validated equation enabling prediction of risks of kidney failure in patients with CKD. However, individual prediction of risks requires uACR data which is not commonly collected in routine care in the general population of patients with CKD. The KFRE is recommended by NICE to guide decision-making for referrals to specialist care;

however, there is a risk of healthcare inequality if only a selected subgroup of patients have the required data to be assessed against these referral criteria.

We developed new equations for kidney failure (requiring KRT) which do not require uACR data, achieving high discrimination which was comparable to KFRE (evaluated in those with uACR results). 10 predictor variables were included in the final model which are likely to be available in routine care including: baseline eGFR, age, history of diabetes, sex, history of atrial fibrillation, use of anti-hypertensive drugs, history of peripheral artery disease, estimated slope of eGFR decline, recent prior acute kidney injury, and history of hypertension.

8.3 Strengths

Uncovering the risks associated with analysis of routinely collected EHRs in studies of CKD progression

This PhD started with a comprehensive review of the literature (paper 1), which opened my eyes to the extent of problems resulting from data quality issues present in EHRs (i.e. substantial risks of selection bias, which were rarely sufficiently acknowledged). It paved the way for me to conduct further research with a deeper understanding of the risks of informatively missing data and recognition of the importance of transparency of reporting and later reflection on biases that may be present in my research. I also consolidated these learnings by proposing advice on transparency of reporting in the context of potential biases related to data quality (*Appendix 3*), which has the potential to improve the quality of future research.

The feasibility study confirmed the reality of informative testing of kidney function in primary care EHRs in the UK, where frequency of availability of eGFR tests and time between tests varied by risk factor, and testing was generally absent in patients without risk factors. Frequency of "loss to eGFR follow-up" also varied by risk factor, but reasons for loss to follow-up (which may include referral to secondary care) were not available in the database and primary care data on initiation of KRT were not complete. Biases in estimation of changes in eGFR due to changes over time in accuracy and completeness of reporting of test results also signalled the challenges around data quality that may be present in EHR data, using data that are predominantly collected to support clinical care. These findings further

emphasised the risks of informative missingness in studies of CKD progression using EHRs, which require consideration in study design, analysis and reporting.

Representative data for the general population

There were 2 research databases used in this PhD: (1) The NCKDA database covering UK primary care, linked to hospital and mortality data; and (2) The SCREAM database covering the entire healthcare system in the region of Stockholm, Sweden.

Both databases hold clinical data for populations with freely accessible universal healthcare coverage. We mentioned in *Chapter 1 (Background)* that the vast majority of citizens in the UK are registered with a GP practice. Hence, study results are likely to be generalisable to patients registered with GPs who stand to benefit from healthcare intervention for CKD. (The details of data available for each patient will of course depend on the extent of patient interaction with the healthcare system.)

This is a major strength over other study designs which are affected by volunteer bias and drop-outs, such as clinical trials, prospective cohort studies and population health surveys. For example, the UK Biobank Study is a large prospective cohort study of 500,000 participants which has been designed to allow research of determinants of disease, with data collected between 2006 and 2010. However, only 5.5% of patients who were invited to participate were finally recruited. A study on the representativeness of the cohort compared to the general population showed that patients had healthier lifestyles than the general population, and approximately half the mortality rate [84].

UK representativeness. The NCKDA database covered 14% of the population of England and Wales. It required GP practices to volunteer for an audit of patient care, and this may have affected generalisability to the overall population to some degree, although the age-sex distribution was the same.

Prevalence of CKD stages 3-5 identified in paper 3 was 3.8% (requiring 2 measures, with most recent measure in last 2 years) compared to 5.1% (requiring a single measure) in the 2016 Health Survey for England [5]. This appears to suggest that we capture approximately 80% of the underlying CKD (stages 3-5) population in our analysis, although detection methods differed.

Sweden representativeness. The SCREAM database is an extremely powerful resource for observational research, which is why we moved to this database for the final research paper.

While data captured in the NCKDA was limited to kidney function measures captured in primary care, the SCREAM database held all routinely collected kidney function test results (and other health data) from all healthcare settings for the entire region (constituting 20-25% of the population of Sweden) [49], and importantly had a functioning linkage to the Swedish Renal Registry. A recent UK study has shown how important such a linkage is to get the correct start-dates for chronic KRT start [85].

Application of chronicity criterion

Rationale for chronicity criterion to reduce misclassification bias. According to international guidelines for the assessment and management of CKD (KDIGO 2024 [9]) introduced in *Chapter 1 (Background)*, detection of CKD stages 3-5 requires evidence of eGFR <60 over at least 3 months to confirm a diagnosis. Evidence supporting development of these guidelines is based on single time-point eGFR from prospective cohort studies [86], rather than evidence of sustained reductions over time. However, concerns around fluctuations in eGFR that may arise in patient presentation in clinical care (due to temporary ill health), and around overdiagnosis which may be upsetting for patients, have led to introduction of the chronicity criterion to confirm a disease diagnosis [24].

We applied the chronicity criterion for CKD in all conducted analyses (papers 2, 3 and 4) to ensure that research findings are applicable to patients with CKD who are detected by GPs according to clinical practice guidelines.

Risks of survivor bias. While the aim is to reduce misclassification bias, it may however contribute to increased risks of survivor bias and immortal time bias in epidemiological research, due to the clinical reality that sometimes the confirmation eGFR test can be delayed by up to 12 months [87]. In our research studies, sizes of CKD analysis populations were substantially reduced by implementation of the chronicity criterion. In paper 3 (using NCKDA), 72% of patients with eGFR<60 were included with confirmed CKD stages 3-5. In paper 4 (using SCREAM), 55% of patients with eGFR<60 were included with confirmed CKD stages 3-4. (Routines for identification differed between studies).

Implications for patient representation. Some patients with CKD may not be identified (by repeat measures) if they do not attend appointments for follow-up tests. This may disproportionately impact those from a poorer demographic, for example due to punitive zero-hours contract working. (On the other hand, more severe test results [very low eGFR readings] are likely to have been picked up and acted upon.) Some patients may also have

died if the GP practice did not have good systems to enforce/remind patients to attend repeat testing appointments. Research findings may therefore only apply for the part of the population who manages to follow relatively rigid NHS care. All in all, this means that some CKD cases may be disproportionately missing from the analysis population, and sicker patients may be over-represented due to more intensive monitoring.

Large sample sizes leading to well-powered studies with precise estimation of results

In the study of CKD progression, one of the most important outcomes is kidney failure requiring KRT, due to burden on patient health and significant financial costs. However, a minority of people with CKD will progress to kidney failure. The prevalence of KRT for ESKD in the UK at the end of 2021 was 1,307 per million population (0.13%) and the incidence of initiation of KRT for ESKD in 2021 was 154 adults per million patients (0.02% initiating KRT) [14]. Large studies are therefore needed to gather sufficient power to evaluate such outcomes. Study of other outcomes such as CV events and mortality will also benefit from large sample sizes, leading to precise estimation of associations.

Our research databases were very large, consisting of over 100,000 confirmed CKD cases, and capable of studying kidney failure outcomes. Analyses of SCREAM (paper 4) included 850 KRT events (although more events were captured over longer time periods than we studied). (We were unfortunately unable to link the NCKDA database to the UK Renal Registry during the course of this PhD.) New risk prediction models for kidney failure (paper 4) were precisely estimated, with relatively narrow confidence intervals for associations between predictors and the outcome, enhancing reliability as a new tool for risk prediction.

NCKDA analyses in paper 3 studied CKD coding at the practice-level rather than patient-level, with the intention to reduce confounding bias. Study of practice-level coding rather than patient-level coding has the effect of diluting the treatment effect, due to the mechanism of improvements through individual coding being “averaged” across CKD patients in the whole practice. Identification of statistically significant associations therefore requires large sample sizes. This study design was therefore only possible due to the large sample size available for analysis, identifying statistically significant hazard ratios for hospitalisation outcomes when comparing different levels of practice CKD coding, with precise estimation of associations.

Breadth of data for detailed evaluation of relevant study variables

A benefit of the use of EHRs for analysis is the breadth of data available. This is particularly true for the SCREAM database, where a major strength is linkage of various healthcare and administrative databases. It is also strengthened by the use of personal identification numbers in Sweden, allowing straightforward and accurate integration of data for individuals [49].

SCREAM holds longitudinal data on serum creatinine tests from 2006 to 2018 for all patients in the region. The volume of data and long duration of data collection gave us choice over time periods for cohort identification, taking into account data completeness in different time periods and relevance to current and future clinical care. By defining cohort entry using creatinine data collected between 2010-2018, we were able to evaluate historical eGFR slopes, computed using 4+ years of prior data (2006+ to cohort entry), without compromising on statistical power for outcomes follow-up. The Swedish Renal Registry provided complete data on KRT outcomes used in analysis, up to 2021, resulting in 5 years of follow-up data for the majority of patients analysed.

Clinically coded diagnoses were captured from healthcare sources across the entire healthcare system, providing abundant information on co-morbidities and adverse events that are relevant to research of CKD. Healthcare utilisation data included clinical diagnoses recorded during consultations in primary and outpatient care, as well as during hospital admissions, with data going back to 1997 when the International Classification of Diseases Version 10 (ICD-10) coding system was implemented [49]. Use of this data allowed us to define and study 14 predictor variables of interest with potential to predict kidney failure risk.

Methodology for tackling missing data (which may be informative)

We have highlighted issues of data completeness and overcome these issues using study methodology.

Proposing methods for handling missing data. In paper 1 (systematic review), we highlighted the problems of informatively missing data that occur naturally in EHR data sources and described the risks of selection bias in such studies. Through this study, we strived to inform about the problems of missing data, which may commonly be informative. We aimed to encourage improvements in study methodology used in future research studies to generate more reliable findings and to better reflect populations of interest.

Capabilities of data depend on data collection strategy. Paper 2 (feasibility analysis) used data in patients who were identified on the basis of being registered and alive at the time of data extraction (2016), looking back at historical eGFR data to summarise frequency of testing and evaluate slopes of decline in eGFR. We acknowledged that use of cross-sectional data captured in this way can lead to survival bias, only evaluating surviving patients, and thereby failing to capture patients who may have progressed rapidly and died.

The primary aim of paper 2 was to assess feasibility for future studies aiming to study CKD progression. If future studies use different primary care databases with longitudinal data capture (such as CPRD), they are likely to benefit from improved data completeness than seen in this study. That is because the current study identified patients with CKD at a single time point and looked back at historically collected eGFR data (when CKD may or may not have been present). A cohort study which identifies CKD at baseline and follows patients forward for eGFR-based outcomes is likely to have higher frequency of eGFR tests, due to patients with established CKD likely having more regular monitoring.

Use of practice-level exposures and adjustment for testing behaviours. In paper 3 (coding analysis), we used a novel approach to deal with (measured and unmeasured) confounding that would have been present at the patient-level, by choosing a practice-level exposure which was not associated with patient risk factors. This avoided the problem of missing data (or incomplete recording) in measurement of patient-level confounders, which were not required for analysis. (There is still, however, the potential for informatively missing data in measurement of practice-level confounders, but which may not be as impactful, due to generally good balance in other practice factors by practice CKD coding levels.) We also sought to capture and adjust for bias due to informative testing that impacted the patient vintage (duration of CKD at detection) and severity of CKD in patients included in analysis.

Utilising data that is routinely available in population of interest. In paper 4 (risk prediction modelling), we highlighted problems due to informatively missing uACR data in routine care that is required by adopted equations (KFRE). We then proposed new equations using data that is routinely available in all patients with CKD, that do not suffer from the same problems. However, we encountered problems with eGFR-based outcomes (secondary outcome), namely eGFR<15 which did not later rebound >= 15, due to incomplete testing over time during follow-up, which was associated with risk factors (ascertainment bias). More detailed thought is therefore needed to carefully define such eGFR-based outcomes for

use in EHR studies. These issues have been thoroughly discussed in recent co-authored work (*Appendix 2*).

Consideration of competing events when studying CKD progression

When developing risk prediction models for kidney failure, we decided against the use of models to account for competing mortality, despite the fact that many patients will not experience kidney failure requiring KRT due to competing risk of death (informative censoring). We instead chose standard Cox models with censoring for death, with predicted risks interpreted as "risk of kidney failure if death could be avoided". The justification for our approach was the concern that patients who die before kidney failure may be on a clinically important trajectory of kidney function decline when they in fact die, and still warrant more focussed clinical care, regardless of whether they die before reaching kidney failure. This is coupled with the fact that those patients who do not die (who may be similar in characteristics to patients that do die) will require further care. We consider this a strength of our analysis but realise that different approaches hold different value.

Pros and cons of different statistical approaches to handling competing risks are discussed further in *Section 8.6 (Implications for research)*.

Enhanced transparency in reporting in the context of high risks of selection biases

We initially emphasised the importance of greater transparency of reporting of research studies using EHRs to study progression of CKD in research paper 1 (systematic review), including greater clarity of how missing data were handled and stating any concerns around selection bias resulting from the data sources. This is important to ensure that any changes in clinical practice based on such research are supported by reliable evidence which is reflective of the populations of interest.

We have strived to clearly describe extent of missingness in our own research analyses (papers 2, 3 and 4) and reflect on the possibility of the impact of selection bias on analysis results.

Co-authored work completed during the course of this PhD *(Appendix 2)* presented detailed interrogation of the biases anticipated in studies of CKD using routine data, with proposed methods to tackle these biases, and pros and cons of different approaches. In addition, I proposed new considerations for study methodology and enhanced reporting in such studies

(*Appendix 3*), which would further support transparency of reporting relating to missing data and potential selection bias.

Striving to tackle healthcare inequalities

A running theme of research conducted in this PhD has been healthcare equality. For example, analyses of completeness of CKD coding (paper 3) highlighted the problems of incomplete coding of CKD which more commonly affects deprived communities and ethnic minorities [19]. Improvements in CKD coding (that we have recommended) have the potential to reduce these inequalities.

Another example is in risk prediction analyses (paper 4), where we identified stark differences in patients who receive uACR tests and those who do not. This is despite the fact that uACR is recommended in patients with CKD. In the UK, the selective group of patients who do receive testing are also more likely to be identified at risk of CKD progression compared to those who do not receive testing, due to NICE referral criteria which requires assessment of KFRE (requiring uACR results). This is another example of potential healthcare inequality, which we aimed to tackle by developing new equations which are suitable for use in all patients, with encouraging results. Further, we evaluated performance of new equations in risk factor subgroups to highlight any risks of instilling inequalities in use of new equations, if accuracy of risk predictions vary between patient groups.

8.4 Limitations

Applicability to ethnic minorities

In UK primary care data (supporting papers 2 and 3), ethnicity coding is poor, and we were unable to establish if results were representative of the general population in terms of ethnicity. In original NCKDA data for patients with CKD or at risk, coded ethnicity was missing for 31% of patients in England and 61% of patients in Wales [16].

A similar problem occurs in Swedish data (supporting paper 4), due to ethnicity data not being available by law. Results of paper 4 are likely to be generalisable to Sweden which has a predominantly white population, but it is unclear if results will be generalisable to the UK or other countries with higher frequency of ethnic minorities. Further validation of risk prediction equations by ethnic groups is needed.

At the time of the audit (data supporting papers 2 and 3), the MDRD study equation [25] was used to adjust for ethnicity (where this was available) in estimation of GFR. This data was used to identify patients with evidence of CKD stages 3-5. The effect of the equation is to multiply estimated GFR levels by approximately 120% in black patients (to account for increased muscle mass), compared to non-black patients with the same serum creatinine levels.

More recently, ethnicity adjustments have been found to be inappropriate, and have been removed from GFR-estimating equations. A recent study in African populations showed that ethnicity co-efficients from MDRD and CKD-EPI 2009 equations led to under-estimation of prevalence of kidney disease and should not be used [88]. It is therefore possible that black patients were under-represented in analyses of patients with CKD (papers 2 and 3). However, this is unlikely to affect the implications of these studies.

Unresolved selection bias

We have highlighted dealing with selection bias (and transparency around this) as a key strength of work presented in this thesis. However, where relevant data were missing in our analyses, we cannot verify if these missing data are reflective of data that are not missing, which is an affliction of the data sources used.

Missing ethnicity and volunteer bias. We have already discussed that we do not know if results are generalisable to different ethnic groups, and the NCKDA only included GP practices who volunteered to participate in an audit of care. This may affect generalisability of results but is unlikely to change the wider implications (for example, of the importance of CKD coding in routine care). One possible impact is that magnitude of association between under-coding of CKD and increased adverse outcomes (in paper 3) may in fact be under-estimated, where practices with lower interest and engagement with the importance of CKD may be less commonly included.

Representativeness of CKD severity. Risk prediction analyses (paper 4) identified patients at first evidence of CKD in the database. This may lead to disproportionate identification of earlier CKD, although there were many prevalent cases captured which will enhance generalisability of findings to CKD patients captured at any stage of disease. There is little impact on study findings expected, due to adjustment for a variety of potentially important risk factors in analysis. Furthermore, model discrimination was very similar in both

development and validation cohorts, despite differences in capture of incident and prevalent CKD cases.

eGFR testing bias in outcome ascertainment. In risk prediction analyses, we sought to study eGFR-based outcomes for kidney failure (eGFR<15, which did not rebound above 15 at any later date) in addition to the primary KRT outcome. eGFR-based outcomes have the potential to better reflect the population burden of people most severely affected by kidney disease due to not being dependent on decisions to start KRT. This is provided eGFR is collected regularly in everybody, with frequency not dependent on risk factors. They also have the potential to strengthen analytical power with larger event numbers expected.

However, evaluation of completeness of eGFR testing in each year following the index date showed only around 70-80% of patients received eGFR tests in each year (and around 8% died per year, precluding eGFR evaluation in some patients). Completeness of testing also varied by risk factors, leading to likely ascertainment bias for eGFR-based outcomes. This led to important concerns about the use of eGFR-based outcomes for reliable predictive inference in this study. The definition we used also did not require any set period of follow-up for identifying "rebounds" >15 to verify chronicity of decline, which may lead to differences in ability to fairly ascertain the outcome between patients who had variable duration of follow-up.

Residual confounding

Problems around confounding are mostly applicable to research paper 3, which explored causal associations.

Identifying all suitable confounders for analysis requires a deep understanding of patient risk factors, healthcare system factors, and understanding of disease mechanisms. It is important to correctly specify causal pathways (usually visualised through use of directed acyclic graphs [DAGs]) to ensure reliable inference. This includes correctly identifying direction of causation and conveying whether variables are confounders or on the causal pathway. Misspecification may result in incorrect interpretation of results. In paper 3, there is potential for residual confounding due to unmeasured (or poorly measured) practice factors, and/or misspecification of causal pathways. By studying practice factors, it was particularly difficult to specify the DAGs reliably.

Unmeasured confounders. We adjusted for practice characteristics, including practice prevalence of risk factors and morbidity in identified CKD populations (which may occur partially because of eGFR testing behaviours). Unmeasured confounding may occur if practices who fail to code CKD also fail to provide care for diabetes and hypertension, which are important causes of CKD, and could explain increased risks of adverse outcomes.

The role of certain practice factors in the underlying causal DAG were not abundantly clear. For example, use of anti-hypertensives to treat hypertension is likely to occur as a consequence of CKD coding and may form part of the causal mechanism of impact of CKD coding on patient outcomes. We therefore did not include treatment for hypertension as a confounder. However, if treatment for hypertension and CKD coding are associated due to occurrence of a common cause (such as practice funding or expertise), this could lead to residual confounding.

Potential for over-adjustment. Adjustment for practice risk factors (such as hypertension prevalence) was performed with the intention of reducing confounding due to differences in underlying practice risk profile which may impact both CKD coding and outcomes. This approach is valid if estimation of hypertension prevalence captures true underlying hypertension prevalence, and practice hypertension prevalence is a factor influencing CKD coding behaviours.

However, our estimation of hypertension prevalence was limited by the use of diagnostic codes to capture hypertension. As a result, this estimate is likely to depend on practice behaviours of coding cases of hypertension, which in turn may be related to practice behaviours of coding CKD. For example, practices with more funding or greater expertise may be better at coding both hypertension and CKD. Alternatively, if the GP's action of coding CKD subsequently leads to improved attentiveness in hypertension coding in the same individuals, this could capture some of the causal pathway and lead to a partial mediating effect of the "practice coded hypertension prevalence" variable when assessing the causal relationship between "practice CKD coding" and patient adverse outcomes. This could lead to over-adjustment bias which has the effect of introducing confounding and may result in bias towards the null, i.e. under-estimating the effect of CKD coding on adverse outcomes.

Poorly measured confounders. Residual confounding may also result from imprecise or incomplete measurement of confounders, or not granular enough levels of measurement. Many variables such as prevalence of coded co-morbidities rely on binary presence of

diagnoses. However, this does not reflect the reality of disease burden which varies between individuals based on disease severity.

Misclassification and measurement error

It is difficult to resolve misclassification errors in the use of EHRs for research, when misclassification is embedded in the available data.

Misclassification of comorbidity. We generally defined comorbidities (covariates/confounders) using diagnostic codes, but we don't know about the accuracy of coding. Recording of diagnoses in EHRs is likely to occur due to a combination of true morbidity and GP behaviour. It is generally likely that recorded diagnoses are specific but not sensitive [87]. In paper 3, this may lead to residual confounding if sensitivity of disease coding is poor.

In paper 4, the main consequence is likely to be excess noise and reduced precision in risk prediction. If the same data used for research that is coded in clinical practice is used for risk prediction in clinical practice, then prediction models should be well-tuned and not impaired by this imprecision. However, if clinicians use different evidence of co-morbidity for risk prediction than is recorded in the healthcare record, or coding accuracy changes over time, this may impact validity of equations.

Misclassification of hospital outcomes. In paper 3, there were some strange trends between CKD coding and AKI risks which were not easily clinically interpretable, which were not resolved after confounding adjustment. One potential reason for this may be misclassification of AKI in hospital records. Diagnostic coding for AKI has been shown to be incomplete [89]. Furthermore, completeness of coding is influenced by patient factors (including age, sex and ethnicity), which may lead to ascertainment bias, and makes interpretation of results more challenging and less reliable.

Measurement of eGFR. Accuracy of measurement of eGFR is changing over time with the pursuit to improve GFR-estimating equations. In conducted research studies, we used GFR-estimating equations which were being implemented in clinical care during data collection and are therefore relevant to conducted studies. One potential limitation is that developed risk prediction equations (which used CKD-EPI 2009) may not remain valid in the future if new GFR-estimating equations are adopted in clinical care, and re-calibration may be required.

8.5 Implications for clinical practice

Need for improvement in repeat testing for CKD, coding of CKD cases, and subsequent management according to NICE guidelines

Coding for CKD in primary care in the UK is highly variable [16] despite financial incentives [75] and is quite poor relative to disease burden. While this is not a new finding, our findings further strengthen the argument of a causal relationship between completeness of coding for CKD and experience of adverse outcomes known to be associated with CKD and its progression.

Improvements in patient outcomes are likely to occur as a result of:

(1) Improved creatinine/eGFR testing for CKD in patients with risk factors, and regular monitoring for established CKD, including improved consistency of repeat testing 3 months following incident eGFR < 60 to confirm diagnosis of CKD stages 3-5;

(2) improved completeness of CKD coding, which is particularly important in more deprived patients who generally suffer worse coding [19];

(3) enhanced patient care in those who are coded according to NICE guidelines [18], including provision of appropriate interventions relevant to causes of CKD (such as management of hypertension, prescription of statins and SGLT2 inhibitors) and better uACR testing.

Regarding (1), complete testing of patients with risk factors has the potential to detect the majority of CKD cases [90]. A previous study comparing prevalence of eGFR < 60 between the Health Survey for England (2009-2010) and CPRD primary care EHRs (most recent eGFR 2009-2014), showed similar prevalence of eGFR < 60 [91], which appears to indicate that testing in those at risk is already strong. (However, eGFR capture in CPRD was accumulated over a 5-year period.) Regular repeat testing is also important for establishing a CKD diagnosis in clinical practice, where delay in testing is likely to delay appropriate care.

Completeness of CKD coding (2) could be improved by automated alerts when laboratory data support a CKD diagnosis.

Enhanced patient care (3) may be aided by better involvement of practice managers in CKD care, who could be responsible for ensuring that all patients with a CKD code receive appropriate care according to NICE guidelines.

An added benefit of improved repeat testing (1) would be strengthening of EHR resources which would gradually become richer, more representative of patient populations of interest, and capable of greater insights in the study of epidemiology of CKD and its progression, which in turn has the potential to further influence patient care for the benefit of patients.

Urine tests

uACR testing in patients with CKD is low [16,33,92], despite its importance in evaluating risk of adverse outcomes associated with CKD, and recommendations for testing in guidelines [9,18]. Incomplete uACR testing in CKD is a problem for identifying patients at risk of kidney failure, in particular due to requirement for evaluation of the KFRE which is recommended by NICE to inform decisions on referral [92].

In Swedish data (paper 4), only 23% of patients identified with CKD stages 3-4 had concurrent uACR results (within the last 12 months or subsequent 3 months of confirming CKD), and testing was more common in those with diabetes (48%), hypertension (26%), CKD stage 4 (37%) and men (29%). Increased uptake of urine tests is needed to support risk stratification, particularly in patients with non-diabetic CKD, who rarely get tested.

Risk prediction models that we have developed offer hope in the absence of uACR testing. If these equations offer similar performance after validation using routinely collected data in the UK, they could be implemented in UK clinical practice. Use of these equations may offer a partial work-around for patients without any uACR test results on file, enabling prediction of risks. If high risks of kidney failure are highlighted by these equations, then perhaps GPs would be prompted to also conduct uACR testing, offer more focussed treatment and monitoring, and in some cases refer to specialist care.

Implementation of new risk prediction models

Previous studies have shown that risk stratification based on risk prediction models (KFRE) offers improvement over use of eGFR thresholds, leading to more referrals of patients who progress and fewer referrals in patients who do not progress [92], therefore offering important value in clinical practice. Our risk prediction equations do not disadvantage patients with poorly collected uACR data required to evaluate risk and have the potential to improve equality of healthcare.

If new equations were to become integrated into clinical practice, through incorporation in NICE guidelines and with integration in software systems for GPs to use, then

implementation would be straightforward and a useful tool to support clinical care. Implementation could be supported by automated risk prediction calculation, helping GPs to identify patients at higher risks of CKD progression. However, this requires working through considerable red tape to get to this stage (due to the Medicines and Healthcare products Regulatory Agency (MHRA) regulations for "medical devices" [93]) and would require engagement with NICE and NHS England to expedite progress. It is likely that the financial risks associated with worsening CKD burden [7] may help with progressing with such initiatives.

Need for patient education

There is a strong need to educate the public about CKD, most of whom do not know what the kidneys do [94]. Patients are not knowledgeable about whether they may be at risk for CKD, and many are not even aware of established CKD diagnoses [95-96]. Presentation to healthcare services is needed as a precursor to any healthcare interventions to protect health, despite no symptoms being present; this leaves a big responsibility on the patient to come forward for testing for risk factors, and continued monitoring after a diagnosis. Improved patient education is therefore needed, especially for patients who are at risk of CKD, if outcomes of CKD are to be prevented in future.

Patient charities such as Kidney Research UK and Kidney Care UK are taking efforts to better educate the public and patients about kidney disease, providing easy to understand resources to educate patients about what kidney disease entails and managing risks, funding patient information events, and ensuring that research projects are patient-centred [97-98].

Meanwhile, greater strides are being made to engage the public with research findings (for example, through patient charity magazines), and it is becoming more commonplace to engage the public earlier in the research process. Patient and public involvement (PPI) is now required to gain funding for research projects [99]. This will hopefully help to encourage patient-centred research, education of the public and knowledge may improve over time.

8.6 Implications for research

There is much still to learn about the progression of CKD in the general population, including research questions which will benefit from the use of EHRs. We have stated the opportunities

of using EHRs earlier in this thesis *(Chapter 1, Section 1.7)* and strengths of research using EHRs in this PhD earlier in the discussion *(Section 8.3)*.

We present here key implications for research based on the research studies conducted as part of this PhD.

Future studies must account for risks of poor data quality and selection bias in study design, analysis and reporting

While researchers are striving to utilise new EHR resources in increasing abundance, there is a concern about using resources mindfully, considering the risks of poor data quality and selection bias. If studies are not conducted appropriately, results may be generated which are biased in magnitude and may reflect unknown sub-populations. At worse, this could lead to healthcare interventions that are not fit for purpose and fail to benefit patients as hoped. There is also the risk of healthcare inequality if research findings presumed to reflect all CKD patients only apply to a select few, yet clinical practice may treat all patients as if research findings apply to all.

Learnings from the systematic review accompanied by knowledge about the potential for biases in EHR research (such as those outlined in the REporting of studies Conducted using Observational Routinely-collected health Data [RECORD] statement [100]) can help to guide researchers endeavouring to utilise EHRs to study CKD progression. To support this aim, we have presented guideline principles to support researchers in designing EHR studies for CKD progression *(Appendix 3)*, to be used alongside existing RECORD guidelines. We hope that this will result in carefully designed studies which are capable of appropriately answering research questions in populations selected for analysis, with full transparency about potential reasons why results may be biased or generalisable.

Key recommendations for study methodology and transparent reporting of future research using EHRs to study CKD progression are as follows (with a more exhaustive list of considerations in *Appendix 3*):

i. Select a **study population** for analysis with high data completeness for exposures, outcomes and relevant covariates or confounders.
   Relevant considerations, ways to reduce likelihood of bias, and transparent reporting: *Consider local healthcare pathways and state in which areas of the healthcare system relevant data are likely to be recorded, whether these areas of the*

healthcare system are captured by data sources, and with what levels of completeness. Where possible, link together multiple data sources to maximise availability of relevant data. Consider any sub-populations which may be under-represented in utilised data sources, reasons for this, and state any anticipated implications on interpretation of resulting analyses. Data completeness criteria for analysis should be clearly reported, as well as the number of patients excluded based on those criteria, and whether patient characteristics differed between those included vs excluded. Clarify whether patients were excluded based on baseline criteria or completeness of follow-up, with reasons. There are trade-offs between requiring lower thresholds for data completeness (which leads to more representative samples being included in the analysis) and ensuring sufficient data is captured to answer research questions (which may lead to results which more reliably reflect those included). If CKD is the population of interest, ensure that disease status is captured through eGFR tests, where diagnostic coding is known to be incomplete.

ii. Select a **study outcome** that can be ascertained fairly in the entire analysis population.
Relevant considerations, ways to reduce likelihood of bias, and transparent reporting: *If studying kidney failure, initiation of KRT is likely to be the most clinically relevant outcome, due to the important burden on patients and healthcare providers. Disease registry data is likely to hold the most complete data on such outcomes, but this may vary between settings and should be verified. Researchers should be aware that individual patient factors may impact decision-making for dialysis uptake, which may impact interpretation of analyses. Significant care should be taken if the decision is made to use eGFR-based outcomes in analysis, where data completeness varies between eligible patients or over time. This may be considered to achieve greater sample sizes or capture CKD progression at earlier stages of disease. In this case, completeness of eGFR testing over time and differences by relevant covariates or confounders (which may lead to ascertainment bias) should be assessed and reported.*

iii. **Statistical methods** should be used to account for missing data that appropriately account for the underlying missing data mechanism.
Relevant considerations, ways to reduce likelihood of bias, and transparent reporting: *Choice of statistical methods will depend on the amount of missing data and anticipated missing data mechanism. Different analytical methods are capable of*

adapting to different levels of completeness of data, and choice of analytical method may also be a key driver in deciding analysis criteria for the study. Complete case analysis is likely to be appropriate if data are likely to be missing completely at random, or if a very high percentage of the target population have sufficient data for analysis. Linear mixed models may be appealing when studying changes in eGFR over time due to low thresholds for data completeness and flexibility in the presence of irregular eGFR testing, but rely on assumptions that missing data are similar to non-missing data (which may be conditional on observed characteristics), i.e. missingness at random. Methods accounting for informative drop-outs (e.g. joint longitudinal survival models, competing risks models) may be appropriate where data are missing not at random, e.g. due to loss to follow-up of patients referred to secondary care, or competing risks of death. Assumptions about the missing data mechanism should be reported alongside statistical methodology and data completeness criteria.

EHR resources are likely to expand and improve in future

The feasibility study (paper 2) demonstrated high frequency of eGFR testing, evaluated retrospectively in patients with risk factors at data extraction, collected over a median period of approximately 6 years. Data capture (repeat eGFR testing) was excellent in patients with CKD. Furthermore, issues around accuracy of slope estimation which occurred due to historically uncalibrated creatinine results are likely to diminish over time if new data are studied (post-2012).

Future studies aiming to study CKD progression using different primary care databases with longitudinal data capture (such as CPRD) are likely to have even higher data completeness than this study. That is because the current study identified patients with CKD at a set time point and looked back at historically collected eGFR data (when CKD may or may not have been present). A cohort study which identifies CKD at baseline and follows patients forward for eGFR-based outcomes is likely to have higher frequency of eGFR tests, due to patients with established CKD likely having more regular monitoring. eGFR slope estimation is also more accurate in patients with established CKD, due to improved accuracy of estimation at lower levels of eGFR.

A challenge not faced in this PhD is changes in data collection behaviours due to the COVID pandemic. Data captured from 2020 onwards may be informatively missing for different

reasons than previously collected data. For example, patients who deem themselves to be at high risk (e.g. due to diabetes, obesity, existing CKD, or general health anxiety) may not seek care or take up recommended blood tests if they are worried about attending medical appointments due to close contact with sick patients. However, this should subside over time.

EHR resources maintained for research purposes such as CPRD are growing to support enhanced research, in terms of coverage of the underlying population, total volume of historical data, and with expanding linkages to other relevant data (such as that relevant to COVID-19) [48,101-102]. EHR resources are therefore likely to improve as time moves forward. More data will gradually be accumulated which can be used to study CKD progression, leading to more patients who can be analysed and longer durations of follow-up. If testing for CKD improves, this will also enhance data availability for studying CKD progression.

Importance of laboratory data for identifying CKD

In the UK, electronic coding terminologies (Read or SNOMED) are used to electronically code diagnoses in primary care [50,103]. However, CKD is under-coded. In the NCKDA, approximately 30% of confirmed CKD cases 3-5 were not coded [16].

While the coding of diagnoses in clinical practice offers great advantages for large-scale research using EHRs, this evidence of lack of coding for CKD in clinical practice highlights the importance of defining CKD based on laboratory test results (including chronicity criteria), to correctly identify CKD cases in research studies using EHRs.

Further research to support implementation of risk prediction strategies

There is a strong basis for further research targeting risk prediction strategies for use in the study of CKD progression, given the importance of disease progression and need for improved individualised care. Guidelines have been developed to support implementation of research studies developing, validating or updating risk prediction models, namely the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) initiative guidelines [104]. These guidelines have been used in this thesis to support risk prediction model development and internal validation, and should be used to support any future risk prediction model development, validation or updating. Similarly to RECORD guidelines [100], the TRIPOD guidelines focus on transparent reporting of key aspects of study conduct and methodology which are likely to impact

reliability of research findings. Improvements in transparent reporting are likely to enable appropriate synthesis of evidence and critical appraisal by the research community, healthcare providers and policy makers, which may lead to improved acceptability of prediction models for use in decision making about prioritisation of care and resource use in clinical practice [105].

Validation in UK. As we have stated, our new risk prediction equations for kidney failure requiring KRT have not been validated in the UK. As has been done for KFRE [80], risk prediction equations need to be validated in populations of intended use, to ensure that they are correctly calibrated for the target population, with re-calibration if necessary. We stated earlier *(Chapter 1, Section 1.3)* that CKD staging criteria (with "traffic lights" for risk status) based on eGFR and uACR are not sufficient as a stand-alone framework for risk stratification for CKD due to significant variability of observed risks within risk categories. Equations such as ours are therefore needed to expand capabilities for risk prediction to the entire CKD population.

Understanding predictive mechanisms, with patient scenarios. Further scrutiny of the mechanisms of our new risk prediction equation may also be beneficial to support implementation. It would be valuable to explore further the predictive value of each predictor in our equation (i.e. which factors hold the most weight in risk prediction). This may simply involve creating (realistic) hypothetical patients with specified predictor values, and exploring the predictive impact of changing each variable, or it could involve sampling typical patients from a database (in which the equation has been validated) and comparing predicted risks between patients.

Model parsimony. It would be useful to explore removing predictors from of our equation, to investigate if a similar level of discrimination can be achieved with reduced data requirements. For example, removing eGFR slopes from the model may be beneficial, where accuracy of slope estimation may vary between patients depending on frequency of historical testing, or may limit implementation due to data requirements.

Further consideration of competing events in risk prediction modelling

Approaches used in previous studies. There are different approaches to statistical methodology for risk prediction of kidney failure in the literature. Development of the KFRE [79] and subsequent validation [80] was based on Cox models with censoring for death (the

same approach as ours). However, studies of more recently developed models argue that competing risks models more accurately predict absolute risks on average [106]. For example, KDpredict provides estimates of risks of both death and kidney failure requiring KRT (accounting for competing risk of death) [106]. This approach has the effect of lowering predicted risks of kidney failure compared to standard models due to any prior death precluding the occurrence of kidney failure.

Rationale for different approaches. While there is no right or wrong modelling approach, it is important to understand that each method has a different interpretation which may hold different value depending on what the results of individual predictions will be used for. This may include: (1) communication of risks to patients; (2) decision-making on treatments to reduce risks of adverse outcomes and/or referral to specialist care; and (3) planning for dialysis initiation. It is important that the end-user of prediction models understands the interpretation of predicted risks, when making clinical decisions.

Value in clinical practice. The use of KDpredict may support clinicians in communicating risks to patients which better reflect the reality of absolute risks faced for an individual (including death and pre-death KRT). This may be helpful for unnecessarily burdening patients with information on high risks of kidney failure, when they are significantly more likely to die before the risk is realised. (This may especially be the case for elderly frail patients who are unlikely to take up KRT.) However, we are concerned that the competing risks approach may "accidentally" infer reduced clinical importance of managing patients who are at high risk of kidney failure, but higher risks of prior death, where many patients stand to benefit from further care. (This may especially be the case in younger patients, who may have a high risk of CKD-related death, but would not wish for a high risk of death to disadvantage access to further renal care to reduce risks of kidney failure requiring KRT.)

Future studies should reflect on the implications of choice of modelling strategies, regarding incorporation of methodologies for competing risks, with full transparency of the interpretation of results, and relevance to patients and clinical practice.

Examining multiplicity in repeated use of risk prediction models, and possibility of unnecessary referrals

Risks of multiplicity. Equations predicting risk of kidney failure are generally developed using data which assesses risk of kidney failure at a single time point with follow-up until the

outcome of interest occurs. In clinical care, patients will have measurements which can be used to assess risks captured repeatedly over time and are likely to be referred (if clinical policy recommends) as soon as data supports referral. In statistics, the concept of multiplicity indicates that the more frequently you look at data, the more likely it is to observe more extreme results by chance simply due to random variation in observed data, and therefore the more likely you are to falsely conclude presence of an association when one does not exist.

Potential for over-referral. This statistical phenomenon deserves consideration when planning clinical referral decisions based on repeatedly observed kidney function test results. According to our prediction equation, a 5 unit increase in eGFR at baseline leads to a 0.53-times reduction in hazard of KRT. However, eGFR is a "noisy" measure. The likely consequence of this is that some patients who are referred based on risk thresholds (e.g. >5%) from prediction models will not represent on average those intended by developed prediction models with certain thresholds of predicted risks (and associated true risk of outcomes), and there may be over-referral, where average risks of those referred are likely to be lower on average than expected.

Clinical implications. The implications of this are unlikely to be grave, where clinical risk thresholds are set by a healthcare system with limits of uncertainty tolerated and indeed expected in clinical decision making. However, it may be a factor to consider in resource allocation and defining referral thresholds. The requirement to ensure chronicity of reduced kidney function (confirming CKD) in patients who are referred will hopefully limit the impact of unnecessary referrals, as patients are confirmed to have established kidney disease, using repeat measures.

Further research. Further research could evaluate the potential impact of multiplicity using the dataset used to develop/validate our equation. For example, we could compute predicted risks at every consecutive eGFR test in the database for each individual to identify when referral criteria are met, and subsequently evaluate if observed and predicted risks remain well aligned in patients evaluated at the time of identification for referral.

More research on utility of prediction models is needed if implemented in clinical practice. This could include comparing observed risks of kidney failure for patients meeting clinical criteria for referral, by whether patients were referred, with matching based on risk factors or using propensity score methods. It will take considerable time to accumulate this data, following clinical implementation.

Extensions of risk prediction strategies

Models which exploit patient clustering. Our risk prediction model and similar models such as KFRE are based on the logic that each predictor has an independent effect on the observed risk of kidney failure. When based on a Cox model, such as ours, it is assumed that a unit change in a predictor variable has a multiplicative impact on the hazard of the outcome, at any given moment. This assumption may or may not be appropriate or "ideal". However, it holds true that discrimination of our model is high, and therefore the model is a useful and reliable tool for risk stratification in patients studied to date. Regardless of this, it would be interesting for further work to consider studying patients by clusters of risk factors or progression trajectories. If patients can be classified into clusters, then this may enable improved prediction conditional on cluster status. This may involve, for example, latent class mixed models, joint latent class mixed models, or machine learning methods. Such an approach may have particular utility in the study of CKD due to known heterogeneity of the condition due to various causes.

Risk prediction for other outcomes associated with CKD progression

Various risk prediction models have been developed for use in prevention and management of CKD [107], including those which predict onset of eGFR <60, 40% decline in eGFR, kidney failure, CV events and mortality. However, these models cannot be implemented in routine care unless they are first validated in UK data and surpass requirements for integration into clinical practice.

Other research priorities to support early diagnosis and management of CKD

Early diagnosis and management are key to tackling the burden of CKD and delaying progression. Further research which reflects the general CKD population (supported by use of EHRs) can support this aim. This may include:

- new studies to provide up-to-date estimates of the incidence and prevalence of CKD (and stages of disease), and population risks of (appropriately defined, clinically important) CKD progression, describing characteristics of suffering patients or clusters of patients
  to improve understanding of underlying disease burden which may enable improved resource allocation; to identify any patient groups which may be overlooked in current care practice

- comparison studies for risks of CKD progression comparing diagnosed and undiagnosed CKD patients
  to identify any patient groups which may be overlooked in current care practice; to provide evidence to support new CKD detection and treatment initiatives

- further research on risk factors for CKD, or clusters of risk factors associated with CKD, and associated CKD-related adverse outcomes
  to support individualised care for CKD; to support hypotheses on biological mechanisms of disease which may aid identification (or development) of new treatments for CKD

- further research on ideal definitions for progression of CKD which represent clinically important decline and can be evaluated within the limitations of EHR data sources
  to support researchers using EHRs to study CKD and its progression to identify relevant exposures and outcomes for research studies, that will ensure current research is relevant to support clinical care; to improve opportunities for future harmonisation of evidence through systematic reviews and meta-analyses

8.7 Personal learning

Working as a statistician supporting reporting of the National Chronic Kidney Disease Audit inspired me to pursue this PhD research. It has been a steep learning curve to step away from being a relatively small cog in a research team carrying out mostly day to day analysis and reporting tasks, to being an independent researcher, conducting entire research projects and writing up results for publication, supported by the expert knowledge and guidance of my supervisors.

Key achievements have included the following:

(1) I have published 3 research papers in respectable journals, with a fourth paper ready to submit. I have become accustomed to the sometimes-frustrating back-and-forth processes of submitting to journals, updating work, and delays in publishing research which was exacerbated by the COVID pandemic.

(2) I have led a fruitful international collaboration with the Karolinska Institute, which included liaising with supervisors and collaborators to agree a research strategy, setting up contracts alongside legal teams, familiarising myself with a never before used database, and conducting analyses independently via remote access.

(3) I undertook a large systematic review of previous research, which required screening of 731 article abstracts and 210 full-texts, and full review of 80 research articles. With the expansion in volume of research in recent years and delays in review of my initial article submission due to the COVID pandemic, by the time my article had been reviewed it was out of date and I was required to subsequently review a large number of further articles before acceptance for publication. The process of conducting the systematic review involved a big learning curve, ensuring that correct processes were followed in identifying relevant research and registering the review at initiation, requiring the skill to scrutinise previous research and extract appropriate data with only basic clinical knowledge, and finally delivering a large body of work, backed up by a huge data extraction spreadsheet, with significant insights into previous work.

(4) Work that I completed on the association between GP practice CKD coding and patient outcomes has helped to inspire further initiatives enforcing recommendations for CKD coding in routine care (CVD Prevention Audit, Renal Services Transformation Program)

(5) During the course of my PhD, I have presented my research locally at LSHTM, nationally at UK Kidney Week and internationally at the European Renal Association and European Dialysis and Transplantation Association (ERA EDTA) conference.

Looking back at the commencement of my PhD studies, I sometimes lacked direction and the confidence to steam ahead. I feel that as I have developed my research skills and knowledge, my confidence has grown significantly, which has enabled me to thrive in my research. I am someone that likes to challenge intricately the things I hear and learn, and am constantly striving to improve my knowledge and understanding, learning from those I am working with and from independent research.

I am now excited for a future research career in statistics, epidemiology and further use of EHRs for research, which will build on the knowledge and skills I have developed during my PhD research.

8.8 Concluding remarks

CKD is a largely preventable and treatable disease, with most cases caused by diabetes and hypertension, which occur predominantly due to unhealthy lifestyles. While early-stage disease tends to be asymptomatic, the consequences of disease progression are undeniable, and increased acknowledgement and action is required to address the growing healthcare burden of CKD. Patient awareness is poor, and better engagement of patients with healthcare services is needed. This includes testing for those at risk, recognition of established diagnoses which is actively communicated to patients, a rigorous individualised monitoring plan, and appropriate use of pharmaceutical interventions and lifestyle modification to delay disease progression, prevent adverse outcomes and reduce the burden on healthcare services.

More research is needed to understand the burden, risk factors and consequences of CKD and its progression in the general population, with a particular focus on strengthening individualisation of care. Prediction tools such as those developed in this PhD and similar tools which have already been developed have potential for use in clinical care, if they are appropriately validated and integrated into clinical guidelines and software systems. Historically, focus has been on end-stage disease requiring dialysis, due to its clear clinical importance. However, enhancing capabilities to study progression at earlier stages of disease is also important, given the marked exposure-response relationships evident between degree of kidney damage (measured by eGFR and uACR) and important adverse outcomes (CV risk, mortality, CKD progression, AKI) preceding kidney failure, and the need to identify CKD progression earlier in the disease course.

We have demonstrated the size, breadth and (potential) generalisability of routinely collected electronic healthcare records that are available to study CKD and its progression, and the benefits of these factors to observational research. Repeat testing of eGFR is captured in the majority of patients with both risk factors for CKD and established disease, and we hope that repeat testing will continue to improve, particularly in patients who are most deprived. At the same time, EHR resources such as CPRD are expanding in population coverage and availability of data linkages to increase capture of data collected in different parts of the healthcare system. If adherence to eGFR testing guidelines improves, then date captured in EHRs will become more complete, and hopefully insights will follow. There is potential to perpetuate a cycle of continuous improvement in clinical care practices, increasing availability of relevant data for research, and applicable research findings.

Future research of CKD progression using EHRs needs to recognise the risks of selection bias that may result from informatively missing (and/or poor quality) data that is collected for purposes of clinical care. Studies must be designed with data quality and completeness in mind, ensuring that research questions can be answered with the available data. Statistical methods should be used to account for missing data, as appropriate, taking into account the likely mechanisms of missingness in the data, which are dependent on clinical context. Reporting must transparently present how patients are selected for analysis, how this impacts which patients are represented by analysis results, and whether estimated results could be biased in magnitude. RECORD guidelines are already available in the public domain to support researchers in conducting and reporting observational studies using EHRs. We also recommend familiarisation with co-authored work *(Appendix 2)* which lays out considerations for studies of CKD, as well as additional guidance for research considerations presented in *Appendix 3* of this thesis. Dealing with data that is missing from 2020 onwards presents new challenges for researchers seeking to study CKD and its progression, where the COVID pandemic has led to additional reasons for missingness and likely reduction in completeness of data in some patients.

References

1. Kovesdy CP. Epidemiology of chronic kidney disease: an update 2022. Kidney Int Suppl (2011). 2022;12(1):7-11.
2. Mills KT, Xu Y, Zhang W, Bundy JD, Chen C-S, Kelly TN, et al. A systematic analysis of worldwide population-based data on the global burden of chronic kidney disease in 2010. Kidney International. 2015;88(5):950-7.
3. GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet. 2015;385(9963):117-71.
4. Rhee CM, Kovesdy CP. Epidemiology: Spotlight on CKD deaths—increasing mortality worldwide. Nat Rev Nephrol. 2015;11(4):199-200.
5. NHS Digital. Health Survey for England 2016, Kidney and liver disease. 2017. Available from https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2016 [Accessed 1 June 2024]
6. Jager KJ, Fraser SDS. The ascending rank of chronic kidney disease in the global burden of disease study. Nephrology Dialysis Transplantation. 2017;32(suppl_2):ii121-ii8.
7. Kidney Research UK. Kidney disease: A UK public health emergency; The health economics of kidney disease to 2033. 2023. Available from https://www.kidneyresearchuk.org/wp-content/uploads/2023/06/Economics-of-Kidney-Disease-full-report_accessible.pdf [Accessed 1 June 2024]
8. Foreman KJ, Marquez N, Dolgert A, Fukutaki K, Fullman N, McGaughey M, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016-40 for 195 countries and territories. The Lancet. 2018;392(10159):2052-90.

9. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney Int. 2024;105(4S):S117–S314.

10. Kidney Research UK. Diabetes and kidney disease. Available from https://www.kidneyresearchuk.org/conditions-symptoms/diabetes/ [Accessed 1 June 2024]

11. Kidney Research UK. Blood pressure and kidney disease. Available from https://www.kidneyresearchuk.org/conditions-symptoms/blood-pressure/ [Accessed 1 June 2024]

12. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med. 2004;351(13):1296-305.

13. Dalrymple LS, Katz R, Kestenbaum B, de Boer IH, Fried L, Sarnak MJ, et al. The risk of infection-related hospitalization with decreased kidney function. Am J Kidney Dis. 2012;59(3):356-63.

14. UK Renal Registry. UK Renal Registry 25th Annual Report – data to 31/12/2021, Bristol, UK. 2023. Available from https://ukkidney.org/audit-research/annual-report [Accessed 1 June 2024]

15. Kerr M, Bray B, Medcalf J, O'Donoghue DJ, Matthews B. Estimating the financial cost of chronic kidney disease to the NHS in England. Nephrol Dial Transplant. 2012;27 Suppl 3(Suppl 3):iii73-80.

16. Nitsch D, Caplin B, Hull S, Wheeler D on behalf of the National CKD Audit and Quality Improvement Programme in Primary Care, First National CKD Audit Report 2017. Available from https://www.lshtm.ac.uk/files/ckd_audit_report.pdf [Accessed 1 June 2024]

17. Kidney Research UK. Chronic kidney disease and prevention: Harnessing the potential of early intervention and disease management. 2023. Available from https://www.kidneyresearchuk.org/wp-content/uploads/2023/03/Kidney-Research-report-on-prevention-and-early-detection.pdf [Accessed 1 June 2024]

18. National Institute for Health and Care Excellence. Chronic kidney disease: assessment and management. 2021. NICE guideline NG203. Available from https://www.nice.org.uk/guidance/ng203/resources/chronic-kidney-disease-assessment-and-management-pdf-66143713055173 [Accessed 1 June 2024]

19. Molokhia M, Okoli GN, Redmond P, Asgari E, Shaw C, Schofield P, et al. Uncoded chronic kidney disease in primary care: a cross-sectional study of inequalities and cardiovascular disease risk management. Br J Gen Pract. 2020;70(700):e785-e92.

20. Levin A, Ahmed SB, Carrero JJ, Foster B, Francis A, Hall RK, et al. Executive summary of the KDIGO 2024 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease: known knowns and known unknowns. Kidney Int. 2024;105(4):684-701.

21. Levin A, Stevens PE. Summary of KDIGO 2012 CKD Guideline: behind the scenes, need for guidance, and a framework for moving forward. Kidney Int. 2014;85(1):49-61.

22. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis. 2002;39(2 Suppl 1):S1-266.

23. Hallan SI, Orth SR. The KDOQI 2002 classification of chronic kidney disease: for whom the bell tolls. Nephrology Dialysis Transplantation. 2010;25(9):2832-6.

24. Levin A, Stevens PE, Bilous RW, Coresh J, De Francisco ALM, De Jong PE et al. Kidney disease: Improving global outcomes (KDIGO) CKD work group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney International Supplements. 2013 Jan 1;3(1):1-150.

25. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. Ann Intern Med. 1999;130(6):461-70.

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

27. Inker LA, Eneanya ND, Coresh J, Tighiouart H, Wang D, Sang Y, et al. New creatinine- and cystatin C–based equations to estimate GFR without race. New England Journal of Medicine. 2021;385(19)1737-1749

28. Kidney Research UK. Understanding test results. Available from https://www.kidneyresearchuk.org/kidney-health-information/living-with-kidney-disease/how-can-i-help-myself/understanding-test-results/ [Accessed 1 June 2024]

29. Kidney Research UK. Stages of kidney disease. Available from https://www.kidneyresearchuk.org/kidney-health-information/stages-of-kidney-disease/ [Accessed 1 June 2024]

30. Sumida K, Nadkarni GN, Grams ME, Sang Y, Ballew SH, Coresh J, et al. Conversion of Urine Protein-Creatinine Ratio or Urine Dipstick Protein to Urine Albumin-Creatinine Ratio for Use in Chronic Kidney Disease Screening and Prognosis : An Individual Participant-Based Meta-analysis. Ann Intern Med. 2020;173(6):426-35.

31. Kim LG, Cleary F, Wheeler DC, Caplin B, Nitsch D, Hull SA. How do primary care doctors in England and Wales code and manage people with chronic kidney disease? Results from the National Chronic Kidney Disease Audit. Nephrol Dial Transplant. 2018;33(8):1373-9.

32. Feakins B, Oke J, McFadden E, Aronson J, Lasserson D, O'Callaghan C, et al. Trends in kidney function testing in UK primary care since the introduction of the quality and outcomes framework: a retrospective cohort study using CPRD. BMJ Open. 2019;9(6):e028062.

33. Shin JI, Chang AR, Grams ME, Coresh J, Ballew SH, Surapaneni A, et al. Albuminuria Testing in Hypertension and Diabetes: An Individual-Participant Data Meta-Analysis in a Global Consortium. Hypertension. 2021;78(4):1042-52.

34. Elendu C, Elendu RC, Enyong JM, Ibhiedu JO, Ishola IV, Egbunu EO, et al. Comprehensive review of current management guidelines of chronic kidney disease. Medicine (Baltimore). 2023;102(23):e33984.

35. UK Kidney Association. The UK eCKD Guide. Updated Feb 2024. Available from https://ukkidney.org/health-professionals/information-resources/uk-eckd-guide#blood-pressure-management [Accessed 1 June 2024]

36. Lees JS, Welsh CE, Celis-Morales CA, Mackay D, Lewsey J, Gray SR, et al. Glomerular filtration rate by differing measures, albuminuria and prediction of cardiovascular disease, mortality and end-stage kidney disease. Nature Medicine. 2019;25(11):1753-60.

37. Gounden V, Bhatt H, Jialal I. Renal function tests. StatPearls [Internet]. 2024. Available from https://www.ncbi.nlm.nih.gov/books/NBK507821/ [Accessed 1 June 2024]

38. UK Kidney Association. The UK eCKD Guide. Updated Feb 2024. Available from https://ukkidney.org/health-professionals/information-resources/uk-eckd-guide#:~:text=The%20CKD%2DEPI%202021%20equation,practical%20considerations%20in%20interpreting%20eGFR. [Accessed 1 June 2024]

39. UK Kidney Association. The UK eCKD Guide. Updated Feb 2024. Available from https://ukkidney.org/health-professionals/information-resources/uk-eckd-guide/measurement-kidney-function#:~:text=Current%20NICE%20guidance%20recommends%20use,been%20validated%20in%20UK%20populations. [Accessed 1 June 2024]

40. Rowe C, Sitch AJ, Barratt J, Brettell EA, Cockwell P, Dalton RN, et al. Biological variation of measured and estimated glomerular filtration rate in patients with chronic kidney disease. Kidney International. 2019;96(2):429-35.

41. Organisation for Economic Co-operation and Development. Progress on implementing and using electronic health record systems. 2023. Available from https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/09/progress-on-implementing-and-using-electronic-health-record-systems_f6c2a59a/4f4ce846-en.pdf [Accessed 1 November 2024]

42. McMillan B, Eastham R, Brown B, Fitton R, Dickinson D. Primary care patient records in the United Kingdom: past, present, and future research priorities. J Med Internet Res 2018;20(12):e11293.

43. Derecho KC, Cafino R, Aquino-Cafino SL, Isla A, Esencia JA, Lactuan NJ, et al. Technology adoption of electronic medical records in developing economies: A systematic review on physicians' perspective. Digital Health. 2024;10:20552076231224605.

44. Akwaowo CD, Sabi HM, Ekpenyong N, Isiguzo CM, Andem NF, Maduka O, et al. Adoption of electronic medical records in developing countries - A multi-state study of the Nigerian healthcare system. Frontiers in Digital Health. 2022;4:1017231.

45. Department of Health & Social Care and NHS England. A plan for digital health and social care. 2022. Available from https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care/a-plan-for-digital-health-and-social-care [Accessed 1 November 2024]

46. World Health Organization. Global strategy on digital health 2020-2025. 2021. Available from https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [Accessed 1 November 2024]

47. Colombijn JMT, Idema DL, van Beem S, Blokland AM, van der Braak K, Handoko ML, et al. Representation of Patients With Chronic Kidney Disease in Clinical Trials of Cardiovascular Disease Medications: A Systematic Review. JAMA Network Open. 2024;7(3):e240427-e.

48. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. International Journal of Epidemiology. 2019;48(6):1740-g.

49. Carrero JJ, Elinder CG. The Stockholm CREAtinine Measurements (SCREAM) project: Fostering improvements in chronic kidney disease care. Journal of Internal Medicine. 2022;291(3):254-68.

50. CPRD. Defining your study population. Last reviewed 2023. Available from https://www.cprd.com/defining-your-study-population [Accessed 1 June 2024]

51. Little RJA, Rubin DB. Statistical analysis with missing data. Wiley series in probability and statistics. 3rd ed. Wiley, Hoboken, NJ: Wiley; 2020.

52. Persson R, Vasilakis-Scaramozza C, Hagberg KW, Sponholtz T, Williams T, Myles P, et al. CPRD Aurum database: Assessment of data quality and completeness of three important comorbidities. Pharmacoepidemiol Drug Saf. 2020;29(11):1456-64.

53. Tate AR, Dungey S, Glew S, Beloff N, Williams R, Williams T. Quality of recording of diabetes in the UK: how does the GP's method of coding clinical data affect

incidence estimates? Cross-sectional study using the CPRD database. BMJ Open. 2017;7(1):e012905.

54. Capita Health and Wellbeing. The quality of clinical coding in the NHS. 2014. Available from https://www.gov.uk/government/publications/payment-by-results-the-quality-of-clinical-coding-in-the-nhs [Accessed 1 June 2024]

55. Abeysekera RA, Healy HG, Wang Z, Cameron AL, Hoy WE. Heterogeneity in patterns of progression of chronic kidney disease. Internal Medicine Journal. 2021;51(2):220-8.

56. Jiang G, Luk AOY, Tam CHT, Xie F, Carstensen B, Lau ESH, et al. Progression of diabetic kidney disease and trajectory of kidney function decline in Chinese patients with Type 2 diabetes. Kidney International. 2019;95(1):178-87.

57. Coresh J, Turin TC, Matsushita K, Sang Y, Ballew SH, Appel LJ, et al. Decline in Estimated Glomerular Filtration Rate and Subsequent Risk of End-Stage Renal Disease and Mortality. JAMA. 2014;311(24):2518-31.

58. Turin TC, Coresh J, Tonelli M, Stevens PE, de Jong PE, Farmer CKT, et al. One-Year Change in Kidney Function Is Associated with an Increased Mortality Risk. American Journal of Nephrology. 2012;36(1):41-9.

59. Schnaper HW, Furth SL, Yao LP. Defining new surrogate markers for CKD progression. Pediatr Nephrol. 2015;30(2):193-8.

60. Coresh J, Turin TC, Matsushita K, Sang Y, Ballew SH, Appel LJ, et al. Decline in Estimated Glomerular Filtration Rate and Subsequent Risk of End-Stage Renal Disease and Mortality. JAMA. 2014;311(24):2518-31.

61. NHS Digital. Attribution Data Set GP-Registered Populations Scaled to ONS Population Estimates – 2011. 2012. Available from http://www.hscic.gov.uk/catalogue/PUB05054 [Accessed 1 June 2024]

62. Department of Health & Social Care. A plan for digital health and social care. 2022. Available from https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care/a-plan-for-digital-health-and-social-care [Accessed 1 June 2024]

63. Edwards L, Pickett J, Ashcroft DM, Dambha-Miller H, Majeed A, Mallen C, et al. UK research data resources based on primary care electronic health records: review and summary for potential users. BJGP Open. 2023;7(3):BJGPO.2023.0057.

64. McMillan B, Eastham R, Brown B, Fitton R, Dickinson D. Primary Care Patient Records in the United Kingdom: Past, Present, and Future Research Priorities. J Med Internet Res. 2018;20(12):e11293.

65. Medicines & Healthcare products Regulatory Agency. Hospital Episode Statistics (HES) Admitted Patient Care and CPRD primary care data documentation. 2021. Available from https://www.cprd.com/sites/default/files/2022-02/Documentation_HES_APC_set22.pdf [Accessed 1 June 2024]

66. NHS England. Commissioning for Quality and Innovation. https://www.england.nhs.uk/nhs-standard-contract/cquin/#:~:text=The%20Commissioning%20for%20Quality%20and,CQUIN%20scheme%20has%20been%20paused. [Accessed 1 June 2024]

67. The Health Improvement Network. Who we are. Available from https://www.the-health-improvement-network.com/ [Accessed 1 June 2024]

68. QResearch. Generating new knowledge to improve patient care. Available from https://www.qresearch.org/ [Accessed 1 June 2024]

69. Kerry-Barnard S, Damiani E, Gomes D, Calderwood L. Next Steps: A guide to the linked health administrative datasets – Hospital Episode Statistics (HES). 2020. London: UCL Centre for Longitudinal Studies. Available from https://doc.ukdataservice.ac.uk/doc/8681/mrdoc/pdf/nextsteps_hes_user_guide_v1.pdf [Accessed 1 June 2024]

70. UK Renal Registry. Acute kidney injury (AKI) in England – a report on the nationwide collection of AKI warning test scores from 2018. 2020. Available from https://ukkidney.org/sites/renal.org/files/AKI_report_FINAL_13072020.pdf [Accessed 1 June 2024]

71. Office for National Statistics. User guide to mortality statistics. Last revised 2023. Available from

https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/userguidetomortalitystatisticsjuly2017 [Accessed 1 June 2024]

72. Socialsystrelsen. About the Swedish healthcare system. Last updated 2020. Available from https://www.socialstyrelsen.se/en/about-us/healthcare-for-visitors-to-sweden/about-the-swedish-healthcare-system/ [Accessed 1 June 2024]

73. Runesson B, Gasparini A, Qureshi AR, Norin O, Evans M, Barany P, et al. The Stockholm CREAtinine Measurements (SCREAM) project: protocol overview and regional representativeness. Clin Kidney J. 2016;9(1):119-27.

74. Socialsystrelsen. National Prescribed Drug Register. Last updated 2024. Available from https://www.socialstyrelsen.se/en/statistics-and-data/registers/national-prescribed-drug-register/ [Accessed 1 June 2024]

75. National Institute for Health and Care Excellence. Clinical Knowledge Summaries; Chronic kidney disease: QOF indicators. Last updated 2024. Available from https://cks.nice.org.uk/topics/chronic-kidney-disease/goals-outcome-measures/qof-indicators/ [Accessed 1 June 2024]

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

77. Cholesterol Treatment Trialists' (CTT) Collaboration. Impact of renal function on the effects of LDL cholesterol lowering with statin-based regimens: a meta-analysis of individual participant data from 28 randomised trials. The Lancet Diabetes & Endocrinology. 2016;4(10):829-39.

78. The SPRINT Research Group. A randomized trial of intensive versus standard blood-pressure control. New England Journal of Medicine. 2015;373(22):2103-2116.

79. Tangri N, Stevens LA, Griffith J, Tighiouart H, Djurdjev O, Naimark D, et al. A Predictive Model for Progression of Chronic Kidney Disease to Kidney Failure. JAMA. 2011;305(15):1553-9.

80. Tangri N, Grams ME, Levey AS, Coresh J, Appel LJ, Astor BC, et al. Multinational Assessment of Accuracy of Equations for Predicting Risk of Kidney Failure: A Meta-analysis. Jama. 2016;315(2):164-74.

81. Shin J-I, Chang AR, Grams ME, Coresh J, Ballew SH, Surapaneni A, et al. Albuminuria Testing in Hypertension and Diabetes: An Individual-Participant Data Meta-Analysis in a Global Consortium. Hypertension. 2021;78(4):1042-52.

82. Van Calster B, Vickers AJ, Pencina MJ, Baker SG, Timmerman D, Steyerberg EW. Evaluation of markers and risk prediction models: Overview of relationships between NRI and decision-analytic measures. Medical Decision Making. 2013;33(4):490-501.

83. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. BMJ. 2016;352:i6.

84. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. American Journal of Epidemiology. 2017;186(9):1026-34.

85. Santhakumaran S, Fisher L, Zheng B, Mahalingasivam V, Plumb L, Parker EP, et al. Identification of patients undergoing chronic kidney replacement therapy in primary and secondary care data: validation study based on OpenSAFELY and UK Renal Registry. BMJ Medicine. 2024;3(1):e000807.

86. Levey AS, de Jong PE, Coresh J, El Nahas M, Astor BC, Matsushita K, et al. The definition, classification, and prognosis of chronic kidney disease: a KDIGO Controversies Conference report. Kidney International. 2011;80(1):17-28.

87. Carrero JJ, Fu EL, Vestergaard SV, Jensen SK, Gasparini A, Mahalingasivam V, et al. Defining measures of kidney function in observational studies using routine health care data: methodological and reporting considerations. Kidney International. 2023;103(1):53-69.

88. Fabian J, Kalyesubula R, Mkandawire J, Hansen CH, Nitsch D, Musenge E, et al. Measurement of kidney function in Malawi, South Africa, and Uganda: a multicentre cohort study. The Lancet Global Health. 2022;10(8):e1159-e69.

89. Savino M, Plumb L, Casula A, Evans K, Wong E, Kolhe N, et al. Acute kidney injury identification for pharmacoepidemiologic studies: Use of laboratory electronic acute kidney injury alerts versus electronic health records in Hospital Episode Statistics. Pharmacoepidemiol Drug Saf. 2021;30(12):1687-95.

90. Hallan SI, Dahl K, Oien CM, Grootendorst DC, Aasberg A, Holmen J, et al. Screening strategies for chronic kidney disease in the general population: follow-up of cross sectional health survey. BMJ. 2006;333(7577):1047.

91. Iwagami M, Tomlinson LA, Mansfield KE, Casula A, Caskey FJ, Aitken G, et al. Validity of estimated prevalence of decreased kidney function and renal replacement therapy from primary care electronic health records compared with national survey and registry data in the United Kingdom. Nephrol Dial Transplant. 2017;32(suppl_2):ii142-ii50.

92. Sullivan MK, Jani BD, Rutherford E, Welsh P, McConnachie A, Major RW, et al. Potential impact of NICE guidelines on referrals from primary care to nephrology: a primary care database and prospective research study. British Journal of General Practice. 2023;73(727):e141-e7.

93. Medicines and Healthcare products Regulatory Agency. Medical devices: software applications (apps). Last updated 2023. Available from https://www.gov.uk/government/publications/medical-devices-software-applications-apps [Accessed 1 June 2024]

94. Think Kidneys. Understanding what the public know about their kidneys and what they do. 2015. Available from https://www.thinkkidneys.nhs.uk/wp-content/uploads/2015/01/Think-Kidneys-Report-Understanding-what-the-public-know-Jan-2015-11.pdf [Accessed 1 June 2024]

95. Abdi Z, Gallagher H, O'Donoghue D. Telling the truth: why disclosure matters in chronic kidney disease. British Journal of General Practice. 2012;62(597):172-3.

96. Chu CD, Chen MH, McCulloch CE, Powe NR, Estrella MM, Shlipak MG, et al. Patient Awareness of CKD: A Systematic Review and Meta-analysis of Patient-Oriented Questions and Study Setting. Kidney Med. 2021;3(4):576-85.e1.

97. Kidney Research UK. Who we are. Available from https://www.kidneyresearchuk.org/about-us/who-we-are/ [Accessed 1 June 2024]

98. Kidney Care UK. Who we are and what we do. Available from https://kidneycareuk.org/about-us/who-we-are-and-what-we-do/ [Accessed 1 June 2024]

99. National Institute for Health and Care Research. PPI (Patient and Public Involvement) resources for applicants to NIHR research programmes. 2019. Available from https://www.nihr.ac.uk/documents/ppi-patient-and-public-involvement-resources-forapplicants-to-nihr-research-programmes/23437 [Accessed 1 June 2024]

100. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLoS Med. 2015;12(10): e1001885.

101. Medicines & Healthcare products Regulatory Agency. CPRD linked data. Last updated 2024. Available from https://www.cprd.com/cprd-linked-data [Accessed 1 June 2024]

102. Medicines & Healthcare products Regulatory Agency. Clinical Practice Research Datalink. Available from https://www.cprd.com/ [Accessed 1 June 2024]

103. NHS Digital. SNOMED CT. Last updated 2024. Available from https://digital.nhs.uk/services/terminology-and-classifications/snomed-ct [Accessed 1 June 2024]

104. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378.

105. Collins GS, Dhiman P, Ma J, Schlussel MM, Archer L, Van Calster B, et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ. 2024;384:e074819.

106. Liu P, Sawhney S, Heide-Jørgensen U, Quinn RR, Jensen SK, Mclean A, et al. Predicting the risks of kidney failure and death in adults with moderate to severe chronic kidney disease: multinational, longitudinal, population based, cohort study. BMJ. 2024;385:e078063.

107. Chronic Kidney Disease Prognosis Consortium. Risk models. Available from https://www.ckdpc.org/risk-models.html [Accessed 1 June 2024]

Appendix 1

Supplementary materials for research paper 1 (systematic review)

See next page for supplementary materials for research paper 1.

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5-6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6-7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supp. Info. S2
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7-8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8-10
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	9-10
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8-10
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	10

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	11
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	12
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Tables 1, 2, 3, Supp. Info. S3
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	15-23; Tables 2 and 3
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Tables 1, 2, 3 and 4, Table S4
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Tables 1, 2, 3 and 4
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	15-23; Tables 2 and 3
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Tables S1, S2, S3
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	30
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	30-32
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	32-33
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Included in PLOS ONE

			application process (funded by MRC PhD programme)

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Supporting Information S2. MEDLINE database search strategy

#	Terms
1	(CKD or CKF or CRD or CRF or "chronic kidney disease" or "kidney disease" or "renal disease" or "kidney failure" of "renal failure").mp.
2	exp Renal Insufficiency, Chronic/
3	1 or 2
4	("kidney function" or "renal function" or GFR or eGFR or mGFR or creatinine or "cystatin C" or "iohexol clearance" or MDRD or "CKD-EPI").mp
5	exp Glomerular Filtration Rate/ or exp Creatinine/ or exp Cystatin C/
6	4 or 5
7	((electronic or computer* or anonymi#ed) adj3 (health* or medical or patient* or GP* or "general practictioner" or "primary care" or "hospital" or "secondary care" or observational or routine*) adj3 (record* or data*)).mp
8	(("primary care" or GP or practice* or "secondary care" or link* or hospital* or clinic* or centre or center) adj3 (data* or record*)).mp
9	(EHR or CPR or EMR or EPR or AMR).mp
10	exp Electronic Health Records/ or exp Medical Record Linkage/ or exp Medical Records Systems, Computerized/
11	7 or 8 or 9 or 10
12	(chang* or declin* or progress* or longitudinal* or trajectory* or slope* or deteriorate* or loss*).mp
13	exp Longitudinal Studies/
14	12 or 13
15	((chang* or declin* or progress* or longitudinal* or trajectory* or slope* or deteriorate* or loss*) adj3 ("kidney function" or "renal function" or GFR or eGFR or mGFR or creatinine or "cystatin C" or "iohexol clearance" or MDRD or "CKD-EPI")).mp
16	5 and 13
17	15 or 16
18	3 and 11 and 17
19	exp clinical trial/ or exp case reports/ or ("clinical trial" or "randomi#ed trial" or RCT* or "case report*").m_titl.
20	("dialysis patient*" or "transplant patient*").m_titl.
21	19 or 20
22	18 not 21

Supporting Information S3. List of reviewed studies *(article numbers match references in manuscript [for articles 11-62] and follow from manuscript reference numbers for articles not referenced in manuscript [articles 69-96])*

11. Chase HS, Hirsch JS, Mohan S, Rao MK, Radhakrishnan J. Presence of early CKD-related metabolic complications predict progression of stage 3 CKD: a case-controlled study. BMC Nephrology. 2014;15:187.

12. Wang Y, Zhao L, Zhang J, Wu Y, Zhang R, Li H, et al. Implications of a Family History of Diabetes and Rapid eGFR Decline in Patients With Type 2 Diabetes and Biopsy-Proven Diabetic Kidney Disease. Frontiers in Endocrinology. 2019;10 (no pagination).

13. Abdelhafiz AH, Tan E, Levett C, Minchin J, Nahas ME. Natural history and predictors of faster glomerular filtration rate decline in a referred population of older patients with type 2 diabetes mellitus. Hospital practice (1995) Hospital practice. 2012;40(4):49-55.

14. Eriksen BO, Ingebretsen OC. The progression of chronic kidney disease: a 10-year population-based study of the effects of gender and age. Kidney International. 2006;69(2):375-82.

15. Jalal K, Anand EJ, Venuto R, Eberle J, Arora P. Can billing codes accurately identify rapidly progressing stage 3 and stage 4 chronic kidney disease patients: a diagnostic test study. Bmc Nephrology. 2019;20.

16. Cabrera CS, Lee AS, Olsson M, Schnecke V, Westman K, Lind M, et al. Impact of CKD Progression on Cardiovascular Disease Risk in a Contemporary UK Cohort of Individuals With Diabetes. Kidney International Reports. 2020;5(10):1651-60.

17. Eriksen BO, Tomtum J, Ingebretsen OC. Predictors of declining glomerular filtration rate in a population-based chronic kidney disease cohort. Nephron. 2010;115(1):c41-50.

18. Annor FB, Masyn KE, Okosun IS, Roblin DW, Goodman M. Psychosocial stress and changes in estimated glomerular filtration rate among adults with diabetes mellitus. Kidney Research and Clinical Practice. 2015;34(3):146-53.

19. Diggle PJ, Sousa I, Asar O. Real-time monitoring of progression towards renal failure in primary care patients. Biostatistics. 2015;16(3):522-36.

20. Butt AA, Ren Y, Puenpatom A, Arduino JM, Kumar R, Abou-Samra AB. Effectiveness, treatment completion and safety of sofosbuvir/ledipasvir and paritaprevir/ritonavir/ombitasvir + dasabuvir in patients with chronic kidney disease: an ERCHIVES study. Alimentary Pharmacology & Therapeutics. 2018;48(1):35-43.

21. Singh A, Nadkarni G, Gottesman O, Ellis SB, Bottinger EP, Guttag JV. Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration. Journal of Biomedical Informatics. 2015;53:220-8.

22. Evans RDR, Cargill T, Goodchild G, Oliveira B, Rodriguez-Justo M, Pepper R, et al. Clinical Manifestations and Long-term Outcomes of IgG4-Related Kidney and Retroperitoneal Involvement in a United Kingdom IgG4-Related Disease Cohort. Kidney International Reports. 2019;4(1):48-58.25. Vejakama P, Ingsathit A, Attia J, Thakkinstian A. Epidemiological Study of Chronic Kidney Disease Progression: A Large-Scale Population-Based Cohort Study. Medicine. 2015;94(4).

23. Jackevicius CA, Lu LY, Ghaznavi Z, Warner AL. Bleeding Risk of Direct Oral Anticoagulants in Patients With Heart Failure And Atrial Fibrillation. Circulation-Cardiovascular Quality and Outcomes. 2021;14(2):155-68.

24. Lai YJ, Lin YC, Peng CC, Chen KC, Chuang MT, Wu MS, et al. Effect of weight loss on the estimated glomerular filtration rates of obese patients at risk of chronic kidney disease: the RIGOR-TMU study. Journal of Cachexia, Sarcopenia and Muscle. 2019;10(4):756-66.

25. Vejakama P, Ingsathit A, Attia J, Thakkinstian A. Epidemiological Study of Chronic Kidney Disease Progression: A Large-Scale Population-Based Cohort Study. Medicine. 2015;94(4).

26. Posch F, Ay C, Stoger H, Kreutz R, Beyer-Westendorf J. Exposure to vitamin k antagonists and kidney function decline in patients with atrial fibrillation and chronic kidney disease. Research and Practice in Thrombosis and Haemostasis. 2019;3(2):207-16.

27. Hsu TW, Hsu CN, Wang SW, Huang CC, Li LC. Comparison of the effects of denosumab and alendronate on cardiovascular and renal outcomes in osteoporotic patients. Journal of Clinical Medicine. 2019;8(7).

28. Inaguma D, Kitagawa A, Yanagiya R, Koseki A, Iwamori T, Kudo M, et al. Increasing tendency of urine protein is a risk factor for rapid eGFR decline in patients with CKD: A machine learning-based prediction model by using a big database. PLoS ONE [Electronic Resource]. 2020;15(9):e0239262.

29. Peng YL, Tain YL, Lee CT, Yang YH, Huang YB, Wen YH, et al. Comparison of uric acid reduction and renal outcomes of febuxostat vs allopurinol in patients with chronic kidney disease. Scientific Reports. 2020;10(1):10734.

30. Yao X, Tangri N, Gersh BJ, Sangaralingham LR, Shah ND, Nath KA, et al. Renal Outcomes in Anticoagulated Patients With Atrial Fibrillation. Journal of the American College of Cardiology. 2017;70(21):2621-32.36. Lv L, Chang DY, Li ZY, Chen M, Hu Z, Zhao MH. Persistent hematuria in patients with antineutrophil cytoplasmic antibody-associated vasculitis during clinical remission: chronic glomerular lesion or low-grade active renal vasculitis? BMC Nephrology. 2017;18(1):354.

31. Lamacchia O, Viazzi F, Fioretto P, Mirijello A, Giorda C, Ceriello A, et al. Normoalbuminuric kidney impairment in patients with T1DM: Insights from annals initiative. Diabetology and Metabolic Syndrome. 2018;10(1).

32. Viazzi F, Greco E, Ceriello A, Fioretto P, Giorda C, Guida P, et al. Apparent treatment resistant hypertension, blood pressure control and the progression of chronic kidney disease in patients with type 2 diabetes. Kidney and Blood Pressure Research. 2018;43(2):422-38.

33. Rej S, Herrmann N, Gruneir A, McArthur E, Jeyakumar N, Muanda FT, et al. Association of Lithium Use and a Higher Serum Concentration of Lithium With the Risk of Declining Renal Function in Older Adults: A Population-Based Cohort Study. The Journal of clinical psychiatry. 2020;81(5).

34. Yoo H, Park I, Kim DJ, Lee S. Effects of sarpogrelate on microvascular complications with type 2 diabetes. International Journal of Clinical Pharmacy. 2019.

35. Tangri N, Reaven NL, Funk SE, Ferguson TW, Collister D, Mathur V. Metabolic acidosis is associated with increased risk of adverse kidney outcomes and mortality in patients with non-dialysis dependent chronic kidney disease: an observational cohort study. BMC Nephrology. 2021;22(1).

36. Lv L, Chang DY, Li ZY, Chen M, Hu Z, Zhao MH. Persistent hematuria in patients with antineutrophil cytoplasmic antibody-associated vasculitis during clinical remission: chronic glomerular lesion or low-grade active renal vasculitis? BMC Nephrology. 2017;18(1):354.

37. Li XM, Rui HC, Liang DD, Xu F, Liang SS, Zhu XD, et al. Clinicopathological characteristics and outcomes of light chain deposition disease: an analysis of 48 patients in a single Chinese center. Annals of Hematology. 2016;95(6):901-9.

38. Gallant JE, Parish MA, Keruly JC, Moore RD. Changes in renal function associated with tenofovir disoproxil fumarate treatment, compared with nucleoside reverse-transcriptase inhibitor treatment. Clinical Infectious Diseases. 2005;40(8):1194-8.

39. Herget-Rosenthal S, Dehnen D, Kribben A, Quellmann T. Progressive chronic kidney disease in primary care: modifiable risk factors and predictive model. Preventive Medicine. 2013;57(4):357-62.

40. Morales-Alvarez MC, Garcia-Dolagaray G, Millan-Fierro A, Rosas SE. Renal Function Decline in Latinos With Type 2 Diabetes. Kidney International Reports. 2019;4(9):1230-4.

41. Nderitu P, Doos L, Strauss VY, Lambie M, Davies SJ, Kadam UT. Analgesia dose prescribing and estimated glomerular filtration rate decline: a general practice database linkage cohort study. BMJ Open. 2014;4(8):e005581.

42. Koraishy FM, Hooks-Anderson D, Salas J, Scherrer JF. Rate of renal function decline, race and referral to nephrology in a large cohort of primary care patients. Family Practice. 2017;34(4):416-22.

43. Johnson F, Phillips D, Talabani B, Wonnacott A, Meran S, Phillips AO. The impact of acute kidney injury in diabetes mellitus. Nephrology. 2016;21(6):506-11.44. Chakera A, MacEwen C, Bellur SS, Chompuk LO, Lunn D, Roberts ISD. Prognostic value of endocapillary hypercellularity in IgA nephropathy patients with no immunosuppression. Journal of Nephrology. 2016;29(3):367-75.

44. Chakera A, MacEwen C, Bellur SS, Chompuk LO, Lunn D, Roberts ISD. Prognostic value of endocapillary hypercellularity in IgA nephropathy patients with no immunosuppression. Journal of Nephrology. 2016;29(3):367-75.

45. Chen H, Liu C, Fu C, Zhang H, Yang H, Wang P, et al. Combined application of eGFR and albuminuria for the precise diagnosis of stage 2 and 3a CKD in the elderly. Journal of Nephrology. 2014;27(3):289-97.

46. Perotte A, Ranganath R, Hirsch JS, Blei D, Elhadad N. Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis. Journal of the American Medical Informatics Association. 2015;22(4):872-80.

47. Cummings DM, Larsen LC, Doherty L, Lea CS, Holbert D. Glycemic Control Patterns and Kidney Disease Progression among Primary Care Patients with Diabetes Mellitus. Journal of the American Board of Family Medicine. 2011;24(4):391-8.

48. Horne L, Ashfaq A, MacLachlan S, Sinsakul M, Qin L, LoCasale R, et al. Epidemiology and health outcomes associated with hyperkalemia in a primary care setting in England. BMC Nephrology. 2019;20(1):85.

49. Robinson DE, Ali MS, Pallares N, Tebe C, Elhussein L, Abrahamsen B, et al. Safety of Oral Bisphosphonates in Moderate-to-Severe Chronic Kidney Disease: A Binational Cohort Analysis. Journal of Bone and Mineral Research. 2021;36(5):820-32.

50. Nichols GA, Deruaz-Luyet A, Brodovicz KG, Kimes TM, Rosales AG, Hauske SJ. Kidney disease progression and all-cause mortality across estimated glomerular filtration rate and albuminuria categories among patients with vs. Without type 2 diabetes. BMC Nephrology. 2020;21(1).

51. Yanagawa T, Koyano K, Azuma K. Retrospective study of factors associated with progression and remission/regression of diabetic kidney disease-hypomagnesemia was associated with progression and elevated serum alanine aminotransferase levels were associated with remission or regression. Diabetology International. 2021;12(3):268-76.

52. Vesga JI, Cepeda E, Pardo CE, Paez S, Sanchez R, Sanabria RM. Chronic kidney disease progression and transition probabilities in a large preventive cohort in colombia. International Journal of Nephrology. 2021;2021 (no pagination).

53. Oetjens M, Bush WS, Birdwell KA, Dilks HH, Bowton EA, Denny JC, et al. Utilization of an EMR-biorepository to identify the genetic predictors of calcineurin-inhibitor toxicity in heart transplant recipients. Pacific Symposium on Biocomputing. 2014:253-64.

54. Neuen BL, Weldegiorgis M, Herrington WG, Ohkuma T, Smith M, Woodward M. Changes in GFR and Albuminuria in Routine Clinical Practice and the Risk of Kidney Disease Progression. American Journal of Kidney Diseases. 2021.

55. Weldegiorgis M, Smith M, Herrington WG, Bankhead C, Woodward M. Socioeconomic disadvantage and the risk of advanced chronic kidney disease: results from a cohort study with 1.4 million participants. Nephrology Dialysis Transplantation. 2020;35(9):1562-70.

56. Niu SF, Wu CK, Chuang NC, Yang YB, Chang TH. Early Chronic Kidney Disease Care Programme delays kidney function deterioration in patients with stage I-IIIa chronic kidney disease: an observational cohort study in Taiwan. Bmj Open. 2021;11(1).

57. O'Riordan A, Dutt N, Cairns H, Rela M, O'Grady JG, Heaton N, et al. Renal biopsy in liver transplant recipients. Nephrology Dialysis Transplantation. 2009;24(7):2276-82.

58. Tsai CW, Lin SY, Kuo CC, Huang CC. Serum Uric Acid and Progression of Kidney Disease: A Longitudinal Analysis and Mini-Review. PLoS ONE [Electronic Resource]. 2017;12(1):e0170393.

59. Leither MD, Murphy DP, Bicknese L, Reule S, Vock DM, Ishani A, et al. The impact of outpatient acute kidney injury on mortality and chronic kidney disease: a retrospective cohort study. Nephrology Dialysis Transplantation. 2019;34(3):493-501.

60. Liu D, You J, Liu Y, Tang X, Tan X, Xia M, et al. Serum immunoglobulin G provides early risk prediction in immunoglobulin A nephropathy. International Immunopharmacology. 2019;66:13-8.

61. Rincon-Choles H, Jolly SE, Arrigain S, Konig V, Schold JD, Nakhoul G, et al. Impact of Uric Acid Levels on Kidney Disease Progression. American Journal of Nephrology. 2017;46(4):315-22.69. Joss N, Paterson KR, Deighan CJ, Simpson K, Boulton-Jones JM. Diabetic nephropathy: how effective is treatment in clinical practice? Qjm. 2002;95(1):41-9.

62. VanWagner LB, Montag S, Zhao L, Allen NB, Lloyd-Jones DM, Das A, et al. Cardiovascular Disease Outcomes Related to Early Stage Renal Impairment After Liver Transplantation. Transplantation. 2018;102(7):1096-107.

69. Joss N, Paterson KR, Deighan CJ, Simpson K, Boulton-Jones JM. Diabetic nephropathy: how effective is treatment in clinical practice? Qjm. 2002;95(1):41-9.

70. Dean BB, Dylan M, Gano A, Knight K, Ofman JJ, Levine BS. Erythropoiesis-stimulating protein therapy and the decline of renal function: a retrospective analysis of patients with chronic kidney disease. Current Medical Research and Opinion. 2005;21(7):981-7.

71. Jones C, Roderick P, Harris S, Rogerson M. An evaluation of a shared primary and secondary care nephrology service for managing patients with moderate to advanced CKD. American Journal of Kidney Diseases. 2006;47(1):103-14.

72. Chen SC, Chang JM, Chou MC, Lin MY, Chen JH, Sun JH, et al. Slowing renal function decline in chronic kidney disease patients after nephrology referral. Nephrology. 2008;13(8):730-6.

73. Boudville N, Kemp A, Moody H, Fassett RG, Pedagogos E, Nelson C, et al. Factors associated with chronic kidney disease progression in Australian nephrology practices. Nephron - Clinical Practice. 2012;121(1-2):c36-c41.

74. Dreyer G, Hull S, Mathur R, Chesser A, Yaqoob MM. Progression of chronic kidney disease in a multi-ethnic community cohort of patients with diabetes mellitus. Diabetic Medicine. 2013;30(8):956-63.

75. Malgor RD, Oderich GS, Vrtiska TJ, Kalra M, Duncan AA, Gloviczki P, et al. A case-control study of intentional occlusion of accessory renal arteries during endovascular aortic aneurysm repair. Journal of Vascular Surgery. 2013;58(6):1467-75.

76. Brosnan EM, Weickhardt AJ, Lu X, Maxon DA, Baron AE, Chonchol M, et al. Drug-induced reduction in estimated glomerular filtration rate in patients with ALK-positive non-small cell lung cancer treated with the ALK inhibitor crizotinib. Cancer. 2014;120(5):664-74.

77. Kose E, An T, Kikkawa A, Matsumoto Y, Hayashi H. Effects on serum uric acid by difference of the renal protective effects with atorvastatin and rosuvastatin in chronic kidney disease patients. Biological & Pharmaceutical Bulletin. 2014;37(2):226-31.

78. Cid Ruzafa J, Paczkowski R, Boye KS, Di Tanna GL, Sheetz MJ, Donaldson R, et al. Estimated glomerular filtration rate progression in UK primary care patients with type 2 diabetes and diabetic kidney disease: a retrospective cohort study. International Journal of Clinical Practice. 2015;69(8):871-82.

79. Kaga M, Utsumi T, Tanaka T, Kono T, Nagano H, Kawamura K, et al. Risk of New-Onset Dyslipidemia After Laparoscopic Adrenalectomy in Patients with Primary Aldosteronism. World Journal of Surgery. 2015;39(12):2935-40.

80. Lai CL, Chou HW, Chan KA, Lai MS. Effects of atorvastatin and rosuvastatin on renal function in patients with type 2 diabetes mellitus. American Journal of Cardiology. 2015;115(5):619-24.

81. Yun WS. Long-Term Follow-Up Results of Acute Renal Embolism after Anticoagulation Therapy. Annals of Vascular Surgery. 2015;29(3).

82. Kim YG, Byun J, Yoon D, Jeon JY, Han SJ, Kim DJ, et al. Renal Protective Effect of DPP-4 Inhibitors in Type 2 Diabetes Mellitus Patients: A Cohort Study. Journal of Diabetes Research. 2016.

83. Mirajkar N, Bellary S, Ahmed M, Singhal R, Daskalakis M, Tahrani AA. The impact of bariatric surgery on estimated glomerular filtration rate in patients with type 2 diabetes: a retrospective cohort study. Surgery for Obesity & Related Diseases. 2016;12(10):1883-9.

84. Nishida Y, Takahashi Y, Tezuka K, Takeuchi S, Nakayama T, Asai S. Comparative Effect of Calcium Channel Blockers on Glomerular Function in Hypertensive Patients with Diabetes Mellitus. Drugs in R&D. 2017;17(3):403-12.

85. Beyer-Westendorf J, Kreutz R, Posch F, Ay C. The CHA2DS2VASc score strongly correlates with glomerular filtration rate and predicts renal function decline over time in elderly patients with atrial fibrillation and chronic kidney disease. International Journal of Cardiology. 2018;253:71-7.

86. Kim WJ, Song JS, Choi ST. The role of a "treat-to-target" approach in the long-term renal outcomes of patients with gout. Journal of Clinical Medicine. 2019;8(7).

87. O'Neill RA, Gallagher P, Douglas T, Little JA, Maxwell AP, Silvestri G, et al. Evaluation of long-term intravitreal anti-vascular endothelial growth factor injections on renal function in patients with and without diabetic kidney disease. BMC Nephrology. 2019;20(1).

88. Park JM, Yang SW, Shin JH, Na YG, Song KH, Lim JS. Oncological and Functional Outcomes of Laparoscopic Radiofrequency Ablation and Partial Nephrectomy for T1a Renal Masses: A Retrospective Single-center 60 Month Follow-up Cohort Study. Urology Journal. 2019;16(1):44-9.

89. Posch F, Ay C, Stoger H, Kreutz R, Beyer-Westendorf J. Longitudinal kidney function trajectories predict major bleeding, hospitalization and death in patients with atrial fibrillation and chronic kidney disease. International Journal of Cardiology. 2019;282:47-52.

90. Spanopoulos D, Okhai H, Zaccardi F, Tebboth A, Barrett B, Busse M, et al. Temporal variation of renal function in people with type 2 diabetes mellitus: A retrospective UK clinical practice research datalink cohort study. Diabetes, Obesity and Metabolism. 2019;21(8):1817-23.

91. Zhao J, Gu S, McDermaid A. Predicting outcomes of chronic kidney disease from EMR data based on Random Forest Regression. Mathematical Biosciences. 2019;310:24-30.

92. Lee JS, Oh JS, Kim YG, Lee CK, Yoo B, Hong S. Recovery of renal function in patients with lupus nephritis and reduced renal function: the beneficial effect of hydroxychloroquine. Lupus. 2020;29(1):52-7.

93. Nakamura A, Miyoshi H, Kameda H, Yamashita K, Kurihara Y. Impact of sodium-glucose cotransporter 2 inhibitors on renal function in participants with type 2 diabetes and chronic kidney disease with normoalbuminuria. Diabetology and Metabolic Syndrome. 2020;12(1).

94. Sise ME, Chute DF, Oppong Y, Davis MI, Long JD, Silva ST, et al. Direct-acting antiviral therapy slows kidney function decline in patients with Hepatitis C virus infection and chronic kidney disease. Kidney International. 2020;97(1):193-201.

95. Cleary F, Prieto-Merino D, Hull S, Caplin B, Nitsch D. Feasibility of evaluation of the natural history of kidney disease in the general population using electronic healthcare records. Clinical Kidney Journal. 2020;14(6):1603-9.

96. Faraj KS, Mi L, Eversman S, Singh R, DeLucia NM, Blodgett G, et al. The effect of urinary diversion on long-term kidney function after cystectomy. Urologic Oncology: Seminars and Original Investigations. 2020;38(10):796.e15-.e21.

Table S1. Summary of study populations, where unclear if EHRs used (N = 10)

Study population characteristics	N (%)
Primary decade of follow up	
2010-2019	3 (30.0%)
2000-2009	6 (60.0%)
1990-1999	0
Not available	1 (10.0%)
Country	
Europe	**1 (10.0%)**
UK	1 (10.0%)
North America	**1 (10.0%)**
USA	1 (10.0%)
Asia	**8 (80.0%)**
South Korea	2 (20.0%)
China	4 (40.0%)
Japan	2 (20.0%)
Oceania	**1 (1.5%)**
Mean age[a]	
Median (IQR)	59 (53, 68)
30-49	1 (10.0%)
50-59	5 (50.0%)
60-69	2 (20.0%)
70-80	2 (20.0%)
Not stated	0
Percent male	
Median (IQR)	57% (48%, 79%)
≤ 34%	0
35-44%	2 (20.0%)
45-54%	3 (30.0%)
55-64%	1 (10.0%)
≥ 65%	4 (40.0%)
Main morbidity /reason for inclusion	
Diabetes	1 (10.0%)
CKD	1 (10.0%)
IgA nephropathy	1 (10.0%)
Other	7 (70.0%)
Data source / clinical setting	
Multiple care settings	1 (10.0%)
Outpatient	1 (10.0%)
Diabetes clinic	1 (9.2%)
Hospital	6 (60.0%)
Tertiary care	1 (10.0%)
Not stated	1 (10.0%)

[a]If mean age unavailable, median used.

Table S2. Study methodology, where unclear if EHRs used (N=10)

Study methodology features	N (%)
Date of publication	
2015-2020	7 (70.0%)
2010-2014	3 (30.0%)
Study design	
Retrospective cohort study	10 (100%)
Research aims	
Risk factor identification / casual inference	8 (80.0%)
Estimation of incidence/prevalence	1 (10.0%)
Descriptive characterisation of changes in renal function	1 (10.0%)
Sample size	
Median (IQR)	101 (41, 247)
≤ 99	5 (50.0%)
100 – 499	5 (50.0%)
Measure of renal function	
eGFR	**8 (80.0%)**
MDRD	3 (37.5%)
CKD-EPI	1 (12.5%)
Japanese formula	1 (12.5%)
Not specified	3 (37.5%)
Serum creatinine	**2 (20.0%)**
Measure of change in renal function over time	
eGFR	**8 (80.0%)**
Rate of change between measures	1 (10.0%)
Rate of percentage change, not clearly defined	1 (10.0%)
Raw absolute change from baseline	3 (30.0%)
Raw percent change from baseline	2 (20.0%)
Binary progression (changes/threshold combination)	1 (10.0%)
Serum creatinine	**2 (20.0%)**
Raw absolute change from baseline	1 (10.0%)
Binary progression to threshold serum creatinine	1 (10.0%)
Change in renal function as outcome or exposure	
Outcome	100 (100%)
Exposure	0
Duration of follow up for renal function changes	
1 – 4.9 years	8 (80.0%)
5 – 9.9 years	2 (20.0%)
Minimum number of renal function measures for inclusion	
2	4 (40.0%)
3	1 (10.0%)
Not stated	5 (50.0%)

Percentage of target population used in analysis	
50% - 75%	2 (20.0%)
75% - 90%	1 (10.0%)
90% - 95%	1 (10.0%)
>95%	1 (10.0%)
Not available	5 (50.0%)
Percentage of study population lost to follow up	
Not available or not relevant	10 (100%)
Statistical tools used[a]	
Descriptive results only	2 (20.0%)
Simple frequentist methods	3 (30.0%)
ANCOVA	1 (10.0%)
Generalised linear models (GLMs)	1 (10.0%)
Cox proportional hazards regression	3 (30.0%)
Statistical model used[a]	
Risk factor identification / casual inference	**N = 8**
Difference in means t-test	1 (10.0%)
Mean difference paired t-test	1 (10.0%)
ANOVA	1 (10.0%)
ANCOVA	1 (10.0%)
Logistic regression	1 (10.0%)
Cox proportional hazards regression	3 (30.0%)
Estimation of incidence/prevalence	**N = 1**
Crude estimation	1 (10.0%)

[a]Multiple items possible for a single study but focus only on main analysis of CKD progression

Table S3. Critique of handling of data quality and methodological challenges, where unclear if EHRs used (N = 10)

Handling of data quality and methodological challenges	N (%)
Handling of sample representativeness of target population used for analysis	
Not mentioned	3 (30.0%)
Mentioned care pathway and inclusion criteria, but not sample completeness	1 (10.0%)
Mentioned sample completeness, but not implications	0
Partially acknowledged implications of sample completeness	5 (50.0%)
Fully acknowledged implications of sample completeness	1 (10.0%)
Tackled methodologically	0
Methods of handling	
None	10 (100%)
Handling of informative drop-outs/censoring	
Not mentioned	8 (80.0%)
Mentioned care pathway follow up, but not losses to follow up (inc. death)	0
Mentioned losses to follow up, but not implications	2 (20.0%)
Partially acknowledged implications of losses to follow up	0
Fully Acknowledged implications of losses to follow up	0
Tackled methodologically	0
Methods of handling	
None	10 (100%)
Handling of missing longitudinal data	
Not mentioned	10 (100%)
Mentioned care pathway follow up, but not data completeness	0
Mentioned data completeness, but not implications	0
Partially acknowledged implications of data completeness	0
Fully acknowledged implications of data completeness	0
Tackled methodologically	0
Methods of handling	
None	10 (100%)
Handling of missing covariate data	
Not mentioned	6 (60.0%)
Mentioned data completeness, but not implications	0
Partially acknowledged implications of data completeness	1 (10.0%)
Fully acknowledged implications of data completeness	0
Tackled methodologically or not an issue	3 (30.0%)
Methods of handling	
None	10 (100%)
Distributional checks/issues	
Not mentioned	10 (100%)
Mentioned or partially addressed	0
Fully Acknowledged	0
Tackled	0
Methods of handling	
None	10 (100%)
Handling of within-patient correlation / variability in kidney function over time	

Not mentioned	7 (70.0%)
Mentioned or partially addressed	3 (30.0%)
Fully Acknowledged	0
Tackled	0
Methods of handling	
None	8 (80.0%)
Outcome likely to identify real change	2 (20.0%)
Handling of population heterogeneity	
Not mentioned	1 (10.0%)
Mentioned or partially addressed	5 (50.0%)
Fully Acknowledged	1 (10.0%)
Tackled	3 (30.0%)
Method of handling	
None	3 (30.0%)
Adjustment for covariates	1 (10.0%)
Stratified or separate/subgroup analysis	5 (50.0%)
ANOVA/ANCOVA	1 (10.0%)
Handling of confounding (risk factor / causal inference analyses only)	N = 8
Not mentioned	4 (50.0%)
Mentioned or partially addressed	2 (25.0%)
Fully Acknowledged	0
Tackled	2 (25.0%)
Methods	
None	4 (50.0%)
Adjustment for baseline confounders	4 (50.0%)

Table S4. Listing of key features of all included studies, sorted by year of publication

Authors [ref][a]	Year	Title	Data collection time-frame	Country	EHRs	Sample size for main analysis	Analysis criteria	Percent of target population analysed	Percent dropped out of follow up	Change in renal function measure	Methods	Average follow up time for renal function
Joss N et al [69]	2002	Diabetic nephropathy: how effective is treatment in clinical practice?	1989 - 1999	UK	Clear EHRs used	125	≥4 x creatinine over ≥6 months; no death/RRT in first year	74%	50%	regression slope of estimated creatinine clearance	linear regression	3 years
Dean BB et al [70]	2005	Erythropoiesis-stimulating protein therapy and the decline of renal function: a retrospective analysis of patients with chronic kidney disease	1998-2002	USA	Clear EHRs used	122	≥4 x creatinine over ≥6 months	not available	not available	regression slope of inverse creatinine	linear regression	1.6 years
Gallant JE et al [38]	2005	Changes in renal function associated with tenofovir disoproxil fumarate treatment, compared with nucleoside reverse-transcriptase inhibitor treatment	2001-2004	USA	Clear EHRs used	658	Unclear; assume ≥3 x creatinine over ≤1 year	not available	not available	percent change in estimated creatinine clearance from baseline	linear regression	1 year
Eriksen BO et al [14]	2006	The progression of chronic kidney disease: a 10-year population-based study of the effects of gender and age	1994-2003	Norway	Clear EHRs used	3047	≥2 x creatinine over ≥3 months	100%	34%	regression slope of eGFR (absolute scale)	linear mixed model	3.7 years

Jones C et al [71]	2006	An evaluation of a shared primary and secondary care nephrology service for managing patients with moderate to advanced CKD	1997-2006	UK	Clear EHRs used	738	≥6 x creatinine (3 x 5-year period pre-referral; 3 x 5-year period post-referral)	78%	not available	regression slope of eGFR (absolute scale)	mean difference paired t-test	Not stated (max 10 years)
Chen SC et al [72]	2008	Slowing renal function decline in chronic kidney disease patients after nephrology referral	2001-2006	Taiwan	Clear EHRs used	213	≥6 x creatinine (3 x 1-year period pre-referral; 3 x 1-year period post-referral); no AKI or dialysis ≤1 year post-referral	not available	not relevant (complete case analysis)	regression slope of eGFR (absolute scale)	generalised estimating equations	Not stated (max 2 years)
O'Riordan A et al [57]	2009	Renal biopsy in liver transplant recipients	1996-? (likely decade 2000-2010)	UK	Clear EHRs used	54	Not stated	3%	not available	binary progression to threshold eGFR	kaplan meier estimation + log-rank test	3.2 years
Eriksen BO et al [17]	2010	Predictors of declining glomerular filtration rate in a population-based chronic kidney disease cohort	1994-2003	Norway	Clear EHRs used	1224	≥2 x creatinine over ≥3 months	88%	not available	regression slope of eGFR (absolute scale)	linear mixed model	4.0 years
Cummings DM et al [47]	2011	Glycemic control patterns and kidney disease progression among primary care patients with diabetes mellitus	1998-2008	USA	Clear EHRs used	791	≥2 x creatinine (and ≥5 x HbA1c)	37%	not available	absolute change in eGFR from baseline	linear regression	7.6 years
Abdelhafiz et al [13]	2012	Natural history and predictors of faster glomerular filtration rate decline in a referred population of older patients with	1993-2010	UK	Clear EHRs used	100	No creatinine criteria; ≥5 years clinic attendance	not available	not available	regression slope of eGFR (absolute scale)	logistic regression	14 years

		type 2 diabetes mellitus										
Boudville N et al [73]	2012	Factors associated with Chronic Kidney Disease Progression in Australian Nephrology Practices	not stated	Australia	Clear EHRs used	1328	≥2 x creatinine over ≥90 days	not available	not available	Rate of change in eGFR (not clearly defined)	linear regression	1.5 years
Dreyer G et al [74]	2013	Progression of chronic kidney disease in a multi-ethnic community cohort of patients with diabetes mellitus	2005-2010	UK	Clear EHRs used	3855	≥3 x creatinine over 5 years	Not available (between 60%-95%)	15%	regression slope of eGFR (absolute scale)	linear mixed model	4.3 years
Herget-Rosenthal S et al [39]	2013	Progressive chronic kidney disease in primary care: modifiable risk factors predictive model	2003-2006	Germany	Clear EHRs used	803	Not clear; assume ≥2 x creatinine (2003 and 2006)	not available	not relevant (complete case analysis)	Rate of change in eGFR (not clearly defined)	logistic regression	3 years
Malgor RD et al [75]	2013	A case-control study of intentional occlusion of accessory renal arteries during endovascular aortic aneurysm repair	1989-2009	USA	Not clear if EHRs used	119	Not stated	not available	not available	absolute change in eGFR from baseline	ANCOVA	3.1 years
Brosnan EM et al [76]	2014	Drug-induced reduction in estimated glomerular filtration rate in patients with ALK-positive non-small cell lung cancer treated with ALK inhibitor crizotinib	2009-2012	USA	Clear EHRs used	38	Not stated	100%	not available	regression slope of eGFR (absolute scale)	linear mixed model	12 weeks

Chase HS et al [11]	2014	Presence of early CKD-related metabolic complications predict progression of stage 3 CKD: a case-controlled study	2006-2012	USA	Clear EHRs used	481	≥4 x creatinine over ≥4 years	69%	not available	regression slope of eGFR (absolute scale)	Naïve Bayes classifier; logistic regression	6 years
Chen H et al [45]	2014	Combined application of eGFR and albuminuria for the precise diagnosis of stage 2 and 3a CKD in the elderly	2000-2012	China	Not clear if EHRs used	365	≥3 x creatinine (and other markers) over 3 years	not available	not relevant (complete case analysis)	Rate of percentage change in eGFR (not clearly defined)	logistic regression	3 years
Kose E et al [77]	2014	Effects on serum uric acid by difference of the renal protective effects with atorvastatin and rosuvastatin in chronic kidney disease patients	2006-2011	Japan	Clear EHRs used	29	≥2 x creatinine over 3 months	83%	not relevant (complete case analysis)	absolute change in eGFR from baseline	mean difference paired t-test	3 months
Nderitu P et al [41]	2014	Analgesia dose prescribing and estimated glomerular filtration rate decline: a general practice database linkage cohort study	2009-2010	UK	Clear EHRs used	4145	≥2 x creatinine over ≥90 days	32%	not available	rate of change in eGFR	logistic regression	9 months
Oetjens M et al [53]	2014	Utilization of an EMR-biorepository to identify the genetic predictors of calcineurin-inhibitor toxicity in heart transplant patients	not stated	USA	Clear EHRs used	115	Not stated	91%	not available	binary progression to threshold eGFR	Cox PH regression	8.8 years
Annor FB et al [18]	2015	Psychosocial stress and changes in estimated glomerular filtration rate among adults with diabetes mellitus	2005-2008	USA	Clear EHRs used	575	Not stated	not available	not available	regression slope of eGFR (absolute scale)	structural equation modelling	4 years

Cid Ruzafa J et al [78]	2015	Estimated glomerular filtration rate progression in UK primary care patients with type 2 diabetes and diabetic kidney disease: a retrospective cohort study	2006-2011	UK	Clear EHRs used	15692	Not stated; assume ≥2 x creatinine over ≥1 year	26%	not available	regression slope of eGFR (absolute scale)	linear mixed model	3.7 years
Diggle PJ et al [19]	2015	Real-time monitoring of progression towards renal failure in primary care patients	1997-2007+	UK	Clear EHRs used	22910	≥1 x creatinine	100%	not available	predicted percent change in eGFR per unit time	linear mixed model	4.5 years
Kaga M et al [79]	2015	Risk of new-onset dyslipidemia after laparoscopic adrenalectomy in patient with primary aldosteronism	1998-2013	Japan	Not clear if EHRs used	57	2 x creatinine over 1 year	not available	not relevant (complete case analysis)	absolute change in eGFR from baseline	mean difference paired t-test	1 year
Lai CL et al [80]	2015	Effects of atorvastatin and rosuvastatin on renal function in patients with type 2 diabetes mellitus	2000-2010	Taiwan	Clear EHRs used	5569	≥2 x creatinine	7%	not available	absolute change in eGFR from baseline	linear regression	7.5 months
Perotte A et al [46]	2015	Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis	up to 2012; approx 12 years prior follow up	USA	Clear EHRs used	2908	≥2 x creatinine over ≥3 months	100%	not available	binary progression to threshold eGFR	Kalman filter time series model; Cox PH regression	Not stated
Singh A et al [21]	2015	Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration	not stated	USA	Clear EHRs used	6435	≥4 x creatinine	not available	not available	percent change in eGFR from baseline	logistic regression	Not stated

Vejakama P et al [25]	2015	Epidemiological study of chronic kidney disease progression: a large-scale population-based cohort study	1997-2011	Thailand	Clear EHRs used	32106	Not stated; assume ≥2 x creatinine over ≥3 months	not available	not available	percent change in eGFR from baseline	competing risks survival models	4.5 years
Yun WS et al [81]	2014	Long-term follow up results of acute renal embolism after anticoagulation therapy	2006-2012	South Korea	Not clear if EHRs used	31	Not stated; assume 2 x creatinine	66%	not available	absolute change in serum creatinine	descriptive result only; no statistical analysis	2.6 years
Chakera A et al [44]	2015	Prognostic value of endocapillary hypercellularity in IgA nephropathy patients with no immunosuppression	not stated	UK	Clear EHRs used	147	Not stated	62%	not available	Rate of change in eGFR (not clearly defined)	logistic regression	7 years
Johnson F et al [43]	2015	The impact of acute kidney injury in diabetes mellitus	2009-2012	UK	Clear EHRs used	200	Not stated; assume ≥2 x creatinine over ≥6 months	41%	not available	Rate of change in eGFR (not clearly defined)	difference in proportions chi-squared test	Not stated
Kim YG et al [82]	2016	Renal protective effect of DPP-4 inhibitors in type 2 diabetes mellitus patients: a cohort study	2010-2015	South Korea	Clear EHRs used	414	≥3 x creatinine over 2 years	not available	not relevant (complete case analysis)	absolute change in eGFR from baseline	mean difference paired t-test	2 years
Li XM et al [37]	2016	Clinicopathological characteristics and outcomes of light chain deposition disease: an analysis of 48 patients in a single Chinese center	2004-2015	China	Not clear if EHRs used	44	Not stated	92%	not available	binary progression to threshold serum creatinine	Cox PH regression	1.8 years

Mirajkar N et al [83]	2016	The impact of bariatric surgery on estimated glomerular filtration rate in patients with type 2 diabetes: a retrospective cohort study	2005-2012	UK	Clear EHRs used	388	≥2 x creatinine	49%	not available	absolute change in eGFR from baseline	simple non-parametric tests (Mann Whitney U)	3 years
Koraishy FM et al [42]	2017	Rate of renal function decline, race and referral to nephrology in a large cohort of primary care patients	2008-2015	USA	Clear EHRs used	2170	≥2 x creatinine over 7 years	7%	not available	rate of change in eGFR	logistic regression	Not stated
Lv L et al [36]	2017	Persistent hematuria in patients with antineutrophil cytoplasmic antibody-associated vasculitis during clinical remission: chronic glomerular lesion or low-grade active renal vasculitis?	1996-2016 (FU 2002-2016)	China	Not clear if EHRs used	208	Not stated; assume ≥2 x creatinine	95%	not available	rate of change in eGFR	Cox PH regression	3.1 years
Nishida Y et al [84]	2017	Comparative effect of calcium channel blockers on glomerular function in hypertensive patients with diabetes mellitus	2004-2012	Japan	Clear EHRs used	1217	Not stated	not available	not available	percent change in eGFR from baseline	linear mixed model	1 year
Rincon-Choles H et al [61]	2017	Impact of uric acid levels on kidney disease progression	2005-2009	USA	Clear EHRs used	1676	≥3 x creatinine; ≥1 x uric acid	6%	not available	Binary progression (changes/threshold combination)	competing risks survival models	2.8 years
Tsai CW et al [58]	2017	Serum Uric Acid and Progression of Kidney Disease: A Longitudinal Analysis and Mini-Review.	2003-2011	Taiwan	Clear EHRs used	739	≥3 x creatinine over 8 years; no RRT in first 30 days	not available	not available	regression slope of eGFR (absolute scale)	linear mixed model	4.3 years

Yao X et al [30]	2017	Renal outcomes in anticoagulated patients with atrial fibrillation	2010-2016	USA	Clear EHRs used	9769	≥2 x creatinine	not available	not available	percent change in eGFR from baseline	Cox PH regression	11 months
Beyer-Westendorf J et al [85]	2018	The CHA2DS2VASc score strongly correlates with glomerular filtration rate and predicts renal function decline over time in elderly patients with atrial fibrillation and chronic kidney disease	2008-2015	multiple european countries	Clear EHRs used	36779	≥1 x creatinine	53% Germany; not available UK	not available	regression slope of eGFR (absolute scale)	joint longitudinal survival model	1.7 years
Butt AA et al [20]	2018	Effectiveness, treatment completion and safety of sofosbuvir/ledipasvir and paritaprevir/ritonavir/ombitasvir + dasabuvir in patients with chronic kidney disease: an ERCHVIES study	2014-2016	USA	Clear EHRs used	17624	≥3 x creatinine over ≥6 months	47%	44%	absolute change in eGFR from baseline	difference in proportions chi-squared test	12 weeks
Lamacchia O et al [31]	2018	Normoalbuminuria kidney impairment in patients with T1DM: insights from annals initiative	2004-2011	Italy	Clear EHRs used	582	≥2 x creatinine over 4 years; ≥1 x albuminuria	42%	not available	percent change in eGFR from baseline	logistic regression	4 years
VanWagner LB et al [62]	2018	Cardiovascular disease outcomes related to early stage renal impairment following liver transplantation	2002-2012	USA	Clear EHRs used	671	≥3 x creatinine over ≤1 year	not available	not available	eGFR trajectory group	trajectory clustering using latent variables	1 year

Viazzi F et al [32]	2018	Apparent treatment resistent hypertension, blood pressure control and the progression of chronic kidney disease in patients with type 2 diabetes	2004-2011	Italy	Clear EHRs used	2312	≥6 x creatinine over ≥4 years; Complete data for BP, eGFR and albuminuria according to treatment protocol	33%	not relevant (complete case analysis)	percent change in eGFR from baseline	logistic regression	4 years
Evans RDR et al [22]	2018	Clinical manifestations and long-term outcomes of IgG4-related kidney and retroperitoneal involvement in a United Kingdom IgG4-related disease cohort	2002-2018	UK	Not clear if EHRs used	24	≥2 x creatinine	86%	not relevant (complete case analysis)	percent change in eGFR from baseline	crude estimation / descriptive results only	5 years
Horne L et al [48]	2019	Epidemiology and health outcomes associated with hyperkalemia in a primary care setting in England	2009-2013	UK	Clear EHRs used	195178	Not stated	not available	not available	transition to CKD stage	Crude estimation / descriptive results only	Not stated
Hsu TW et al [27]	2019	Comparison of the effects of dsnosumab and alendronate on cardiovascular and renal outcomes in osteoporotic patients	2005-2017	Taiwan	Clear EHRs used	5046	≥2 x creatinine	not available	not available	percent change in eGFR from baseline	Cox PH regression	Max 5 years
Jalal K et al [15]	2019	Can billing codes accurately identify rapidly progressing stage 3 and stage 4 chronic kidney disease patients: a diagnostic test study	2007-2017	USA	Clear EHRs used	10927	≥5 x creatinine over ≥3 years	38%	not available	regression slope of eGFR (absolute scale)	linear mixed model	≥3 years

Kim WJ et al [86]	2019	The role of a treat-to-target approach in the long-term renal outcomes of patients with gout	2007-2018	South Korea	Clear EHRs used	244	≥2 x creatinine over ≥1 year	93%	not available	absolute change in eGFR from baseline	logistic regression	2 years
Lai YJ [24]	2019	Effect of weight loss on the estimated glomerular filtration rates of obese patients at risk of chronic kidney disease: the RIGOR-TMU study	2008-2016+	Taiwan	Clear EHRs used	1620	≥3 x creatinine over ≥3 months; Propensity score matched only	37%	not available	percent change in eGFR from baseline	Cox PH regression	Max 1 year
Leither MD et al [59]	2019	The impact of outpatient acute kidney injury on mortality and chronic kidney disease: a retrospective cohort study	not stated	USA	Clear EHRs used	196209	Not clear; assume ≥1 x creatinine	51%	not available	Binary progression (eGFR changes/threshold combination)	Cox PH regression	5.3 years
Liu D et al [60]	2019	Serum immunoglobin G provides early risk prediction in immunoglobin A nephropathy	2009-2014	China	Not clear if EHRs used	455	Not clear but ≥2 years; 1 x serum IgG	73%	not available	Binary progression (eGFR changes/threshold combination)	Cox PH regression	3.7 years
Morales-Alvarez MC et al [40]	2019	Renal function decline in latinos with type 2 diabetes	2002-2015	USA	Clear EHRs used	594	≥2 x creatinine	65%	not available	rate of change in eGFR	difference in means t-test	Not stated
O'Neill RA et al [87]	2019	Evaluation of long-term intravitreal antivascular endothelial growth factor injections on renal function in patients with and without diabetic kidney disease	2012-2018	UK	Clear EHRs used	85	Not stated	not available	not available	regression slope of eGFR (absolute scale)	linear regression	2.6 years

Park JM et al [88]	2018	Oncological and functional outcomes of laparoscopic radiofrequency ablation and partial nephrectomy for T1a renal masses: a retrospecive single-center 60 month follow up cohort study	2005-2014	South Korea	Not clear if EHRs used	115	Not stated but ≥2 years	not available	not available	absolute change in eGFR from baseline	difference in means t-test	5.3 years
Posch F et al [89]	2019	Longitudinal kidney function trajectories predict major bleeding, hospitalisation and death in patients with atrial fibrillation and chronic kidney disease	2009-2015	UK	Clear EHRs used	18240	Not clear but ≥6 months; assume ≥1 x creatinine	88%	not available	rate of change in eGFR	joint longitudinal survival model	1.9 years
Posch F et al [26]	2019	Exposure to vitamin k antagonists and kidney function decline in patients with atrial fibrillation and chronic kidney disease	2009-2015	Germany	Clear EHRs used	14432	≥1 x creatinine; ≥1 x CHA2DS2VASc score	39%	not available	regression slope of eGFR (absolute scale and percent scale)	linear mixed model	1.4 years
Spanopoulos D et al [90]	2019	Temporal variation of renal function in people with type 2 diabetes mellitus: a retrospective UK clinical practice research datalink cohort study	2009-2016	UK	Clear EHRs used	7766	≥6 x creatinine over 5 years	17%	not available	transition to CKD stages	descriptive result only; no statistical analysis	5 years
Wang Y et al [12]	2019	Implications of a family history of diabetes and rapid eGFR decline in patients with type 2 diabetes and biopsy-	2007-2017	China	Clear EHRs used	128	≥3 x creatinine over ≥1 year	not available	not available	regression slope of eGFR (absolute scale)	logistic regression	2 years

		proven diabetic kidney disease										
Yoo H et al [34]	2019	Effects of sarpogrelate on microvascular complications with type 2 diabetes	2010-2015	South Korea	Clear EHRs used	478	Not clear; assume ≥2 x creatinine; Propensity score matched only	9%	not available	percent change in eGFR from baseline	kaplan meier estimation + log-rank test	5.7 years
Zhao J et al [91]	2019	Predicting outcomes of chronic kidney disease from EMR data based on random forest regression	2009-2017	USA	Clear EHRs used	61740	≥3 x creatinine over 7 years	51%	not available	eGFR prediction at multiple time points	random forest regression	Not stated
Lee JS et al [92]	2020	Recovery of renal function in patients with lupus nephritis and reduced renal function: the beneficial effect of hydroxychloroquine	1995-2018	South Korea	Clear EHRs used	90	≥2 x creatinine over 6 months	100%	not relevant (complete case analysis)	binary progression to threshold eGFR	logistic regression	6 months
Nakamura A et al [93]	2020	Impact of sodium-glucose cotransporter 2 inhibitors on renal function in participants with type 2 diabetes and chronic kidney disease with normoalbuminuria	not stated	Japan	Not clear if EHRs used	87	≥2 x creatinine over 2 years	not available	not relevant (complete case analysis)	percent change in eGFR from baseline	ANOVA	2 years
Sise ME et al [94]	2019	Direct-acting antiviral therapy slows kidney function decline in patients with Hepatitis C virus infection and chronic kidney disease	2013-2017	USA	Clear EHRs used	1178	≥4 x creatinine over ≤6 years	60%	not available	regression slope of eGFR (absolute scale)	generalised estimating equations	1.6 years (post-therapy)

Weldegiorgis M et al [55]	2020	Socioeconomic disadvantage and the risk of advanced chronic kidney disease: results from a cohort study with 1.4 million participants	2020-2014+	UK	Clear EHRs used	1,397,573	≥2 years "data" before baseline; ≥3 years follow-up; no specific creatinine requirement	not available	not available	Binary progression to threshold eGFR	Cox PH regression	7.5 years
Cabrera CS et al [16]	2020	Impact of CKD Progression on Cardiovascular Disease Risk in a Contemporary UK Cohort of Individuals With Diabetes	2005-2015	UK	Clear EHRs used	3,022	≥2 x creatinine at baseline; ≥1 x creatinine follow-up	57%	Not available	regression slope of eGFR (absolute scale)	Cox PH regression	4.3 years
Cleary F et al [95]	2020	Feasibility of evaluation of the natural history of kidney disease in the general population using electronic healthcare records	2008-2016	UK	Clear EHRs used	1,597,629	≥3 x creatinine	25%	2.4%	regression slope of eGFR (absolute scale)	Linear regression	5.7 years
Faraj KS et al [96]	2020	The effect of urinary diversion on long-term kidney function after cystectomy	2007-2018+	USA	Clear EHRs used	563	≥1 x creatinine follow-up; unclear if baseline requirements	98%	Not available	regression slope of eGFR (absolute scale)	Linear mixed model	3.9 years
Inaguma D et al [28]	2020	Increasing tendency of urine protein is a risk factor for rapid eGFR decline in patients with CKD: A machine learning-based prediction model by using a big database	2004-2019	Japan	Clear EHRs used	9,911	Unclear	Not available	Not relevant (complete case analysis)	Rate of percent change in eGFR, not clearly defined	Logistic regression ; Random forest regression	Not stated
Nichols GA et al [50]	2020	Kidney disease progression and all-cause mortality across estimated glomerular filtration rate and	2006-2016	USA	Clear EHRs used	Approx. 36,727	≥1 x creatinine baseline; ≥1 x creatinine follow-up; ≥1 x proteinuria/	47%	Not available	Transition between CKD stages	Life-table analysis	5 years

		albuminuria categories among patients with vs. Without type 2 diabetes					ACR at baseline					
Peng YL et al [29]	2020	Comparison of uric acid reduction and renal outcomes of febuxostat vs allopurinol in patients with chronic kidney disease	2010-2015	Taiwan	Clear EHRs used	1,050	≥2 x creatinine baseline; Creatinine, SUA follow-up (unclear)	Not available	Not available	regression slope of eGFR (absolute scale)	Linear mixed model	1.5 years
Rej S et al [33]	2020	Association of Lithium Use and a Higher Serum Concentration of Lithium With the Risk of Declining Renal Function in Older Adults: A Population-Based Cohort Study	2007-2015	Canada	Clear EHRs used	6,226	≥1 x creatinine; ≥1 x lithium	Not available	Not available	percent change in eGFR from baseline	Cox PH regression	3.1 years
Jackeviciu s CA et al [23]	2021	Bleeding Risk of Direct Oral Anticoagulants in Patients With Heart Failure And Atrial Fibrillation. Circulation-Cardiovascular Quality and Outcomes	2010-2018	USA	Clear EHRs used	49,458	Unclear	92%	Not available	Rate of percent change in eGFR, not clearly defined	Cox PH regression	1.4 years
Neuen BL et al [54]	2021	Changes in GFR and Albuminuria in Routine Clinical Practice and the Risk of Kidney Disease Progression	2000-2015	UK	Clear EHRs used	91,319	≥2 x creatinine; ≥2 x UACR	1%	Not available	Percent change in eGFR between measures	Cox PH regression	2.9 years
Niu SF et al [56]	2021	Early Chronic Kidney Disease Care Programme delays kidney function deterioration in patients with stage I-IIIa chronic kidney disease: an	2012-2017+	Taiwan	Clear EHRs used	3,114	Unclear; ≥2 x medical visits; possibly ≥1 x creatinine	Not available	Not available	Binary progression to threshold eGFR	Cox PH regression	3.0 years

		observational cohort study in Taiwan										
Robinson DE et al [49]	2021	Safety of Oral Bisphosphonates in Moderate-to-Severe Chronic Kidney Disease: A Binational Cohort Analysis. Journal of Bone and Mineral Research	1997-2016	Multiple European countries	Clear EHRs used	19,324	≥2 x creatinine at baseline; ≥1 x creatinine follow-up	Not available	Not available	Transition between CKD stages	Competing risks survival models	3.7 years
Tangri N et al [35]	2021	Metabolic acidosis is associated with increased risk of adverse kidney outcomes and mortality in patients with non-dialysis dependent chronic kidney disease: an observational cohort study	2007-2017	USA	Clear EHRs used	32,007	≥3 x creatinine; ≥3 x serum bicarbonate	Not available	Not available	percent change in eGFR from baseline	Cox PH regression	3.9 years
Vesga JI et al [52]	2021	Chronic kidney disease progression and transition probabilities in a large preventive cohort in colombia	2009-2018	Colombia	Clear EHRs used	2,752	≥2 x creatinine	90%	65%	Transition between CKD stages	Crude estimation	Not stated
Yanagawa et al [51]	2021	Retrospective study of factors associated with progression and remission/regression of diabetic kidney disease-hypomagnesemia was associated with progression and elevated serum alanine aminotransferase	2003-2019	Japan	Clear EHRs used	681	Not stated	99%	Not relevant (complete case analysis)	Transition between CKD stages	Cox PH regression	6.2 years

		levels were associated with remission or regression										

[a][ref] refers to reference number in supplementary listing of reviewed studies (which also match those listed in manuscript Table 4)

Appendix 2

Co-authored research: Defining measures of kidney function in observational studies using routine health care data – methodological and reporting considerations

See next page for co-authored published research paper.

Defining measures of kidney function in observational studies using routine health care data: methodological and reporting considerations

OPEN

Juan Jesus Carrero[1,24], Edouard L. Fu[1,2,3,24], Søren V. Vestergaard[4,5], Simon Kok Jensen[4,5], Alessandro Gasparini[1], Viyaasan Mahalingasivam[6], Samira Bell[7], Henrik Birn[5,8,9], Uffe Heide-Jørgensen[4,5], Catherine M. Clase[10,11], Faye Cleary[6], Josef Coresh[12], Friedo W. Dekker[3], Ron T. Gansevoort[13], Brenda R. Hemmelgarn[14], Kitty J. Jager[15,16], Tazeen H. Jafar[17], Csaba P. Kovesdy[18], Manish M. Sood[19], Bénédicte Stengel[20], Christian F. Christiansen[4,5], Masao Iwagami[6,21,25] and Dorothea Nitsch[6,22,23,25]

[1]*Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Solna, Sweden;* [2]*Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA;* [3]*Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands;* [4]*Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark;* [5]*Department of Clinical Medicine, Aarhus University, Aarhus, Denmark;* [6]*Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK;* [7]*Division of Population Health and Genomics, University of Dundee, Dundee, UK;* [8]*Department of Biomedicine, Aarhus University, Aarhus, Denmark;* [9]*Department of Renal Medicine, Aarhus University Hospital, Aarhus, Denmark;* [10]*Department of Medicine, McMaster University, Hamilton, Ontario, Canada;* [11]*Department of Health Research and Methodology, McMaster University, Hamilton, Ontario, Canada;* [12]*Department of Epidemiology, Johns Hopkins University, Baltimore, Maryland, USA;* [13]*Department of Nephrology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands;* [14]*Department of Medicine, University of Alberta, Edmonton, Alberta, Canada;* [15]*ERA Registry, Amsterdam UMC location University of Amsterdam, Medical Informatics, Meibergdreef, Amsterdam, Netherlands;* [16]*Amsterdam Public Health Research Institute, Quality of Care, Amsterdam, the Netherlands;* [17]*Program in Health Services and Systems Research, Duke-NUS Medical School, Singapore;* [18]*Division of Nephrology, Department of Medicine, University of Tennessee Health Science Center, Memphis, Tennessee, USA;* [19]*Department of Medicine, the Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada;* [20]*CESP (Center for Research in Epidemiology and Population Health), Clinical Epidemiology Team, University Paris-Saclay, University Versailles-Saint Quentin, Inserm U1018, Villejuif, France;* [21]*Department of Health Services Research, University of Tsukuba, Ibaraki, Japan;* [22]*Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand; and* [23]*UK Renal Registry, UK Kidney Association, Bristol, UK*

The availability of electronic health records and access to a large number of routine measurements of serum creatinine and urinary albumin enhance the possibilities for epidemiologic research in kidney disease. However, the frequency of health care use and laboratory testing is determined by health status and indication, imposing certain challenges when identifying patients with kidney injury or disease, when using markers of kidney function as covariates, or when evaluating kidney outcomes. Depending on the specific research question, this may influence the interpretation, generalizability, and/or validity of study results. This review illustrates the heterogeneity of working definitions of kidney disease in the scientific literature and discusses advantages and limitations of the most commonly used approaches using 3 examples. We summarize ways to identify and overcome possible biases and conclude by proposing a framework for reporting definitions of exposures and outcomes in studies of kidney disease using routinely collected health care data.

Kidney International (2023) **103,** 53–69; https://doi.org/10.1016/j.kint.2022.09.020

KEYWORDS: albuminuria; chronic kidney disease (CKD); creatinine; epidemiology; estimated glomerular filtration rate (eGFR); routinely collected health care data

Routinely collected health care data from registries, electronic health records, and claims databases are increasingly used for research purposes. The availability of laboratory-based kidney function markers, such as serum creatinine and albuminuria, in these data sources increases the opportunities for research in kidney disease. Carefully conducted epidemiologic studies are critical to address the burden, incidence, and prevalence of kidney disease, and to identify mechanisms of action, optimal mitigation/treatment strategies, and gaps in health care processes that collectively have the potential to improve care and, ultimately, outcomes.

Correspondence: *Juan Jesus Carrero, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Nobels väg 12A, Box 281, 171 77 Stockholm, Sweden. E-mail: juan.jesus.carrero@ki.se; and Dorothea Nitsch, Department of Non-Communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. E-mail: Dorothea.Nitsch@lshtm.ac.uk*

[24]JJC and ELF shared first-author position.

[25]MI and DN shared last-author position.

Received 10 November 2021; revised 31 August 2022; accepted 9 September 2022; published online 22 October 2022

Using routinely collected health care data poses specific challenges when defining chronic kidney disease (CKD),[1] acute kidney injury (AKI),[2] the recently proposed entity, acute kidney disease (AKI that is still evolving),[3] and CKD progression. There are ongoing initiatives to harmonize efforts for establishing diagnoses of CKD, AKI, and acute kidney disease,[4] identify reproducible and valid end points in clinical trials,[5–7] and define outcomes that are important to people living with kidney disease.[8] However, concurrent efforts to harmonize definitions in epidemiologic studies that rely on routine clinical data have been lacking.

The aim of this review is to highlight potential challenges of working definitions of measurements of kidney health in studies of routine care. We start by discussing general issues when working with routine care data: routine care data sources may be fragmented and capture sicker patients. We then discuss pros and cons of the most commonly used definitions and suggest ways to identify and overcome potential biases introduced by using these definitions (Table 1).[9–14] Throughout the article, we use 3 exemplar research questions as illustration. We specifically focus on the following causal questions, for which biases (confounding, selection bias, and measurement bias) are well defined:

(i) What is the causal effect of receiving a CKD diagnosis on mortality risk in older patients with 2 estimated glomerular filtration rate (eGFR) values <60 ml/min per 1.73 m^2 >3 months apart?
(ii) Among people with CKD, what is the causal effect of initiating sodium-glucose cotransporter-2 inhibitors versus dipeptidyl peptidase 4 inhibitors on the risk of CKD progression, heart failure admissions, and all-cause mortality?
(iii) After AKI, what is the causal effect of stopping versus continuing renin-angiotensin system inhibitors (RASi) on the risk of recurrent AKI?

We will not cover cohort studies with prospective recruitment and follow-up: considerations for this research design are different from those when using routinely collected data. We conclude by proposing a framework for consensus efforts on reporting definitions of exposures and outcomes in observational studies addressing kidney disease using routinely collected data.

General issues when working with routine care data

Data fragmentation affects who is captured and followed up. In many countries, routinely collected health care data are captured in disjointed software systems, which are not necessarily integrated.[15] For instance, laboratory information may be only captured in a specific clinical setting,[16] such as ambulatory care, or in hospitals, leading to fragmentation of information and follow-up in the data set. Other databases may include patients on enrollment in an insurance plan (e.g., in the United States), or when they become aged 65 years (e.g., Medicare in the United States and Ontario Drug Benefits): when this happens, data before cohort entry are usually not available. Similarly, patients may exit the database when they move to another general practitioner, or when switching from insurance, not contributing further to the database.[17] Commonly used health care databases in kidney research are summarized in Supplementary Table S1 and Figure 1.[18,19]

The completeness of the data capture may influence the interpretation, generalizability, and internal validity of study results.[20] For instance, when studying the causal effect of initiating sodium-glucose cotransporter-2 inhibitors versus dipeptidyl peptidase 4 inhibitors on CKD progression in claims data sources (question 2), unavailability of data before enrollment may lead to misclassification of conditions, and this may bias effect estimates (information bias).[21] Furthermore, selection bias due to informative censoring[22] will occur when using data sources in which patients with advanced kidney disease get referred from primary care to secondary care and are not followed up thereafter, because those developing advanced kidney disease will drop out from the database.

Sicker patients have more tests on file than healthy patients. Routinely collected health care data do not capture a random sample of the population, but a subgroup of patients who interact with the health care system. As an example, the Stockholm CREAtinine Measurements (SCREAM) cohort in Sweden described that during 2006 to 2011 roughly 67% of the Stockholm population underwent creatinine testing at least once.[23] This nongeneralizability is unlikely to be important for studies that focus on drug effectiveness and safety (such as questions 2 and 3), because the target population for such studies is usually the population eligible for receiving these drugs (i.e., those interacting with the health care system), and not the complete population. However, if the interest would lie in the estimation of the CKD prevalence in Stockholm, the investigator would need to account for the fact that healthier individuals are underrepresented in the data set.[24]

The presence and frequency of a certain laboratory measurement reflect aspects of disease (e.g., albuminuria testing in routine care is mainly directed to specific populations [people with diabetes, hypertension, pregnancy, and known CKD]).[25–27] For question 2, it would be useful to include albuminuria as a potential confounder, but the issue of missing data must be addressed. Excluding individuals without these data may have 2 consequences. First, the cohort may not be similar anymore to the target population, because sicker subgroups are oversampled, which may affect generalizability of study findings (e.g., the medication may be more beneficial in the study population because it oversampled patients with macroalbuminuria, for whom the absolute benefit is larger). It is therefore good practice to report how representative the study population is compared with the target population (e.g., by comparing baseline characteristics or incidence of outcomes).[28] Complete case analysis can lead to bias when data are not missing completely at random (Table 1), although several exceptions exist.[29] We caution against the uncritical use of multiple imputation methods using electronic health data as they can worsen bias if the models are misspecified.[29–32]

Table 1 | Glossary of terminology associated with bias and examples of research questions

		1. What is the causal effect of receiving a CKD diagnosis on mortality risk in older patients with 2 eGFR values <60 ml/min per 1.73 m^2 >3 months apart?		2. Among people with CKD, what is the causal effect of initiating SGLT2i vs. DPP4i on the risk of CKD progression, heart failure admissions, and all-cause mortality?		3. After AKI, what is the causal effect of stopping vs. continuing RASi on the risk of recurrent AKI?	
		Exposure: receiving a CKD diagnosis vs. not receiving a CKD diagnosis Outcome: mortality Population: people aged ≥65 yr with 2 eGFR values <60 ml/min per 1.73 m^2 >3 months apart		Exposure: SGLT2i vs. DPP4i Outcomes: CKD progression, heart failure admissions, and all-cause mortality Population: people with CKD		Exposure: stopping vs. continuing RASi Outcome: recurrent AKI Population: people with AKI	
Bias	**Definition**	**Example of how bias arises**	**Potential solution**	**Example of how bias arises**	**Potential solution**	**Example of how bias arises**	**Potential solution**
Selection bias	Bias in the estimated association or effect of an exposure on an outcome that arises from the procedures used to select individuals into the study or the analysis. It arises when conditioning on a common effect. Examples of selection bias include depletion of susceptibles (or survivor bias), prevalent user bias (those who did not tolerate drug use or died are excluded at baseline), informative censoring or loss to follow up, and missing data. Collider bias is a special case of selection bias where the analysis conditions (either by statistical adjustment or restriction of the study population) on a collider (a variable that is affected by 2 other variables [e.g., exposure or outcome or related variables, which then introduce a spurious association]).	1. Assessing the relationship between CKD diagnosis and mortality in routinely collected health care data implicitly restricts to the subset of people with 2 eGFR measurements. As both eGFR level and health status influence availability of test results in the study, collider bias is introduced (Supplementary Figure S1A). 2. When follow-up is started from the second measurement, depletion of susceptibles may lead to selection bias due to 2 colliders (Supplementary Figure S1B).	1. Conduct study in a setting/population where kidney function is measured in everybody at baseline (e.g., by restricting to a certain subpopulation). 2. Selection bias due to depletion of susceptibles may be small when the window between 2 measurements is short and low-risk populations are studied.	1. Loss to follow-up or dropout that is differential with respect to the exposure leads to selection bias (e.g., this occurs when data are fragmented [only primary care data are available] and more people in the DPP4i arm are lost to follow-up as they transition to specialized care). Differential dropout due to death also biases estimation of kidney function slopes (Supplementary Figure S1C). 2. Missing data: excluding individuals without eGFR or UACR measurements introduces selection bias when the missingness depends on the exposure (either directly or indirectly) and health status (Supplementary Figure S1D).	1. Loss to follow-up or dropout can be handled using inverse probability of censoring weighting or by using joint models, which explicitly model the dropout process and longitudinal outcome simultaneously through shared random effects. 2. Provide clarity as to who the population with available measurements (eGFR/UACR measurements) was and do not extrapolate further. Multiple imputation can be attempted but may be misspecified.	Studying recurrent events is susceptible to selection bias when prior treatment influences the risk of AKI[9] (Supplementary Figure S1E).	Adjusting for prior treatment, or restricting to people who have not been treated previously, removes the selection bias.

(*Continued on following page*)

Table 1 | (Continued) **Glossary of terminology associated with bias and examples of research questions**

		1. What is the causal effect of receiving a CKD diagnosis on mortality risk in older patients with 2 eGFR values <60 ml/min per 1.73 m^2 >3 months apart?		2. Among people with CKD, what is the causal effect of initiating SGLT2i vs. DPP4i on the risk of CKD progression, heart failure admissions, and all-cause mortality?		3. After AKI, what is the causal effect of stopping vs. continuing RASi on the risk of recurrent AKI?	
		Exposure: receiving a CKD diagnosis vs. not receiving a CKD diagnosis Outcome: mortality Population: people aged ≥65 yr with 2 eGFR values <60 ml/min per 1.73 m^2 >3 months apart		Exposure: SGLT2i vs. DPP4i Outcomes: CKD progression, heart failure admissions, and all-cause mortality Population: people with CKD		Exposure: stopping vs. continuing RASi Outcome: recurrent AKI Population: people with AKI	
Bias	**Definition**	**Example of how bias arises**	**Potential solution**	**Example of how bias arises**	**Potential solution**	**Example of how bias arises**	**Potential solution**
Information bias	Bias in an estimate arising from measurement errors or misclassification (the erroneous classification of an individual, a value, or an attribute into a category other than that to which it should be assigned).	Immortal time bias: immortal time is introduced if follow-up is started at the moment eGFR decreases to <60 ml/min per 1.73 m^2, but patients who receive a CKD diagnosis later during follow-up are classified into the CKD diagnosis group. Immortal time bias can be considered a form of misclassification because unexposed person-time is incorrectly considered exposed person-time.	Causal study designs, such as clone-censor-weight or sequential trials, appropriately align the start of follow-up with the start of exposure and mitigate immortal time bias.	1. More kidney function measurements are taken during follow-up in the DPP4i arm, leading to differential measurement error in the outcome. CKD progression is therefore more likely to be picked up in the DPP4i arm, biasing the effect estimates. 2. Admissions of patients with advanced CKD who have volume overload but normal cardiac function on echocardiography are miscoded as admissions with heart failure. 3. Classifying transient decreases in GFR as CKD progression leads to bias in the estimation of the incidence rate of the outcome and of the absolute risk differences.	1. To detect differential outcome ascertainment bias, the number of kidney function measurements can be compared between exposure groups. More kidney function measurements in 1 exposure group than the other point toward differential outcome ascertainment bias. Differential outcome ascertainment bias is unlikely to occur for hard end points that do not depend on testing, such as kidney replacement therapy. 2. Perform validation study of heart failure codes for population with CKD to quantify whether this bias is common or rare. The obtained sensitivity and specificity can be used in quantitative bias analyses to provide adjusted effect estimates. 3. Identifying sustained declines in eGFR by using a linear mixed model helps to appropriately classify transient decreases in GFR as nonevent.	1. Using diagnosis codes to ascertain AKI may miss many AKI events (high specificity and low sensitivity), [10–13] leading to an underestimate of the incidence rate and bias in absolute risk differences 2. Differential outcome ascertainment may occur if more creatinine measurements are performed during follow-up in 1 exposure group and hospitalizations without a baseline creatinine measurement are not considered for AKI events.	Clearly report the definition used. Use different definitions to assess their influence on point estimates. To detect differential outcome ascertainment bias, the number of kidney function measurements can be compared between exposure groups.
Confounding bias	Bias of the estimated effect of an exposure on an outcome because of the presence of common causes of the exposure and the outcome.	Many risk factors for receiving a CKD diagnosis are also risk factors for all-cause mortality. For instance, physicians may give a coded CKD diagnosis to sicker patients who have more comorbidities and a lower eGFR.	Measuring and appropriately adjusting for all confounders. Alternatively, quasi-experimental designs, such as regression discontinuity, can be used to study the effect of receiving a CKD diagnosis on outcomes by using a "threshold" (i.e., the probability to receive	SGLT2i are more likely to be prescribed in people with CKD (study population) but with prescriptions depending on kidney function itself, and also in people with atherosclerotic cardiovascular disease, or heart failure. Corresponding diagnosis codes (which often have high specificity but low sensitivity) may lead to residual confounding.	Measure and adjust for all confounders. Whenever available, adjust for measurements of kidney function, such as eGFR and UACR, and metrics of heart failure/volume overload at baseline (e.g., LVEF and NT-proBNP). Be aware of fragmentation of data. Negative or positive control outcomes can be used to detect and adjust for residual confounding. Quantitative bias analysis can be used to assess the influence of residual confounding on effect estimates.	The severity of AKI may be an important confounder because it will influence the decision of whether to stop RASi and is associated with the likelihood of having a repeated AKI.	Adjust for the severity of AKI, taking into account the magnitude of creatinine elevations as well as whether kidney replacement therapy was needed.

(*Continued on following page*)

Table 1 | (Continued) **Glossary of terminology associated with bias and examples of research questions**

	1. What is the causal effect of receiving a CKD diagnosis on mortality risk in older patients with 2 eGFR values <60 ml/min per 1.73 m^2 >3 months apart?		2. Among people with CKD, what is the causal effect of initiating SGLT2i vs. DPP4i on the risk of CKD progression, heart failure admissions, and all-cause mortality?		3. After AKI, what is the causal effect of stopping vs. continuing RASi on the risk of recurrent AKI?	
	Exposure: receiving a CKD diagnosis vs. not receiving a CKD diagnosis Outcome: mortality Population: people aged ≥65 yr with 2 eGFR values <60 ml/min per 1.73 m^2 >3 months apart		Exposure: SGLT2i vs. DPP4i Outcomes: CKD progression, heart failure admissions, and all-cause mortality Population: people with CKD		Exposure: stopping vs. continuing RASi Outcome: recurrent AKI Population: people with AKI	
Bias	Example of how bias arises	Potential solution	Example of how bias arises	Potential solution	Example of how bias arises	Potential solution
Definition	a CKD diagnosis is much higher <60 ml/min per 1.73 m^2 than >60 ml/min per 1.73 m^2, but people closely around the threshold have a similar prognosis).					

AKI, acute kidney injury; CKD, chronic kidney disease; DPP4i, dipeptidyl peptidase 4 inhibitors; eGFR, estimated glomerular filtration rate; GFR, glomerular filtration rate; LVEF, left ventricular ejection fraction; NT-proBNP, N-terminal prohormone of brain natriuretic peptide; RASi, renin-angiotensin system inhibitors; SGLT2i, sodium-glucose cotransporter-2 inhibitors; UACR, urinary albumin–to–creatinine ratio.
Definitions were sourced from *Last's Dictionary of Epidemiology*[14] and adapted to the current context.

Laboratory testing may also be influenced by external factors, such as financial incentivization. For instance, there was a notable increase in serum creatinine testing among patients with diabetes attending primary care in the United Kingdom following the implementation of the Quality and Outcomes Framework in 2004.[33]

Considerations when using CKD as exposure or population

Various algorithms have been used to identify persons with CKD in health care databases[34]: Table 2 describes those most commonly used, along with identified merits and caveats. Figure 2 graphically shows an example of how different algorithms may identify the same patient at different points during the disease course. This means that for research question 2, different CKD populations will be identified, depending on the definition used, which affects generalizability and interpretation of study results.

Diagnostic coding of CKD. In settings without laboratory data, diagnosis codes (e.g., *International Classification of Diseases, Ninth Revision* [*ICD-9*] or *International Classification of Disease, Tenth Revision* [*ICD-10*]) are commonly used to identify patients with CKD.[34] Diagnostic codes have high specificity for CKD, and can detect patients with structural abnormalities not recognized by laboratory-based algorithms.[33–35] However, relying on recorded clinical diagnoses of CKD often fails to identify a large proportion of patients with CKD due to limited awareness of kidney disease, meaning a low sensitivity.[36,37] The consequences of using diagnostic codes to identify patients with CKD also depends on coding practices: increasing awareness resulting from system changes, such as automatic eGFR, can lead to changes in the completeness of data over time.[38] In studies with cohort identification periods spanning many years, underlying morbidity or severity of diagnosed CKD in selected patients may vary over time.[39–42] For questions 1 and 2, studies should therefore take account of calendar year and health provider (e.g., different general practitioners in the United Kingdom) to address temporal and health provider variation in CKD identification, which is likely nonrandom, and potentially associated with health outcomes. We suggest using the term "diagnosed CKD" when detection is limited to *International Classification of Diseases* codes.

For question 1, "What is the causal effect of receiving a CKD diagnosis on mortality risk in older patients with 2 eGFR values <60 ml/min per 1.73 m^2 >3 months apart?," receiving a CKD diagnosis is the exposure, but not the population. Those who have biochemical evidence for a CKD diagnosis but have no formal diagnosis on file are the comparison group.

Laboratory variables and equations to estimate glomerular filtration rate. The laboratory assay used for quantifying serum or plasma creatinine, and its traceability to the isotope-dilution mass spectrometry international standard, as well as the equation used for estimating glomerular filtration rate should be clearly reported in research.[43] Researchers need to be aware that eGFR may not reflect true kidney function; and

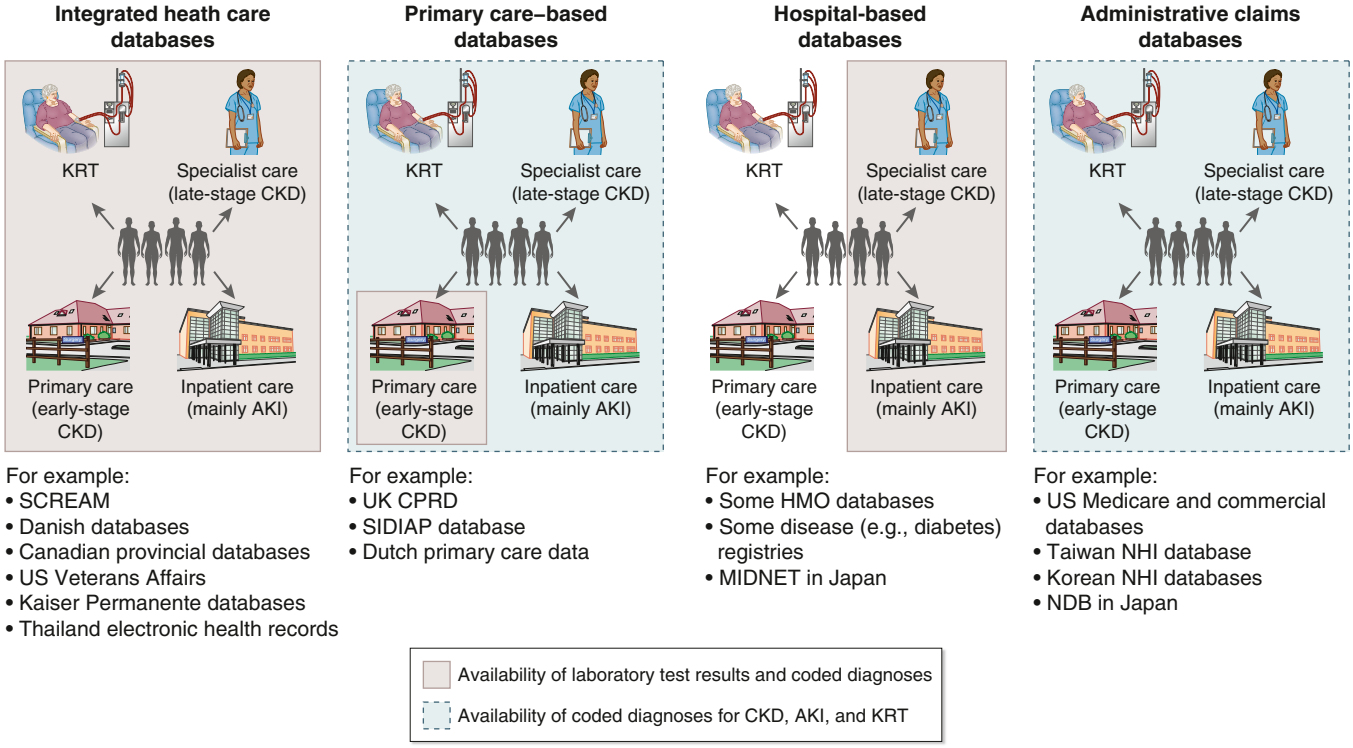


Figure 1 | Overview of different routinely collected health care databases used in kidney disease research, illustrating data fragmentation. AKI, acute kidney injury; CKD, chronic kidney disease; CPRD, Clinical Practice Research Datalink; HMO, health maintenance organization; KRT, kidney replacement therapy; MIDNET, Medical Information Database NETwork; NDB, National Database of Health Insurance Claims and Specific Health Checkups of Japan; NHI, National Health Insurance; SCREAM, Stockholm CREAtinine Measurements; SIDIAP, Information System for Research in Primary Care in Catalonia, Spain.

depending on the study question, this can lead to bias. An additional issue, if the data are available, is the type of health care encounter in which the test took place (i.e., outpatient vs. inpatient creatinine test). We note that eGFR equations have not been validated when kidney function is not stable.

Consequences of using the chronicity criterion. Kidney Disease: Improving Global Outcomes (KDIGO) developed a consensus definition for diagnosing CKD in clinical practice.[1] This definition is based on the presence of reduced eGFR or albuminuria for at least 3 months or structural abnormalities of the kidneys, and requires repeated testing, if the first screening result of eGFR or albuminuria is abnormal. Ensuring chronicity is essential in establishing a CKD diagnosis and has become routine worldwide. This approach has also frequently been applied in observational studies evaluating the incidence, prevalence, risk factors, and outcomes of CKD.[44–47] However, requiring 2 consecutive eGFR measurements 3 months apart in routinely collected data can lead to a selective population, because it requires that the patient is sick enough to seek health care twice or be recalled for a confirmatory test, which may vary between clinicians and based on patient characteristics. Furthermore, if the time for the diagnosis of CKD is defined by the second low eGFR beyond 3 months, the identification with CKD will be delayed at least 3 months. In a recent study, using the definition for CKD of 2 eGFR measurements of <60 ml/min per 1.73 m^2 at least 90 days apart (with no upper limit) resulted in a "delay," with more than half of patients being recognized as having CKD >1 year after the first low eGFR (median, 13 months [interquartile range, 6–35 months]).[37]

When answering research question 1, care should be taken to appropriately align the start of follow-up with the start of the exposure to prevent immortal time bias or depletion of susceptibles bias.[48] If follow-up is started at the time eGFR decreases below 60 ml/min per 1.73 m^2, but patients receive a diagnosis of CKD only later during follow-up, immortal time will be introduced.[49–51] Patients in the CKD diagnosis group cannot die during the period between eGFR <60 ml/min per 1.73 m^2 and the CKD diagnosis. After all, they would have been assigned to the "no CKD diagnosis" group if they had died during this period. This gives an artificial survival advantage to the CKD diagnosis group. If receiving a CKD diagnosis truly has a causal effect (either beneficial or harmful) on mortality, and follow-up is started some period *after* patients received a CKD diagnosis, "depletion of susceptibles" bias is introduced,[52] which is a form of selection or

Table 2 | Advantages and disadvantages of different definitions of CKD used in previous studies based on routinely collected data

Definition	Advantages	Disadvantages
Diagnosis codes	- High specificity, because clinically verified - Usually available in data sources without laboratory measurements (e.g., claims databases) - May pick up structural changes that are not picked up by eGFR and/or albuminuria definitions	- Low sensitivity - Considerable delay in identification - Sensitive to changes in testing and coding practices - Misclassification influenced by coding practices and purpose (e.g., reimbursement, pay for performance, and documentation in routine practice)
Single eGFR <60 ml/min per 1.73 m^2 Single UACR >30 mg/g	- High sensitivity - Minimal delay in identification	- Sensitive to changes in testing practices - Loss of information associated with dichotomizing the outcome by a certain threshold - Lacks confirmation of chronicity - May identify AKI or AKD instead of CKD - Testing for albuminuria is less frequent and may vary between specific patient groups; selected patient groups tested for UACR will be overrepresented
Two eGFRs <60 ml/min per 1.73 m^2 and/or 2 UACRs >30 mg/g at least 90 d apart	- Ensures chronicity - In accordance with guidelines - Acknowledges the criteria of kidney damage	- Delay/missed identification (requires regular testing in study population) - Sensitive to changes in testing practices - May identify patients with 2 episodes of AKI or dehydration. Additional condition "no eGFR >60 ml/min per 1.73 m^2 or UACR <30 mg/g during the CKD-defining period of at least 90 d" could minimize the risk of including such patients - A time limit (e.g., no more than 365 d apart) may need to be defined to target well-observed patients with CKD, in return for higher risk of missing patients with infrequent tests - Baseline for follow-up can only start at second measurement, resulting in survivor bias - Testing for albuminuria is less frequent and may vary between specific patient groups; selected patient groups tested for UACR will be overrepresented

AKD, acute kidney disease; AKI, acute kidney injury; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; UACR, urine albumin–to–creatinine ratio.

survivorship bias (Table 1). Analyses will need to use appropriate statistical methods (e.g., time-dependent exposure variables) or comparison groups (e.g., allowing for the same period of "run-in" immortal time for the non-CKD cohort) to align the start of follow-up and start of exposure to prevent immortal time bias and depletion of susceptibles bias.[16,48,53,54]

Adding albuminuria to eGFR to classify CKD. Most observational studies identify CKD cases on the basis of eGFR only.[55] At least a fifth of the populations with CKD remain understudied and uncharacterized because they have CKD category G1 and G2 and require either A2 or greater urine albumin–to–creatinine ratio (UACR) or other signs of renal damage (e.g., structural kidney disease) for identification.[56,57] Recent initiatives have been taken to improve patient identification in routine databases by developing conversion formulas between urinary protein-creatinine ratio or urinary dipstick protein to UACR.[58,59] However, even these tests are not universally performed, and even if such conversion is introduced, the tests are not fully comparable. Notably, urine dipstick analysis, which measures protein, but not creatinine, has generally high false-negative rates, and can also have a high false-positive rate in the general community setting when compared with more quantitative tests.[60] Finally, researchers have used different strategies for classification of CKD (e.g., the least severe, the most severe, the most recent, or the mean or median of eGFR or UACR level during the period used to define CKD).[61,62] Being transparent about and justifying the chosen definition are essential for the reader to understand the study's strengths, limitations, generalizability, and likely reproducibility.

Defining CKD progression

There is ample heterogeneity in how CKD progression is defined in epidemiologic studies, including both claims-related end points (kidney replacement therapy [KRT][53] or death attributed to CKD[63]), time to laboratory-based percentages of eGFR decline relative to baseline (typically 30%, 40%, 50%, or 57%),[64] time to doubling of serum creatinine,[65] eGFR values below a certain threshold (e.g., incident <60 or <15 ml/min per 1.73 m^2),[66] diagnostic coding for CKD,[67] longitudinal eGFR decline, and combinations of these in a composite outcome. Table 3 lists some of the methods used to define CKD progression and discusses pros and cons. The same challenges that apply to CKD ascertainment also apply herein. Because of space limitations, we will not discuss definitions of albuminuria progression, which can be ascertained by transition to a different "A" category or changes in continuous UACR over time. As explained earlier, the capacity to detect these outcomes depends on the type of testing

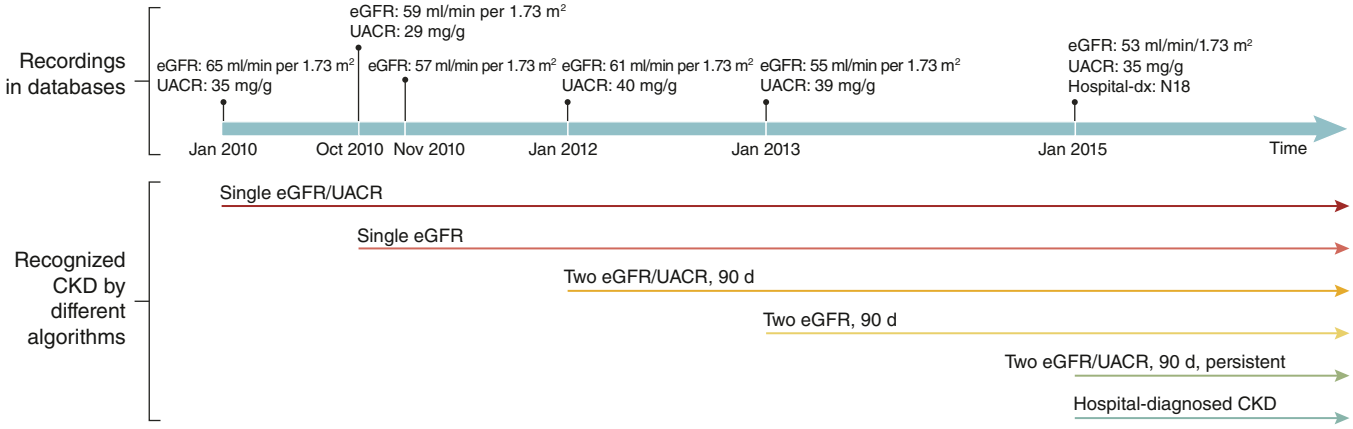


Figure 2 | The algorithm used to identify chronic kidney disease (CKD) influences when patients are included in the study. This is an example of a patient with recorded estimated glomerular filtration rate (eGFR) and urine albumin–to–creatinine ratio (UACR) tests and hospital diagnoses (dx; top), and common algorithms that would recognize the patients as having CKD based on the recordings (below). In this case example, there is a 5-year time gap between identification of CKD by the first and the last defining algorithm, during which the patients need to survive.

(categories of albuminuria by dipstick or continuous albuminuria concentration by UACR) and frequency of testing and is limited by the fact that testing tends to be directed toward people at higher risk.

Kidney replacement therapy. KRT has historically been the preferred outcome for observational studies, because it is assumed that the incidence of KRT is not affected by the frequency of laboratory testing or health care use, and it is clearly an outcome of great importance to patients.[8] Thus, bias due to differential outcome ascertainment is unlikely to occur when using KRT.[68] However, KRT is not the same as eGFR <15 ml/min per 1.73 m^2—many, especially older, patients with CKD G5 survive for a prolonged period without requiring dialysis, or choose not to have dialysis, either temporarily or as an enduring decision (which in clinical practice is sometimes termed conservative care). Hence, those who access KRT are a selected group of patients, and, depending on health system funding of KRT, may not represent the population burden of people most severely affected by kidney disease. Ascertainment of KRT episodes in administrative data requires algorithms to identify the date of the first chronic dialysis. Alternatively, data sources in selected countries often have linkages with national KRT registries. Finally, a potential disadvantage is that evaluating risks of KRT may require large sample size and long follow-up for sufficient power in low-risk populations.

Laboratory-based definition of CKD progression. Most studies use composite outcomes that incorporate creatinine- or eGFR-based definitions. As discussed above, the frequency, indication, and location of testing may pose a risk for differential outcome ascertainment. For question 2, this differential outcome ascertainment occurs when there are more creatinine measurements in the dipeptidyl peptidase 4 inhibitor arm than in the sodium-glucose cotransporter-2 inhibitor arm. The current consensus is to consider a 30% to 40% glomerular filtration rate decline as a surrogate end point for kidney failure for clinical trials of CKD progression,[69,70] and these are also often applied in observational studies. Finally, these surrogate outcomes are chosen with the idea of reflecting a clinically important event, but the dichotomization comes at the expense of loss of information and loss of power.[71]

Sustained declines of eGFR over time. Figure 3 illustrates the trajectory of outpatient eGFR measurements over time from selected participants in the SCREAM project. It can be easily observed that in some cases, reaching a certain threshold does not necessarily align with the behavior of the rest of kidney function measurements throughout the patient's journey. Nonrenal determinants of eGFR, intense periods of disease and testing, or even an AKI episode may falsely be identified as a doubling of creatinine or as a 30% eGFR decline from baseline. For instance, the incidence rate of CKD progression may be overestimated in question 2, when only 1 measurement below a certain threshold is required to be considered as the occurrence of an outcome.

A confirmation measurement (i.e., a decline in eGFR that is sustained over time) will improve the positive predictive value of the outcome, at a cost. How the scientific literature addresses this is variable, and often not reported. The considerations around confirmation, mentioned earlier in the indentation patients with CKD, apply herein; requiring the presence of consecutive measurements of a similar magnitude or relative eGFR reduction depends on health care access and testing,[72] and is not possible in the case of death.

In practice, researchers sometimes fit a linear regression line through the eGFR measurements that are available per individual to confirm a sustained decline and ascertain when a certain eGFR threshold is reached.[73,74] However, linear regression cannot be estimated well if only few measurements are available, and often patients with only 1 eGFR measurement during follow-up are excluded.[74] Furthermore, people may drop out owing to KRT or death. A better alternative is to

Table 3 | Advantages and disadvantages of different definitions of CKD progression when using routinely collected health care data

Definition	Advantages	Disadvantages
Diagnosis codes for more severe CKD stages	- High specificity, because clinically verified - Available in settings without laboratory data	- Low sensitivity - Considerable delay in identification (codes may not be updated regularly to reflect kidney function change) - Changes as coding practices/incentives change - Dependent on physician awareness, likely to be highest in patients who seek care more often - May distort measures of inequality if a particular group is less likely to be diagnosed (e.g., women and ethnic minorities).
Initiation of KRT	- Hard end point and of great importance to patients - Strongly related to cost of care - Low likelihood of differential outcome ascertainment	- May not be available without linkage to national registry - Subject to clinical judgement/practice variation - National registry may not capture all acute dialysis starters (typically only KRT rates for 90-d survivors are reported) - Will only capture those who are offered and elect to undergo dialysis/transplantation - Not valid in settings where economic inequalities and absence of funding make KRT unaffordable to many patients - In low-risk populations, too few events, resulting in underpowered study - In view of high competing mortality, less informative for early prevention efforts
eGFR <15 ml/min per 1.73 m^2 without confirmation	- Better proxy for kidney failure than KRT - Many patients with this level of kidney function will present to health services because of symptoms	- To distinguish new decline from undetected long-standing CKD, this can only be used in a population who undergoes repeated kidney function testing - Depends on who has access to test (setting and funding) - Susceptible to measurement error - May identify AKI instead of CKD - Interpretation in terms of cost implications/health burden can be different from the interpretation of KRT, particularly at older age
eGFR <15 ml/min per 1.73 m^2 with confirmation	- Better proxy for kidney failure than KRT - Includes conservative care - Applies in LMIC, where KRT may not be available or universally accessible	- To define an incident event, this requires a population that undergoes repeated kidney function testing and depends on who has access to test (setting and funding). - Competing mortality (high risk of death after first eGFR <15ml/min per 1.73 m^2). - Interpretation in terms of cost implications/health burden can be different from the interpretation of KRT, particularly at older age
Time to % eGFR decline (30%, 40%, 50%, or 57%) without confirmation	- More power and greater relevance for early prevention at higher CKD GFR stages - Larger eGFR declines better surrogate measure for kidney failure	- May identify AKI instead of CKD - Susceptible to measurement error - Some events are transient because of eGFR fluctuations - Loss of information associated with dichotomizing the outcome by a certain threshold
Time to % eGFR decline (30%, 40%, 50%, or 57%) with confirmation	- More power and greater relevance for early prevention at higher CKD GFR stages - Larger eGFR declines better proxy for kidney failure - More robust to transient changes in eGFR	- Same as above and also: - Delay in identification or failure to identify in case of death - Immortal time - Informative visit process as timing to next test driven by patient characteristics/comorbidity status
Linear interpolation and smoothing of eGFR slopes with linear regression	- Uses all measurements, so less sensitive to AKI or measurement error - Easy to implement - May be accurate when using prospective data with no dropout and at least 3 measurements per person	- Performance likely to be worse than linear mixed models in routinely collected data because of few measurements and dropout - If only few measurements are available, the slope cannot be estimated well and hence the time point of crossing the threshold cannot be precisely determined - Patients with only 1 measurement during follow-up are excluded - Gives biased estimates in the case of dropout due to kidney failure with replacement therapy or death
Longitudinal eGFR decline with linear mixed model	- Superior performance to linear regression - Uses all measurements, so less sensitive to AKI or measurement error - Can account for data missing at random - Can include patients with only 1 measurement or few measurements - Fitted model can be used to ascertain when a certain decline threshold was reached (sustained decline)	- Reasons for repeated testing can bias coefficients associated with random effects (severe bias only when all measurements are irregular); explicit modeling assumptions required to address competing mortality and informative censoring in joint models - Assumes linear eGFR decline, but the linearity assumptions can be relaxed by including appropriate transformations of time in the model.
Progression of albuminuria	- Part of KDIGO CKD definition - Often assessed in clinical trials - Formulas have been developed to convert urinary PCR or dipstick measurements to ACR	- Substantial bias by reasons for urine testing - High variability of albuminuria introduces substantial measurement error (difficult to interpret small changes at the individual level)

ACR, albumin-to-creatinine ratio; AKI, acute kidney injury; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; GFR, glomerular filtration rate; KDIGO, Kidney Disease: Improving Global Outcome; KRT, kidney replacement therapy; LMIC, low- and middle-income countries; PCR, protein-to-creatinine ratio.

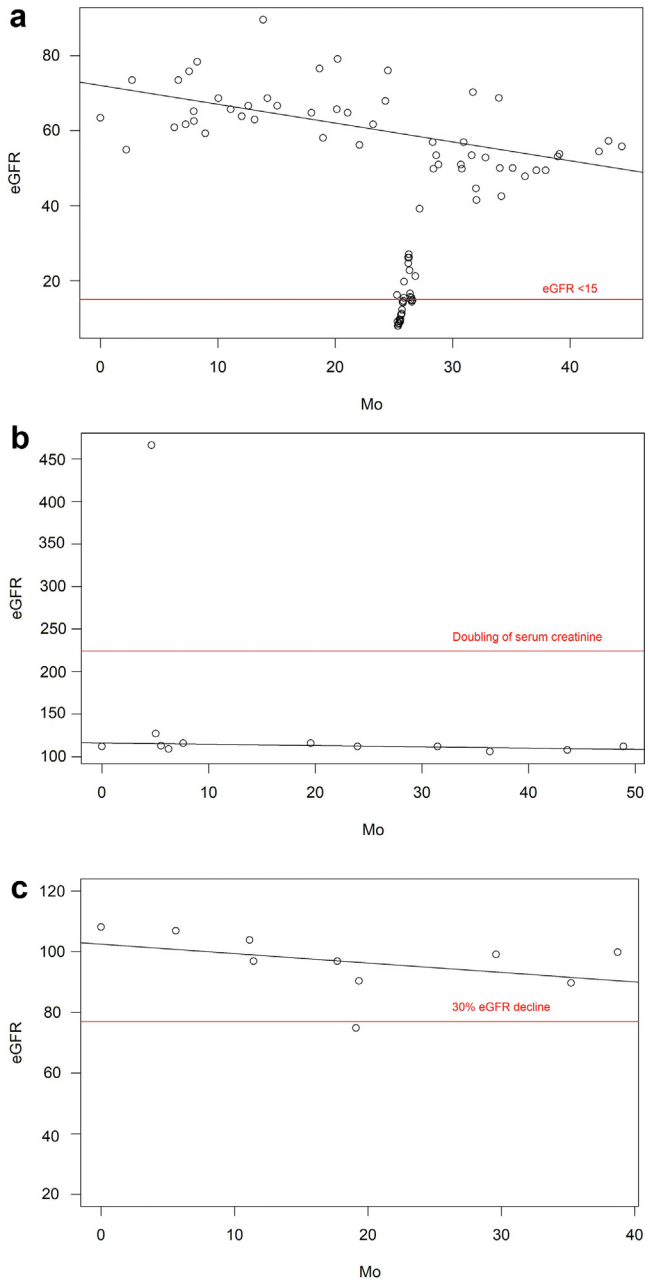


Figure 3 | Plots of outpatient estimated glomerular filtration rate (eGFR) or creatinine measurements from 3 individuals in the Stockholm Creatinine Measurements database. (**a**) The red line depicts an eGFR <15 ml/min per 1.73 m^2. This individual has many measurements during follow-up. At 25 months, there are many low eGFR measurements, which may represent acute illness or an acute kidney injury (AKI). Identifying simply a decline of >30% from baseline as study outcome would misclassify this AKI as a chronic kidney disease (CKD) progression event, leading to a biased estimate of the incidence of CKD progression in the study population. (**b**) The individual has many creatinine measurements after baseline, with 1 creatinine measurement surpassing the threshold of doubling of serum creatinine (red line). However, this likely does not reflect a "true" doubling of serum creatinine. Note that the y-axis suggests serum creatinine (umol/L). (**c**) The individual has 10 eGFR measurements, with 1 measurement below the threshold of 30% eGFR decline. On the basis of the global information that we have for this patient, it seems a random observation possibly influenced by disease or hydration status.

use a linear mixed model with random intercepts and slopes.[74–76] The use of random effects allows for "borrowing" of information across individuals. These models increase stability for patients for whom only few measurements are available and can include information from individuals with only 1 follow-up eGFR assessment. Furthermore, under certain conditions, mixed models can even handle informative missingness if the predictors of missingness are included as covariates in the model.[77] It is recommended to add the baseline eGFR value to the outcome vector. Linearity assumptions can be easily relaxed by including appropriate transformations of time in the model, such as quadratic terms or splines. The fitted mixed model can then be used to ascertain when a certain decline threshold was reached. In this definition, reaching the interpolated threshold will be sustained over time, and may identify outcomes that theoretically occurred earlier, during a period when laboratory testing was not frequent. One can easily see how interpolation with smoothing of the eGFR slope can also serve to improve the identification of patients with "confirmed" CKD.

Considerations for choosing AKI definitions

The scientific literature reports varying algorithms to define AKI in health care databases (Table 4).[78] Below, we summarize common features of these definitions that require attention as they can affect the generalizability or validity of study findings. We illustrate this using our third example question: "After AKI, what is the causal effect of stopping versus continuing renin-angiotensin system inhibitors on the risk of recurrent AKI?"[79] For this particular research question, the inception episode of AKI defines the population of interest, the next episode is the outcome of interest, and history of AKI before inception could be used as a covariate to adjust for confounding. Most of the considerations discussed below likely also apply when evaluating the newly defined entity of acute kidney disease,[3,4] but few studies to date have explored acute kidney disease in health care databases.[80]

Diagnostic coding of AKI. Hospital-recorded AKI diagnoses are often included in health care databases and coded with *International Classification of Diseases* codes, and these are often used to identify AKI populations, as an outcome or as a covariate. Although the specificity of the hospital-recorded diagnoses is high (>95%), the coding is incomplete and may only identify a quarter to a third of all AKI episodes identified by changes in serum creatinine,[10–13] even fewer when considering all cases, including those defined by oliguria.[13] The reason for coding (e.g., reimbursement, pay for performance, or documentation in routine practice) may also impact the validity of codes. For our specific example, using AKI diagnosis codes will lead to a selective population of more severe AKI cases, which may impact the generalizability of results: findings may not be necessarily generalized to the complete AKI population, and would also include more severe AKI cases (i.e., stage 3 AKIs are more likely to lead to a diagnostic code compared with stage 1 AKI).

Table 4 | Advantages and disadvantages of different definitions of AKI used in previous studies based on routine care data

Definition	Advantages	Disadvantages
Diagnosis codes	- Available in settings without laboratory registries - High specificity for severe AKI and AKI requiring dialysis	- Low sensitivity for AKI, especially for less severe stages - Quality of coding relies on the specific health care setting, changes in diagnostic criteria, and coding practices over time - AKI during elective admissions is less likely to be captured compared with admissions where AKI was the reason for hospitalization[78] - Misclassification influenced by coding practices and purpose (e.g., reimbursement, pay for performance, and documentation in routine practice)
KDIGO serum creatinine criteria	- Possible to separate AKI from prevalent CKD when a valid baseline serum creatinine is available - When definitions are harmonized, comparable standardized incidence rates of AKI across populations, allowing for direct comparison between studies	- Inpatient tests cannot distinguish AKI from preexisting CKD - Outpatient tests may be missing - Choice of numerous baseline serum creatinine definitions - Sensitive to changes in testing practices
KDIGO urine output criteria	- Research indicates that short- and long-term risk of death or KRT is greatest when patients meet both serum creatinine and urine output criteria for AKI	- Seldom captured in administrative data, and rarely available outside the ICU

AKI, acute kidney injury; CKD, chronic kidney disease; ICU, intensive care unit; KDIGO, Kidney Disease: Improving Global Outcomes; KRT, kidney replacement therapy.

When AKI is the outcome, using AKI diagnoses may impact the validity of findings. Whether bias occurs depends on the specificity and sensitivity of the outcome definition, as well as whether the interest lies in relative or absolute risks. When the specificity is high (i.e., the probability of not having AKI among those who truly do not have AKI equals 1 for both the exposed and unexposed), as is the case when using AKI diagnoses, relative risk estimates will not be biased, even if the sensitivity is low (i.e., probability of recorded AKI among those who truly have AKI).[81] However, absolute risk estimates will be biased, leading to an underestimation of the absolute risk difference of all AKI cases. As discussed earlier in the section on CKD progression, bias will also occur if the measurement error in the outcome is differential with respect to the exposure (e.g., in question 3, this will occur if physicians suspect that RASi use causes recurrent AKI and therefore monitor patients who continue RASi more closely than patients who stop RASi).[82] Such differential measurement error across exposure groups may be less likely for severe AKI (as ascertained by diagnosis codes), because these will be recorded regardless of exposure status.

When history of AKI is a confounder, as in the example question 3, using diagnosis codes may lead to residual confounding.[21] Whenever a patient had an AKI that was not severe enough to be coded, this measurement error leads to residual confounding when the prescriber was aware of the history of AKI, and bases his/her prescribing decision (stopping vs. continuing RASi) on this.

Defining AKI cases by urine output. The 2012 KDIGO classification of AKI is currently widely used for both clinical and research purposes.[2] Using these criteria fully (i.e., considering both changes in serum creatinine and urine output) is recommended in clinical practice, because short- and long-term risk of death and KRT is greatest when patients meet both criteria.[83] However, this level of detail (e.g., hourly urine output) is not easily accessible in many routine health care databases, limiting their use in epidemiologic studies.[84,85] In most electronic health care records, urine output and point-of-care creatinine measurements are added to the electronic health records as unstructured text, which will hamper accurate extraction, although this problem may be mitigated by using natural language processing to extract and classify this information from unstructured texts. The same considerations regarding generalizability and bias as discussed above for AKI diagnosis codes apply for urine output.

AKI based on creatinine. Although creatinine measurements are preferred to diagnosis codes, certain challenges arise when using routinely collected data sources. The KDIGO criteria for diagnosis of AKI in clinical practice refer to a relative increase in serum or plasma creatinine of $\geq$1.5, known or presumed to have occurred within the prior 7 days, or an absolute increase of $\geq$0.3 mg/dl ($\geq$26.5 μmol/l) within 48 hours. To avoid the influence of acute illness, outpatient serum creatinine tests are preferred sources to establish the baseline creatinine. Ideally, serum creatinine would need to have been measured within 7 days before AKI onset for detection of AKI.[62] However, this is seldom the case, except for situations such as planned surgeries. Researchers are then left with the option of applying different windows to identify prior serum creatinine measurements to define as to who had AKI.[62,86–89] Nevertheless, for a proportion of patients, a creatinine test within the specified period will be lacking.

A recent scoping review confirmed a lack of consistency in how KDIGO definitions for AKI were used in epidemiologic studies; for instance, the window to ascertain the baseline creatinine ranged from 0 days to more than a year before the AKI. More concerning, however, was the absence of description of the process used in 33% of the identified studies.[85]

If >1 eligible creatinine test is available per patient, it is unclear whether the preferred approach would be to select the most recent serum creatinine,[86,90,91] the median,[92,93] or the mean,[94] of all eligible tests, or to model the slope of creatinine and select its intersection with the 7-day period of interest. In 1 study, the mean outpatient serum creatinine measured in the year before hospitalization most closely approximated nephrologist-adjudicated "baseline" serum creatinine values.[62]

Because of lack of testing, most studies in the literature (71%) opt to exclude individuals who lack a baseline creatinine test.[85] Strategies used in previous literature to estimate baseline serum creatinine, when not measured, include simple or multiple imputation, using the serum creatinine at admission,[95] assuming an eGFR of 75 ml/min per 1.73 m^2,[96] or using a post-AKI nadir value.[97] Studies comparing these approaches suggest that multiple imputation is superior to simple imputation or assuming an eGFR of 75 ml/min per 1.73 m^2.[98] Using a nadir serum creatinine during hospitalization as baseline may lead to incorrect detection of AKI, because serum creatinine in the inpatient setting is influenced by nonrenal factors, such as fluid accumulation and loss of muscle mass.[99,100] Using the first serum creatinine on admission could result in AKI episodes being missed if serum creatinine was already elevated on admission. However, time lag between a kidney insult (due to an acute illness) and serum creatinine elevation should be acknowledged: it may take up to 48 to 72 hours after the kidney insult happened for creatinine to increase.[101] Indeed, a US study showed that the first inpatient serum creatinine was not higher than the most recent outpatient serum creatinine in a large proportion of hospitalized patients with AKI.[62] However, this may vary by cause of hospitalization. For example, in health systems with rapid admission for an acute ST-elevation myocardial infarction within hours of onset of chest pain, serum creatinine elevation would be only visible after admission. However, for admissions with infections and other conditions that gradually develop over several days, serum creatinine may be already elevated on admission.

Box 1 summarizes recommendations for clearly reporting the time frame for eligible baseline creatinine values and the rationale for doing so, how missing baseline creatinine values are handled, and the method chosen to select the baseline eGFR when there are multiple eligible values within the defining window, with reference to a recent consensus by a Delphi panel composed of nephrologists and epidemiologists with experience in AKI research.[85] A recent study showed that harmonizing AKI definitions across 4 population-based databases produced comparable standardized incidence rates of AKI.[102]

Moving forward: toward more robust estimations

The longitudinal analysis of routinely collected health care data relies on the assumption that the timing and frequency of the measurement of longitudinal outcomes should be independent of the value of the outcome itself.[103–105] Understanding the extent to which this assumption is violated is important; patients will visit the physician when they have been feeling ill and hence have worse biomarker values; patients with comorbidities are likely to have more health care visits than patients without comorbidities. It becomes apparent that observations and outcomes are dependent, and thus missing laboratory tests are not completely at random. This has been referred to as "informative presence," or alternatively, "informative visit process," "dynamic observation plans,"[106] or "outcome-dependent visits" and is an aspect often ignored in research practice and can be considered a form of information bias.[107,108]

Relatively simple analyses can be performed to assess the magnitude of effects owing to informative visits in the data set. *First*, when the data set contains information on whether a visit is scheduled or unscheduled, the longitudinal eGFR slope can be calculated separately for scheduled and unscheduled visits. A substantial difference between slopes is suggestive of an informative visit process.[105] *Second*, one can calculate the correlation between a subject's eGFR value at a certain time point and the time between this measurement and the next, for all measurements.[105] Alternatively, the number of visits can be compared between individuals with a high or low eGFR. *Third*, when comparing 2 different interventions (e.g., question 2, dipeptidyl peptidase 4 inhibitors vs. sodium-glucose cotransporter-2 inhibitors), differential outcome ascertainment may be assessed by comparing the proportion of individuals with at least 1 creatinine measurement, the rate of creatinine testing during periods of treatment,[109,110] or the average time gap between tests. *Finally*, recurrent events models (such as the Andersen-Gill model) could be used to quantify the association between study covariates and the rate of observation.[111] Overall, simply reporting the number of visits per patient, gaps between visits, and potential predictors of visit time can give the reader an indication of the extent of irregularity and its informativeness.[107]

Some strategies may serve to mitigate the bias introduced by outcome-dependent visits; applying an active comparator design might yield a reference group with similar observation and dropout patterns, as described elsewhere, provided that testing rates are similar.[112–114] Bias can be attenuated when a certain proportion of the sample contains noninformative, regularly planned visits.[77,104,105,115,116] In many cohorts, at least part of the visits will be regular. If information is available on whether visits are planned or unplanned, the analysis could be restricted to preplanned visits to yield a cohort of subjects where the information process is independent of disease severity. Another option is to restrict the analytical sample to a population with an indication for regular kidney function monitoring (e.g., patients with diabetes),[33] at the expense of the external validity or generalizability of the study findings. The large sample sizes of health

Box 1 | Reporting recommendations when studying AKI/AKD

1. Studies should describe the intended target population (all patients with AKI? only diagnosed/severe AKI?), and whether study results are generalizable to that target population.
2. Studies should clarify how populations with/without baseline creatinine results differ (sample size, characteristics, and setting of testing [i.e., outpatient vs. inpatient results]), and the timing of the baseline creatinine relative to the AKI precipitating event.
3. Studies should clearly report the AKI definition used (e.g., whether 0.3 mg/dl increase over 48 hours is included [required for full alignment with the KDIGO AKI definition], whether staging criteria for stages 1, 2, and 3 are used, and whether urine output criteria were included).
4. Studies should clarify the definition of a baseline creatinine if multiple baseline creatinine results were available (e.g., was the mean of measurements used or the latest measurement? were measurements <7 days before AKI discarded?).
5. Studies should clearly report what was done whenever baseline results were not available. If studies impute missing baseline creatinine tests, they should specify methods used and discuss the implications of this imputation on study findings.

AKD, acute kidney disease; AKI, acute kidney injury; KDIGO, Kidney Disease: Improving Global Outcomes.

care data sets often allow for this type of selection, while still retaining power. Comparing results across different health systems with different testing indications is also helpful.

Alternatively, various methods have been proposed to accommodate an informative observation process or dropout in the study, such as, for example, due to referral to specialist care not covered by the database studied[103,117]; herein, we will discuss briefly approaches based on inverse probability weighting and approaches that aim to fully model the different processes. Methods based on inverse probability weighting rely on the idea of weighting each observation by the inverse probability of each measurement to be recorded or for the inverse probability of nondropout from the study; consequently, this approach creates a pseudopopulation in which the observation or the dropout process is static (rather than dynamic); that is, the process is completely at random and can, therefore, be ignored.[106,118,119] This can be implemented in practice using standard, off-the-shelf statistical software, or using user-contributed packages, such as the IrregLong package in R.[120]

Furthermore, it is possible to fully model the observation and/or dropout process within the joint longitudinal-survival modeling framework.[77,121–123] Joint models consist of 2 "submodels": 1 to model the survival outcome (i.e., the observation/dropout process), and the other to model the longitudinal outcome (i.e., the longitudinal kidney function measurements). A survival model is used to model the survival outcome, and a linear mixed model is used for the longitudinal outcome. The submodels are generally linked via shared random effects and estimated jointly. Focusing on dropout, the joint modeling approach can accommodate informative truncation of longitudinal trajectories due to dropout (e.g., death). Similar to linear mixed models, the baseline value should be part of the outcome vector as it contributes to estimating the measurement error. Compared with the inverse probability weighting approach, the joint modeling approach has the advantage of explicitly modeling all the processes of interest, allowing joint inference on the different aspects of the problem under study. However, this approach has the disadvantage of being more computationally intensive, limiting its applicability (especially with large data sets) as well as needing to specify the shared random effects correctly; the joint modeling approach can be implemented using readily available statistical software in R and Stata (e.g., the merlin package[124]).

Future directions

Observational research of kidney disease has made great strides with studies including larger populations, more sophisticated

Box 2 | Key points a causal study should consider discussing when using routinely collected health care data to study populations with CKD or CKD progression as an outcome

- Investigate and discuss to what extent study results are generalizable to the target population in the context of the definition used to identify populations with CKD (e.g., based on diagnosis codes, eGFR measurements, and UACR measurements).
- Investigate and discuss the potential for differential outcome ascertainment (e.g., check whether more kidney function measurements are performed in 1 exposure group).
- Investigate the impact of exposure misclassification in the context of the definition used (diagnosis codes and eGFR based on serum creatinine).
- Discuss the key potential confounders of the exposure-outcome relationship, and discuss potential residual confounding (e.g., owing to disease severity or misclassification). Investigate the presence of residual confounding through positive or negative control outcomes, and its impact with quantitative bias analysis.
- When using eGFR measurements to classify CKD, discuss how multiple measurements are handled (mean, median, and most recent).
- When using eGFR or UACR as adjustment variables for confounding, discuss how patients with missing data were handled (complete case analysis, multiple imputation, and weighting). Discuss the possibility for selection bias when a complete case analysis is performed, also in light of the pattern of missingness and the proportion of patients with missing data.
- What are the data sources from which kidney function information for individuals was obtained? Does the chosen database adequately capture all kidney function measurements? Discuss consequences of data fragmentation on study results, including loss to follow-up.
- The use of a diagram is recommended to illustrate key aspects of the study design(s), including study entry, exposure, confirmation of exposure, comparison groups, lag and observation periods, and covariate definitions as relevant.

CKD, chronic pkidney disease; eGFR, estimated glomerular filtration rate; UACR, urine albumin–to–creatinine ratio.

analytic methods, individual-level meta-analysis, and sensitivity analyses to help gauge the validity of the results. However, as shown in this review, the field will benefit from more transparent and structured reporting, and thoughtful acknowledgement and discussion of potential biases.

Given that the suitability of each data set will depend on the research question and local structural factors, it may not be possible to impose a single strict definition that suits all studies that use routine health care data. However, validation studies are helpful within specific health systems to investigate the local sensitivity and specificity of these definitions. We advocate for concerted efforts to encourage improved reporting practices for routinely collected data on kidney exposures and outcomes.

The REporting of studies Conducted using Observational Routinely collected health Data for PharmacoEpidemiology (RECORD-PE) guidelines were produced as part of an international collaboration to improve such practice, building on the existing Strengthening the Reporting of Observational Studies in Epidemiology and RECORD guidelines.[17] We encourage researchers and editors of scientific journals to use this template to guide reporting of definitions of kidney exposures and outcomes (Supplementary Table S2). Key information that an article should address is outlined in Box 2.

Several specific approaches that are referred to are worth highlighting:

- Study design diagrams may serve to effectively illustrate algorithms used to define exposure or outcome for kidney function (including any sensitivity analyses).[125]
- Sensitivity analyses can be used to examine the assumptions underlying the chosen analytical approaches.
- Triangulation approaches to address biases can enhance causal inference.[126] This necessitates analyses in a range of settings and the integration of results from several approaches, each prone to their own different and unrelated sources of potential bias, to qualitatively determine and explicitly articulate the strength of evidence; examples include cross-context comparisons, such as different study populations, which would be expected to introduce their own inherent biases.
- Use of directed acyclic graph is recommended to consider bias, such as selection bias, and confounding.[127]
- Open working methods mandate open sharing of all analysis codes to encourage a culture of external review, reuse, and collaboration using a given source of data.[128] Similar approaches could be used for other large routinely collected data in other settings to enhance transparency and replication of analyses to enhance trust in research findings.

In conclusion, the perfect definition of kidney exposures, covariates, or outcomes using routinely collected data depends on the research question and availability of data, but clearer and more transparent reporting of these decisions in observational research is necessary to move the field forward.

DISCLOSURE

All the authors declared no competing interests.

ACKNOWLEDGMENTS

JJC acknowledges support from the Swedish Research Council. ELF acknowledges support by a Rubicon Grant of the Netherlands Organization for Scientific Research (NWO). MMS is funded by the Jindal Research Chair for the Prevention of Kidney Disease. CFC acknowledges support from the Independent Research Fund Denmark. VM is funded by the National Institutes for Health and Care Research (NIHR). FC is funded by the Medical Research Council (MR/N013638/1). JC acknowledges support from the National Institutes of Health and US National Kidney Foundation.

SUPPLEMENTARY MATERIAL

Supplementary File (PDF)

Table S1. Fragmentation of care in different health care settings of the world.

Table S2. Adapted REporting of studies Conducted using Observational Routinely collected health Data for PharmacoEpidemiology (RECORD-PE) checklist for electronic health record studies involving measurement and/or classification of kidney function for exposure or outcome.

Figure S1. Directed acyclic graphs showing selection bias due to conditioning on a collider for the different example questions.

Supplementary References.

REFERENCES

1. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl*. 2013;3:1–150.
2. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO clinical practice guideline for acute kidney injury. *Kidney Int Suppl*. 2012;2:1–138.
3. Chawla LS, Bellomo R, Bihorac A, et al. Acute kidney disease and renal recovery: consensus report of the Acute Disease Quality Initiative (ADQI) 16 Workgroup. *Nat Rev Nephrol*. 2017;13:241–257.
4. Lameire NH, Levin A, Kellum JA, et al. Harmonizing acute and chronic kidney disease definition and classification: report of a Kidney Disease: Improving Global Outcomes (KDIGO) Consensus Conference. *Kidney Int*. 2021;100:516–526.
5. Stevens LA, Greene T, Levey AS. Surrogate end points for clinical trials of kidney disease progression. *Clin J Am Soc Nephrol*. 2006;1:874–884.
6. Perkovic V, Koitka-Weber A, Cooper ME, et al. Choice of endpoint in kidney outcome trials: considerations from the EMPA-REG OUTCOME® trial. *Nephrol Dial Transplant*. 2020;35:2103–2111.
7. Levin A, Agarwal R, Herrington WG, et al. International consensus definitions of clinical trial outcomes for kidney failure: 2020. *Kidney Int*. 2020;98:849–859.
8. Standardised Outcomes in Nephrology (SONG). The SONG initiative. Accessed November 10, 2021. https://songinitiative.org/
9. Shrier I, Stovitz SD, Wang C, Steele RJ. Beware of collider stratification bias when analyzing recurrent injuries. *Scand J Med Sci Sports*. 2022;32:270–272.
10. Logan R, Davey P, De Souza N, et al. Assessing the accuracy of *ICD-10* coding for measuring rates of and mortality from acute kidney injury and the impact of electronic alerts: an observational cohort study. *Clin Kidney J*. 2020;13:1083–1090.
11. Hwang YJ, Shariff SZ, Gandhi S, et al. Validity of the *International Classification of Diseases, Tenth Revision* code for acute kidney injury in elderly patients at presentation to the emergency department and at hospital admission. *BMJ Open*. 2012;2:e001821.
12. Waikar SS, Wald R, Chertow GM, et al. Validity of *International Classification of Diseases, Ninth Revision, Clinical Modification* codes for acute renal failure. *J Am Soc Nephrol*. 2006;17:1688–1694.
13. Grams ME, Waikar SS, MacMahon B, et al. Performance and limitations of administrative data in the identification of AKI. *Clin J Am Soc Nephrol*. 2014;9:682–689.
14. Last JM. *A Dictionary of Epidemiology*. 4th ed. Oxford University Press; 2000.
15. Heinze G, Wallisch C, Kainz A, et al. Chances and challenges of using routine data collections for renal health care research. *Nephrol Dial Transplant*. 2015;30(suppl 4):iv68–iv75.

16. Fu EL, Evans M, Carrero JJ, et al. Timing of dialysis initiation to reduce mortality and cardiovascular events in advanced chronic kidney disease: nationwide cohort study. *BMJ*. 2021;375:e066306.
17. Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018;363:k3532.
18. Toh ST, Andrade SE, Raebel MA, et al. Examples of existing automated databases. In: Strom BL, Kimmel SE, Hennessy S, eds. *Textbook of Pharmacoepidemiology*. 2nd ed. Wiley Blackwell; 2013:123–177.
19. Gerhard T, Moride Y, Pottegård A, Pratt N. Encounter databases. In: Strom BL, Kimmel SE, Hennessy S, eds. *Pharmacoepidemiology*. 6th ed. Wiley Blackwell; 2019:211–240.
20. Wang SV, Pinheiro S, Hua W, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ*. 2021;372:m4856.
21. Funk MJ, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. *Curr Epidemiol Rep*. 2014;1:175–185.
22. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.
23. Runesson B, Gasparini A, Qureshi AR, et al. The Stockholm CREAtinine Measurements (SCREAM) project: protocol overview and regional representativeness. *Clin Kidney J*. 2016;9:119–127.
24. Chesnaye NC, Stel VS, Tripepi G, et al. An introduction to inverse probability of treatment weighting in observational research. *Clin Kidney J*. 2022;15:14–20.
25. Shin JI, Chang AR, Grams ME, et al. Albuminuria testing in hypertension and diabetes: an individual-participant data meta-analysis in a global consortium. *Hypertension*. 2021;78:1042–1052.
26. Bello AK, Ronksley PE, Tangri N, et al. Quality of chronic kidney disease management in Canadian primary care. *JAMA Netw Open*. 2019;2:e1910704.
27. Nitsch D, Caplin B, Hull S, Wheeler DC. National chronic kidney disease audit national report. Accessed November 10, 2021. https://www.lshtm.ac.uk/files/ckd_audit_report.pdf
28. Fu EL, Clase CM, Evans M, 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*. 2021;77:719–729.e1.
29. Hughes RA, Heron J, Sterne JAC, Tilling K. Accounting for missing data in statistical analyses: multiple imputation is not always the answer. *Int J Epidemiol*. 2019;48:1294–1304.
30. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338:b2393.
31. Madley-Dowd P, Hughes R, Tilling K, Heron J. The proportion of missing data should not be used to guide decisions on multiple imputation. *J Clin Epidemiol*. 2019;110:63–73.
32. Blazek K, van Zwieten A, Saglimbene V, Teixeira-Pinto A. A practical guide to multiple imputation of missing data in nephrology. *Kidney Int*. 2021;99:68–74.
33. McDonald HI, Shaw C, Thomas SL, et al. Methodological challenges when carrying out research on CKD and AKI using routine electronic health records. *Kidney Int*. 2016;90:943–949.
34. Anderson J, Glynn LG. Definition of chronic kidney disease and measurement of kidney function in original research papers: a review of the literature. *Nephrol Dial Transplant*. 2011;26:2793–2798.
35. Paik JM, Patorno E, Zhuo M, et al. Accuracy of identifying diagnosis of moderate to severe chronic kidney disease in administrative claims data. *Pharmacoepidemiol Drug Saf*. 2022;31:467–475.
36. Gasparini A, Evans M, Coresh J, et al. Prevalence and recognition of chronic kidney disease in Stockholm healthcare. *Nephrol Dial Transplant*. 2016;31:2086–2094.
37. Vestergaard SV, Christiansen CF, Thomsen RW, et al. Identification of patients with CKD in medical databases: a comparison of different algorithms. *Clin J Am Soc Nephrol*. 2021;16:543–551.
38. Tuttle KR, Alicic RZ, Duru OK, et al. Clinical characteristics of and risk factors for chronic kidney disease among adults and children: an analysis of the CURE-CKD Registry. *JAMA Netw Open*. 2019;2:e1918169.
39. van Oosten MJM, Brohet RM, Logtenberg SJJ, et al. The validity of Dutch health claims data for identifying patients with chronic kidney disease: a hospital-based study in the Netherlands. *Clin Kidney J*. 2020;14:1586–1593.
40. van Oosten MJM, SJJ Logtenberg, Edens MA, et al. Health claims databases used for kidney research around the world. *Clin Kidney J*. 2021;14:84–97.
41. Roy L, Zappitelli M, White-Guay B, et al. Agreement between administrative database and medical chart review for the prediction of chronic kidney disease G category. *Can J Kidney Health Dis*. 2020;7:2054358120959908.
42. Castro AF, Coresh J. CKD surveillance using laboratory data from the population-based National Health and Nutrition Examination Survey (NHANES). *Am J Kidney Dis*. 2009;53(suppl 3):S46–S55.
43. Brück K, Jager KJ, Dounousi E, et al. Methodology used in studies reporting chronic kidney disease prevalence: a systematic literature review. *Nephrol Dial Transplant*. 2015;30(suppl 4):iv6–iv16.
44. Tuot DS, McCulloch CE, Velasquez A, et al. Impact of a primary care CKD registry in a US public safety-net health care delivery system: a pragmatic randomized trial. *Am J Kidney Dis*. 2018;72:168–177.
45. de Boer IH, Kovesdy CP, Navaneethan SD, et al. Pragmatic clinical trials in CKD: opportunities and challenges. *J Am Soc Nephrol*. 2016;27:2948–2954.
46. Navaneethan SD, Jolly SE, Sharp J, et al. Electronic health records: a new tool to combat chronic kidney disease? *Clin Nephrol*. 2013;79:175–183.
47. Navaneethan SD, Jolly SE, Schold JD, et al. Development and validation of an electronic health record-based chronic kidney disease registry. *Clin J Am Soc Nephrol*. 2011;6:40–49.
48. Hernan MA, Sauer BC, Hernandez-Diaz S, et al. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70–75.
49. Shariff SZ, Cuerden MS, Jain AK, Garg AX. The secret of immortal time bias in epidemiologic studies. *J Am Soc Nephrol*. 2008;19:841–843.
50. Tam-Tham H, Quinn RR, Weaver RG, et al. Survival among older adults with kidney failure is better in the first three years with chronic dialysis treatment than not. *Kidney Int*. 2018;94:582–588.
51. Fu EL, van Diepen M. Comment on Kwon et al: the long-term effects of metformin on patients with type 2 diabetic kidney disease: *Diabetes Care* 2020;43:948–955. *Diabetes Care*. 2020;43:e190.
52. Stovitz SD, Banack HR, Kaufman JS. "Depletion of the susceptibles" taught through a story, a table and basic arithmetic. *BMJ Evid Based Med*. 2018;23:199.
53. 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;32:424–435.
54. Hernan MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183:758–764.
55. Delanaye P, Glassock RJ, De Broe ME. Epidemiology of chronic kidney disease: think (at least) twice! *Clin Kidney J*. 2017;10:370–374.
56. Ocak G, Khairoun M, Khairoun O, et al. Chronic kidney disease and atrial fibrillation: a dangerous combination. *PLoS One*. 2022;17:e0266046.
57. Hill NR, Fatoba ST, Oke JL, et al. Global prevalence of chronic kidney disease - a systematic review and meta-analysis. *PLoS One*. 2016;11:e0158765.
58. Sumida K, Nadkarni GN, Grams ME, et al. Conversion of urine protein-creatinine ratio or urine dipstick protein to urine albumin-creatinine ratio for use in chronic kidney disease screening and prognosis: an individual participant-based meta-analysis. *Ann Intern Med*. 2020;173:426–435.
59. Weaver RG, James MT, Ravani P, et al. Estimating urine albumin-to-creatinine ratio from protein-to-creatinine ratio: development of equations using same-day measurements. *J Am Soc Nephrol*. 2020;31:591–601.
60. White SL, Yu R, Craig JC, et al. Diagnostic accuracy of urine dipsticks for detection of albuminuria in the general community. *Am J Kidney Dis*. 2011;58:19–28.
61. Mendu ML, Ahmed S, Maron JK, et al. Development of an electronic health record-based chronic kidney disease registry to promote population health management. *BMC Nephrol*. 2019;20:72.
62. Siew ED, Ikizler TA, Matheny ME, et al. Estimating baseline kidney function in hospitalized patients with impaired kidney function. *Clin J Am Soc Nephrol*. 2012;7:712–729.
63. Pasternak B, Wintzell V, Melbye M, et al. Use of sodium-glucose co-transporter 2 inhibitors and risk of serious renal events: Scandinavian cohort study. *BMJ*. 2020;369:m1186.
64. Hundemer GL, Knoll GA, Petrcich W, et al. Kidney, cardiac, and safety outcomes associated with alpha-blockers in patients with CKD: a population-based cohort study. *Am J Kidney Dis*. 2021;77:178–189.e1.

65. Fu EL, Franko MA, Obergfell A, et al. High-sensitivity C-reactive protein and the risk of chronic kidney disease progression or acute kidney injury in post-myocardial infarction patients. *Am Heart J*. 2019;216:20–29.
66. Novak M, Mucsi I, Rhee CM, et al. Increased risk of incident chronic kidney disease, cardiovascular disease, and mortality in patients with diabetes with comorbid depression. *Diabetes Care*. 2016;39:1940–1947.
67. Wetmore JB, Yan H, Herzog CA, et al. CKD progression in Medicare beneficiaries with nonvalvular atrial fibrillation treated with apixaban versus warfarin. *Am J Kidney Dis*. 2021;78:180–189.
68. Hodlmoser S, Carrero JJ, Kurnikowski A, et al. Kidney function, kidney replacement therapy, and mortality in men and women. *Kidney Int Rep*. 2022;7:444–454.
69. Levey AS, Inker LA, Matsushita K, et al. GFR decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration. *Am J Kidney Dis*. 2014;64:821–835.
70. Levey AS, Gansevoort RT, Coresh J, et al. Change in albuminuria and GFR as end points for clinical trials in early stages of CKD: a scientific workshop sponsored by the National Kidney Foundation in collaboration with the US Food and Drug Administration and European Medicines Agency. *Am J Kidney Dis*. 2020;75:84–104.
71. Ragland DR. Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power on the cutpoint. *Epidemiology*. 1992;3:434–440.
72. Neuen BL, Weldegiorgis M, Herrington WG, et al. Changes in GFR and albuminuria in routine clinical practice and the risk of kidney disease progression. *Am J Kidney Dis*. 2021;78:350–360.e1.
73. Zee J, Mansfield S, Mariani LH, Gillespie BW. Using all longitudinal data to define time to specified percentages of estimated GFR decline: a simulation study. *Am J Kidney Dis*. 2019;73:82–89.
74. Leffondre K, Boucquemont J, Tripepi G, et al. Analysis of risk factors associated with renal function trajectory over time: a comparison of different statistical approaches. *Nephrol Dial Transplant*. 2015;30:1237–1243.
75. Heinze G, Christensen J, Haller MC. Modeling pulse wave velocity trajectories-challenges, opportunities, and pitfalls. *Kidney Int*. 2022;101:459–462.
76. Janmaat CJ, van Diepen M, Tsonaka R, et al. Pitfalls of linear regression for estimating slopes over time and how to avoid them by using linear mixed-effects models. *Nephrol Dial Transplant*. 2019;34:561–566.
77. Gasparini A, Abrams KR, Barrett JK, et al. Mixed-effects models for health care longitudinal data with an informative visiting process: a Monte Carlo simulation study. *Stat Neerl*. 2020;74:5–23.
78. Savino M, Plumb L, Casula A, et al. Acute kidney injury identification for pharmacoepidemiologic studies: use of laboratory electronic acute kidney injury alerts versus electronic health records in hospital episode statistics. *Pharmacoepidemiol Drug Saf*. 2021;30:1687–1695.
79. Janse RJ, Fu EL, Clase CM, et al. Stopping versus continuing renin–angiotensin–system inhibitors after acute kidney injury and adverse clinical outcomes: an observational study from routine care data. *Clin Kidney J*. 2022;15:1109–1119.
80. James MT, Levey AS, Tonelli M, et al. Incidence and prognosis of acute kidney diseases and disorders using an integrated approach to laboratory measurements in a universal health care system. *JAMA Netw Open*. 2019;2:e191795.
81. Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*. Springer; 2009.
82. Bidulka P, Fu EL, Leyrat C, et al. Stopping renin-angiotensin system blockers after acute kidney injury and risk of adverse outcomes: parallel population-based cohort studies in English and Swedish routine care. *BMC Med*. 2020;18:195.
83. Kellum JA, Sileanu FE, Murugan R, et al. Classifying AKI by urine output versus serum creatinine level. *J Am Soc Nephrol*. 2015;26:2231–2238.
84. Levey AS, Eckardt KU, Dorman NM, et al. Nomenclature for kidney function and disease: report of a Kidney Disease: Improving Global Outcomes (KDIGO) Consensus Conference. *Kidney Int*. 2020;97:1117–1129.
85. Guthrie G, Guthrie B, Walker H, et al. Developing an AKI consensus definition for database research: findings from a scoping review and expert opinion using a Delphi process. *Am J Kidney Dis*. 2022;79:488–496.e1.
86. Liu KD, Hsu CY, Yang J, et al. Acute kidney injury ascertainment is affected by the use of first inpatient versus outpatient baseline serum creatinine. *Kidney Int Rep*. 2018;3:211–215.
87. Cooper DJ, Plewes K, Grigg MJ, et al. An evaluation of commonly used surrogate baseline creatinine values to classify AKI during acute infection. *Kidney Int Rep*. 2021;6:645–656.
88. Bouchard J. Estimating baseline serum creatinine for assessing acute kidney injury: not a one size fits all approach. *Kidney Int Rep*. 2021;6:562–564.
89. Flynn N. Estimating baseline kidney function in hospitalized patients with impaired kidney function. *Ann Clin Biochem*. 2013;50:89.
90. Hsu CY, Chinchilli VM, Coca S, et al. Post-acute kidney injury proteinuria and subsequent kidney disease progression: the Assessment, Serial Evaluation, and Subsequent Sequelae in Acute Kidney Injury (ASSESS-AKI) Study. *JAMA Intern Med*. 2020;180:402–410.
91. Siew ED, Peterson JF, Eden SK, et al. Outpatient nephrology referral rates after acute kidney injury. *J Am Soc Nephrol*. 2012;23:305–312.
92. Sawhney S, Marks A, Fluck N, et al. Intermediate and long-term outcomes of survivors of acute kidney injury episodes: a large population-based cohort study. *Am J Kidney Dis*. 2017;69:18–28.
93. Kashani K, Shao M, Li G, et al. No increase in the incidence of acute kidney injury in a population-based annual temporal trends epidemiology study. *Kidney Int*. 2017;92:721–728.
94. Siew ED, Parr SK, Abdel-Kader K, et al. Predictors of recurrent AKI. *J Am Soc Nephrol*. 2016;27:1190–1200.
95. Liaño F, Pascual J, Madrid Acute Renal Failure Study Group. Epidemiology of acute renal failure: a prospective, multicenter, community-based study. *Kidney Int*. 1996;50:811–818.
96. Bagshaw SM, Uchino S, Cruz D, et al. A comparison of observed versus estimated baseline creatinine for determination of RIFLE class in patients with acute kidney injury. *Nephrol Dial Transplant*. 2009;24:2739–2744.
97. Hsu RK, Hsu CY, McCulloch CE, et al. Research-based versus clinical serum creatinine measurements and the association of acute kidney injury with subsequent kidney function: findings from the Chronic Renal Insufficiency Cohort study. *Clin Kidney J*. 2020;13:55–62.
98. Siew ED, Peterson JF, Eden SK, et al. Use of multiple imputation method to improve estimation of missing baseline serum creatinine in acute kidney injury research. *Clin J Am Soc Nephrol*. 2013;8:10–18.
99. Prowle JR, Kolic I, Purdell-Lewis J, et al. Serum creatinine changes associated with critical illness and detection of persistent renal dysfunction after AKI. *Clin J Am Soc Nephrol*. 2014;9:1015–1023.
100. Liu KD, Thompson BT, Ancukiewicz M, et al. Acute kidney injury in patients with acute lung injury: impact of fluid accumulation on classification of acute kidney injury and associated outcomes. *Crit Care Med*. 2011;39:2665–2671.
101. Koyner JL. Assessment and diagnosis of renal dysfunction in the ICU. *Chest*. 2012;141:1584–1594.
102. Sawhney S, Bell S, Black C, et al. Harmonization of epidemiology of acute kidney injury and acute kidney disease produces comparable findings across four geographic populations. *Kidney Int*. 2022;101:1271–1281.
103. Pullenayegum EM, Lim LS. Longitudinal data subject to irregular observation: a review of methods with a focus on visit processes, assumptions, and study design. *Stat Methods Med Res*. 2016;25:2992–3014.
104. Neuhaus JM, McCulloch CE, Boylan RD. Analysis of longitudinal data from outcome-dependent visit processes: failure of proposed methods in realistic settings and potential improvements. *Stat Med*. 2018;37:4457–4471.
105. McCulloch CE, Neuhaus JM, Olin RL. Biased and unbiased estimation in longitudinal studies with informative visit processes. *Biometrics*. 2016;72:1315–1324.
106. Hernan MA, McAdams M, McGrath N, et al. Observation plans in longitudinal studies with time-varying treatments. *Stat Methods Med Res*. 2009;18:27–52.
107. Farzanfar D, Abumuamar A, Kim J, et al. Longitudinal studies that use data collected as part of usual care risk reporting biased results: a systematic review. *BMC Med Res Methodol*. 2017;17:133.
108. McGee G, Haneuse S, Coull BA, et al. On the nature of informative presence bias in analyses of electronic health records. *Epidemiology*. 2022;33:105–113.
109. Iskander C, Cherney DZ, Clemens KK, et al. Use of sodium-glucose cotransporter-2 inhibitors and risk of acute kidney injury in older adults

with diabetes: a population-based cohort study. *CMAJ*. 2020;192:E351–E360.
110. Xu Y, Fu EL, Clase CM, et al. GLP-1 receptor agonist versus DPP-4 inhibitor and kidney and cardiovascular outcomes in clinical practice in type-2 diabetes. *Kidney Int*. 2022;101:360–368.
111. Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. *Int J Epidemiol*. 2015;44:324–333.
112. Rassen JA, Murk W, Schneeweiss S. Real-world evidence of bariatric surgery and cardiovascular benefits using electronic health records data: a lesson in bias. *Diabetes Obes Metab*. 2021;23:1453–1462.
113. Fu EL, van Diepen M, Xu Y, et al. Pharmacoepidemiology for nephrologists (part 2): potential biases and how to overcome them. *Clin Kidney J*. 2021;14:1317–1326.
114. Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. *Nat Rev Rheumatol*. 2015;11:437–441.
115. Geskus RB. Which individuals make dropout informative? *Stat Methods Med Res*. 2014;23:91–106.
116. Goldstein BA, Phelan M, Pagidipati NJ, Peskoe SB. How and when informative visit processes can bias inference when using electronic health records data for clinical research. *J Am Med Inform Assoc*. 2019;26:1609–1617.
117. Sisk R, Lin L, Sperrin M, et al. Informative presence and observation in routine health data: a review of methodology for clinical risk prediction. *J Am Med Inform Assoc*. 2021;28:155–166.
118. Kurland BF, Heagerty PJ. Directly parameterized regression conditioning on being alive: analysis of longitudinal data truncated by deaths. *Biostatistics*. 2005;6:241–258.
119. Buzkova P, Brown ER, John-Stewart GC. Longitudinal data analysis for generalized linear models under participant-driven informative follow-up: an application in maternal health epidemiology. *Am J Epidemiol*. 2010;171:189–197.
120. Pullenayegum E. IrregLong: analysis of longitudinal data with irregular observation times: R package version 0.3.3. Accessed November 10, 2021. https://CRAN.R-project.org/package=IrregLong
121. Lawrence Gould A, Boye ME, Crowther MJ, et al. Joint modeling of survival and longitudinal non-survival data: current methods and issues: report of the DIA Bayesian joint modeling working group. *Stat Med*. 2015;34:2181–2195.
122. Harhay MO, Gasparini A, Walkey AJ, et al. Assessing the course of organ dysfunction using joint longitudinal and time-to-event modeling in the Vasopressin and Septic Shock Trial. *Crit Care Explor*. 2020;2:e0104.
123. Chesnaye NC, Tripepi G, Dekker FW, et al. An introduction to joint models-applications in nephrology. *Clin Kidney J*. 2020;13:143–149.
124. Crowther MJ. merlin—A unified modeling framework for data analysis and methods development in Stata. *Stata J*. 2020;20:763–784.
125. Schneeweiss S, Rassen JA, Brown JS, et al. Graphical depiction of longitudinal study designs in health care databases. *Ann Intern Med*. 2019;170:398–406.
126. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol*. 2016;45:1866–1886.
127. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10:37–48.
128. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020;584:430–436.

Appendix 3

Considerations for planning a research study evaluating CKD progression using EHRs

Background and aims

The below recommendations were developed following completion of the systematic review of previous research (paper 1), which identified considerable issues in study design, reporting and analyses, which failed to recognise, evaluate and account for statistical biases that are common in studies evaluating CKD progression using EHRs. Developing an understanding of what is done badly, combined with knowledge of epidemiological biases (understanding what should be done ideally) has allowed us to come up with recommendations for planning of research studies, which we present here.

It is about quality of research, with an emphasis on transparency of data used for analysis, in areas which may implicate bias. It is limited to advice for cohort studies, which is the optimal study design for research of CKD progression using EHRs. CKD progression (or changes in kidney function) may be evaluated as exposure or outcome, using eGFR-based or non-eGFR based measures.

It is recommended to be used in conjunction with RECORD guidelines. Some elements of RECORD will be repeated here or elaborated upon, to aid emphasis of certain points.

Recommendations

METHODS

- ### Setting of data sources
    - State the data sources used in the study, including any linkages, and which part(s) of the healthcare system study data are captured in
    - State the care pathway and intended follow-up framework in the healthcare system that the data are captured in
    - Consider whether the data sources capture sufficient data in the end-to-end patient healthcare experience to answer the research question of interest (for example, in identifying a representing sample of the target population and in evaluating exposures, outcomes, and relevant confounders or covariates).
    - State any anticipated challenges faced in completeness of data collection, if any, which result from limitations in data sources available for the study. Specifically state any processes which may lead to loss to follow up, or differential follow-up of patients for outcomes. If applicable, state how data sources or linkages may address losses to follow-up.
    - Consider a diagram to show periods of recruitment, timings of evaluation of exposures and duration of follow-up
- ### Participant selection
    - State the data sources used to identify patients for inclusion in the study, and which part(s) of the healthcare system these data are captured in
    - Clearly state the inclusion/exclusion criteria for participation in the study, including what data is used *(e.g. diagnostic codes / eGFR results)* to support identification
    - State whether considerations about data completeness affected the choice of study population (e.g. choosing to study a population with good data completeness, to enhance reliability of study results)

- If the target study population is patients with CKD, note that electronic diagnostic codes for identification of CKD are rarely complete, and it is preferable to use repeated eGFR test results to identify and confirm evidence of CKD for study inclusion
- Clarify whether data completeness criteria apply to collection of data for evaluation of exposures, outcomes or both.
- Clearly state any specific data completeness criteria for different statistical analyses, if different from the main inclusion/exclusion criteria of the study, including which analyses they correspond to.
- If there are multiple different data completeness criteria used in the study, a table may aid clarification of criteria used and analyses applied to.
- Consider and state the rationale for data completeness criteria (perhaps with pros and cons), with implications for representativeness of results, and implications that results may be generalisable to different populations
- Note that different analytical methods are capable of adapting to different levels of completeness of data, and this may be a key driver in deciding analysis criteria for the study. For example, linear mixed models have low threshold for data completeness, but rely on assumptions that those patients with missing data are similar in some ways to patients without data. There are trade-offs between requiring lower thresholds for data completeness (which leads to more representative samples being included in the analysis) and ensuring sufficient data is captured to answer research questions (which may lead to results which are more reliably reflective of those included).

• Variables

- Describe methods to define and evaluate exposures, outcomes, and any other relevant study variables, such as confounders or covariates, including specific diagnostic codes used, if any.
- State any anticipated challenges faced in accurately measuring study variables using available data, e.g. incomplete recording of diagnostic codes, testing not indicated in patient population, variation in healthcare behaviours.

- Consider the trade-offs between inclusion of certain variables in analysis (such as specific confounders) versus risks of poor data completeness.
- State any imputation methods for variables with missing data, or whether complete case analysis was used.
- Note that complete case analysis may be a good option in the case of small amounts of missing data, as results are anticipated to have good generalisability to the population of interest. Multiple imputation may be reasonable if missing data are expected to be missing at random (MAR), i.e. missing results are expected to be similar in nature to observed results, given observed characteristics. Benefits of multiple imputation over simple imputation methods (e.g. last observation carried forward, imputation of mean) are preservation of variation in the data. Risks are missingness not at random (MNAR) where missing data may be different in nature to non-missing data, even after accounting for observed data (i.e. missing data due to non-observed factors). It is rarely possible to test for violation of assumptions of missingness mechanisms, as values of missing data cannot be compared to values of non-missing data. However, limited checks can still be performed on data that is available, e.g. comparing distribution of variable values and degree of missingness by risk factors.
- If CKD progression was defined in some way, state any reasoning for choice of definition, e.g. use of validated definitions, clinical importance, ability to accurately/reliably identify a clinically important change, capturing chronicity of decline, data requirements to evaluate change

- **Data sources/measurement**
  - Note that this section may not be needed if all items are captured in previous sections.
  - If not previously stated, clarify the data sources used for each study variable, and what area of the healthcare system is covered by the data source.
  - If not previously stated, state any relevant data that may be missing due to limitations in data capture due to the data source, *e.g. eGFR tests are captured*

in primary care, but not in secondary care, despite relevance of secondary care eGFR tests to the research question

- Comment on any anticipated issues in data quality impacting accurate measurement of study variables, due to limitations of data sources, *e.g. poor sensitivity (or specificity) of diagnostic coding for comorbidities, variation in creatinine calibration practices, lack of information on GFR-estimating equations used for laboratory reported eGFR, needing to approximate uACR using PCR or dipstick results due to missing uACR data*
- Comment on any changes over time in availability or quality of data used for study variables, *e.g. due to improvements in disease coding over time, due to reduced uACR testing due to the COVID pandemic, changes in GFR-estimating equations used by laboratories reporting eGFR results*
- If data sources involve multiple healthcare systems or regions, with anticipated differences in data quality or availability, this should be acknowledged with consideration for implications

- **Bias**
  - State any concerns about anticipated biases in the analysis and any efforts to mitigate biases. Specifically refer to risks of selection bias and ascertainment bias, which may result from data completeness criteria for variables evaluated at baseline, and incomplete data for follow-up assessments (or variable data completeness, which may be associated with variables relevant to the study [observed or unobserved]). Other biases that may impact analyses (that may be related to selection bias and ascertainment bias) may be mentioned or referenced, e.g. informative censoring (in some cases, due to competing events), survival bias, misclassification bias, unobserved or residual confounding

- **Sample size**
  - Sample sizes for EHR studies may be planned with consideration of analytical power, or may be based on feasibility if availability of data sources are set, with minimal power to control sample sizes for analysis. When planning a

study, consider the impact of data completeness criteria on the final sample sizes likely to be available for analysis. State any decisions that were made in order to achieve required sample sizes due to concerns about data availability or completeness in final analysis datasets, *e.g. imputation of missing covariate data, choice of eGFR-based study outcome instead of RRT outcome, decisions to study broader populations with less complete data*

- **Missing data**
    - If not already stated, state how missing data were handled for each study variable (e.g. complete case analysis, single imputation, multiple imputation, linear mixed models). Choice of missing data strategies will be influenced by assumptions about mechanisms of missingness in the observed data (missingness completely at random (MCAR), missingness at random (MAR), missingness not at random (MNAR)). Definitions of missingness mechanisms with examples of choices of analytical methods were introduced in **Chapter 2 (Background), Section 2.6**.
    - Note that large amounts of missing data are unlikely to be solved by analytical techniques for handling missing data, with substantial risks of missingness being due to unobserved factors, leading to biased study results. It may be more appropriate to study populations for which there is good data completeness, investigate alternative data sources with more complete data, or seek to link complementary data sources.
- **Statistical methods**
    - State how any concerns about missing data influenced the choice of statistical analyses, specifically stating any methods to account for missingness at baseline or during follow-up, *e.g. linear mixed models to account for irregular eGFR testing with varying frequency of measures, sensitivity analyses in groups with different data completeness, competing risks models, joint longitudinal survival models to account for cause-specific drop-outs, adjustment for indicators of missingness [as a proxy for unobserved confounders]*

- State any subgroup or sensitivity analyses used for the purposes of checking robustness of analyses to concerns around missing data, *e.g. diabetes only analysis (due to more complete data), sensitivity analysis in patients with more (or less) complete data*
- Describe any data cleaning methods relevant to handling of missing or poor quality data

RESULTS

- **Participants**
    - A flow chart is recommended to show how patients were identified, including impact on sample size due to application of data completeness criteria. Ideally shows details of number of patients excluded due to specific reasons. Study outcomes should be shown for patients identified as the study population. May consider producing multiple flow charts for development of outcomes, before and after data completeness criteria (in relation to target study population) are applied, if relevant
- **Descriptive data**
    - State the percentage of the target population excluded due to application of data completeness criteria
    - State the percentage of patients lost to follow-up (with causes if known), if possible to evaluate this and if applicable to the study question
    - Describe patient characteristics and completeness of data for study variables in those included vs excluded (relative to the target study population)
    - Describe frequency and completeness of eGFR testing in patients who are followed-up, overall and stratified by risk factors (for eGFR-based outcomes)
    - Summarise frequency or rates of study outcome and competing events over time during follow-up
    - Provide similar descriptive data where applicable for subgroup analyses or sensitivity analyses
- **Analysis results**
    - Report results of main analyses, as well as any subgroup or sensitivity analyses, which may indicate robustness of methods to missing data

DISCUSSION

- **Key findings**
    - State study results in the context of any problems encountered due to missing or poor quality data
- **Strengths**
    - Discuss strengths of methods (which may include data sources, study design, reporting and analysis) to handle missing data which are likely to occur in analysis of EHRs, and how risks of bias were mitigated due to chosen methods, in the context of data completeness observed in the study
- **Limitations**
    - Discuss limitations of the study resulting from issues of data quality or completeness that could not be resolved by study methodology, and how these biases may impact generalisability of results or bias in observed results
- **Interpretation**
    - Evaluate implications of study results in the context of potential biases, which may or may not be partly tackled by use of study methodology, including implications for clinical care and researchers

Appendix 4

Supplementary materials for research paper 2 (feasibility analysis)

See next page for supplementary materials for research paper 2.

Supplementary Materials

Tables and figures referred to in the manuscript are provided in the supplementary materials, including descriptive data checks, sensitivity analyses, analysis population characteristics and numerical figures to support graphics provided in the manuscript.

Table 1 Availability of repeat creatinine tests in primary care in adults with coded diabetes

Risk factor	Patients with ≥3 tests (N, %)	Test frequency* (median + IQR)	Duration of test coverage, years* (median + IQR)	No test in last 3 years* (N, %)	Coded RRT if no test in last 3 years* (N,%)
All diabetes	366,098 (92.5%)	10 (7, 14)	6.2 (5.1, 6.7)	2,154 (0.6%)	13.7%
Age:					
18-40	1,3871 (69.8%)	6 (4, 8)	4.6 (2.8, 6)	263 (1.9%)	5.3%
40-60	96,139 (89.2%)	8 (6, 11)	5.8 (4.1, 6.5)	899 (0.9%)	10.5%
60-80	193,926 (95.3%)	11 (8, 14)	6.3 (5.5, 6.7)	750 (0.4%)	21.1%
80+	62,162 (95.9%)	12 (9, 17)	6.4 (5.9, 6.8)	242 (0.4%)	12.4%
Sex:					
Male	205,549 (92.1%)	10 (7, 14)	6.2 (5, 6.7)	1,257 (0.6%)	15.1%
Female	160,549 (92.9%)	10 (7, 14)	6.2 (5.2, 6.7)	897 (0.6%)	11.8%
Hypertension	222,608 (95.7%)	11 (8, 15)	6.3 (5.6, 6.8)	1065 (0.5%)	23.8%
CVD	92,641 (95.9%)	12 (9, 16)	6.4 (5.7, 6.8)	396 (0.4%)	33.3%
CKD code	69,095 (97.2%)	14 (10, 19)	6.5 (6, 6.9)	437 (0.6%)	62.7%
Confirmed CKD	67,105 (98.5%)	14 (10, 19)	6.5 (5.9, 6.9)	N/A**	N/A**
CKD stage*** (last GFR):					
1 (90+)	100,569 (89.8%)	9 (6, 12)	5.9 (4.3, 6.6)	775 (0.8%)	0%
2 (60-90)	181,391 (94.1%)	10 (7, 13)	6.2 (5.1, 6.7)	843 (0.5%)	0.8%
3 (30-60)	75,590 (97.1%)	13 (9, 18)	6.4 (5.8, 6.8)	238 (0.3%)	10.1%
4 (15-30)	7,032 (97.7%)	18 (13, 25)	6.6 (5.9, 6.9)	88 (1.3%)	77.3%
5 (<15)	1,516 (93.2%)	16 (10, 24)	6.0 (3.9, 6.7)	210 (13.9%)	93.8%

*in patients with ≥3 tests

**Loss to follow up not evaluable in confirmed CKD since group definition requires creatinine measure in last 2 years

***CKD stage evaluated in all diabetes patients

Table 2 Duration of coverage of tests for patients with at least 3 test results (and hence meeting slope analysis criteria)

Duration of coverage of tests	≥ 3 valid reported eGFR results [Reported GFR; MDRD (1)] N (%)	≥3 valid MDRD re-calculated eGFR results [MDRD (2)] N (%)	≥ 3 creatinine results N (%)
	N = 1,100,460	N = 1,594,629	N = 1,597,629
<90 days	5,136 (0.5%)	5,033 (0.3%)	5,038 (0.3%)
90 days – 1 year	28,117 (2.6%)	28,883 (1.8%)	29,014 (1.8%)
1-2 years	67,739 (6.2%)	74,653 (4.7%)	74,791 (4.7%)
2-4 years	224,007 (20.4%)	258,104 (16.2%)	258,470 (16.2%)
4-6 years	410,626 (37.3%)	573,623 (36.0%)	574,211 (35.9%)
6-8.5 years	364,835 (33.1%)	654,378 (41.0%)	656,105 (41.1%)

*Valid refers to GFR results in the 0-150 range

Figure 1 Percentage agreement (+/-1 and +/-3 ml/min/1.73m^2) between reported eGFR and corresponding re-calculated MDRD eGFR by calendar year and eGFR CKD stage

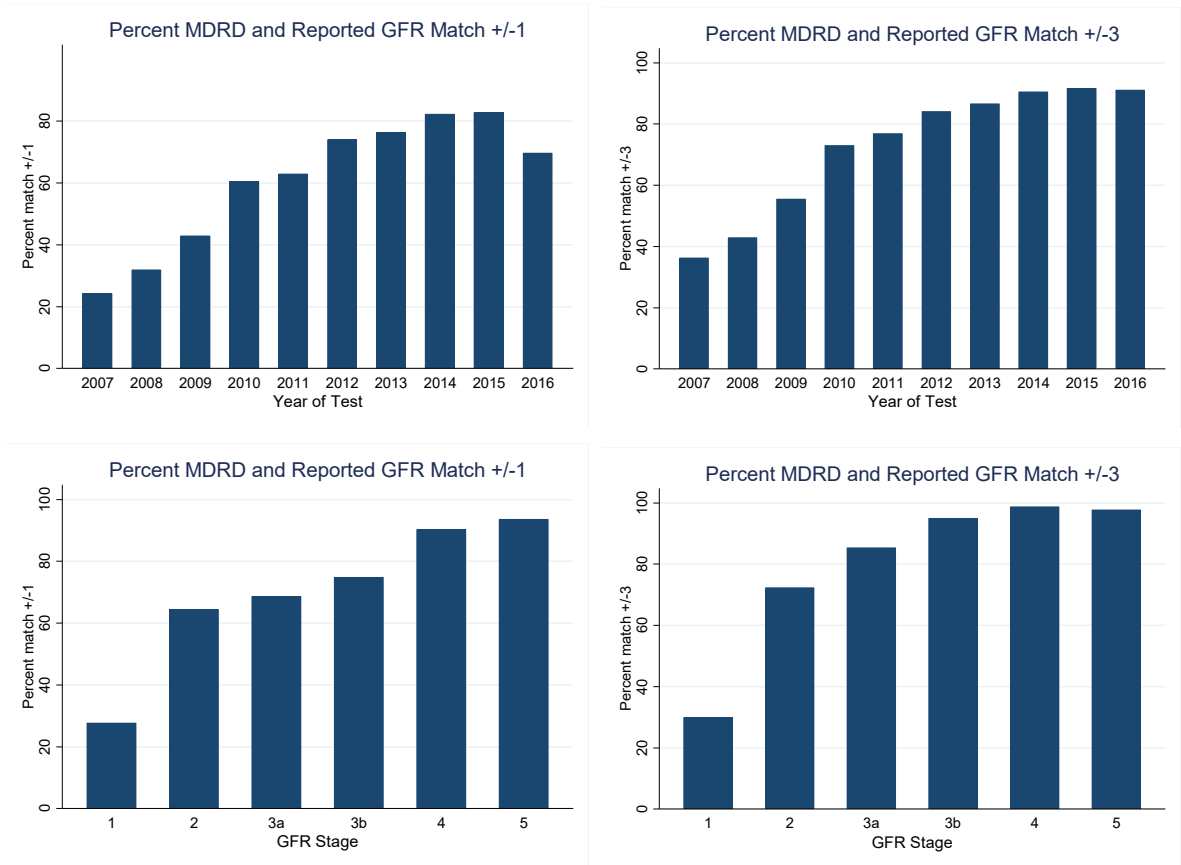


Table 3 eGFR regression slopes by risk factor and coverage of underlying population

A – Re-calculated eGFR slopes (in all patients with at least 3 valid re-calculated eGFR results)

Risk factor	Number and percent of patients analysed from full extracted dataset	Change in eGFR per year (median + IQR) [MDRD-2]
Diabetes	363,626 (92.2%)	+0.10 (-1.61, 1.84)
Hypertension	737,472 (84.7%)	+0.30 (-1.23, 1.94)
CVD	350,938 (89.9%)	+0.07 (-1.49, 1.68)
CKD code	251,784 (94.5%)	-0.47 (-1.90, 0.85)
Confirmed CKD	247,351 (96.4%)	-1.04 (-2.48, 0.11)

B – All slopes (in all patients with at least 3 REPORTED eGFR results)

Risk factor	Number and percent of patients analysed from full extracted dataset	Change in eGFR per year (median + IQR) [REPORTED GFR]	Change in eGFR per year (median + IQR) [MDRD-1]	Change in eGFR per year (median + IQR) [MDRD-2]
Diabetes	253,974 (64.4%)	-0.60 (-2.07, 0.68)	-0.11 (-1.62, 1.33)	-0.13 (-1.72, 1.41)
Hypertension	523,638 (60.1%)	-0.39 (-1.70, 0.84)	+0.09 (-1.28, 1.50)	+0.12 (-1.29, 1.58)
CVD	272,624 (69.8%)	-0.56 (-1.97, 0.70)	-0.09 (-1.54, 1.33)	-0.09 (-1.58, 1.38)
CKD code	244,127 (91.7%)	-0.79 (-2.23, 0.45)	-0.45 (-1.87, 0.86)	-0.49, (-1.91, 0.81)
Confirmed CKD	241,047 (94.0%)	-1.32 (-2.84, -0.16)	-0.98 (-2.44, 0.18)	-1.04, (-2.48, 0.11)

C – Difference in slopes (in all patients with at least 3 REPORTED GFR results)

Risk factor	Number and percent of patients analysed from full extracted dataset	Difference in slopes (median + IQR) [MDRD (1) - REPORTED GFR]	Difference in slopes (median + IQR) [MDRD (2) - REPORTED GFR]
Diabetes	253,974 (64.4%)	+0.24 (-0.01, 0.94)	+0.36 (-0.06, 1.27)
Hypertension	523,638 (60.1%)	+0.18 (-0.02, 0.94)	+0.32 (-0.04, 1.22)
CVD	272,624 (69.8%)	+0.25 (-0.01, 0.90)	+0.34 (-0.04, 1.15)
Confirmed CKD	241,047 (94.0%)	+0.19 (0.00, 0.61)	+0.17 (-0.03, 0.65)

Table 4 Age, sex, and ethnicity breakdown by underlying health condition in patients with at least 3 REPORTED GFR results (reflects population of boxplots in manuscript Figure 2), contextualised by age, sex, ethnicity breakdown of the underlying population, for patients with at least 3 creatinine tests, and among all patients with at least 3 reported GFR results

Risk factor	Underlying population under study*	All adults ≥3 creatinine	All adults ≥3 reported GFR	CKD	Diabetes	Hypertension	CVD
Number of adults	N = 6,513,000	N = 1,597,629	N = 1,100,460	N = 241,047	N = 253,974	N = 523,638	N = 272,624
Median + IQR last re-calculated eGFR	N/A	73 (62, 86)	68 (58, 78)	48 (40, 54)	69 (55, 81)	69 (57, 79)	65 (53, 77)
Age							
18-39	2,301,700 (35.3%)	59,187 (3.7%)	16,340 (1.5%)	945 (0.4%)	2,733 (1.1%)	3,012 (0.6%)	314 (0.1%)
40-59	2,214,100 (34.0%)	419,144 (26.2%)	224,735 (20.4%)	14,036 (5.8%)	47,329 (18.6%)	87,394 (16.7%)	22,318 (8.2%)
60-79	1,578,600 (24.2%)	824,468 (51.6%)	605,442 (55.0%)	115,208 (47.8%)	148,065 (58.3%)	292,425 (55.8%)	150,435 (55.2%)
80+	418,600 (6.4%)	294,830 (18.4%)	253,943 (23.1%)	110,858 (46.0%)	55,847 (22.0%)	140,807 (26.9%)	99,557 (36.5%)
Sex:							
Male	3,200,400 (49.1%)	765,907 (47.9%)	503,054 (45.7%)	98,402 (40.8%)	134,891 (53.1%)	227,455 (43.4%)	156,648 (57.5%)
Female	3,312,600 (50.9%)	831,715 (52.1%)	597,402 (54.3%)	142,645 (59.2%)	119,083 (46.9%)	296,182 (56.6%)	115,975 (42.5%)
Ethnicity:							
Black	111,300 (1.7%)	17,917 (1.1%)	12,638 (1.2%)	1,026 (0.4%)	4,457 (1.8%)	5,312 (1.0%)	1,466 (0.5%)
Non-black	6,401,700 (98.3%)	1,579,712 (98.9%)	1,087,822 (98.8%)	240,021 (99.6%)	249,517 (93.2%)	518,326 (99.0%)	271,158 (99.5%)

IMPORTANT: Percentages are of column headers (in this case, population under study or underlying health condition)

*Population age, sex and ethnicity breakdown is estimated based on aggregate data provided at the practice level

In patients with confirmed CKD that have at least 3 reported eGFR results, only 0.4% are aged under 40 and 93.8% are aged 60 and over. There is under-representation of males and black ethnicity in this group compared to the underlying population. In diabetes, 1.1% are aged under 40 and 80.3% are aged 60 and over, covering a broader population than for confirmed CKD. Black ethnicity appears to be under-represented across all risk

factor subgroups, when compared to the underlying population and considering that co-morbidity prevalence is likely truly higher in black ethnicity than non-black ethnicity. The same may be true for males, who suffer higher burden of co-morbidities than females.

Table 5 Age, sex, ethnicity breakdown by CKD stage (1-5) at last GFR in patients with at least 3 REPORTED GFR results (reflects population of boxplots in manuscript Figure 3), contextualised by age, sex, ethnicity breakdown of the underlying population, for patients with at least 3 creatinine tests, and among all patients with at least 3 reported GFR results

Risk factor	Underlying population under study*	All adults ≥3 creatinine	All adults ≥3 reported GFR	Last GFR CKD stage 1	Last GFR CKD stage 2	Last GFR CKD stage 3	Last GFR CKD stage 4	Last GFR CKD stage 5
	N = 6,513,000	N = 1,597,629	N = 1,100,460	N = 75,435	N = 701,579	N = 302,199	N = 17,751	N = 3,471
Last re-calculated eGFR result (median + IQR)	N/A	73 (62, 86)	68 (58, 78)	96 (92, 102)	72 (67, 79)	52 (45, 56)	25 (22, 28)	11 (8, 13)
Age								
18-39	2,301,700 (35.3%)	59,187 (3.7%)	16,340 (1.5%)	1,788 (2.4%)	12,799 (1.8%)	1,403 (0.5%)	199 (1.1%)	144 (4.2%)
40-59	2,214,100 (34.0%)	419,144 (26.2%)	224,735 (20.4%)	24,320 (32.2%)	175,110 (25.0%)	23,444 (7.8%)	1,158 (6.5%)	696 (20.1%)
60-79	1,578,600 (24.2%)	824,468 (51.6%)	605,442 (55.0%)	40,526 (53.7%)	404,968 (57.7%)	151,978 (50.3%)	6,328 (35.7%)	1,638 (47.2%)
80+	418,600 (6.4%)	294,830 (18.4%)	253,943 (23.1%)	8,801 (11.7%)	108,702 (15.5%)	125,374 (41.4%)	10,066 (56.7%)	993 (28.6%)
Sex:								
Male	3,200,400 (49.1%)	765,907 (47.9%)	503,054 (45.7%)	41,861 (55.5%)	329,663 (47.0%)	121,692 (40.3%)	7,815 (44.0%)	2,014 (58.0%)
Female	3,312,600 (50.9%)	831,715 (52.1%)	597,402 (54.3%)	33,574 (44.5%)	371,916 (53.0%)	180,507 (59.7%)	9,936 (56.0%)	1,457 (42.0%)
Ethnicity:								
Black	111,300 (1.7%)	17,917 (1.1%)	12,638 (1.2%)	3,701 (4.9%)	7,443 (1.1%)	1,297 (0.4%)	109 (0.6%)	84 (2.4%)
Non-black	6,401,700 (98.3%)	1,579,712 (98.9%)	1,087,822 (98.8%)	71,734 (95.1%)	694,136 (98.9%)	300,902 (99.6%)	17,642 (99.4%)	3,387 (97.6%)

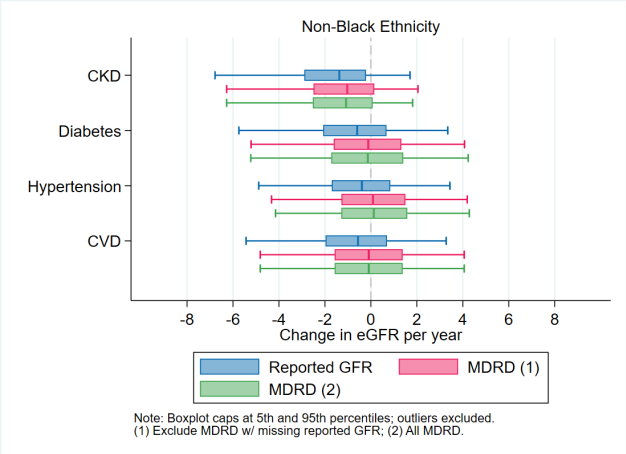
IMPORTANT: Percentages are of column headers (in this case, population under study or CKD stage indicated by last re-calculated GFR result)
*Population age, sex and ethnicity breakdown is estimated based on aggregate data provided at the practice level

There appears to be under-representation of males, black ethnicity and young adults for last GFR results that indicate CKD stage 3. While under-representation of young adults may be partially due to low prevalence of kidney disease, this is unlikely the case for males and black ethnicity.

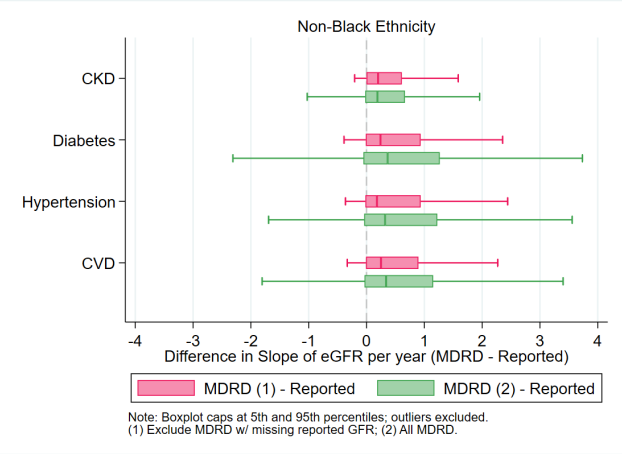
Under-representation of these groups appears to diminish as CKD progresses. Some of these groups appear to be over-represented at CKD stage 5, possibly suggesting worse outcomes for groups that are not identified at earlier stages of disease.

Figure 2 Distribution of slopes of change in eGFR (A) and distribution of differences between re-calculated and reported GFR slopes (B) in patients with at least 3 reported GFR results, by risk factors and method of estimation of slope of eGFR, separate plots for non-black (1) and black (2) ethnicity

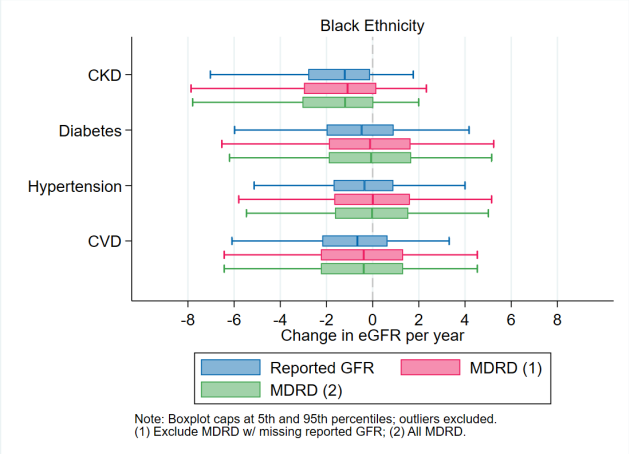
1A



1B



2A



2B

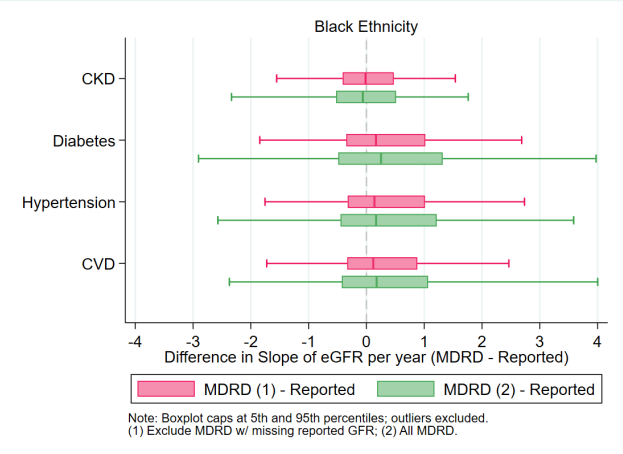
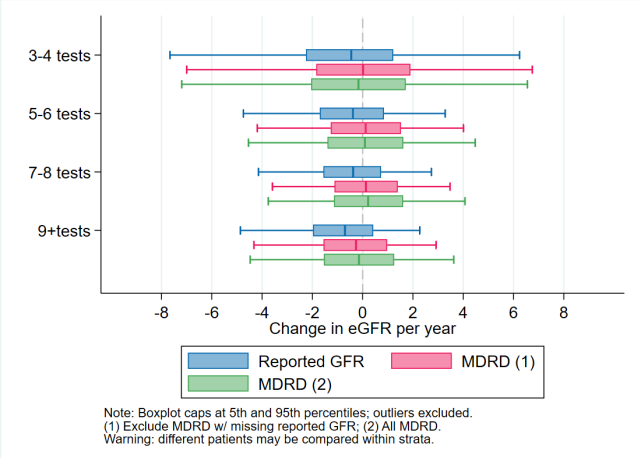


Figure 3 Distribution of slope of change in eGFR by test frequency



Appendix 5

Supplementary materials for research paper 3 (coding analysis)

See next page for supplementary materials for research paper 3.

SUPPLEMENTARY MATERIALS

Supplementary Table 1. Outcome definitions

Outcome	Definition	Details	ICD-10 codes
AKI	AKI at admission	Any diagnosis (HES) of AKI recorded in the first episode of care	N17
HF	Admission for HF	A primary diagnosis (HES) of HF recorded in the first episode of care	I50, I11.0, I13, I97.1
CV event	Admission for CV event	A primary diagnosis (HES) of HF, CHD, stroke/TIA, PAD or AAA recorded in the first episode of care	HF (I50, I11.0, I13, I97.1) CHD (I20-I25, I51.6) Stroke/TIA (G45-G46) PAD (I79.0, I79.2, I73.8, I73.9, I74.3, I74.4, I74.5, I70.2) AAA (I71, I74.0) Cerebrovascular disease (I60-I69)
All-cause mortality	All-cause mortality	Any death (ONS)	N/A

Supplementary Table 2. Percentage of practices and patients eligible for analysis after application of sample size eligibility criteria (minimum of 50 CKD cases) and comparison of coding rates in the original database and analysis database, in the CKD population and in subgroups

Population	Total practices	Total patients	Total patients coded	Eligible practices	Eligible patients	Eligible patients coded
All CKD	695	169,002	119,248 (70.6%)	637 (91.7%)	167,208 (98.9%)	117,932 (70.5%)
CKD stage 3a	695	106,981	66,514 (62.2%)	580 (83.5%)	103,615 (96.9%)	64,398 (62.2%)
CKD stages 3b-5	695	62,021	52,734 (85.0%)	477 (68.6%)	56,122 (90.5%)	47,700 (84.5%)
Diabetes	695	42,063	32,099 (76.3%)	362 (52.1%)	33,065 (78.6%)	25,273 (76.4%)
No diabetes	695	126,939	87,149 (68.7%)	605 (87.1%)	124,364 (98.0%)	85,324 (68.6%)

Supplementary Information 1. Methods for estimating the percentage of CV and HF hospitalisation events that are preventable among CKD patients in practices coding 55% of CKD cases if practice coding improved to 88% (attributable fraction for first events)

In order to estimate the percentage of first hospitalisation events that could be prevented over a period of 3.8 years (median follow-up duration) among patients in practices coding 55% of CKD cases (13th practice coding percentile, lower boundary of sextile 2), if such practices instead coded 88% of CKD cases (83rd practice coding percentile, upper boundary of sextile 5), we adopted the Austin method in [1], which we adapted for use in STATA statistical software. We make an assumption that after adjustment for model covariates, the difference in expected event rates between practice coding groups estimated using the methods listed below is caused by practice coding performance. We followed the following steps to estimate the percentage of first events attributable to lower practice coding (55%), when compared to higher practice coding (88%), for both CV and HF hospitalisations:

1. Fit the fully adjusted Cox regression model for time to first event with 5-knot spline for practice percent coded CKD and other practice covariates (model 2), first centring all continuous covariates to generate a sensible baseline group

2. Estimate the baseline survival function [2] using: "predict s, basesurv", sort data by analysis time and extract the baseline survival probability at t = 3.8 years (median follow-up)

3. Predict the estimated survival probability at t = 3.8 years for every patient in database assuming 54.8% practice percent coded, based on true values of all covariates except with practice percent coded forced to 54.8%, as follows:

    a. Recode data to set value of practice coding variable to 54.8% coded in all patients. (Spline variable values recoded accordingly.)

    b. Use equation $\hat{S}_i(t) = \hat{S}_0(t)^{\exp(\boldsymbol{x}_i\widehat{\boldsymbol{\beta}})}$ [3] to estimate individual survival probabilities $\hat{S}_i(t)$ at time t = 3.8 years, where $\hat{S}_0(t)$ is the estimated baseline survival at time t and $\boldsymbol{x}_i\widehat{\boldsymbol{\beta}}$ is the individual prediction of the linear predictor evaluated at individual true covariate values $\boldsymbol{x}_i$ (but specified practice percent coded of 54.8%) based on coefficient estimates $\widehat{\boldsymbol{\beta}}$

4. Repeat step 3 for practice percent coded forced to 87.5% coded, to obtain individual predicted survival probabilities at t = 3.8 years assuming 87.5% practice percent coded

5. Estimate the expected number of first events occurring over 3.8 years if all practices coded at 54.8% (with other practice characteristics unchanged) by taking the sum of individual probabilities of an event (1 - $\hat{S}_i(t = 3.8)$, using results from (3b)). Repeat for assumption of all practices coding at 87.5% by taking the sum of individual probabilities of an event (1 - $\hat{S}_i(t = 3.8)$, from (4b)).

6. Using results from step 5, compute the percentage of events preventable among individuals in practices coding 54.8% of CKD cases if practices instead coded 87.5% of CKD cases (attributable fraction) as:

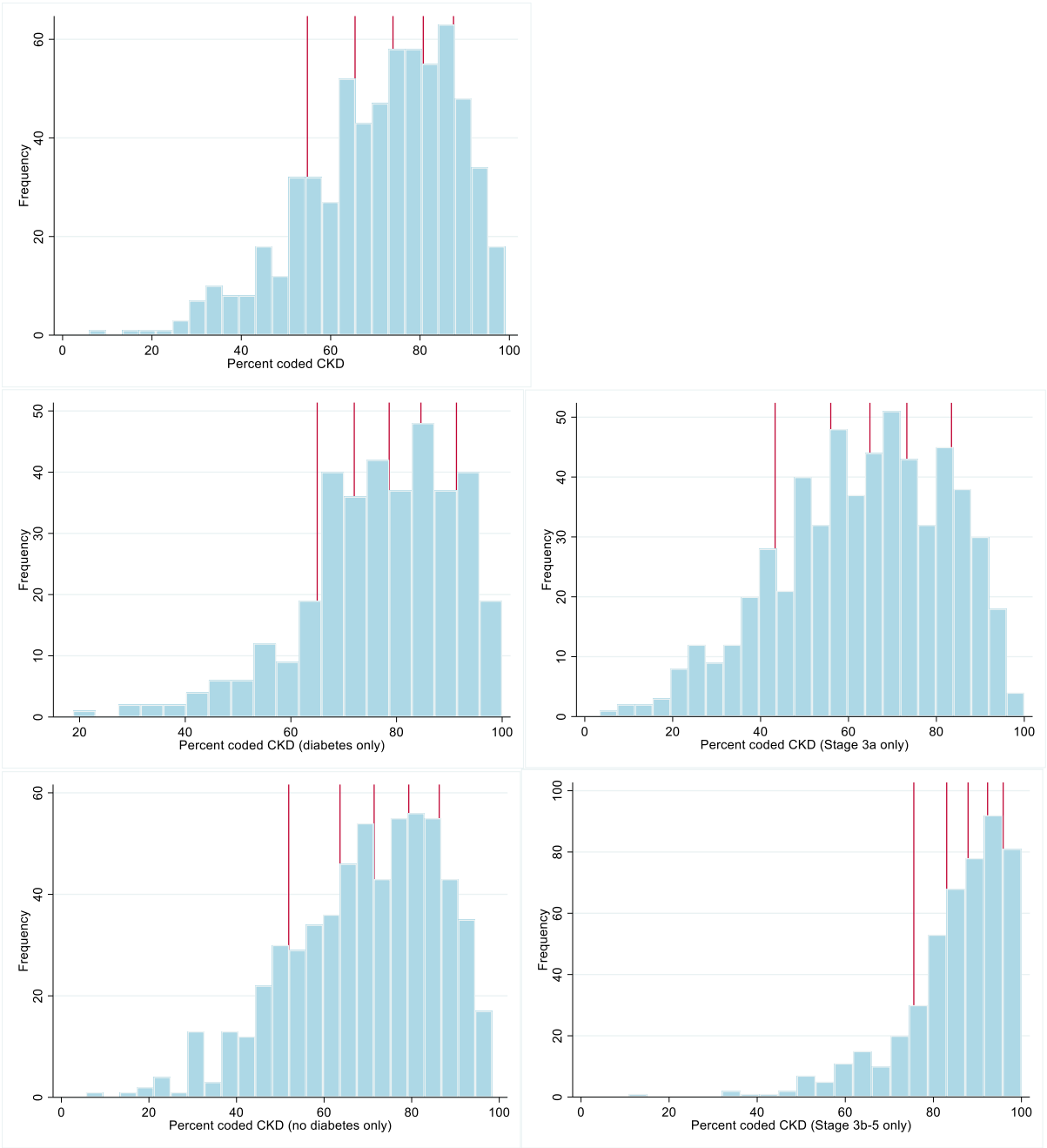
$$\%\ events\ preventable\ in\ practices\ coding\ 54.8\%\ of\ CKD\ cases$$

$$= \frac{expected\ events\ assuming\ 54.8\%\ coding - expected\ events\ assuming\ 87.5\%\ coding}{expected\ events\ assuming\ 54.8\%\ coding} \times 100\%$$

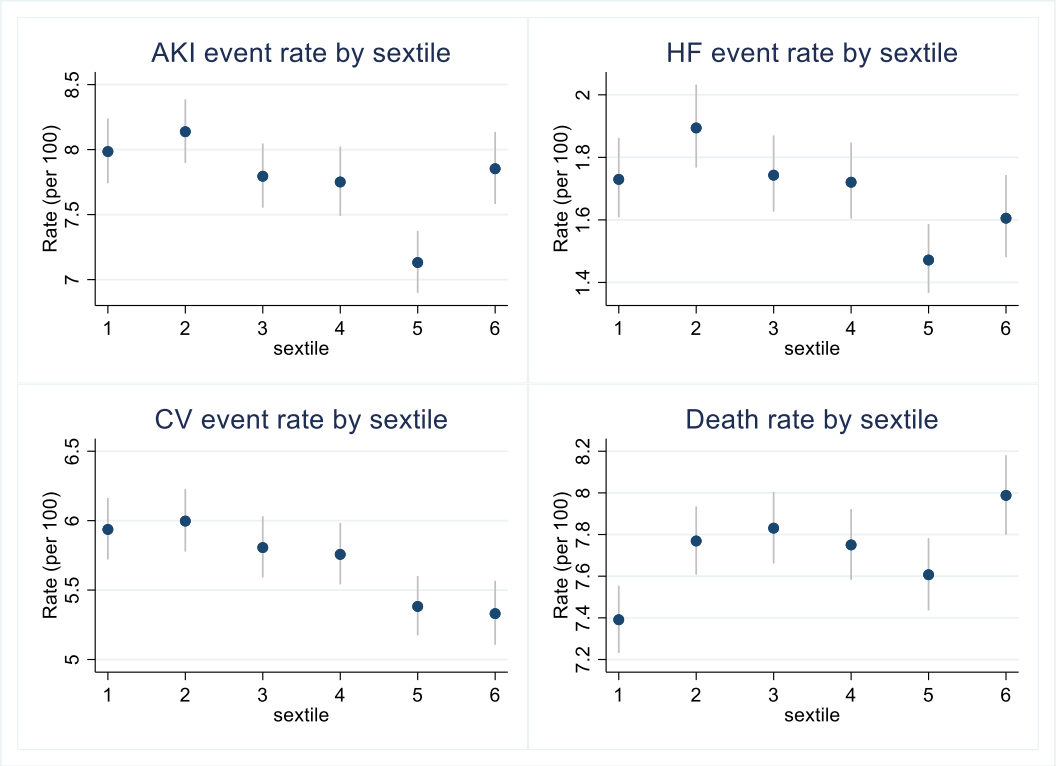
References:

1. Zhang Z, Ambrogi F, Bokov AF, et al. Estimate risk difference and number needed to treat in survival analysis. *Ann Transl Med* 2018;6(7):120.
2. StataCorp. STATA Survival Analysis Reference Manual Release 17. 2021. StataCorp LLC. https://www.stata.com/manuals/st.pdf. Accessed 17 Feb 2022.
3. Rodriguez, G. Approaches to Survival Modelling. In: Lecture Notes on Generalized Linear Models. Princeton University. 2007. https://data.princeton.edu/wws509/notes/c7.pdf. Accessed 17 Feb 2022

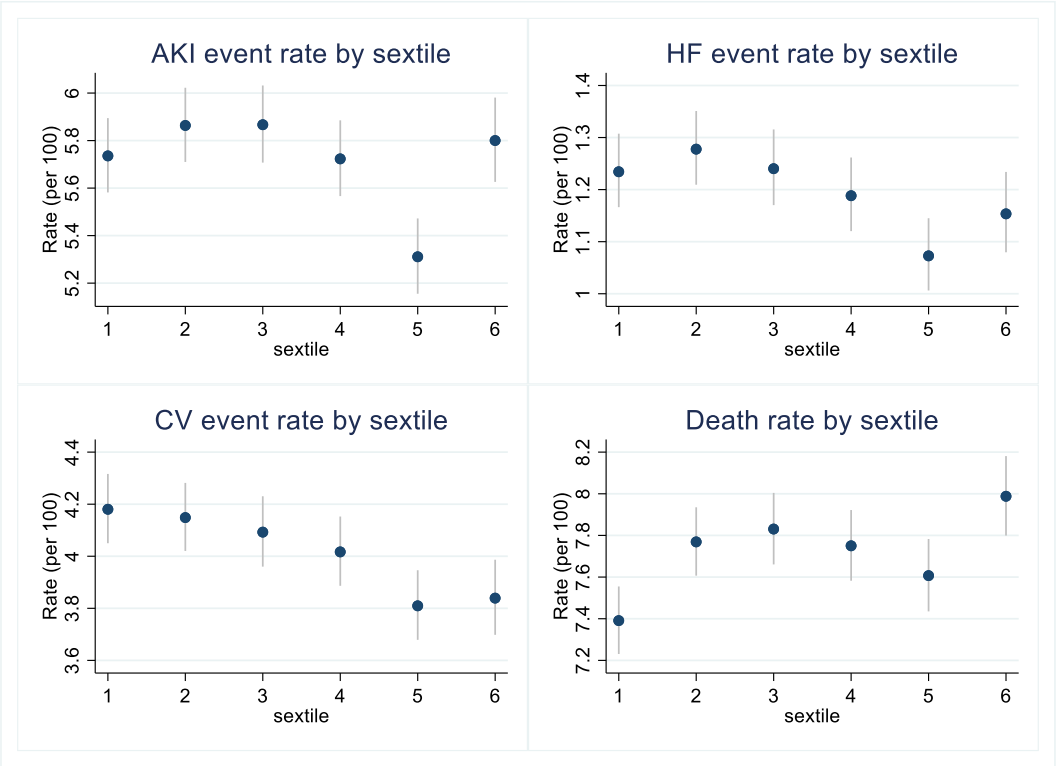
Supplementary Figure 1. Distribution of practice CKD coding performance in the CKD analysis population and in subgroups, with red lines depicting sextile boundaries



Supplementary Figure 2. Event rates per 100 patient years and 95% confidence intervals, by practice coding sextile in **all CKD** patients (including **recurring events**)



Supplementary Figure 3. Event rates per 100 patient years and 95% confidence intervals, by practice coding sextile in **all CKD** patients (**first events** only)



Supplementary Table 3. Event rates (including recurring events) per 100 patient years and 95% confidence intervals, by practice coding sextile

Practice coding sextile (percent coded CKD)	**S1 (<54.8%)**	**S2 (54.8% - 65.5%)**	**S3 (65.5% - 73.9%)**	**S4 (73.9% - 80.7%)**	**S5 (80.7% - 87.5%)**	**S6 (≥87.5%)**
CV event	N_E = 5766 5.94 (5.72, 6.16)	N_E = 5980 6.00 (5.78, 6.23)	N_E = 5326 5.81 (5.59, 6.03)	N_E = 5340 5.76 (5.54, 5.98)	N_E = 4677 5.38 (5.17, 5.60)	N_E = 4037 5.33 (5.11, 5.57)
HF	N_E = 1680 1.73 (1.61, 1.86)	N_E = 1889 1.89 (1.77, 2.03)	N_E = 1599 1.74 (1.63, 1.87)	N_E = 1596 1.72 (1.60, 1.85)	N_E = 1279 1.47 (1.37, 1.59)	N_E = 1216 1.61 (1.48, 1.74)
AKI	N_E = 7755 7.99 (7.74, 8.24)	N_E = 8114 8.14 (7.90, 8.39)	N_E = 7150 7.80 (7.55, 8.05)	N_E = 7189 7.75 (7.49, 8.02)	N_E = 6197 7.13 (6.90, 7.38)	N_E = 5947 7.85 (7.58, 8.14)
All-cause mortality	N_E = 7985 7.39 (7.23, 7.55)	N_E = 8612 7.77 (7.61, 7.93)	N_E = 7994 7.83 (7.66, 8.00)	N_E = 8014 7.75 (7.58, 7.92)	N_E = 7362 7.61 (7.43, 7.78)	N_E = 6740 7.99 (7.80, 8.18)

N_E = number of events

Supplementary Table 4 Event rates (first events only) per 100 patient years and 95% confidence intervals, by practice coding sextile

Practice coding sextile (percent coded CKD)	**S1 (<54.8%)**	**S2 (54.8% - 65.5%)**	**S3 (65.5% - 73.9%)**	**S4 (73.9% - 80.7%)**	**S5 (80.7% - 87.5%)**	**S6 (≥87.5%)**
CV event	N_E = 3834 4.18 (4.05, 4.32)	N_E = 3911 4.15 (4.02, 4.28)	N_E = 3553 4.09 (3.96, 4.23)	N_E = 3532 4.02 (3.89, 4.15)	N_E = 3144 3.81 (3.68, 3.95)	N_E = 2758 3.84 (3.70, 3.96)
HF	N_E = 1183 1.23 (1.17, 1.31)	N_E = 1257 1.28 (1.21, 1.35)	N_E = 1123 1.24 (1.17, 1.31)	N_E = 1089 1.19 (1.12, 1.26)	N_E = 922 1.07 (1.01, 1.14)	N_E = 863 1.15 (1.08, 1.23)
AKI	N_E = 5228 5.74 (5.58, 5.89)	N_E = 5480 5.86 (5.71, 6.02)	N_E = 5055 5.87 (5.71, 6.03)	N_E = 4993 5.72 (5.57, 5.88)	N_E = 4359 5.31 (5.16, 5.47)	N_E = 4135 5.80 (5.63, 5.98)
All-cause mortality	N_E = 7985 7.39 (7.23, 7.55)	N_E = 8612 7.77 (7.61, 7.93)	N_E = 7994 7.83 (7.66, 8.00)	N_E = 8014 7.75 (7.58, 7.92)	N_E = 7362 7.61 (7.43, 7.78)	N_E = 6740 7.99 (7.80, 8.18)

N_E = number of events

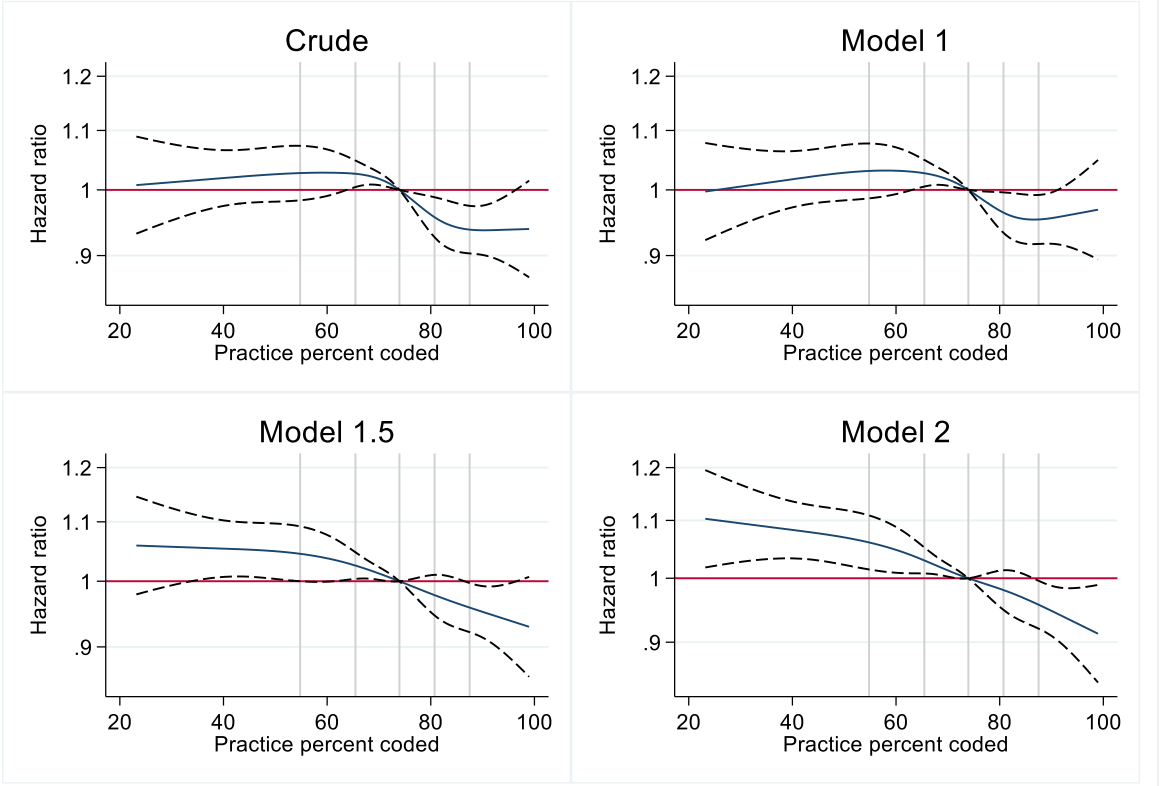
Supplementary Information 2. Methods for sequentially adjusted Cox regression analyses shown in supplementary analyses

Main Cox regression analyses, models 1 and 2, adjusted for practice characteristics are described in the main methods (repeated here for completeness) with results for all studied outcomes shown in the main results. Supplementary figures 4-23 show additional sequential adjustments, demonstrating the role of confounding as different variables were incorporated in analyses, as follows:

- Crude analysis: unadjusted
- Model 1: adjusted for ***practice characteristics reflecting practice risk profile*** (mean age, percent male, median rank of IMD, diabetes prevalence, hypertension prevalence, CVD prevalence)
- Model 1.5: adjusted for ***practice characteristics reflecting practice risk profile*** (model 1 variables), as well as ***practice characteristics of the detected CKD population*** (percent of CKD cases stages 3b-5 ["CKD severity"]), percent admitted for COPD in last 3 years, percent admitted for cancer in last 3 years)
- Model 2: adjusted for ***practice characteristics relating to overall practice risk profile*** and ***of the detected CKD population*** (model 1.5 variables), as well as ***testing biases*** which may result in confounding due to different vintages (i.e. duration of underlying disease) (percent GFR test in last year in diabetes, percent GFR test in last year in CKD, percent of adult population with detected CKD)

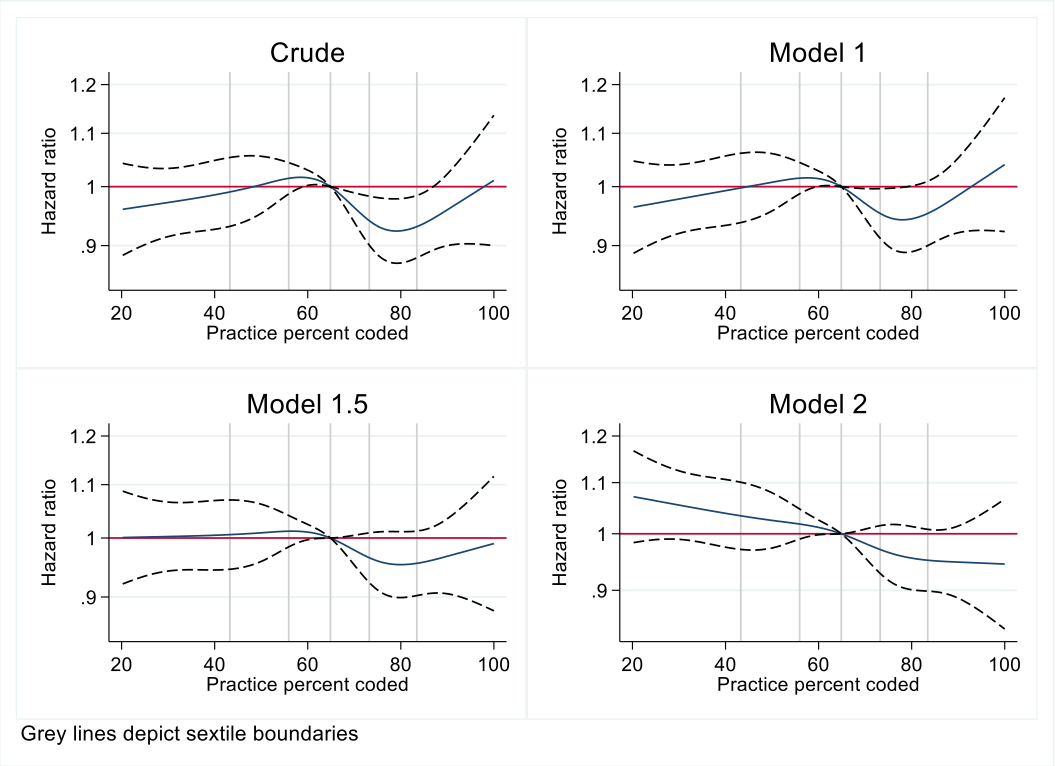
As in main analyses, adjusted hazard ratios (HRs) for outcomes with 95% confidence intervals were plotted across the spectrum of practice CKD coding, compared to average practice CKD coding.

Supplementary Figure 4. Hazard ratio splines for time to first **CV event**, compared to median (73.9%) practice percent coded (sequential adjustments)

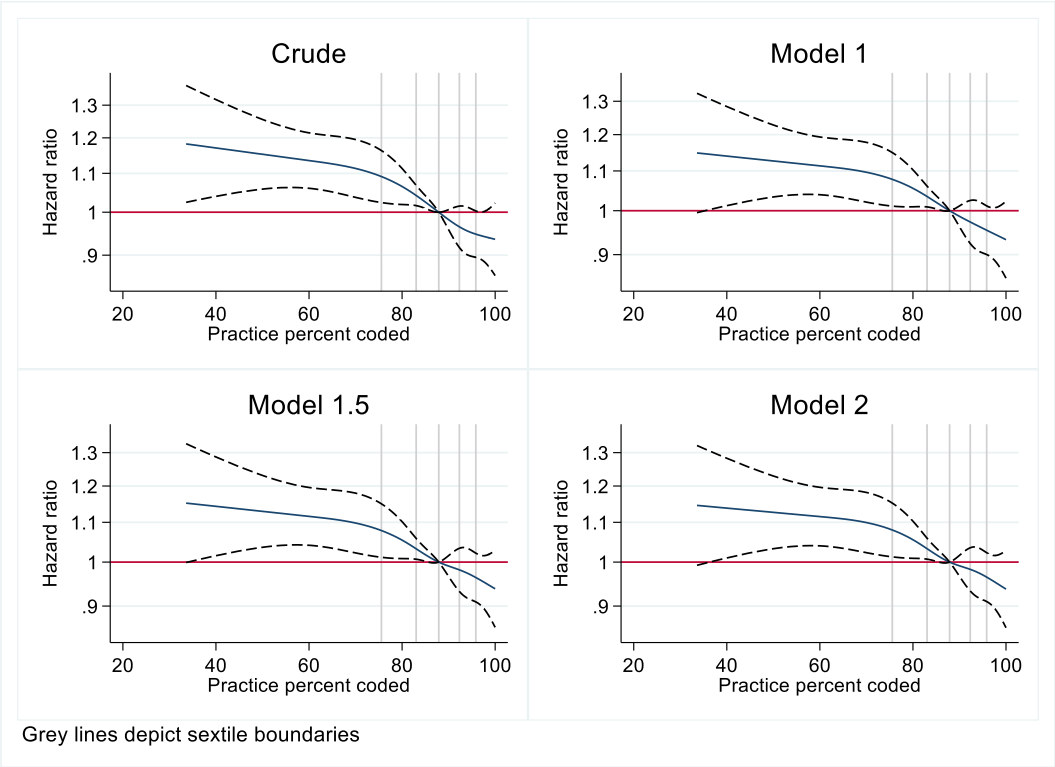


Grey lines depict sextile boundaries

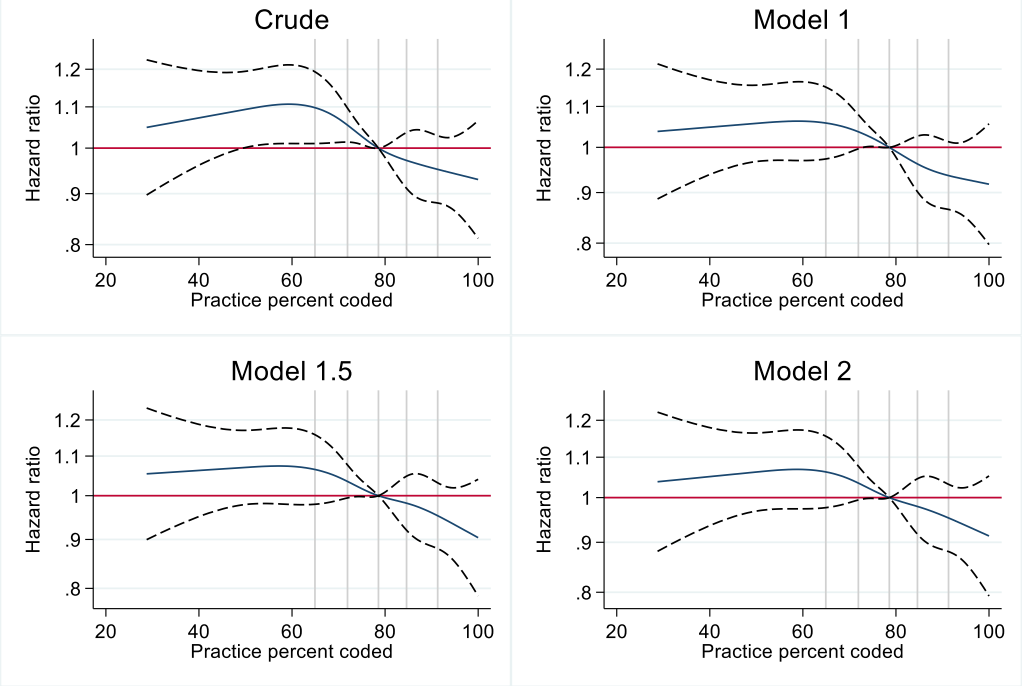
Supplementary Figure 5. Hazard ratio splines for time to first **CV event**, compared to median (64.9%) practice percent coded, in **CKD stage 3a only** (sequential adjustments)



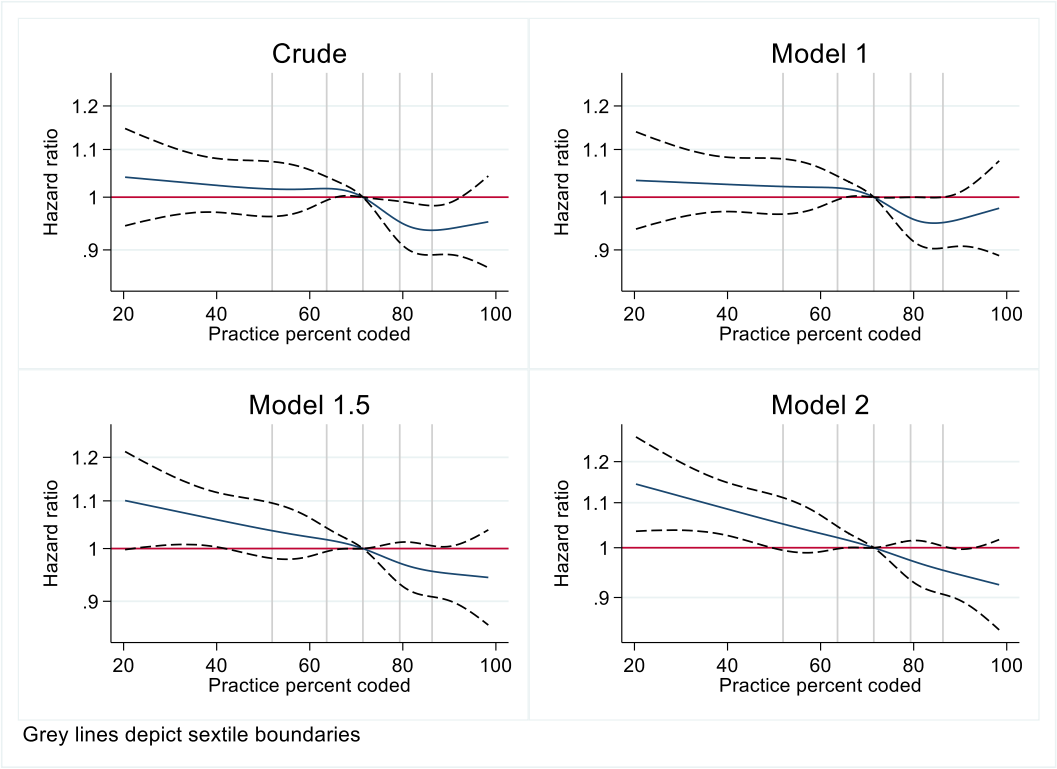
Supplementary Figure 6. Hazard ratio splines for time to first **CV event**, compared to median (87.9%) practice percent coded, in **CKD stages 3b-5 only** (sequential adjustments)



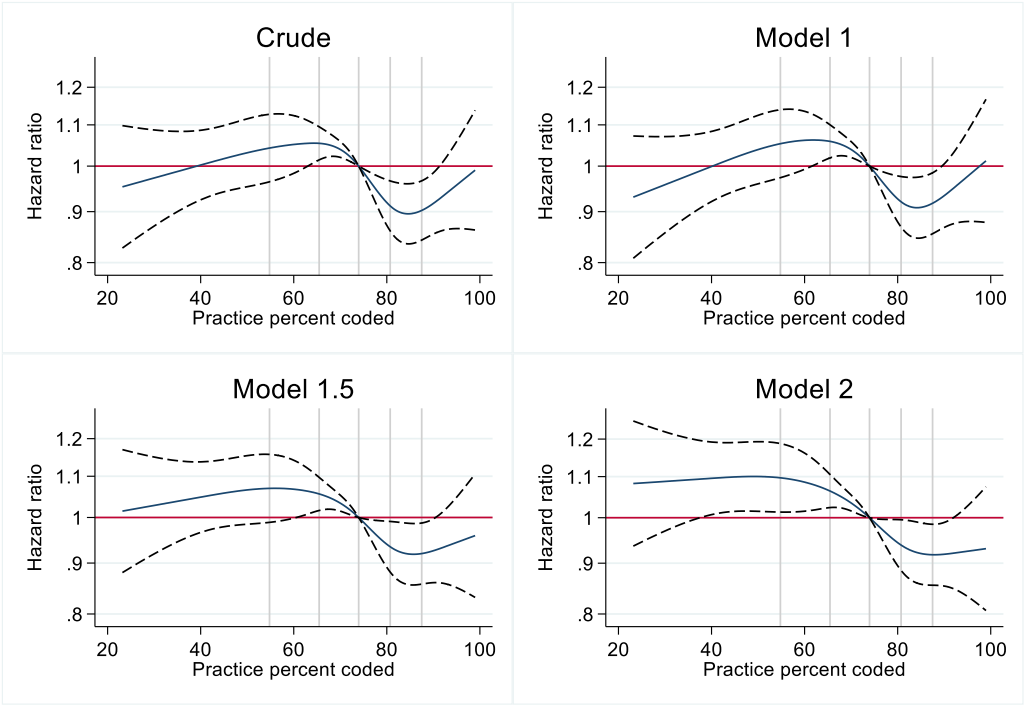
Supplementary Figure 7. Hazard ratio splines for time to first **CV event**, compared to median (78.6%) practice percent coded, in **diabetes only** (sequential adjustments)



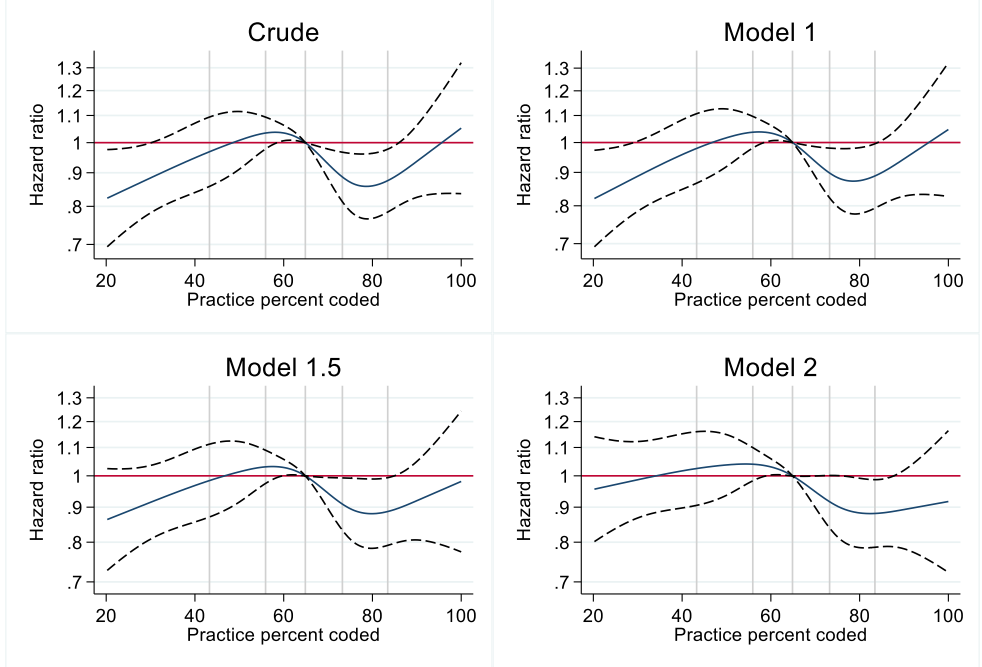
Supplementary Figure 8. Hazard ratio splines for time to first **CV event**, compared to median (71.4%) practice percent coded, in **no diabetes only** (sequential adjustments)



Supplementary Figure 9. Hazard ratio splines for time to first **HF** event, compared to median (73.9%) practice percent coded (sequential adjustments)

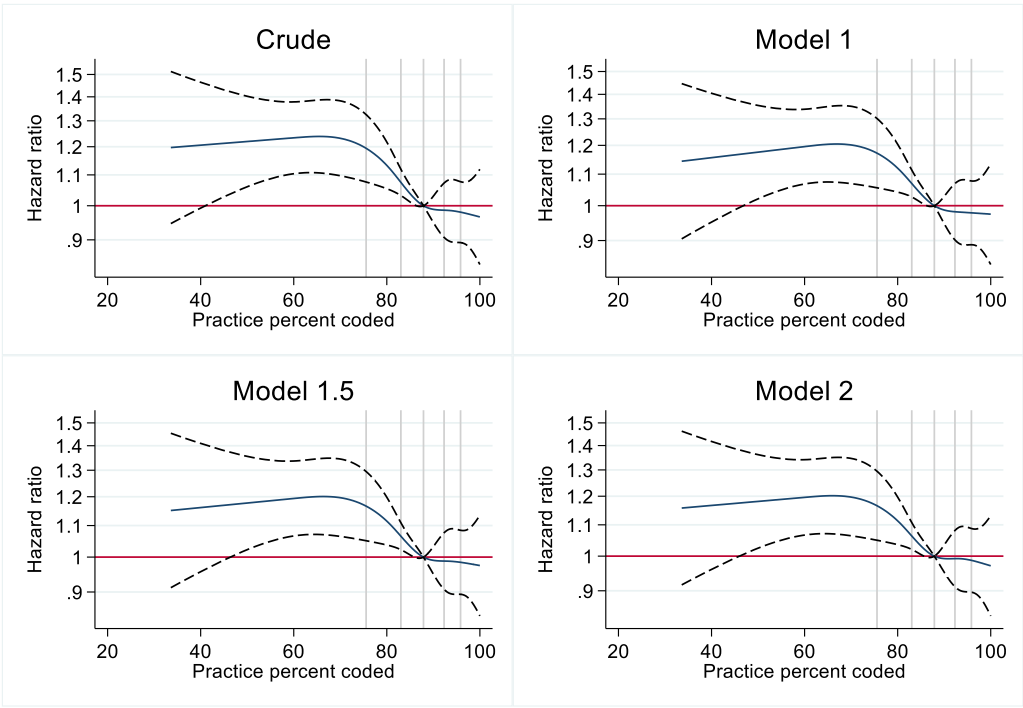


Supplementary Figure 10. Hazard ratio splines for time to first **HF** event, compared to median (64.9%) practice percent coded, in **CKD stage 3a only** (sequential adjustments)



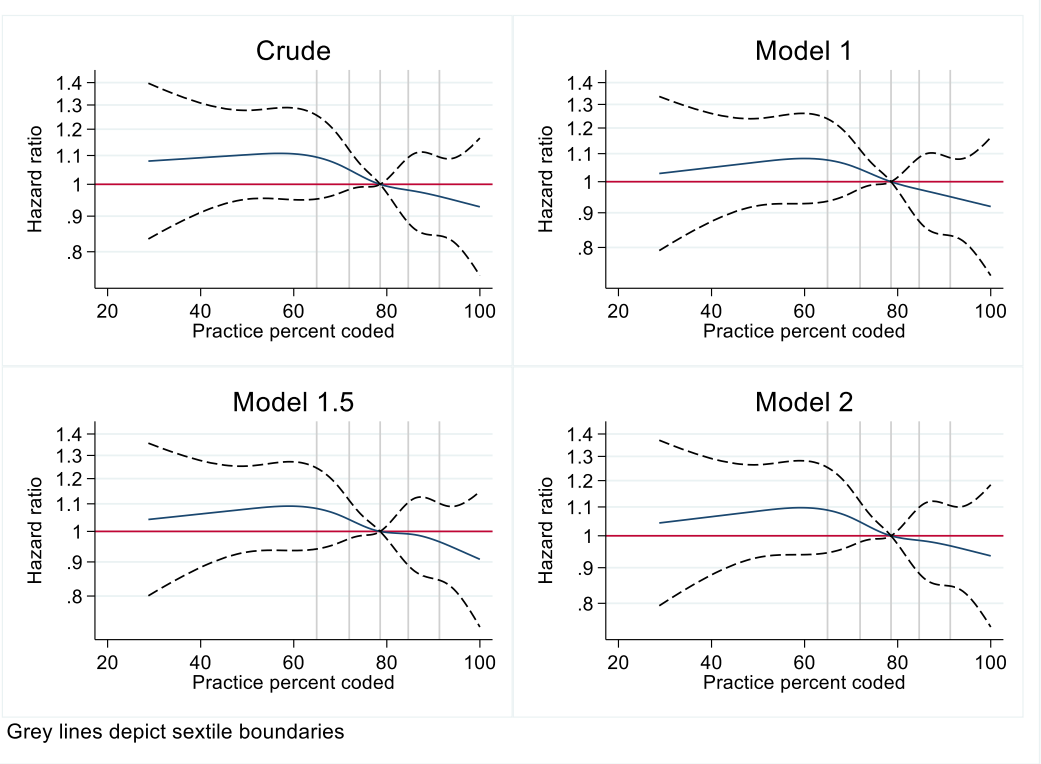
Grey lines depict sextile boundaries

Supplementary Figure 11. Hazard ratio splines for time to first **HF** event, compared to median (87.9%) practice percent coded, in **CKD stages 3b-5** only (sequential adjustments)

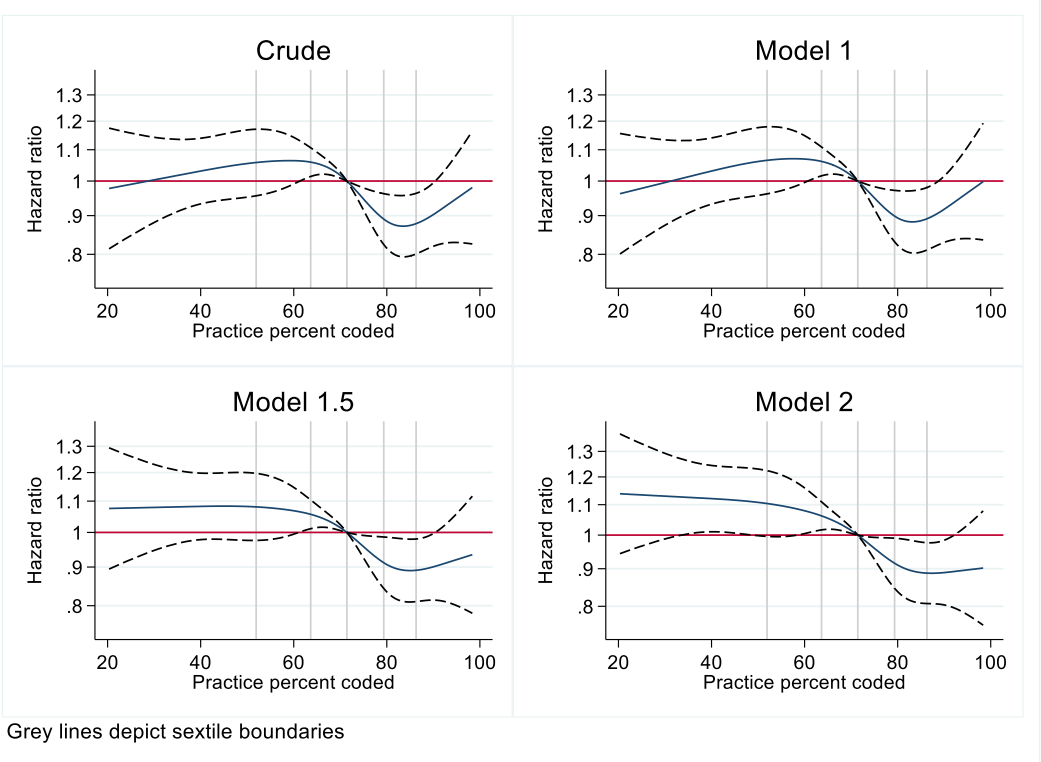


Grey lines depict sextile boundaries

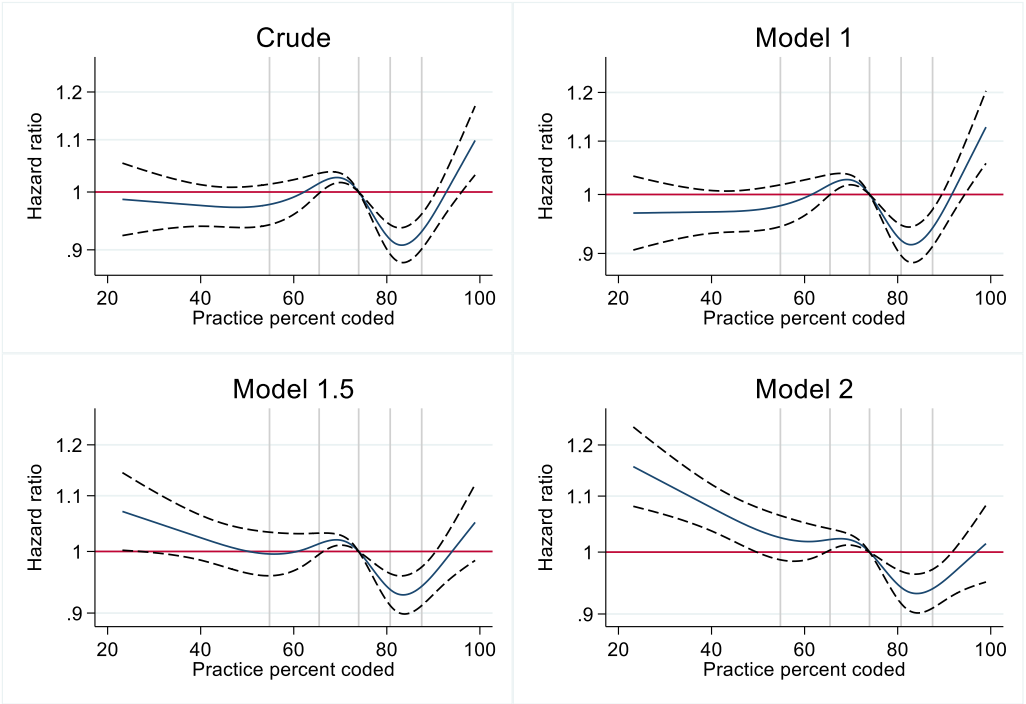
Supplementary Figure 12. Hazard ratio splines for time to first **HF** event, compared to median (78.6%) practice percent coded, in **diabetes only** (sequential adjustments)



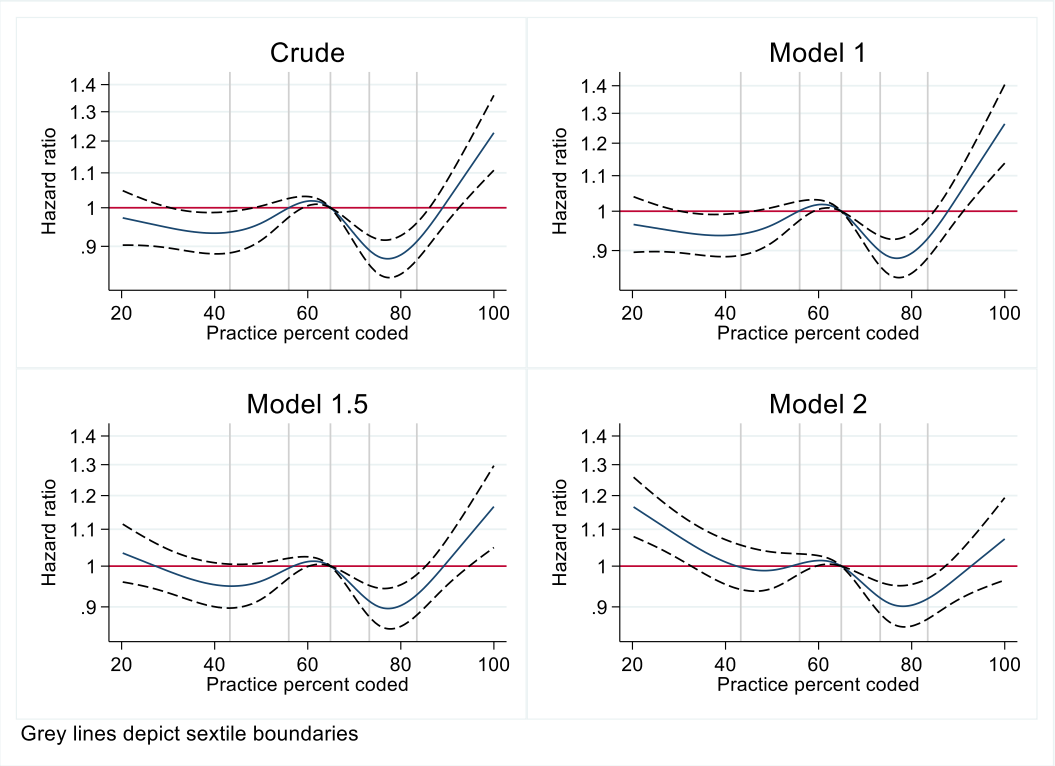
Supplementary Figure 13. Hazard ratio splines for time to first **HF** event, compared to median (71.4%) practice percent coded, in **no diabetes only** (sequential adjustments)



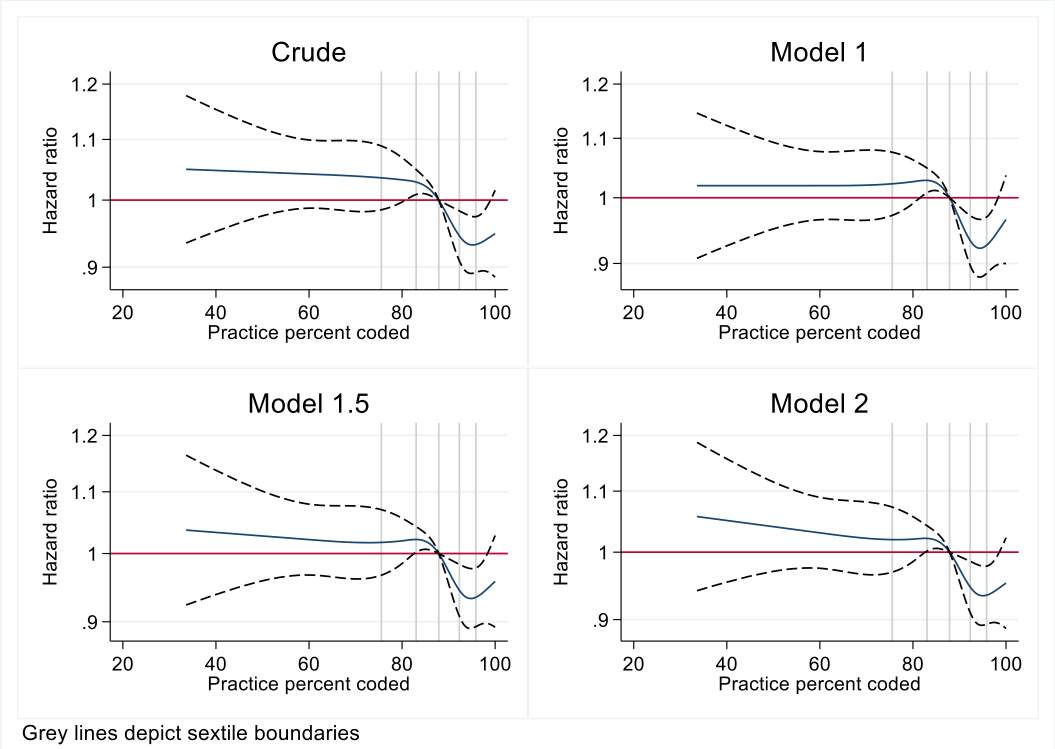
Supplementary Figure 14. Hazard ratio splines for time to first **AKI** event, compared to median (73.9%) practice percent coded (sequential adjustments)



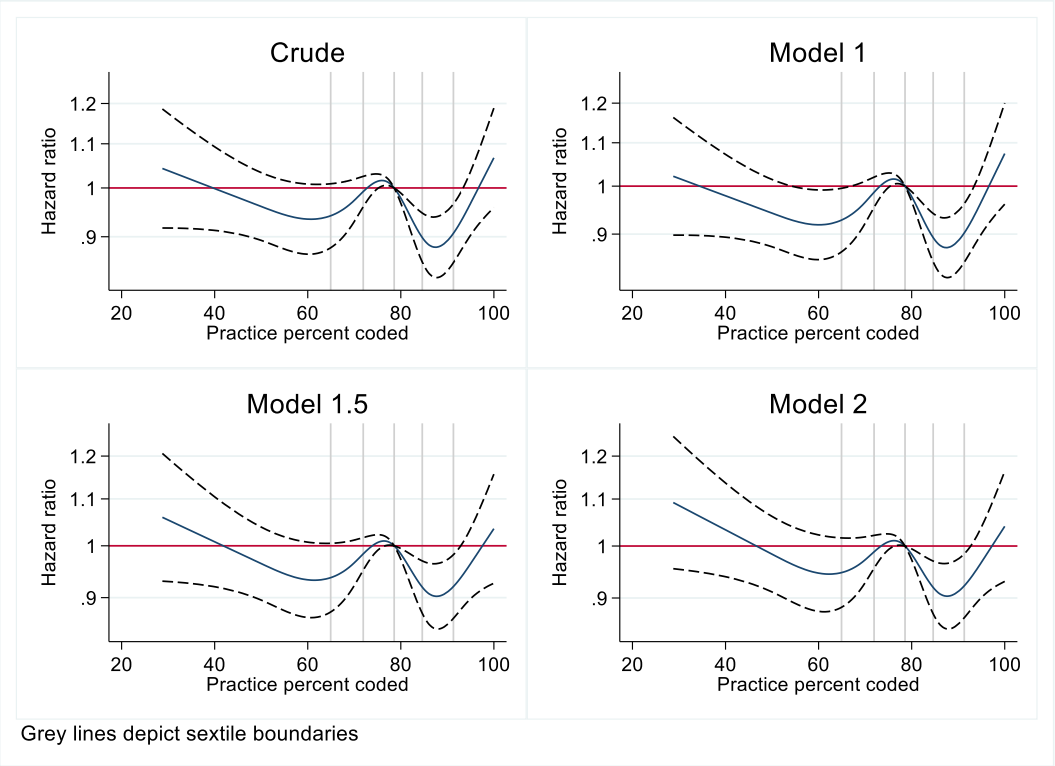
Supplementary Figure 15 Hazard ratio splines for time to first **AKI** event, compared to median (64.9%) practice percent coded, in **CKD stage 3a only** (sequential adjustments)



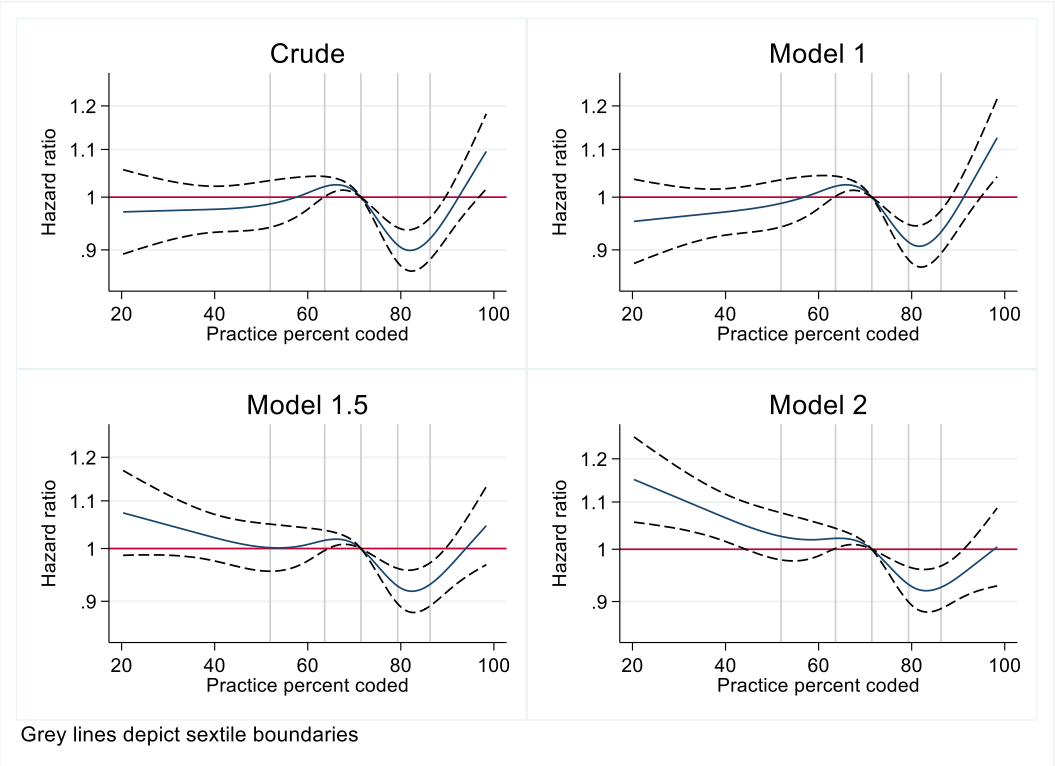
Supplementary Figure 16. Hazard ratio splines for time to first **AKI** event, compared to median (87.9%) practice percent coded, in **CKD stages 3b-5 only** (sequential adjustments)



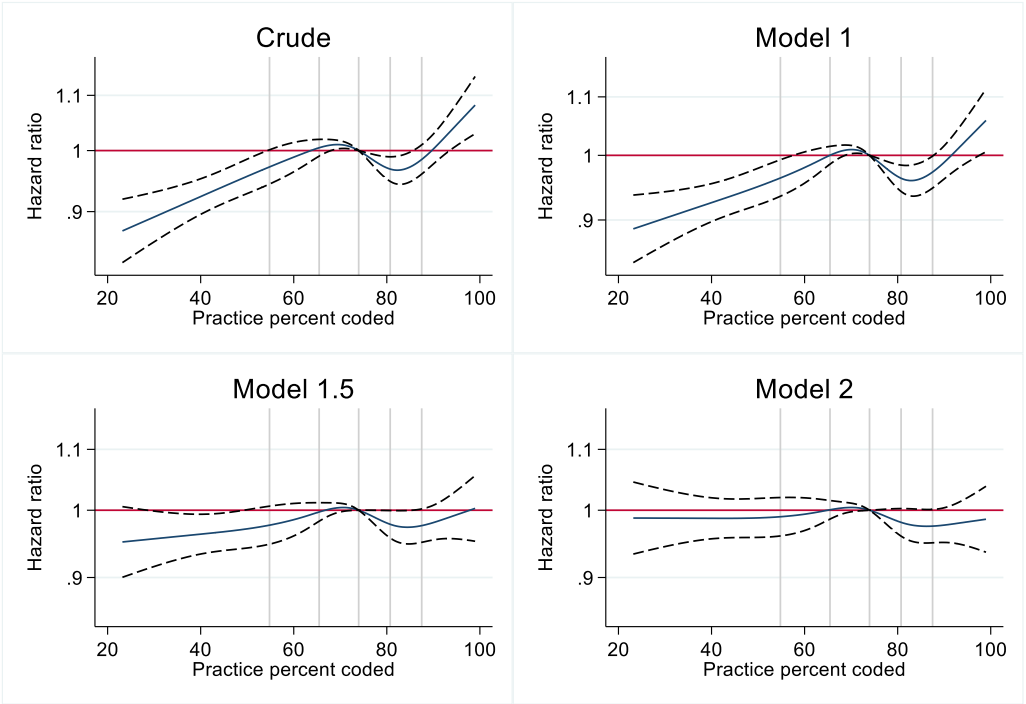
Supplementary Figure 17. Hazard ratio splines for time to first **AKI** event, compared to median (78.6%) practice percent coded, in **diabetes only** (sequential adjustments)



Supplementary Figure 18. Hazard ratio splines for time to first **AKI** event, compared to median (71.4%) practice percent coded, in **no diabetes only** (sequential adjustments)

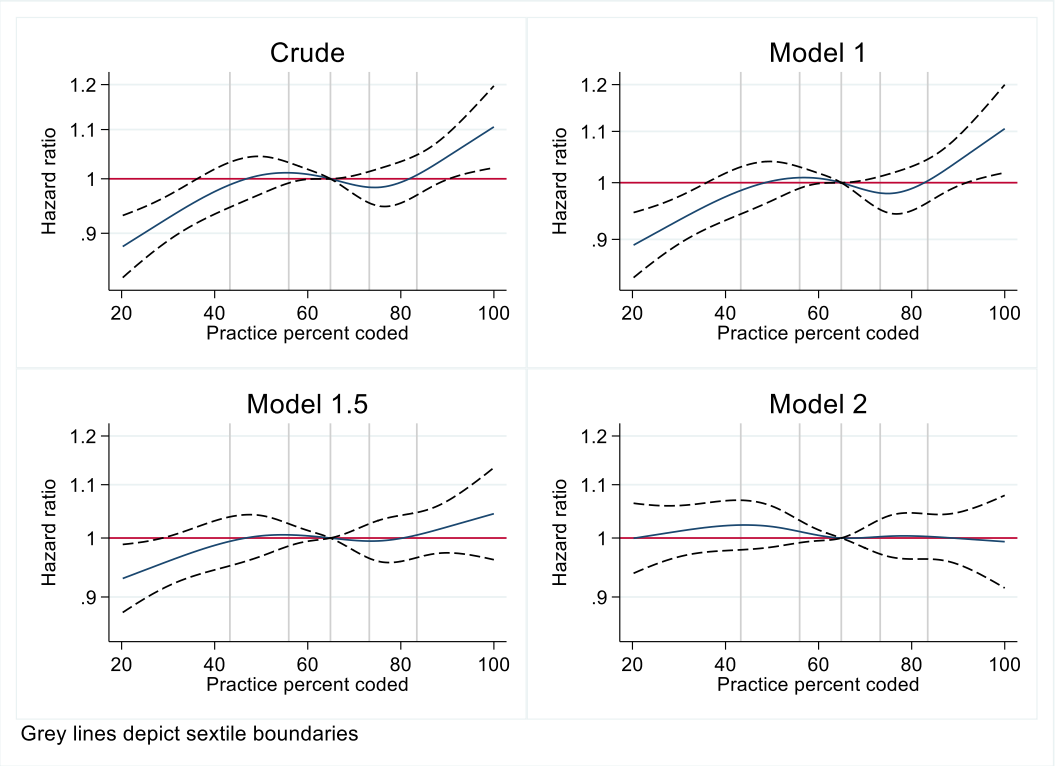


Supplementary Figure 19. Hazard ratio splines for time to **death**, compared to median (73.9%) practice percent coded (sequential adjustments)

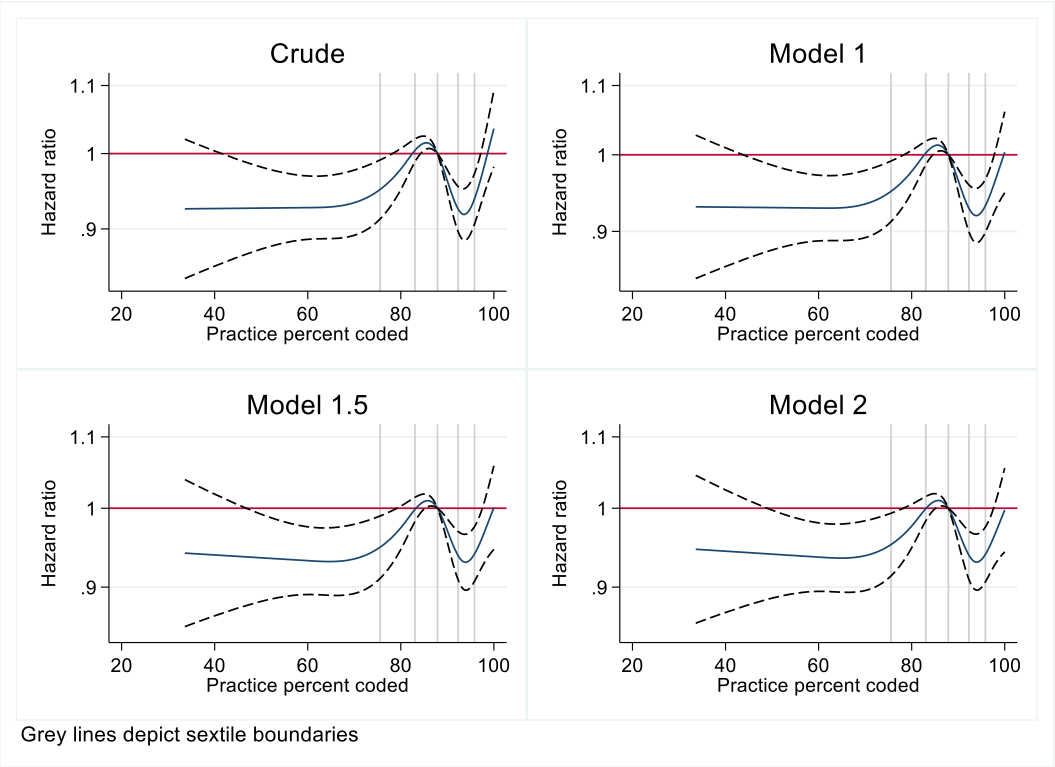


Grey lines depict sextile boundaries

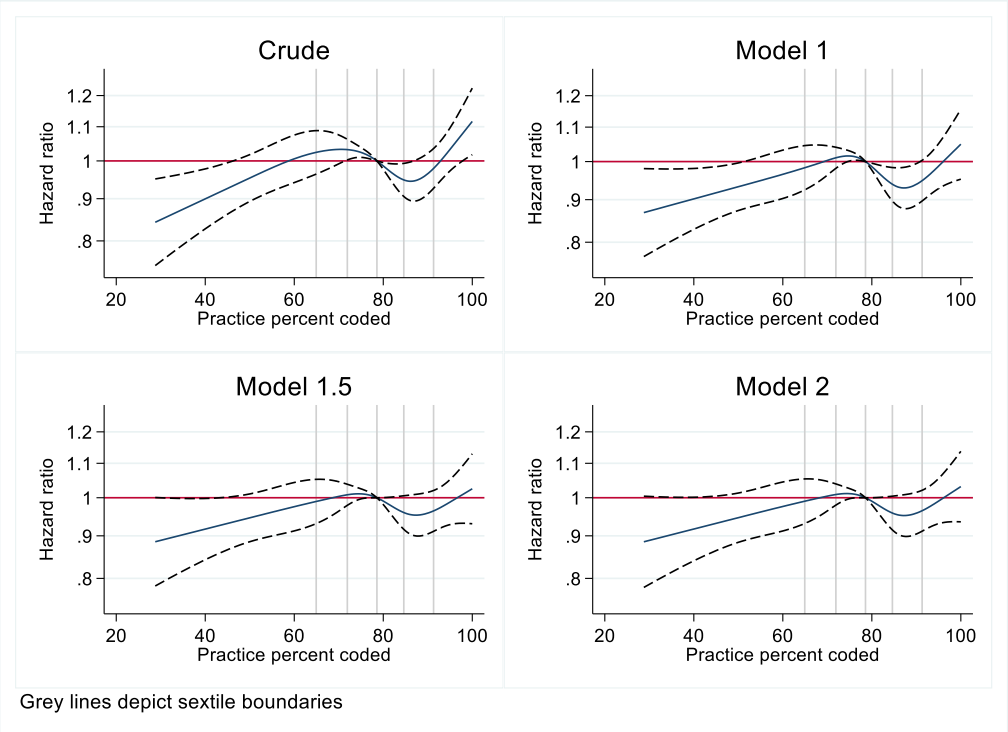
Supplementary Figure 20. Hazard ratio splines for time to **death**, compared to median (64.9%) practice percent coded, in **CKD stage 3a only** (sequential adjustments)



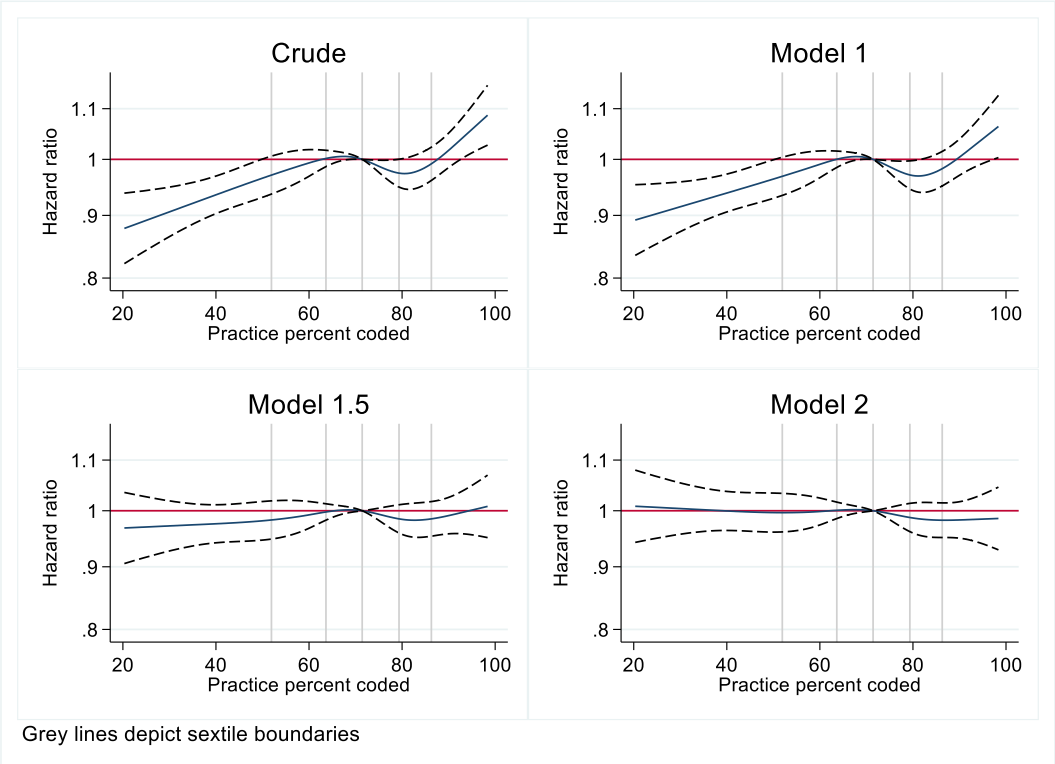
Supplementary Figure 21. Hazard ratio splines for time to **death**, compared to median (87.9%) practice percent coded, in **CKD stages 3b-5 only** (sequential adjustments)



Supplementary Figure 22. Hazard ratio splines for time to **death**, compared to median (78.6%) practice percent coded, in **diabetes only** (sequential adjustments)



Supplementary Figure 23. Hazard ratio splines for time to **death**, compared to median (71.4%) practice percent coded, in **no diabetes** only (sequential adjustments)

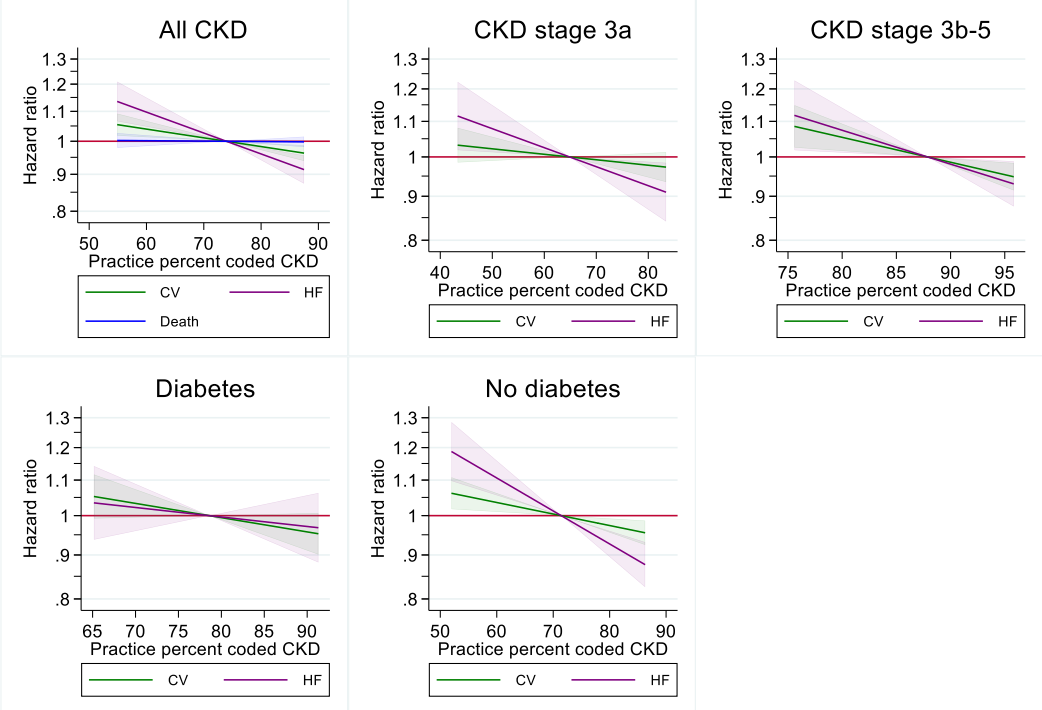


Supplementary Table 5. Fully adjusted hazard ratios (95% CI) per 10% increase in practice percent coded CKD for each outcome analysed, overall and by subgroup

	CV events	HF	Death
All CKD	0.973 (0.955, 0.991)*	0.936 (0.905, 0.967)*	0.999 (0.986, 1.011)
CKD stage 3a	0.985 (0.964, 1.007)	0.950 (0.911, 0.992)*	N/A**
CKD stage 3b-5	0.935 (0.893, 0.980)*	0.913 (0.846, 0.986)*	N/A**
Diabetes	0.963 (0.921, 1.006)	0.975 (0.906, 1.049)	N/A**
No diabetes	0.970 (0.948, 0.991)*	0.915 (0.879, 0.953)*	N/A**

Analysis not performed for outcomes and subgroups where assumption of linearity not supported by visual checks. (Descriptive likelihood ratio tests confirmed improved fit for linear term over 5-knot spline model in model 2 adjusted analyses, based on p-value threshold of 0.05, for all reported results, providing further justification for linearity assumption.)
Example interpretation: An adjusted hazard ratio of 0.973 for CV events represents a 6.4% reduction in the rate of CV hospitalisations for each 10% increase in practice CKD coding.

Supplementary Figure 24. Hazard ratio slope estimates (and 95% CI) for CV hospitalisations, HF hospitalisations and deaths, assuming linear differences in practice CKD coding performance compared to median practice coding, in the overall CKD population and within subgroups. *Analyses were carried out in the middle 4 sextiles of practices only (the two-thirds of most typically performing practices), with adjustment for all practice factors (model 2)*



Note: X-scales differ between subgroups, covering middle 4 sextiles within each subgroup

Supplementary Table 6. Adjusted hazard ratios for the association between practice behaviour variables and **AKI** events, sorted by point estimate

Practice behaviour	**HR (95% CI)**
Percent ACR/PCR test in last year in CKD (>58.7%)	0.959 (0.935, 0.984)*
Percent flu vaccination in last year in CKD (>78.8%)	0.965 (0.940, 0.990)*
Percent usage of ACEi/ARBs in hypertension (>76.6%)	0.967 (0.944, 0.990)*
Percent usage of statins in CVD (>93.0%)	0.969 (0.944, 0.995)*
Percent usage of statins in diabetes (>84.1%)	0.989 (0.964, 1.015)
Percent pneumococcus vaccination in past 5 years in CKD stages 4-5 (>12.5%)	1.008 (0.985, 1.032)
Percent meeting blood pressure target in last year in CKD (>57.8%)	1.030 (1.004, 1.055)*

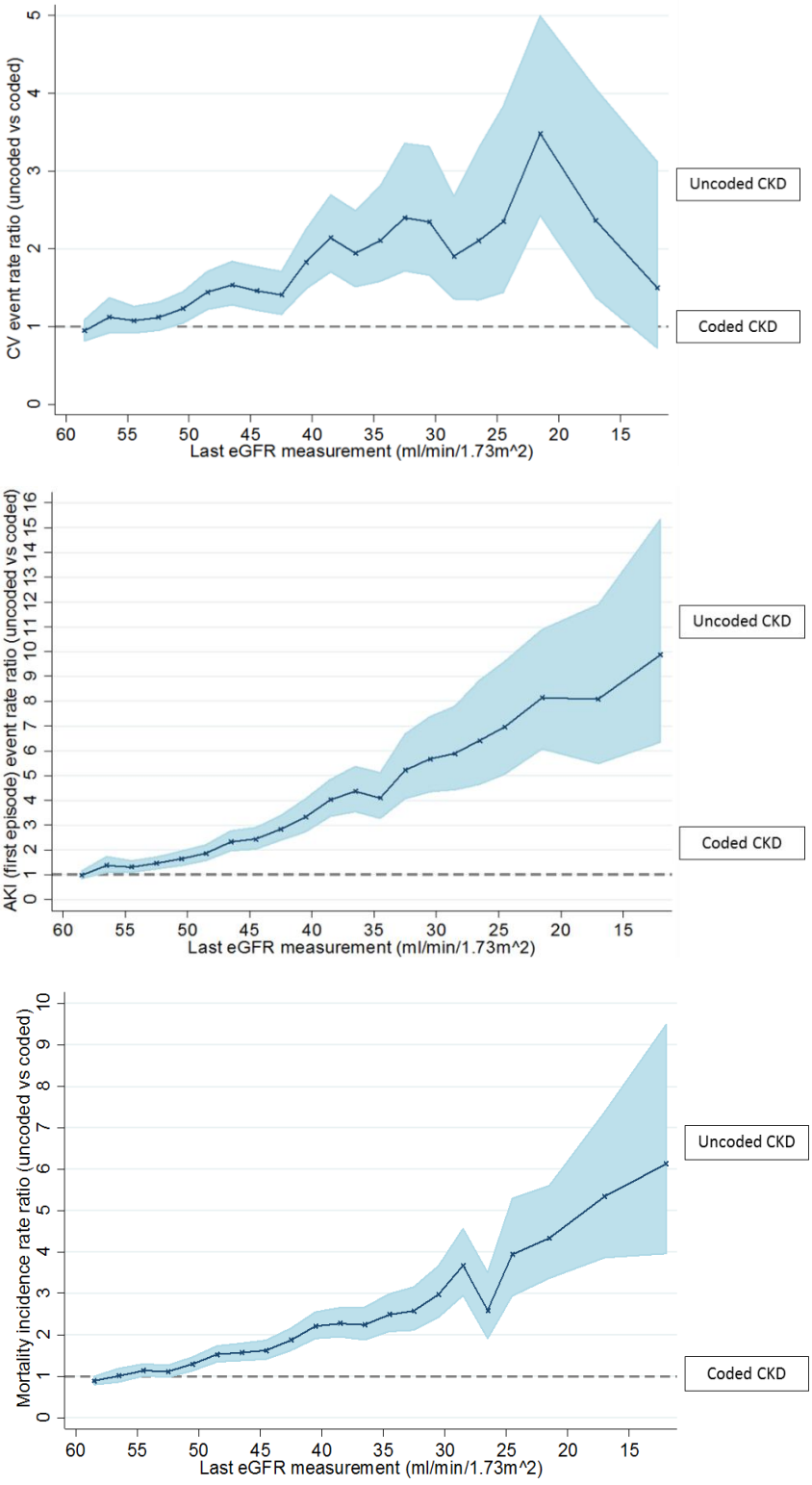
Supplementary Table 7. Adjusted hazard ratios for the association between practice behaviour variables and **HF** events, sorted by point estimate

Practice behaviour	**HR (95% CI)**
Percent usage of ACEi/ARBs in hypertension (>76.6%)	0.935 (0.888, 0.984)*
Percent ACR/PCR test in last year in CKD (>58.7%)	0.951 (0.901, 1.004)
Percent flu vaccination in last year in CKD (>78.8%)	0.959 (0.908, 1.013)
Percent meeting blood pressure target in last year in CKD (>57.8%)	0.976 (0.926, 1.028)
Percent usage of statins in CVD (>93.0%)	0.985 (0.932, 1.042)
Percent pneumococcus vaccination in past 5 years in CKD stages 4-5 (>12.5%)	1.023 (0.973, 1.075)
Percent usage of statins in diabetes (>84.1%)	1.040 (0.984, 1.100)

Supplementary Table 8. Adjusted hazard ratios for the association between practice behaviour variables and **deaths**, sorted by point estimate with descriptive p-values

Practice behaviour	**HR (95% CI)**
Percent usage of statins in CVD (>93.0%)	0.950 (0.930, 0.970)*
Percent usage of ACEi/ARBs in hypertension (>76.6%)	0.965 (0.947, 0.984)*
Percent usage of statins in diabetes (>84.1%)	0.971 (0.951, 0.991)*
Percent pneumococcus vaccination in past 5 years in CKD stages 4-5 (>12.5%)	1.002 (0.983, 1.020)
Percent flu vaccination in last year in CKD (>78.8%)	1.007 (0.987, 1.027)
Percent meeting blood pressure target in last year in CKD (>57.8%)	1.010 (0.992, 1.030)
Percent ACR/PCR test in last year in CKD (>58.7%)	1.012 (0.992, 1.033)

Supplementary Figure 25. Results from prior NCKDA analyses, copied from manuscript reference [12]: Hazard ratios for CV events, AKI and mortality in uncoded vs coded patients, adjusted for patient age, sex, and presence of coded diabetes, hypertension and CVD, stratified by latest eGFR measurement, among patients with biochemical evidence of CKD



Supplementary Information 3. Reporting of studies Conducted using Observational Routinely-collected Data (RECORD) checklist

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1, 2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
		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.	2
		RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract.	2
		RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-8
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	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.	7
		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.	N/A
		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.	6 (described only, no diagram)

Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-10; Supplementary Table 1
		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 reported, an explanation should be provided.	7-10; Supplementary Table 1
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	6-10
Bias	9	Describe any efforts to address potential sources of bias	10-11
Study size	10	Explain how the study size was arrived at	7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8-12
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10-12
		(*b*) Describe any methods used to examine subgroups and interactions	11
		(*c*) Explain how missing data were addressed	7-11
		(*d*) If applicable, explain how loss to follow-up was addressed	8, 10
		(*e*) Describe any sensitivity analyses	11
		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.	6-7
		RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	7-10
		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.	6
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	12, 14, 17
		(b) Give reasons for non-participation at each stage	12, 17; Also see methods p.6-8
		(c) Consider use of a flow diagram	12

		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.	7, 12
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	14
		(b) Indicate number of participants with missing data for each variable of interest	See methods p.6-10
		(c) Summarise follow-up time (eg, average and total amount)	14, 17
Outcome data	15*	Report numbers of outcome events or summary measures over time	17; Also Sup. Figures 2-3, Sup. Table 2
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	17-19
		(*b*) Report category boundaries when continuous variables were categorized	14, 17-19
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	17-18
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	17-19
Discussion			
Key results	18	Summarise key results with reference to study objectives	19
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	19, 21
		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.	20-21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	20-22
Generalisability	21	Discuss the generalisability (external validity) of the study results	20-21
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	23
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.	24; Supplementary materials

*Information should be provided separately for exposed and unexposed groups in cohort studies

Appendix 6

Supplementary materials for research paper 4 (risk prediction modelling)

See next page for supplementary materials for research paper 4.

Supplementary Table 1: Frequency of identifiable CKD cases (2 x eGFR<60 over 90+ days) and availability of eGFR and albuminuria data at the time point of identification, over various time periods indicating possible analysis cohorts

	2006-2014	**2010-2014**	**2006-2018**	**2010-2018** (Final cohort)
Number of CKD patients identified in period	N = 104,048	N = 76,897	N = 140,991	N = 116,158
Number of eGFR tests up to and including date of confirmed CKD, median (IQR)	4 (2, 8)	9 (6, 14)	6 (3, 11)	10 (7, 16)
Number of patients with ≥ 3 eGFR tests as of index date	72,127 (69.3%)	75,238 (97.8%)	107,802 (76.5%)	113,350 (97.6%)
Number of patients with albuminuria testing within +/- 3 months of date of confirmed CKD				
Any albuminuria test[a]	**23,544 (22.6%)**	**17,340 (22.6%)**	**35,816 (25.4%)**	**30,304 (26.1%)**
uACR	10,164 (9.8%)	9,621 (12.5%)	18,245 (12.9%)	18,100 (15.6%)
PCR	1,514 (1.5%)	1,239 (1.6%)	2,247 (1.6%)	2,001 (1.7%)
Dipstick albumin	16,293 (15.7%)	11,217 (14.6%)	23,079 (16.4%)	18,395 (15.8%)
Number of patients with albuminuria testing within +/- 6 months of date of confirmed CKD				
Any albuminuria test[a]	**31,395 (30.2%)**	**23,276 (30.3%)**	**47,484 (33.7%)**	**40,296 (34.7%**)
uACR	14,036 (13.5%)	13,373 (17.4%)	24,706 (17.5%)	24,549 (21.1%)
PCR	2,353 (2.3%)	1,928 (2.5%)	3,515 (2.5%)	3,130 (2.7%)
Dipstick albumin	22,555 (21.7%)	15,627 (20.3%)	32,186 (22.8%)	25,851 (22.3%)
Number of patients with albuminuria testing within +/- 12 months of date of confirmed CKD				
Any albuminuria test[a]	**40,494 (38.9%)**	**30,465 (39.6%)**	**61,011 (43.3%)**	**52,173 (44.9%)**
uACR	19,037 (18.3%)	17,957 (23.4%)	32,739 (23.2%)	32,311 (27.8%)
PCR	3,394 (3.3%)	2,857(3.7%)	5,031 (3.6%)	4,552 (3.9%)
Dipstick albumin	30,013 (28.8%)	21,426 (27.9%)	43,455 (30.8%)	35,711 (30.7%)
Number of patients with albuminuria testing within **- 12 months to +3 months** of date of confirmed CKD				
Any albuminuria test[a]	**33,530 (32.2%)**	**25,473 (33.1%)**	**51,071 (36.2%)**	**43,974 (37.9%)**
uACR	14,571 (14.0%)	14,538 (18.9%)	26,164 (18.6%)	26,658 (23.0%)
PCR	2,373 (2.3%)	2,037 (2.7%)	3,602 (2.6%)	3,307 (2.9%)
Dipstick albumin	24,137 (23.2%)	17,140 (22.3%)	34,833 (24.7%)	28,457 (24.5%)

Age[b] 18-39 40-59 60-79 80+	 539 (0.5%) 4,591 (4.4%) 46,506 (44.7%) 52,412 (50.4%)	 364 (0.5%) 3,099 (4.0%) 33,409 (43.5%) 40,025 (52.1%)	 762 (0.5%) 6,923 (4.9%) 67,356 (47.8%) 65,950 (46.8%)	 590 (0.5%) 5,487 (4.7%) 55,161 (47.5%) 54,920 (47.3%)
Female	61,233 (58.9%)	44,746 (58.2%)	81,343 (57.7%)	66,429 (57.2%)
CKD stage[b] 3a 3b 4	 73,095 (70.3%) 24,935 (24.0%) 6,018 (5.8%)	 51,133 (66.5%) 20,379 (26.5%) 5,385 (7.0%)	 104,263 (74.0%) 29,971 (21.3%) 6,757 (4.8%)	 84,065 (72.4%) 25,873 (22.3%) 6,220 (5.4%)
Diabetes[b]	23,447 (22.5%)	18,993 (24.7%)	31,944 (22.7%)	27,953 (24.1%)
Hypertension[b]	69,236 (66.5%)	57,865 (75.3%)	96,689 (68.6%)	87,125 (75.0%)
Total number of **KRT** events [and event rate]	1,972 (1.9%) [2.78 per 1000 py]	1,412 (1.8%) [2.98 per 1000 py]	2,170 (1.5%) [2.51 per 1000 py]	1,618 (1.4%) [2.53 per 1000 py]
Number of **KRT** events in next 5 years[c] [and event rate]	787 (0.8%) [1.90 per 1000 py]	695 (0.9%) [2.28 per 1000 py]	956 (0.7%) [1.71 per 1000 py]	870 (0.7%) [1.90 per 1000 py]
Total number of **eGFR < 15** events[d] [and event rate]	6,897 (6.6%) [9.90 per 1000 py]	4,985 (6.5%) [10.7 per 1000 py]	7,594 (5.4%) [8.91 per 1000 py]	5,729 (4.9%) [9.10 per 1000 py]
Number of **eGFR < 15** events in next 5 years[c,d] [and event rate]	3,781 (3.6%) [9.21 per 1000 py]	3,115 (4.1%) [10.3 per 1000 py]	4,415 (3.1%) [7.98 per 1000 py]	3,794 (3.3%) [8.38 per 1000 py]
Total number of **eGFR<15 events with no subsequent eGFR >= 15**[d] [and event rate]	4,600 (4.4%) [6.52 per 1000 py]	3,356 (4.4%) [7.13 per 1000 py]	5,020 (3.6%) [5.83 per 1000 py]	3,796 (3.3%) [5.97 per 1000 py]
Number of **eGFR<15 events in next 5 years, with no subsequent eGFR >= 15**[c,d] [and event rate]	2,159 (2.1%) [5.22 per 1000 py]	1,842 (2.4%) [6.06 per 1000 py]	2,524 (1.8%) [4.53 per 1000 py]	2,227 (1.9%) [4.89 per 1000 py]
Total number of **deaths** [and event rate]	72,644 (69.8%) [101 per 1000 py]	48,513 (63.1%) [101 per 1000 py]	83,161 (59.0%) [95.1 per 1000 py]	60,005 (51.7%) [93.1 per 1000 py]
Number of **deaths** in next 5 years[c] [and event rate]	38,835 (37.3%) [93.2 per 1000 py]	29,227 (38.0%) [95.4 per 1000 py]	48,479 (34.4%) [86.5 per 1000 py]	39,743 (34.2%) [86.6 per 1000 py]

[a]Number of patients with any test (uACR, PCR or dipstick). Note: This is not the sum of different types of albuminuria tests, as patients can have multiple different tests within the time window.

[b]Risk factors are defined at first biochemical evidence of CKD (2 x eGFR <60, separated by ≥90 days).

[c]Note: Outcomes censored at: 5 years since index date; end of follow up for outcomes (31st December 2021); or death, whichever occurs first. Note: Shorter follow up times will be available for cohort entry periods covering more recent years.

[d]Assessed using pre-KRT eGFR results only

Supplementary Table 2: ICD-10 codes (and ATC codes, where specified) for candidate predictor variables

Comorbidity	**ICD-10 codes (and additional ATC codes, where specified)**
Diabetes mellitus	E10-14
Hypertension	I10-15
Heart failure (HF)	I50
Coronary heart disease (CHD)	I21, I22
Atrial fibrillation (AF)	I48
Stroke (all-cause)	I60-64
Peripheral artery disease (PAD)	I70-73
Chronic obstructive pulmonary disorder (COPD)	I278, I279
Acute kidney injury (AKI)	N17
Use of ACEi/ARBs	ATC codes: C09A, C09B, C09C, C09D

Supplementary Table 3: Availability of creatinine data and follow up for outcomes in years following index date (excluding index date), in CKD analysis cohort, overall and stratified by factors which may predict likelihood of creatinine testing (leading to ascertainment bias)

	Year 1	Year 2	Year 3	Year 4	Year 5
Number of patients starting follow up: N = 116,158					
eGFR test recorded in year, N (%)	91,832 (79.1%)	87,581 (75.4%)	79,488 (68.4%)	70,930 (61.1%)	58,271 (50.2%)
Pre-KRT eGFR test recorded in year, N (%)	91,829 (79.1%)	87,498 (75.3%)	79,295 (68.3%)	70,617 (60.8%)	57,840 (49.8%)
Number at risk at beginning of year (following index date) [a]	**N = 116,158**	**N = 106,482**	**N = 97,809**	**N = 89,840**	**N = 82,347**
eGFR test recorded in year if at risk at beginning of year, N (%)	91,832 (79.1%)	87,509 (82.2%)	79,305 (81.1%)	70,630 (78.6%)	57,852 (70.3%)
Pre-KRT eGFR test recorded in year if at risk at beginning of year, N (%)	91,829 (79.1%)	87,498 (82.2%)	79,295 (81.1%)	70,615 (78.6%)	57,840 (70.2%)
eGFR test frequency per year if at risk at beginning of year, Median (IQR)[b]	2 (1, 3)	2 (1, 3)	2 (1, 3)	2 (1, 3)	1 (0, 3)
KRT event in year, N (%)	96 (0.08%)	143 (0.13%)	177 (0.18%)	214 (0.24%)	240 (0.29%)
Non-rebounding eGFR<15 event in year, N (%)	409 (0.35%)	504 (0.47%)	456 (0.47%)	442 (0.49%)	416 (0.51%)
Deaths in year, N (%)	9,600 (8.3%)	8,551 (8.0%)	7,845 (8.0%)	7,342 (8.2%)	6,405 (7.8%)
Diabetes: Number at risk	**N = 27,953**	**N = 25,557**	**N = 23,272**	**N = 21,215**	**N = 19,179**
Pre-KRT eGFR test recorded	24,004 (85.9%)	22,732 (88.9%)	20,396 (87.6%)	18,203 (85.8%)	14,609 (76.2%)
KRT event	48 (0.17%)	77 (0.30%)	83 (0.36%)	93 (0.44%)	103 (0.54%)

uACR tested:					
Number at risk	**N = 26,658**	**N = 25,206**	**N = 23,698**	**N = 22,200**	**N = 20,669**
Pre-KRT eGFR test recorded	23,247 (87.2%)	22,594 (89.6%)	20,872 (88.1%)	18,758 (84.5%)	14,674 (71.0%)
KRT event	77 (0.29%)	115 (0.46%)	140 (0.59%)	155 (0.70%)	179 (0.87%)
uACR not tested:					
Number at risk	**N = 89,500**	**N = 81,276**	**N = 74,111**	**N = 67,640**	**N = 61,678**
Pre-KRT eGFR test recorded	68,582 (76.6%)	64,904 (79.9%)	58,423 (78.8%)	51,857 (76.7%)	43,166 (70.0%)
KRT event	19 (0.02%)	28 (0.03%)	37 (0.05%)	59 (0.09%)	61 (0.10%)
CKD stage 3a:					
Number at risk	**N = 84,065**	**N = 79,094**	**N = 73,980**	**N = 69,111**	**N = 64,408**
Pre-KRT eGFR test recorded	65,113 (77.5%)	64,828 (82.0%)	59,870 (80.9%)	53,936 (78.9%)	44,224 (68.7%)
KRT event	14 (0.02%)	15 (0.02%)	29 (0.04%)	41 (0.06%)	65 (0.10%)
CKD stage 3b:					
Number at risk	**N = 25,873**	**N = 22,507**	**N = 19,836**	**N = 17,495**	**N = 15,326**
Pre-KRT eGFR test recorded	21,331 (82.4%)	18,502 (82.2%)	16,060 (81.0%)	14,005 (80.1%)	11,547 (75.3%)
KRT event	15 (0.06%)	30 (0.13%)	58 (0.29%)	77 (0.44%)	74 (0.48%)
CKD stage 4:					
Number at risk	**N = 6,220**	**N = 4,881**	**N = 3,993**	**N = 3,234**	**N = 2,613**
Pre-KRT eGFR test recorded	5,385 (86.6%)	4,168 (85.4%)	3,365 (84.3%)	2,674 (82.7%)	2,069 (79.2%)
KRT event	67 (1.08%)	98 (2.01%)	90 (2.25%)	96 (2.97%)	101 (3.87%)

[a]To be at risk at beginning of year, patient must be alive and not have started KRT

[b]Results shown for all patients, including those with no eGFR tests in year

Supplementary Table 4: Patient characteristics in development and validation cohorts

Patient characteristics	Development cohort	Validation cohort
N	**N = 85,012**	**N = 28,338**
Age, median (IQR)	80 (72, 86)	76 (70, 83)
Sex female, n (%)	49,158 (57.8%)	15,548 (54.9%)
Diabetes, n (%)	21,117 (24.8%)	6,587 (23.2%)
Hypertension, n (%)	64,316 (75.7%)	21,378 (75.4%)
CHD, n (%)	9,596 (11.3%)	2,572 (9.1%)
HF, n (%)	21,790 (25.6%)	5,329 (18.8%)
AF, n (%)	21,072 (24.8%)	6,750 (23.8%)
Stroke, n (%)	10,924 (12.9%)	3,055 (10.8%)
PAD, n (%)	8,875 (10.4%)	2,469 (8.7%)
COPD, n (%)	15,567 (18.3%)	5,662 (20.0%)
eGFR, median (IQR)	50 (42, 55)	54 (49, 57)
CKD stage, n (%) 3a 3b 4	 58,155 (68.4%) 21,353 (25.1%) 5,504 (6.5%)	 24,054 (84.9%) 3,727 (13.2%) 557 (2.0%)
Prior decline in eGFR slope (units per year)[a]	2.62 (1.09, 4.61)	2.33 (1.42, 3.46)
Recent AKI	1,134 (1.3%)	624 (2.2%)
Use of ACEi/ARBs	51,857 (61.0%)	18,062 (63.7%)
KRT, N (rate)	736 (2.17 per 1000 py)	114 (1.07 per 1000 py)
Deaths, N (rate)	31,536 (92.5 per 1000 py)	7,289 (68.5 per 1000 py)

[a]Includes all measures between 2006-2018 prior to (and including) the index date, excluding inpatient measures

Supplementary Table 5: Discrimination statistics for new risk models, by sup-population, evaluated in all patients (development and validation cohorts combined).

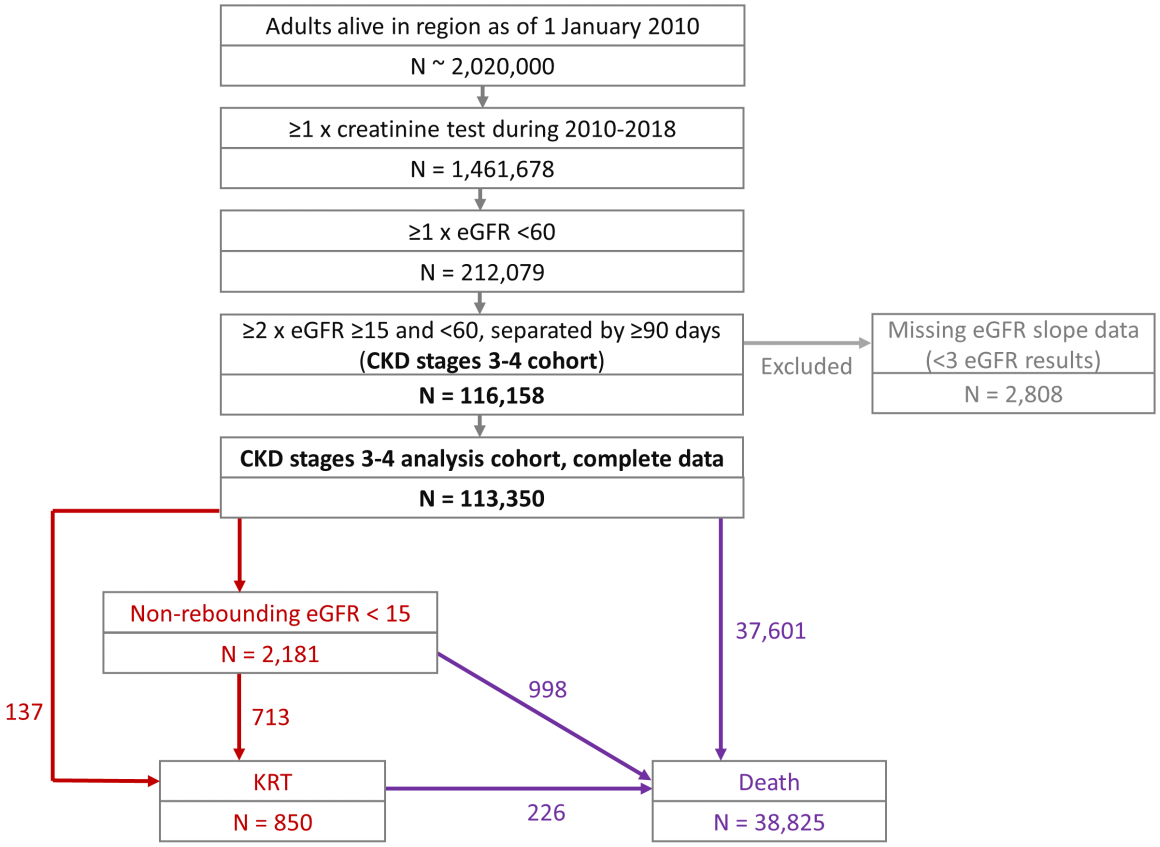
Patient population	C statistic (95% CI)[a]	
	KRT model	**eGFR outcome model**
Entire CKD cohort	0.942 (0.933, 0.951)	0.880 (0.871, 0.888)
CKD stage		
3a	0.855 (0.821, 0.889)	0.691 (0.664, 0.719)
3b	0.898 (0.879, 0.917)	0.737 (0.716, 0.757)
4	0.865 (0.850, 0.880)	0.762 (0.748, 0.776)
Diabetes		
Yes	0.938 (0.924, 0.952)	0.873 (0.861, 0.884)
No	0.930 (0.918, 0.942)	0.875 (0.863, 0.888)
Albuminuria data		
Yes	0.926 (0.915, 0.936)	0.893 (0.883, 0.902)
No	0.897 (0.872, 0.922)	0.836 (0.821, 0.851)

[a]Discrimination statistics by subgroup should be interpreted with caution. Lower C statistics are expected within subgroups due to less variation (less heterogeneity in predictor values) which makes it more difficult for the model to discriminate between outcomes. Calibration plots are more useful for subgroup comparisons of model performance.

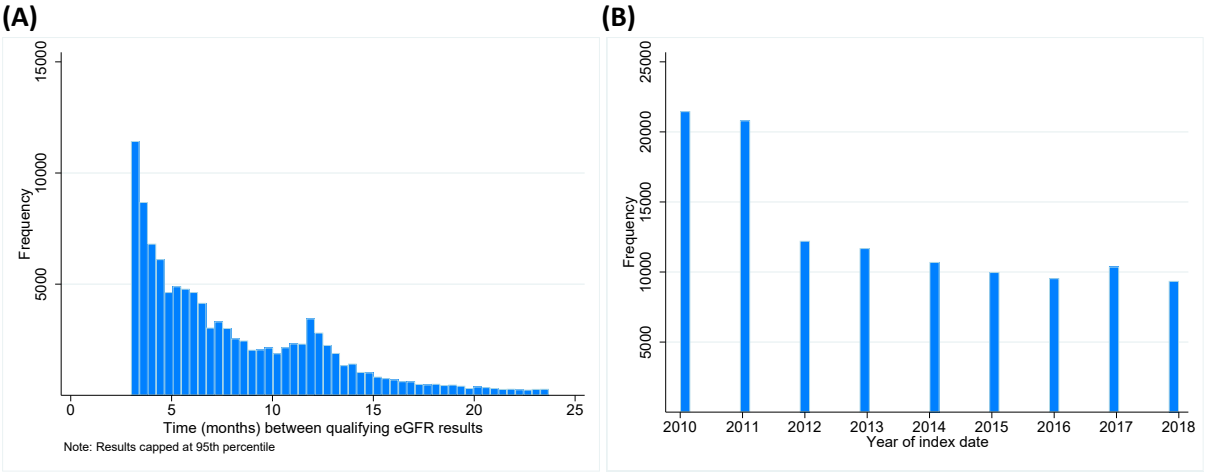
Supplementary Table 6: Patient characteristics by quintile of model predicted risk, separately for new equation and KFRE, in the subset of patients with uACR data within -12 months to +3 months of the index date

	Q1: New equation (P, 0 – 0.001) KFRE (P not shown; miscalibrated)		Q2: New equation (P, 0.001 – 0.003) KFRE (P not shown; miscalibrated)		Q3: New equation (P, 0.003 – 0.006) KFRE (P not shown; miscalibrated)		Q4: New equation (P, 0.006 – 0.018) KFRE (P not shown; miscalibrated)		Q5: New equation (P, 0.018 - 100) KFRE (P not shown; miscalibrated)	
New equation	Age, median (IQR)	81 (76,86)	Age, median (IQR)	76 (71, 81)	Age, median (IQR)	73 (68, 79)	Age, median (IQR)	71 (64, 77)	Age, median (IQR)	67 (57, 76)
	Female, N (%)	3,649 (69.6%)	Female, N (%)	2,717 (51.8%)	Female, N (%)	2,131 (40.6%)	Female, N (%)	1,912 (36.5%)	Female, N (%)	1,656 (31.6%)
	Diabetes, N (%)	1,400 (26.7%)	Diabetes, N (%)	2,581 (59.2%)	Diabetes, N (%)	3,164 (60.3%)	Diabetes, N (%)	3,285 (62.6%)	Diabetes, N (%)	2,993 (57.1%)
	Hyp, N (%)	4,111 (78.4%)	Hyp, N (%)	4,418 (84.2%)	Hyp, N (%)	4,560 (86.9%)	Hyp, N (%)	4,501 (86.8%)	Hyp, N (%)	4,452 (84.9%)
	CKD stage, N (%)		CKD stage, N (%)		CKD stage, N (%)		CKD stage, N (%)		CKD stage, N (%)	
	3a	5,158 (89.3%)	3a	4,861 (92.7%)	3a	4,343 (82.8%)	3a	3,012 (57.4%)	3a	713 (13.6%)
	3b	88 (1.7%)	3b	382 (7.3%)	3b	887 (16.9%)	3b	2,062 (39.3%)	3b	2,467 (47.0%)
	4	0	4	4 (0.1%)	4	16 (0.3%)	4	172 (3.3%)	4	2,065 (39.3%)
	uACR (mg/g)		uACR (mg/g)		uACR (mg/g)		uACR (mg/g)		uACR (mg/g)	
	median (IQR)	0.90 (0.50, 2.40)	median (IQR)	0.90 (0.45, 2.90)	median (IQR)	1.20 (0.50, 18.3)	median (IQR)	1.80 (0.60, 7.90)	median (IQR)	6.90 (1.38, 39.5)
	<30	5,130 (97.8%)	<30	5,029 (95.9%)	<30	4,887 (93.2%)	<30	4,649 (88.6%)	<30	3,754 (71.6%)
	30-299	108 (2.1%)	30-299	203 (3.9%)	30-299	342 (6.5%)	30-299	541 (10.3%)	30-299	1,317 (25.1%)
	≥300	8 (0.2%)	≥300	14 (0.3%)	≥300	17 (0.3%)	≥300	56 (1.1%)	≥300	174 (3.3%)
KFRE	Age, median (IQR)	76 (71, 81)	Age, median (IQR)	75 (69, 80)	Age, median (IQR)	75 (68, 81)	Age, median (IQR)	74 (66, 80)	Age, median (IQR)	70 (60, 79)
	Female, N (%)	3,323 (63.3%)	Female, N (%)	2,645 (50.4%)	Female, N (%)	2,244 (42.8%)	Female, N (%)	2,047 (39.0%)	Female, N (%)	1,806 (34.4%)
	Diabetes, N (%)	2,481 (47.3%)	Diabetes, N (%)	2,684 (51.2%)	Diabetes, N (%)	2,861 (54.5%)	Diabetes, N (%)	2,877 (54.8%)	Diabetes, N (%)	2,520 (48.1%)
	Hyp, N (%)	4,439 (84.6%)	Hyp, N (%)	4,418 (84.2%)	Hyp, N (%)	4,450 (84.8%)	Hyp, N (%)	4,452 (84.9%)	Hyp, N (%)	4,283 (81.7%)
	CKD stage, N (%)		CKD stage, N (%)		CKD stage, N (%)		CKD stage, N (%)		CKD stage, N (%)	
	3a	5,235 (99.8%)	3a	5,200 (99.1%)	3a	4,350 (82.9%)	3a	2,617 (49.9%)	3a	685 (13.1%)
	3b	11 (0.2%)	3b	46 (0.9%)	3b	895 (17.1%)	3b	2,567 (48.9%)	3b	2,367 (45.1%)
	4	0	4	0	4	1 (0.02%)	4	62 (1.2%)	4	2,193 (41.8%)
	uACR (mg/g)		uACR (mg/g)		uACR (mg/g)		uACR (mg/g)		uACR (mg/g)	
	median (IQR)	0.45 (0.30, 0.70)	median (IQR)	0.80 (0.48, 1.55)	median (IQR)	1.70 (0.70, 4.00)	median (IQR)	3.90 (1.30, 12.7)	median (IQR)	18.1 (4.60, 73.0)
	<30	5,246 (100%)	<30	5,246 (100%)	<30	5,215 (99.4%)	<30	4,631 (88.3%)	<30	3,111 (59.3%)
	30-299	0	30-299	0	30-299	31 (0.6%)	30-299	607 (11.6%)	30-299	1,873 (35.7%)
	≥300	0	≥300	0	≥300	0	≥300	8 (0.2%)	≥300	261 (5.0%)

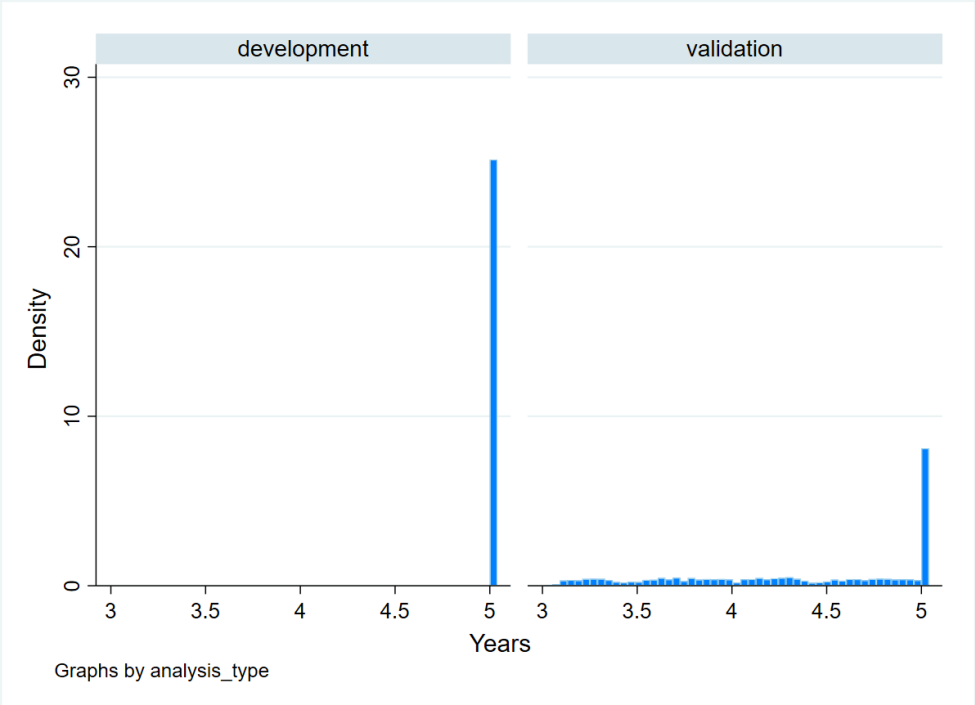
Supplementary Figure 1. Flow chart of analysis cohort identification and number of outcome events, with event numbers shown in the analysis population after removal of those with missing eGFR slope data. Outcome events and deaths include those recorded within the next 5 years post index date. Breakdown of deaths after KRT in those with complete eGFR slope data: 165 had eGFR < 15 outcome; 61 did not have eGFR < 15 outcome.



Supplementary Figure 2: Distribution of (A) time between qualifying eGFR results and (B) year of index date, in entire CKD cohort

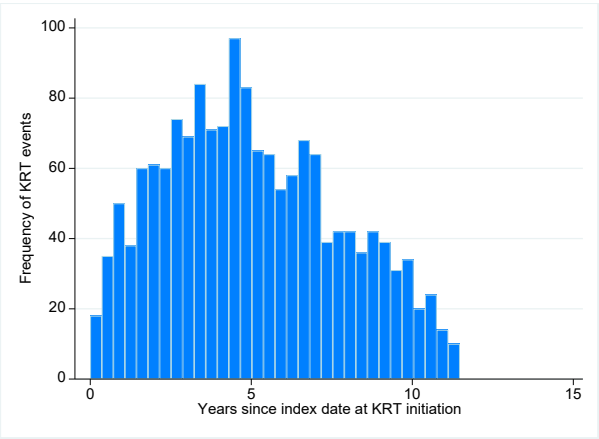


Supplementary Figure 3: Available follow-up time (years) between index date and end of outcomes data collection capped at 5 years in entire CKD cohort, by analysis cohort (development, validation)

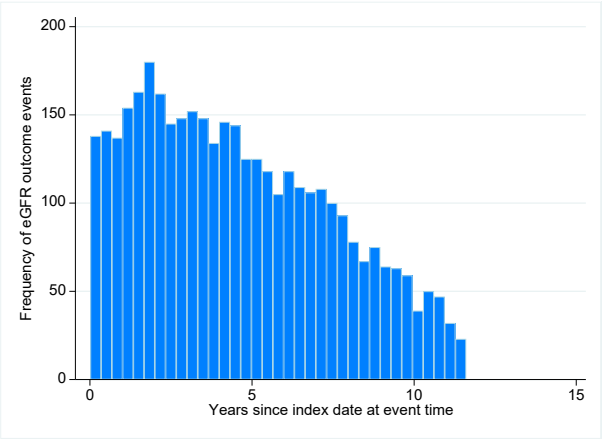


Supplementary Figure 4: Histograms of outcome events (and censoring variables) by year: (A) KRT; (B) non-rebounding eGFR<15; (C) Pre-KRT deaths. Analysis includes only those events up to 5 years post index date.

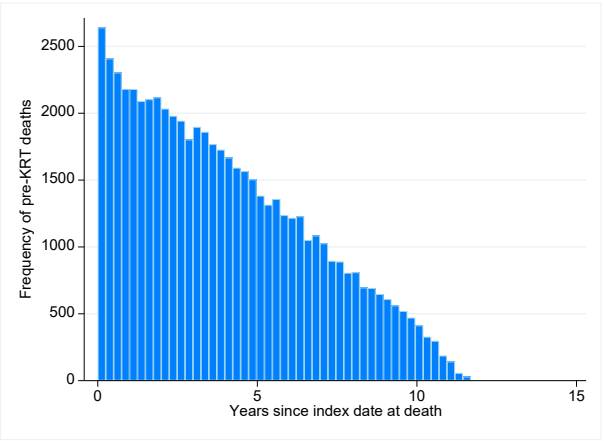
(A)



(B)

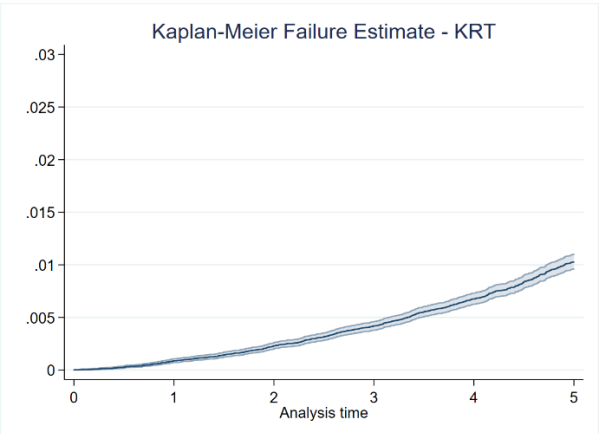


(C)

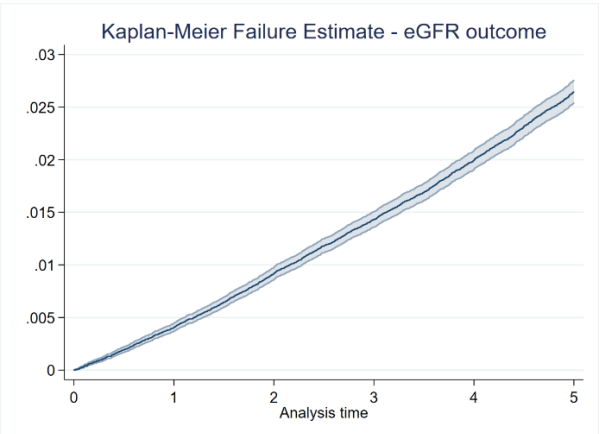


Supplementary Figure 5: Kaplan Meier failure curves for (A) main outcome (KRT) and (B) sensitivity outcome (non-rebounding eGFR < 15 or KRT), in entire CKD cohort

(A)

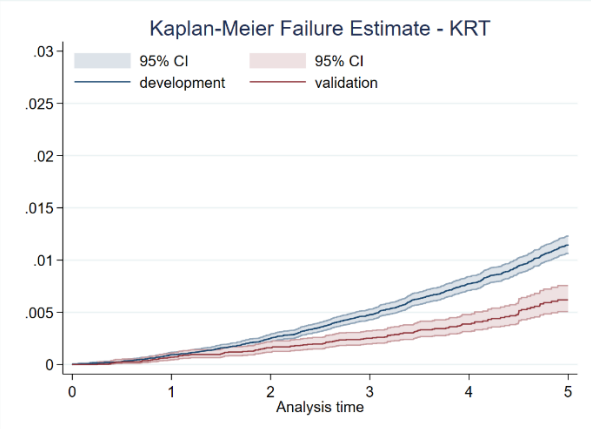


(B)

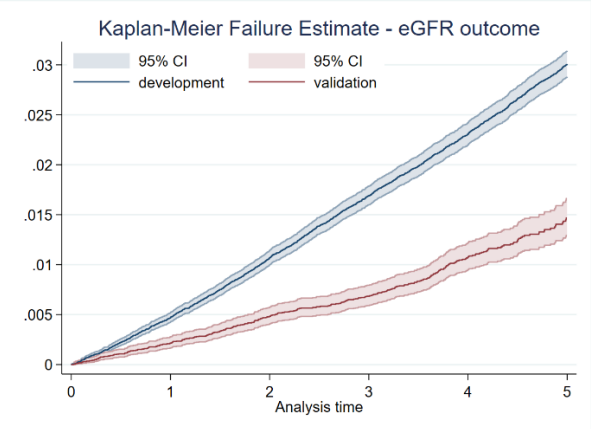


Supplementary Figure 6: Kaplan Meier failure curves for (A) main outcome (KRT) and (B) sensitivity outcome (non-rebounding eGFR < 15 or KRT), by analysis cohort (development, validation)

(A)

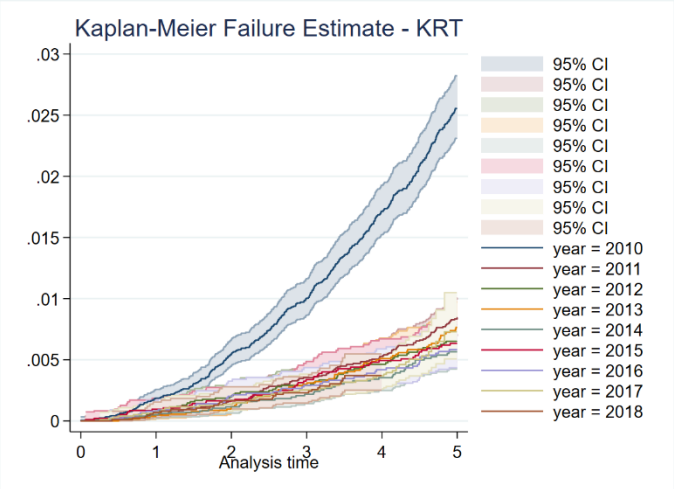


(B)

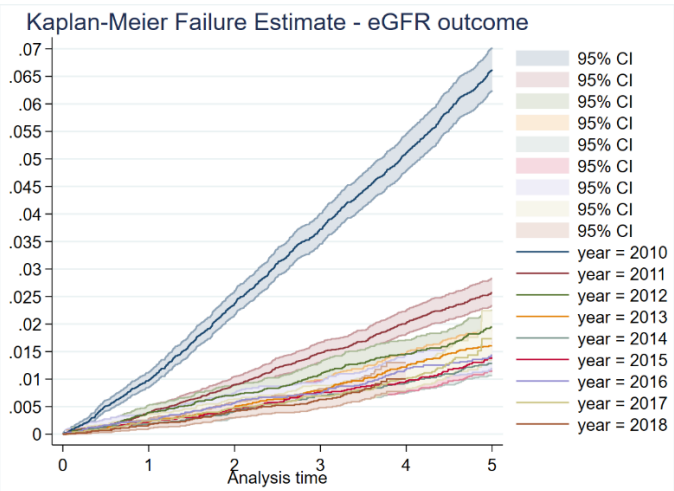


Supplementary Figure 7: Kaplan Meier failure curves for (A) main outcome (KRT) and (B) sensitivity outcome (non-rebounding eGFR < 15 or KRT), by year of index date

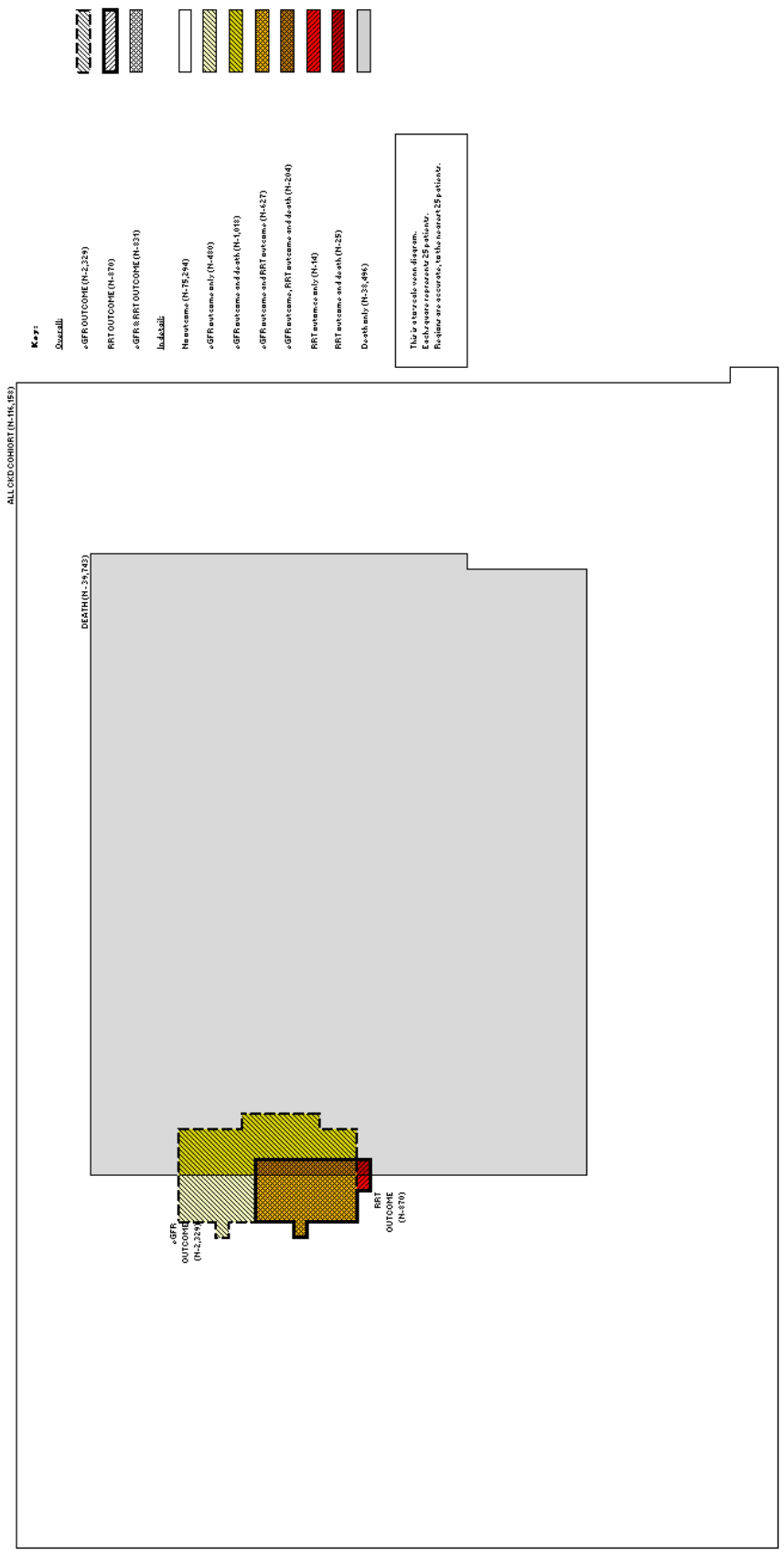
(A)



(B)

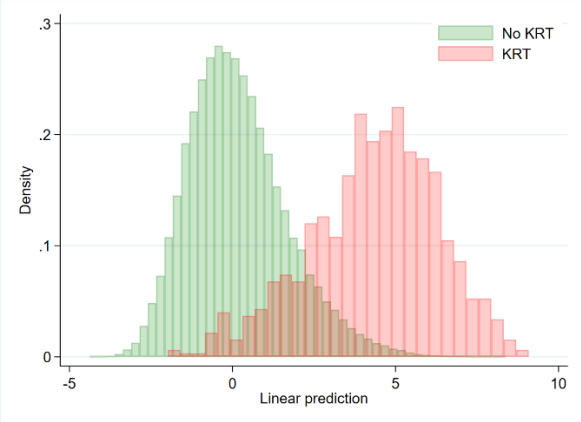


Supplementary Figure 8: To-scale venn diagram displaying overlap in outcomes experienced by the CKD cohort in a 5 year period

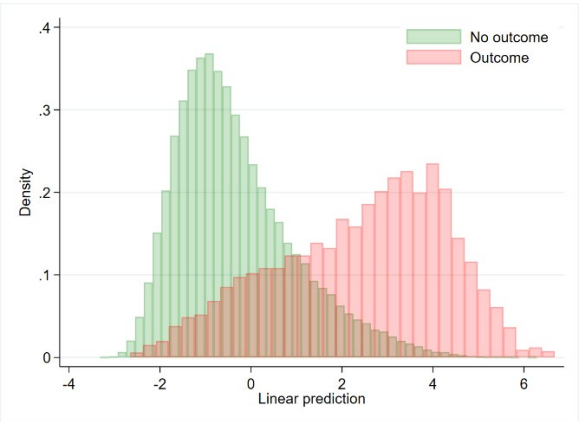


Supplementary Figure 9: Histograms (densities) of linear predictor by outcome status using **new risk equations** in all patients in the entire CKD cohort: (A) KRT, (B) eGFR outcome. Densities are presented for each group (outcome, no outcome) and are overlaid in one plot. Densities should not be confused with frequencies (outcome events appear over inflated compared to non-outcomes, due to rare events). Note that some patients without outcomes are censored before 5 years.

(A)

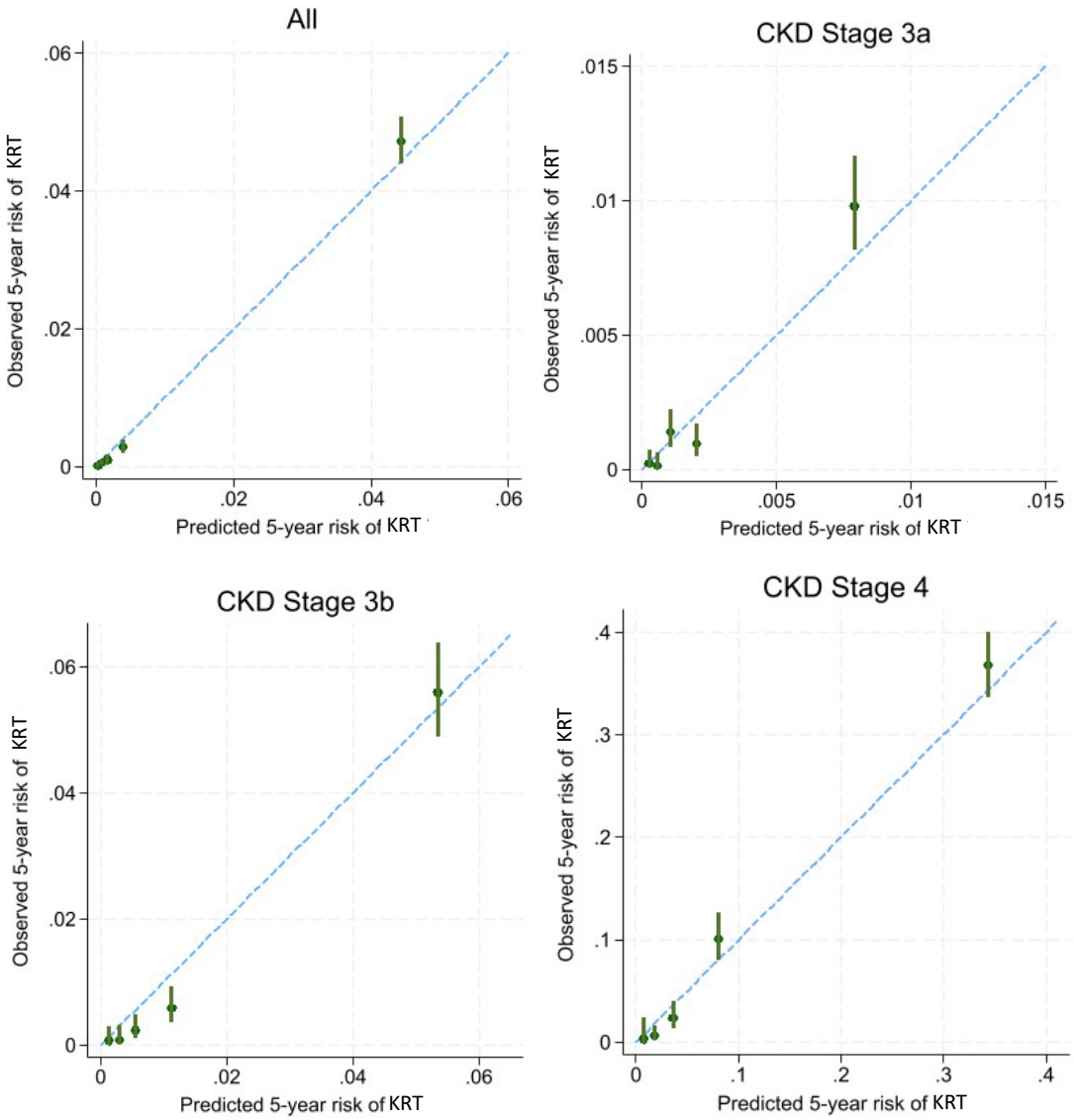


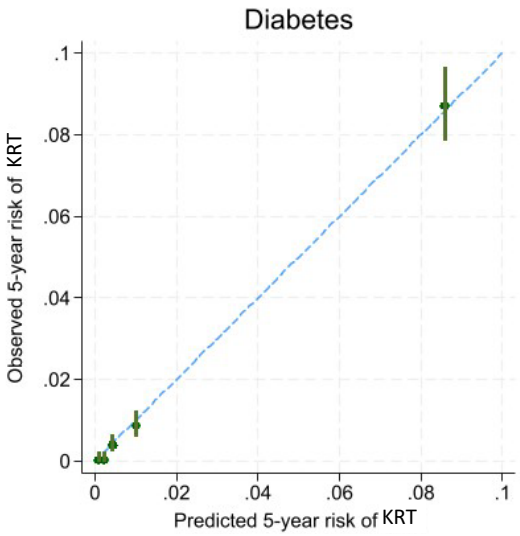
(B)

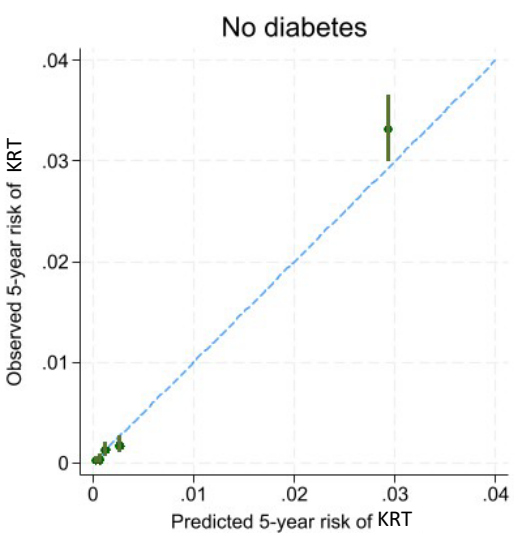


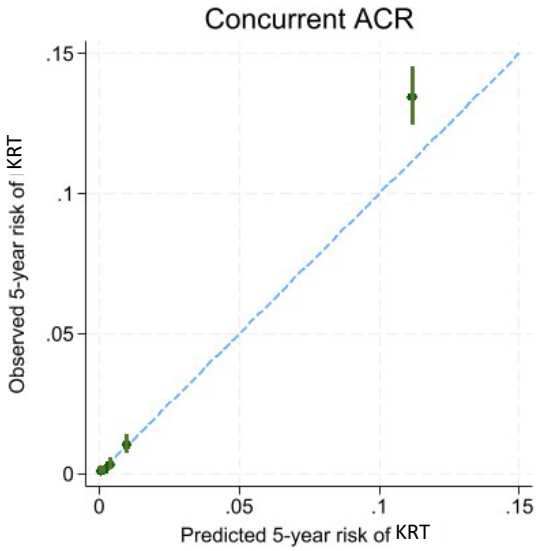
Supplementary Figure 10: Calibration graphics, by subgroup (in entire cohort).
Subgroup plots are shown in the entire cohort, rather than validation cohort, for sample size reasons, and the interest being in heterogeneity in model performance. The all patients plot is shown for completeness, in which case calibration is expected to be almost perfect (overall), with calibration slope close to 1.

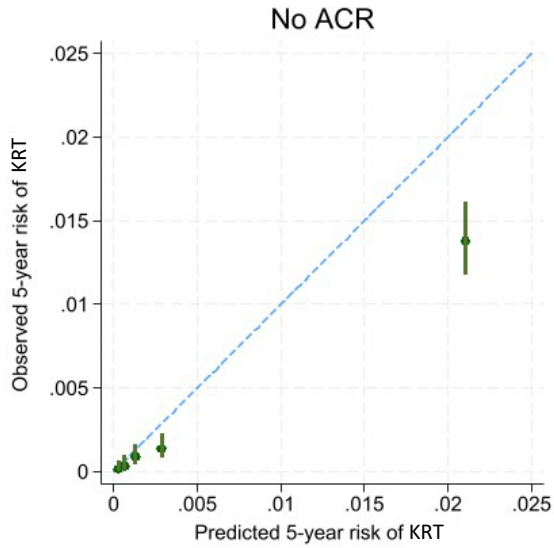
KRT:



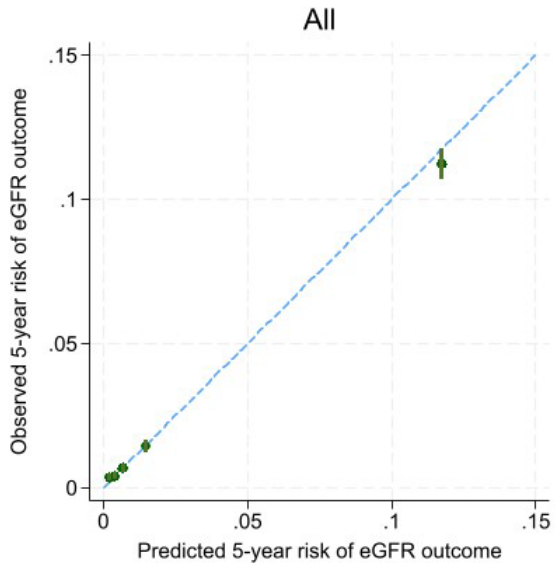


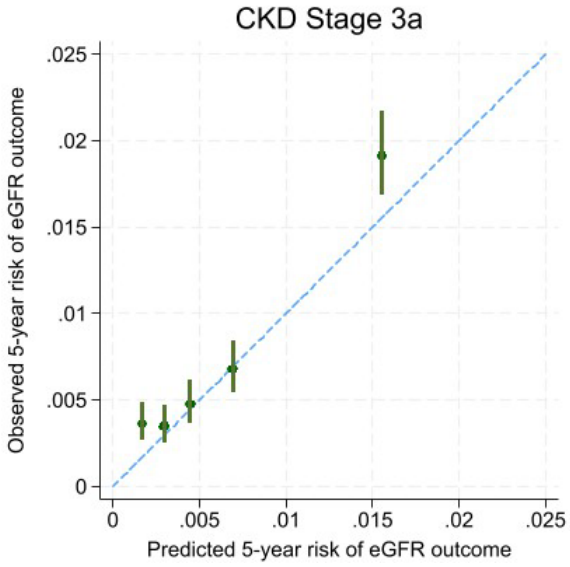


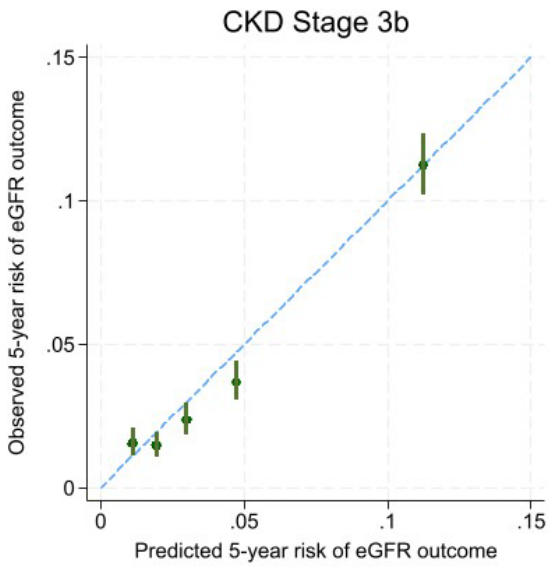


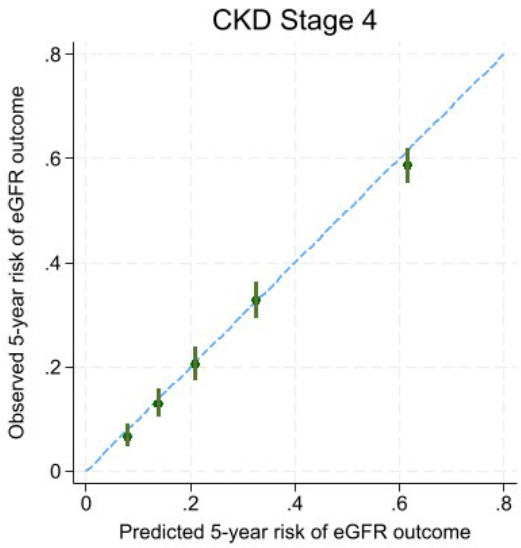


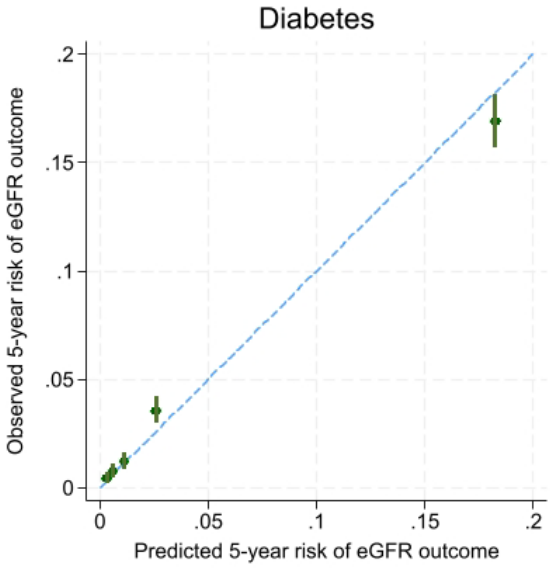
eGFR outcome:

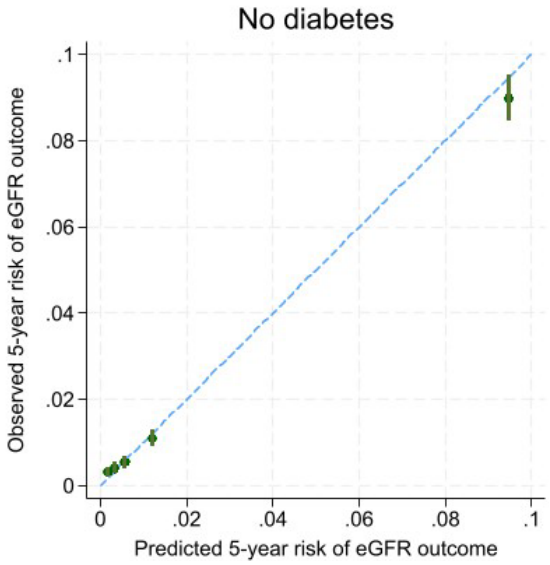


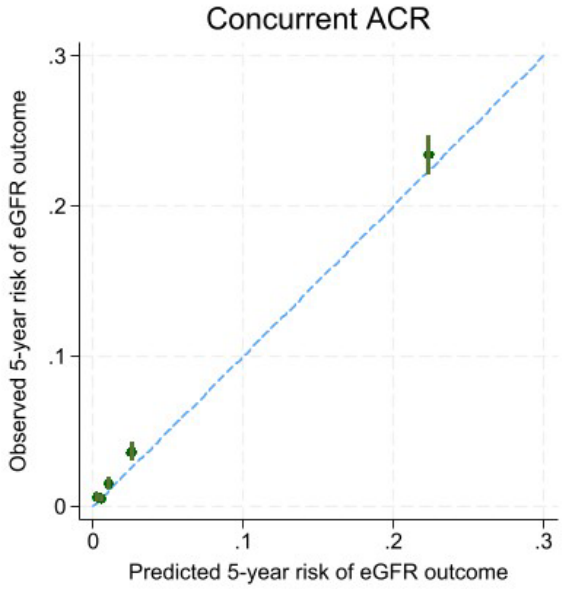


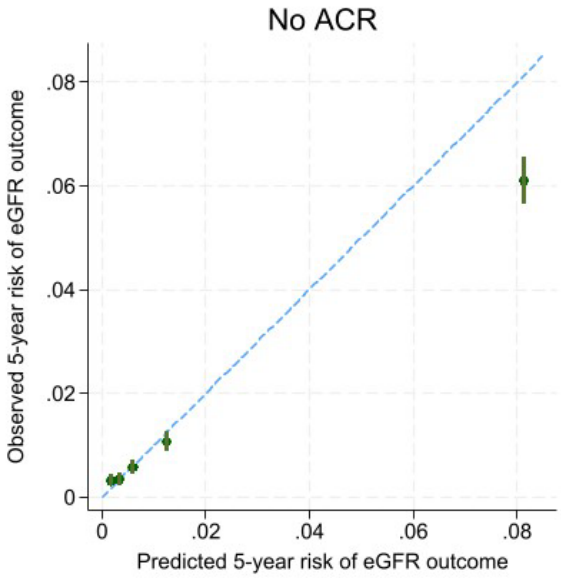

CKD Stage 3b
Observed 5-year risk of eGFR outcome
Predicted 5-year risk of eGFR outcome
0
.05
.1
.15


CKD Stage 4
Observed 5-year risk of eGFR outcome
Predicted 5-year risk of eGFR outcome
0
.2
.4
.6
.8

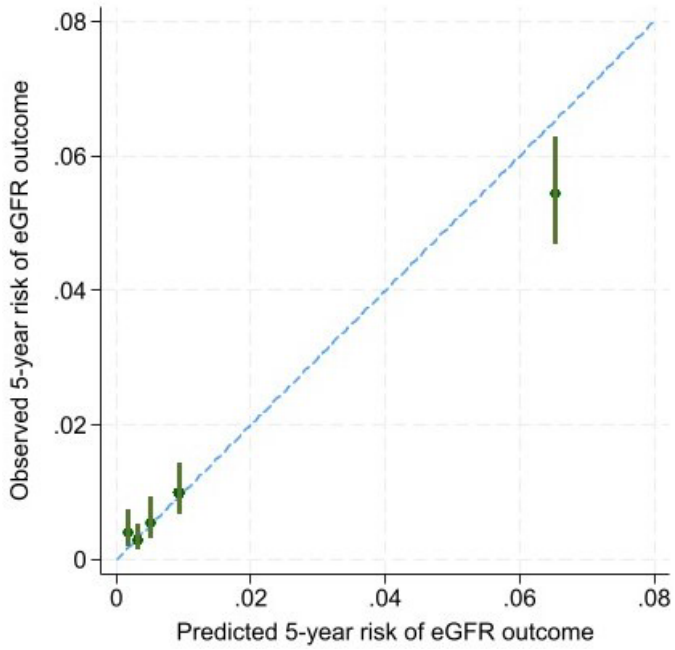

Diabetes
Observed 5-year risk of eGFR outcome
Predicted 5-year risk of eGFR outcome
0
.05
.1
.15
.2


No diabetes
Observed 5-year risk of eGFR outcome
Predicted 5-year risk of eGFR outcome
0
.02
.04
.06
.08
.1


Concurrent ACR
Observed 5-year risk of eGFR outcome
Predicted 5-year risk of eGFR outcome
0
.1
.2
.3

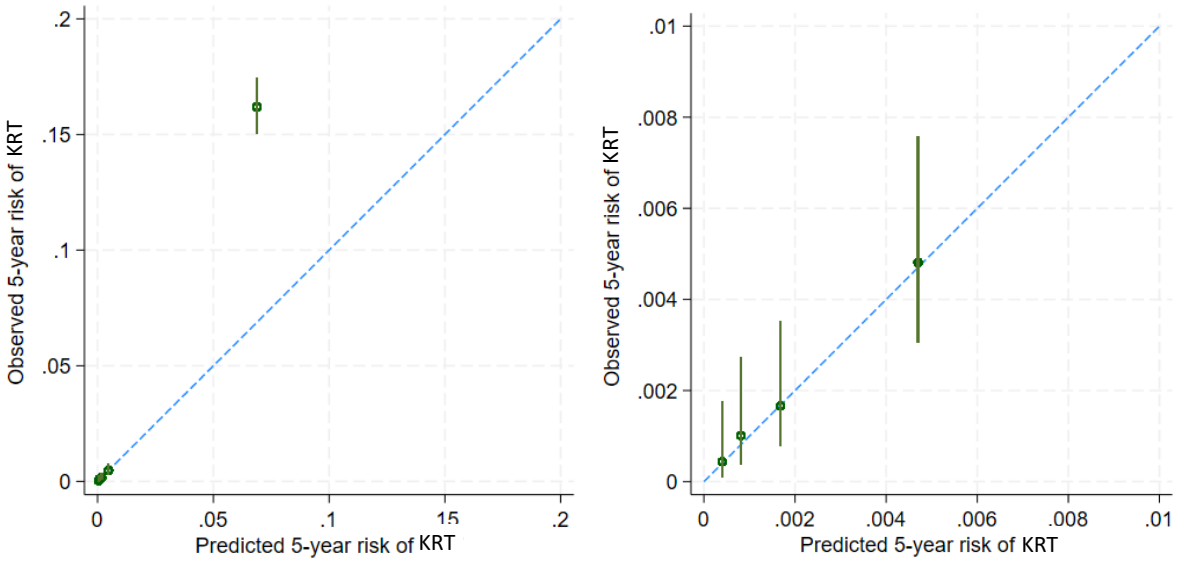

No ACR
Observed 5-year risk of eGFR outcome
Predicted 5-year risk of eGFR outcome
0
.02
.04
.06
.08

Supplementary Figure 11: Observed vs predicted probability of composite outcome non-rebounding eGFR<15 or KRT at 5 years, by quintile of predicted risk, in validation cohort. Quintiles of predicted risk: 0% - 0.246%; 0.246% - 0.398%; 0.398% - 0.654%; 0.654% - 1.357%; 1.357% - 100%.



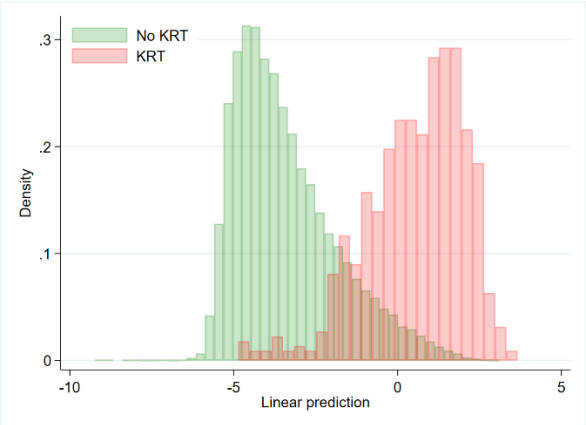
Supplementary Figure 12: Calibration plots for estimation of risks using the regional calibrated 4-variable KFRE in all patients with uACR data: (A) All patients, 5 quintiles; (B) Subset of first 4 quintiles. (Predicted risk quintiles: 0.01% - 0.07%, 0.07% - 0.14%, 0.14% - 0.30%, 0.30% - 1.07%, 1.07% - 95.3%)

(A) **(B)**

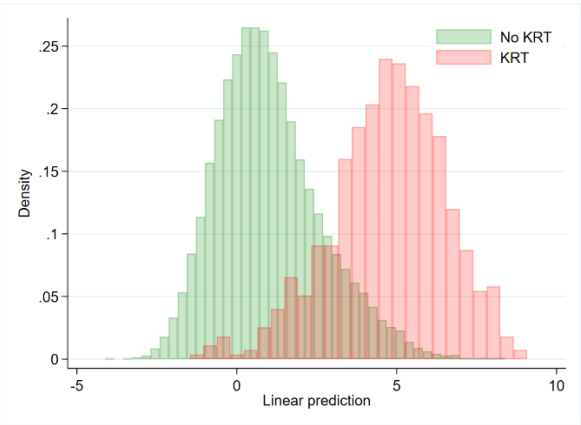


Supplementary Figure 13: Overlaid histograms (densities, not frequencies) of linear predictor by KRT outcome status in all patients in the entire cohort with an uACR result available: (A) 4-variable KFRE; (B) New equation

(A)

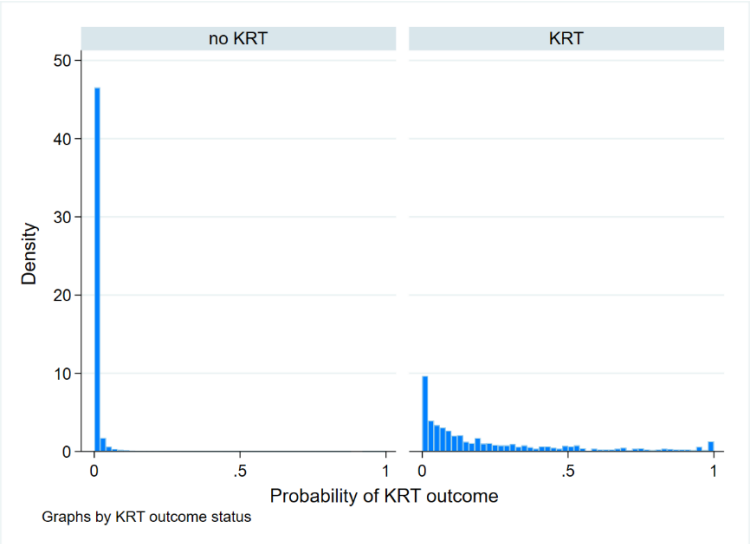


(B)

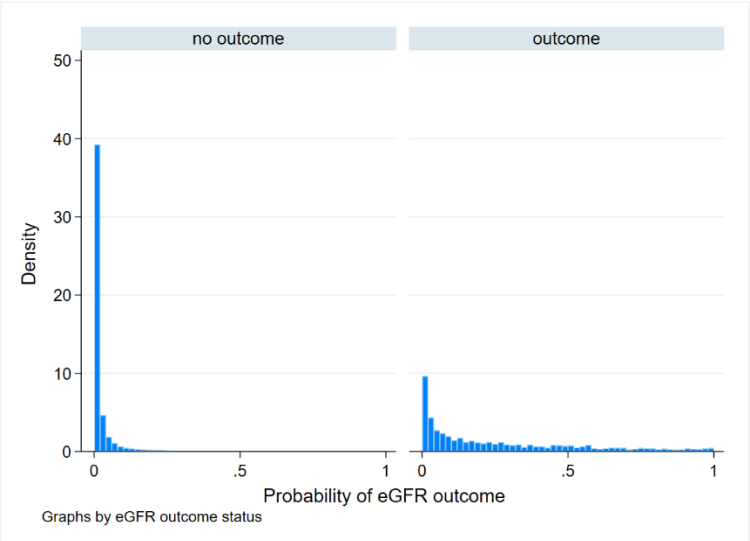


Supplementary Figure 14: Histograms of predicted risks by outcome status, in entire CKD cohort: (A) KRT; (B) non-rebounding eGFR<15 or KRT

(A)

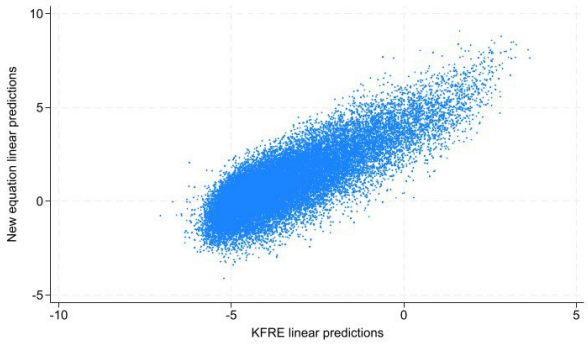


(B)

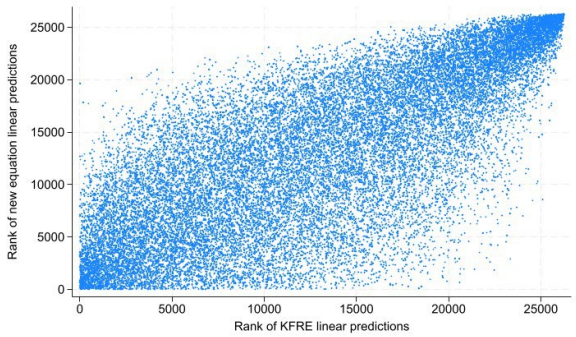


Supplementary Figure 15: Scatter plots comparing (A) linear predictions and (B) ranks of linear predictions, by equation, in patients with uACR data

(A)

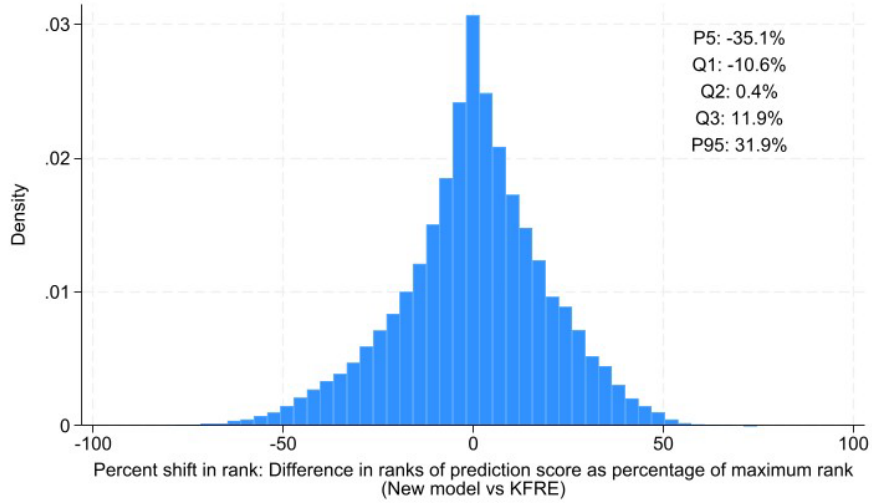


(B)

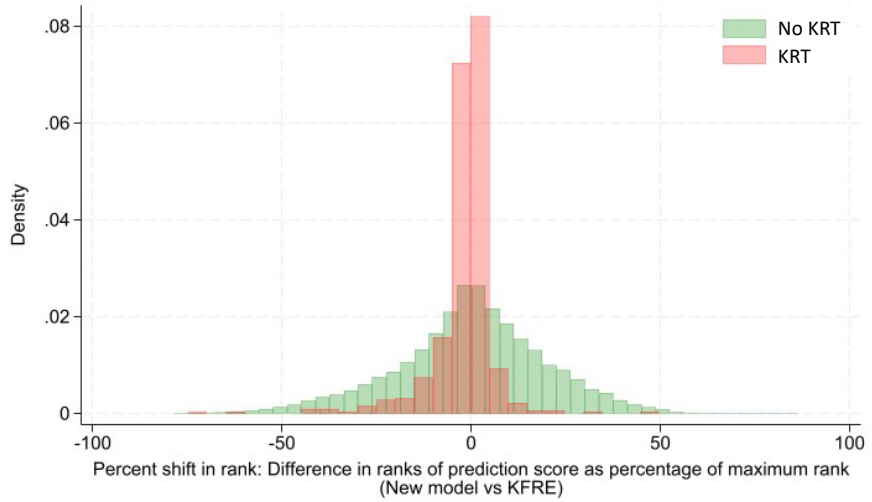


Supplementary Figure 16: Percent shift in rank of linear prediction score, comparing new equation with 4-variable KFRE, in all patients in CKD cohort with an uACR result: (A) All patients; (B) Stratified by KRT outcome status

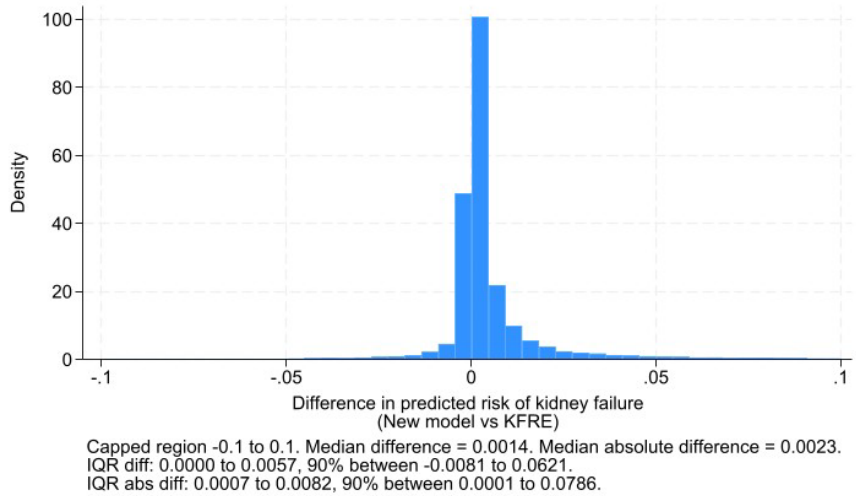
(A)



(B)



Supplementary Figure 17: Histogram of difference in predicted risks of KRT comparing new equation with KFRE



Capped region -0.1 to 0.1. Median difference = 0.0014. Median absolute difference = 0.0023.
IQR diff: 0.0000 to 0.0057, 90% between -0.0081 to 0.0621.
IQR abs diff: 0.0007 to 0.0082, 90% between 0.0001 to 0.0786.

Supplementary Information 1: Details of new prediction equations

The following details show how we arrive at an equation for risk of kidney failure within 5 years of patient identification using results of the final Cox regression model.

Probability of survival at time t years, $\hat{S}_i(t)$, is estimated using the probability of survival at time t years in the baseline group, $\hat{S}_0(t)$, and the value of the linear predictor for an individual patient, $\boldsymbol{x}_i\widehat{\boldsymbol{\beta}}$, as follows:

$$\hat{S}_i(t) = \hat{S}_0(t)^{\exp(\boldsymbol{x}_i\widehat{\boldsymbol{\beta}})}$$

The baseline survival function, $\hat{S}_0(t)$, is estimated from the analysis dataset. We estimate the result at 5 years for KRT outcome as $\hat{S}_0(t = 5\ years)$ = 0.99839.

Coefficient estimates $\widehat{\boldsymbol{\beta}}$ are the log(hazard ratio) estimates for each covariate from the regression model (presented as hazard ratio estimates in Table 2).

The final prediction equation is as follows:

P(RRT within 5 years of identification) = 1 – 0.99839 x exp{
(-0.64373) x ((baseline eGFR – 47.851)/5)
+ (-0.71971) x ((age – 78.287)/10)
+ (0.76776) x (1 if diabetes, 0 otherwise)
+ (-0.61870) x (1 if female, 0 if male)
+ (-0.57871) x (1 if AF, 0 otherwise)
+ (0.48235) x (1 if ACEiARB, 0 otherwise)
+ (0.40786) x (1 if PAD, 0 otherwise)
+ (0.034136) x ((eGFR decline per year – 3.3343)/5)
+ (-0.56595) x (1 if AKI, 0 otherwise)
+ (0.23353) x (1 if hypertension, 0 otherwise)
}

For precise details of variable definitions, see methods section.

Similarly, the final prediction equation for eGFR outcome is:

P(non-rebounding eGFR<15 or RRT within 5 years of identification) = 1 – 0.98717 x exp{
(-0.70161) x ((baseline eGFR – 47.851)/5)
+ (-0.36603) x ((age – 78.287)/10)
+ (-0.58579) x (1 if female, 0 if male)
+ (0.37122) x (1 if diabetes, 0 otherwise)
+ (0.22808) x (1 if PAD, 0 otherwise)
+ (0.12733) x (1 if HF, 0 otherwise)
+ (-0.17506) x (1 if CHD, 0 otherwise)
+ (-0.24789) x (1 if AF, 0 otherwise)
+ (-0.33448) x (1 if AKI, 0 otherwise)
}

Supplementary Information 2. Equation for 4-variable KFRE

The regional calibrated 4-variable KFRE (non-North America) for predicting 5 year risk was used in model comparisons, as follows:

Predicted risk = 1 – 0.9365 ^ exp (-0.2201 × (age/10 – 7.036) + 0.2467 × (male – 0.5642) – 0.5567 × (eGFR/5 – 7.222) + 0.4510 × (log(uACR) – 5.137))

Section/Topic	Item	Development / evaluation[1]	Checklist item	Reported on page
TITLE				
Title	1	D;E	Identify the study as developing or evaluating the performance of a multivariable prediction model, the target population, and the outcome to be predicted	1
ABSTRACT				
Abstract	2	D;E	See TRIPOD+AI for Abstracts checklist	2
INTRODUCTION				
Background	3a	D;E	Explain the healthcare context (including whether diagnostic or prognostic) and rationale for developing or evaluating the prediction model, including references to existing models	4-5
	3b	D;E	Describe the target population and the intended purpose of the prediction model in the context of the care pathway, including its intended users (e.g., healthcare professionals, patients, public)	5
	3c	D;E	Describe any known health inequalities between sociodemographic groups	4
Objectives	4	D;E	Specify the study objectives, including whether the study describes the development or validation of a prediction model (or both)	5
METHODS				
Data	5a	D;E	Describe the sources of data separately for the development and evaluation datasets (e.g., randomised trial, cohort, routine care or registry data), the rationale for using these data, and representativeness of the data	6
	5b	D;E	Specify the dates of the collected participant data, including start and end of participant accrual; and, if applicable, end of follow-up	6
Participants	6a	D;E	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including the number and location of centres	6
	6b	D;E	Describe the eligibility criteria for study participants	6
	6c	D;E	Give details of any treatments received, and how they were handled during model development or evaluation, if relevant	7
Data preparation	7	D;E	Describe any data pre-processing and quality checking, including whether this was similar across relevant sociodemographic groups	8
Outcome	8a	D;E	Clearly define the outcome that is being predicted and the time horizon, including how and when assessed, the rationale for choosing this outcome, and whether the method of outcome assessment is consistent across sociodemographic groups	6, 8
	8b	D;E	If outcome assessment requires subjective interpretation, describe the qualifications and demographic characteristics of the outcome assessors	N/A
	8c	D;E	Report any actions to blind assessment of the outcome to be predicted	N/A
Predictors	9a	D	Describe the choice of initial predictors (e.g., literature, previous models, all available predictors) and any pre-selection of predictors before model building	7
	9b	D;E	Clearly define all predictors, including how and when they were measured (and any actions to blind assessment of predictors for the outcome and other predictors)	7
	9c	D;E	If predictor measurement requires subjective interpretation, describe the qualifications and demographic characteristics of the predictor assessors	N/A
Sample size	10	D;E	Explain how the study size was arrived at (separately for development and evaluation), and justify that the study size was sufficient to answer the research question. Include details of any sample size calculation	7
Missing data	11	D;E	Describe how missing data were handled. Provide reasons for omitting any data	7
Analytical methods	12a	D	Describe how the data were used (e.g., for development and evaluation of model performance) in the analysis, including whether the data were partitioned, considering any sample size requirements	8
	12b	D	Depending on the type of model, describe how predictors were handled in the analyses (functional form, rescaling, transformation, or any standardisation).	7
	12c	D	Specify the type of model, rationale[2], all model-building steps, including any hyperparameter tuning, and method for internal validation	8
	12d	D;E	Describe if and how any heterogeneity in estimates of model parameter values and model performance was handled and quantified across clusters (e.g., hospitals, countries). See TRIPOD-Cluster for additional considerations[3]	N/A
	12e	D;E	Specify all measures and plots used (and their rationale) to evaluate model performance (e.g., discrimination, calibration, clinical utility) and, if relevant, to compare multiple models	8-9
	12f	E	Describe any model updating (e.g., recalibration) arising from the model evaluation, either overall or for particular sociodemographic groups or settings	N/A
	12g	E	For model evaluation, describe how the model predictions were calculated (e.g., formula, code, object, application programming interface)	9
Class imbalance	13	D;E	If class imbalance methods were used, state why and how this was done, and any subsequent methods to recalibrate the model or the model predictions	N/A
Fairness	14	D;E	Describe any approaches that were used to address model fairness and their rationale	9
Model output	15	D	Specify the output of the prediction model (e.g., probabilities, classification). Provide details and rationale for any classification and how the thresholds were identified	8

[1] D=items relevant only to the development of a prediction model; E=items relating solely to the evaluation of a prediction model; D;E=items applicable to both the development and evaluation of a prediction model
[2] Separately for all model building approaches.
[3] TRIPOD-Cluster is a checklist of reporting recommendations for studies developing or validating models that explicitly account for clustering or explore heterogeneity in model performance (eg, at different hospitals or centres). Debray et al, BMJ 2023; 380: e071018 [DOI: 10.1136/bmj-2022-071018]

Training versus evaluation	16	D;E	Identify any differences between the development and evaluation data in healthcare setting, eligibility criteria, outcome, and predictors	6, 8, 14
Ethical approval	17	D;E	Name the institutional research board or ethics committee that approved the study and describe the participant-informed consent or the ethics committee waiver of informed consent	25-26
OPEN SCIENCE				
Funding	18a	D;E	Give the source of funding and the role of the funders for the present study	25
Conflicts of interest	18b	D;E	Declare any conflicts of interest and financial disclosures for all authors	25
Protocol	18c	D;E	Indicate where the study protocol can be accessed or state that a protocol was not prepared	10
Registration	18d	D;E	Provide registration information for the study, including register name and registration number, or state that the study was not registered	10
Data sharing	18e	D;E	Provide details of the availability of the study data	26
Code sharing	18f	D;E	Provide details of the availability of the analytical code[4]	N/A
PATIENT & PUBLIC INVOLVEMENT				
Patient & Public Involvement	19	D;E	Provide details of any patient and public involvement during the design, conduct, reporting, interpretation, or dissemination of the study or state no involvement.	10
RESULTS				
Participants	20a	D;E	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	11
	20b	D;E	Report the characteristics overall and, where applicable, for each data source or setting, including the key dates, key predictors (including demographics), treatments received, sample size, number of outcome events, follow-up time, and amount of missing data. A table may be helpful. Report any differences across key demographic groups.	Table 1, Supp. Tables 1 & 4
	20c	E	For model evaluation, show a comparison with the development data of the distribution of important predictors (demographics, predictors, and outcome).	Supp. Table 4
Model development	21	D;E	Specify the number of participants and outcome events in each analysis (e.g., for model development, hyperparameter tuning, model evaluation)	14
Model specification	22	D	Provide details of the full prediction model (e.g., formula, code, object, application programming interface) to allow predictions in new individuals and to enable third-party evaluation and implementation, including any restrictions to access or re-use (e.g., freely available, proprietary)[5]	Supp. Information 1
Model performance	23a	D;E	Report model performance estimates with confidence intervals, including for any key subgroups (e.g., sociodemographic). Consider plots to aid presentation.	Table 2, Figure 2, Supp. Table 5
	23b	D;E	If examined, report results of any heterogeneity in model performance across clusters. See TRIPOD Cluster for additional details[3].	N/A
Model updating	24	E	Report the results from any model updating, including the updated model and subsequent performance	N/A
DISCUSSION				
Interpretation	25	D;E	Give an overall interpretation of the main results, including issues of fairness in the context of the objectives and previous studies	17
Limitations	26	D;E	Discuss any limitations of the study (such as a non-representative sample, sample size, overfitting, missing data) and their effects on any biases, statistical uncertainty, and generalizability	18
Usability of the model in the context of current care	27a	D	Describe how poor quality or unavailable input data (e.g., predictor values) should be assessed and handled when implementing the prediction model	18-19
	27b	D	Specify whether users will be required to interact in the handling of the input data or use of the model, and what level of expertise is required of users	20
	27c	D;E	Discuss any next steps for future research, with a specific view to applicability and generalizability of the model	21

From: Collins GS, Moons KGM, Dhiman P, et al. *BMJ* 2024;385:e078378. doi:10.1136/bmj-2023-078378

[4] This relates to the analysis code, for example, any data cleaning, feature engineering, model building, evaluation.
[5] This relates to the code to implement the model to get estimates of risk for a new individual.