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

Check for updates

© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY. Published by BMJ.

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

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

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

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

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

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

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

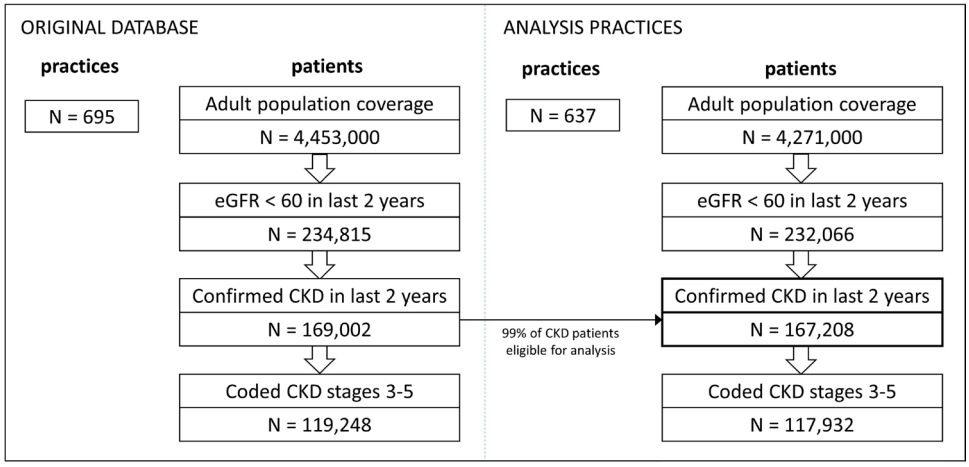


Figure 1 Flow chart of selection of study population (confirmed CKD in last 2 years). CKD, chronic kidney disease.

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

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.

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

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]*	18 236 (9,574, 23,480)	17 887 (11,835, 22,921)	17 363 (9,987, 24,460)	17 114 (10,392, 21,711)	15 793 (10,946, 21,345)	16 555 (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.

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

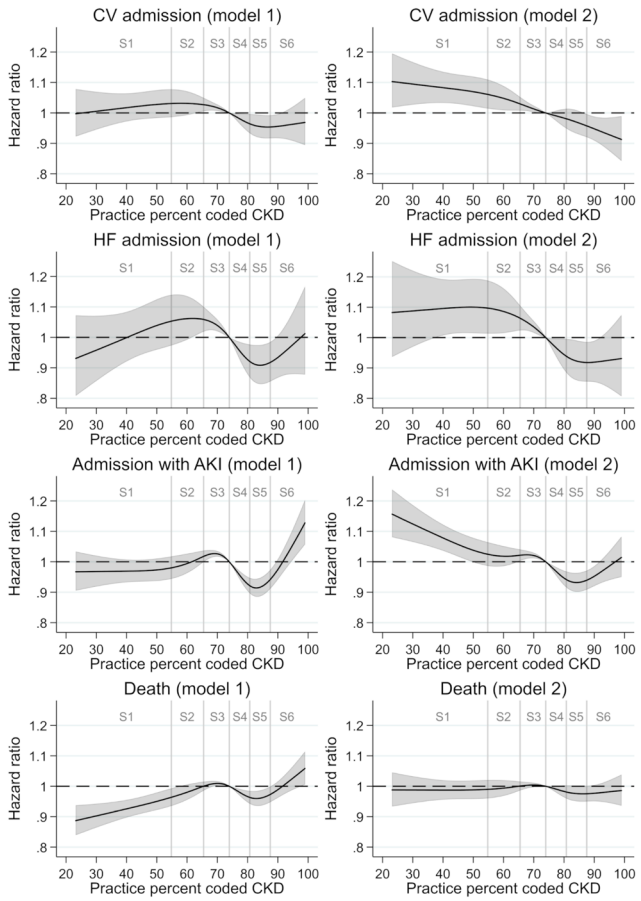


Figure 2 HR curves for all outcomes according to practice percent coded CKD. Analysis includes all patients with confirmed CKD. Primary exposure is (continuous) practice percent coded CKD, with median practice coding (74% coded) as the reference group. Model 1 adjusts for practice risk profile characteristics: 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).

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

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

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

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.

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

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.

BMJ Open: first published as 10.1136/bmjopen-2022-064513 on 11 October 2022. Downloaded from http://bmjopen.bmj.com/ on October 26, 2022 at UCL Library Services. Protected by copyright.

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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%)

Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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:

$$\%\ events\ preventable\ in\ practices\ coding\ 54.8\%\ of\ CKD\ cases$$
$$= \frac{expected\ events\ assuming\ 54.8\%\ coding - expected\ events\ assuming\ 87.5\%\ coding}{expected\ events\ assuming\ 54.8\%\ coding} \times 100\%$$

Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

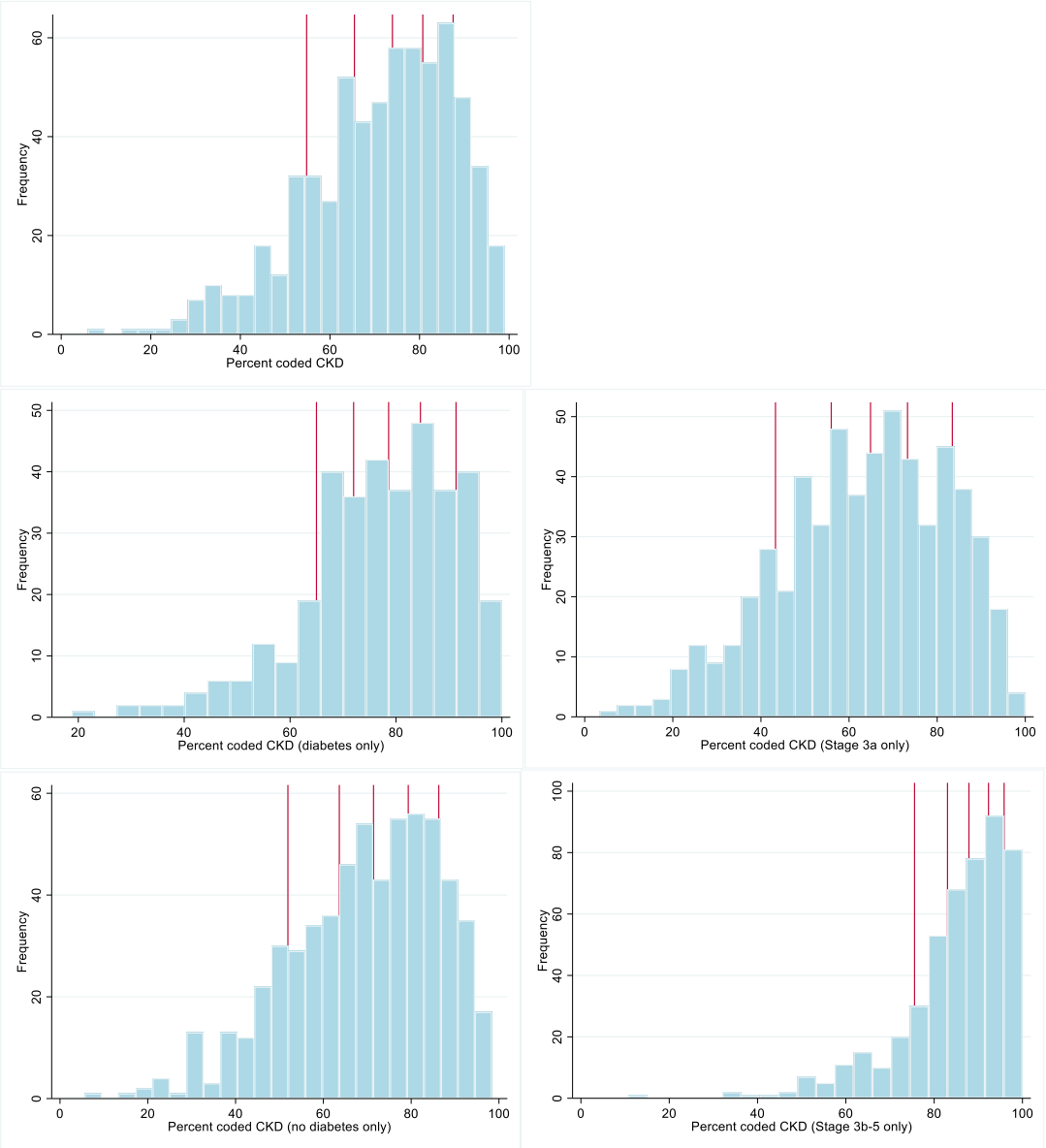
BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

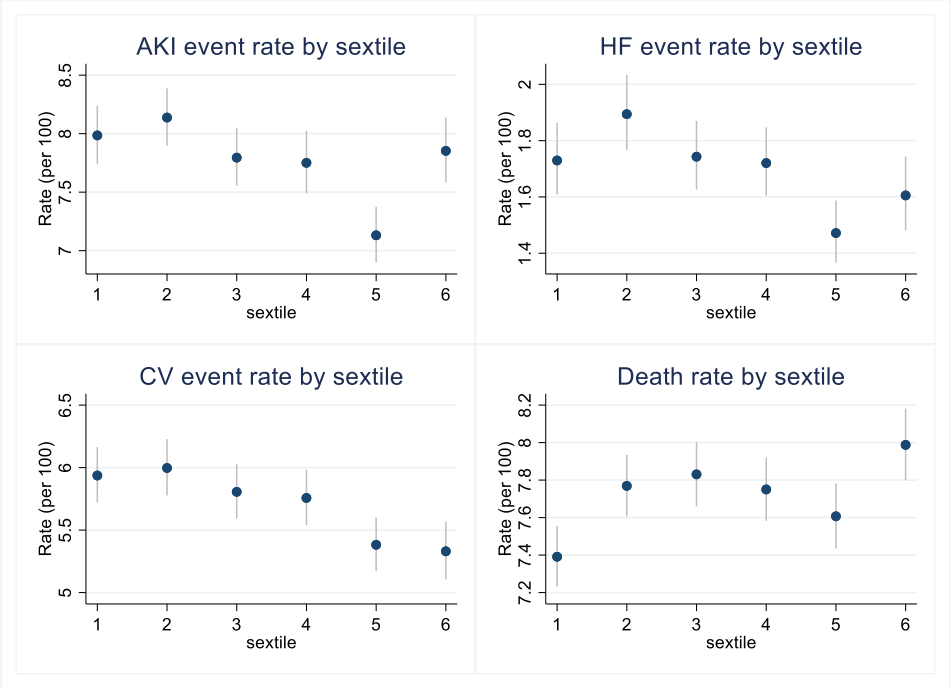
Supplementary Figure 1. Distribution of practice CKD coding performance in the CKD analysis population and in subgroups, with red lines depicting sextile boundaries



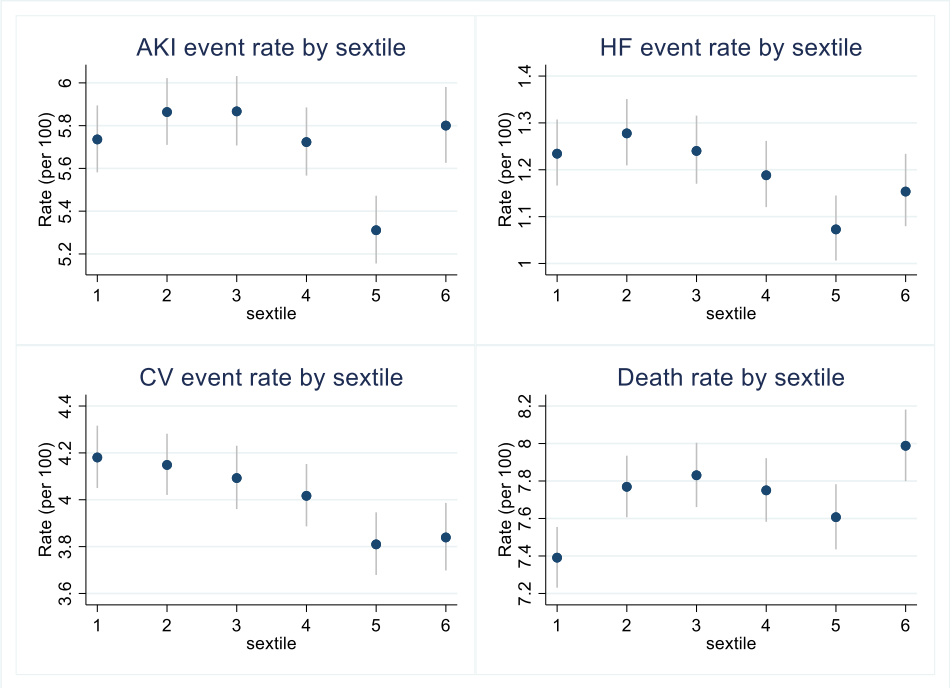
Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)



Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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

Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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:

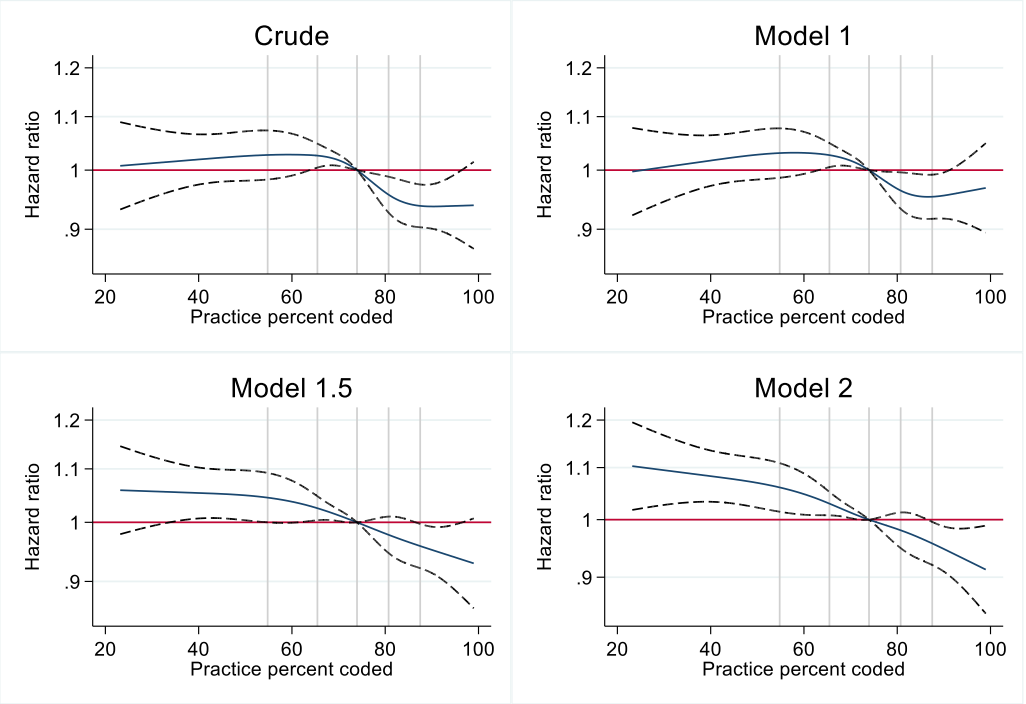
- Crude analysis: unadjusted
- Model 1: adjusted for ***practice characteristics reflecting practice risk profile*** (mean age, percent male, median rank of IMD, diabetes prevalence, hypertension prevalence, CVD prevalence)
- Model 1.5: adjusted for ***practice characteristics reflecting practice risk 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.

Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

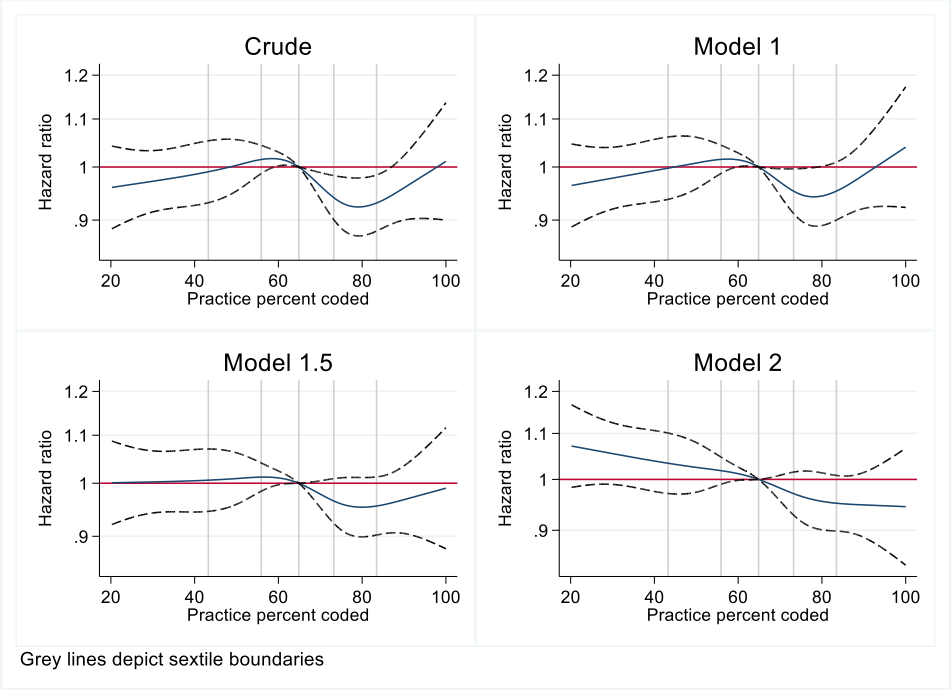
Supplementary Figure 4. Hazard ratio splines for time to first **CV event**, compared to median (73.9%) practice percent coded (sequential adjustments)



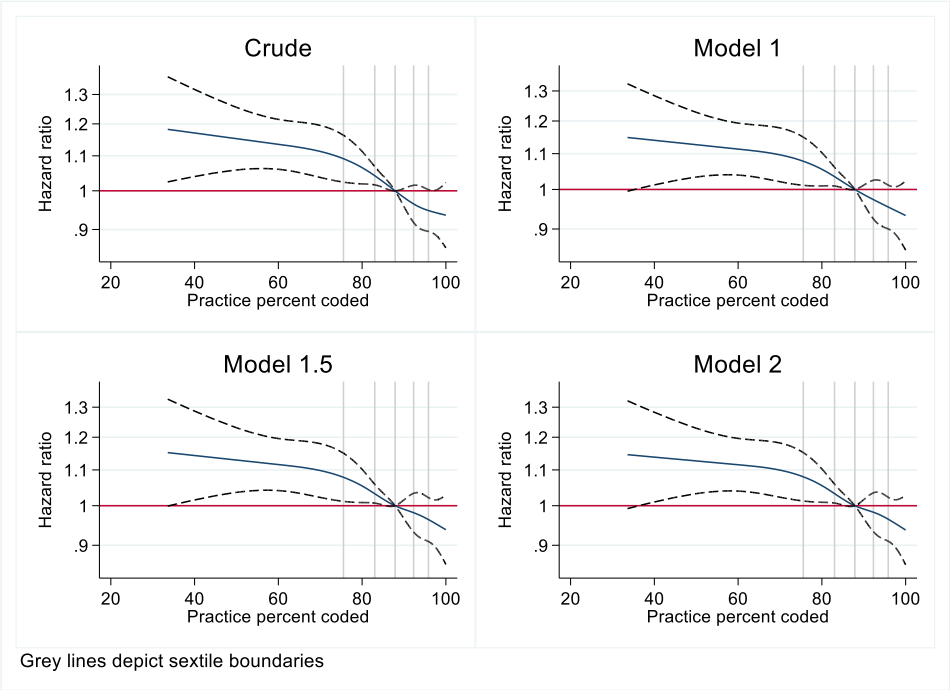
Grey lines depict sextile boundaries

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)

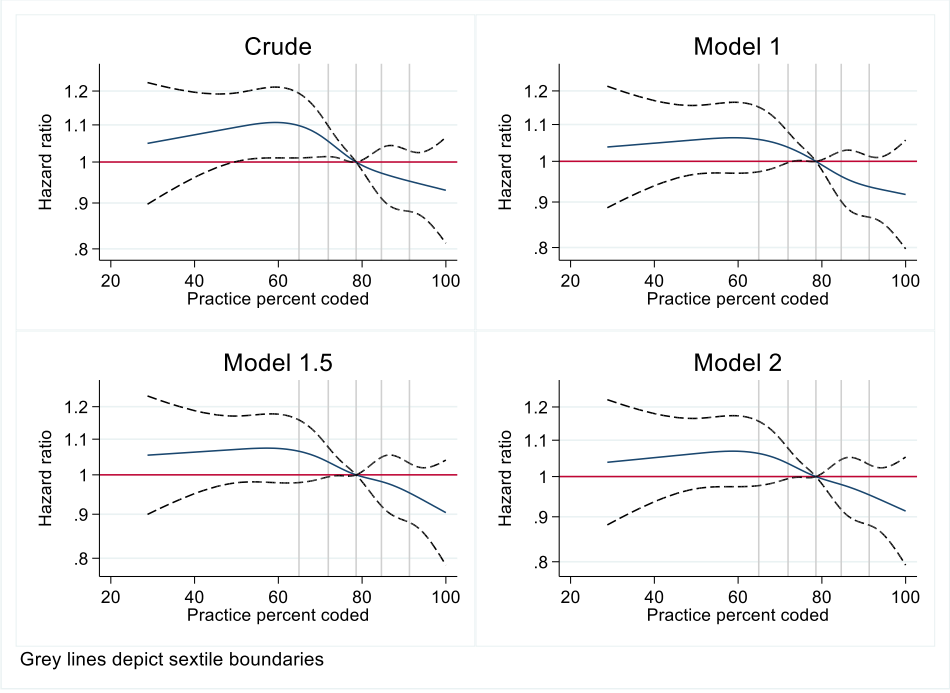


Supplementary Figure 6. Hazard ratio splines for time to first **CV event**, compared to median (87.9%) practice percent coded, in **CKD stages 3b-5 only** (sequential adjustments)

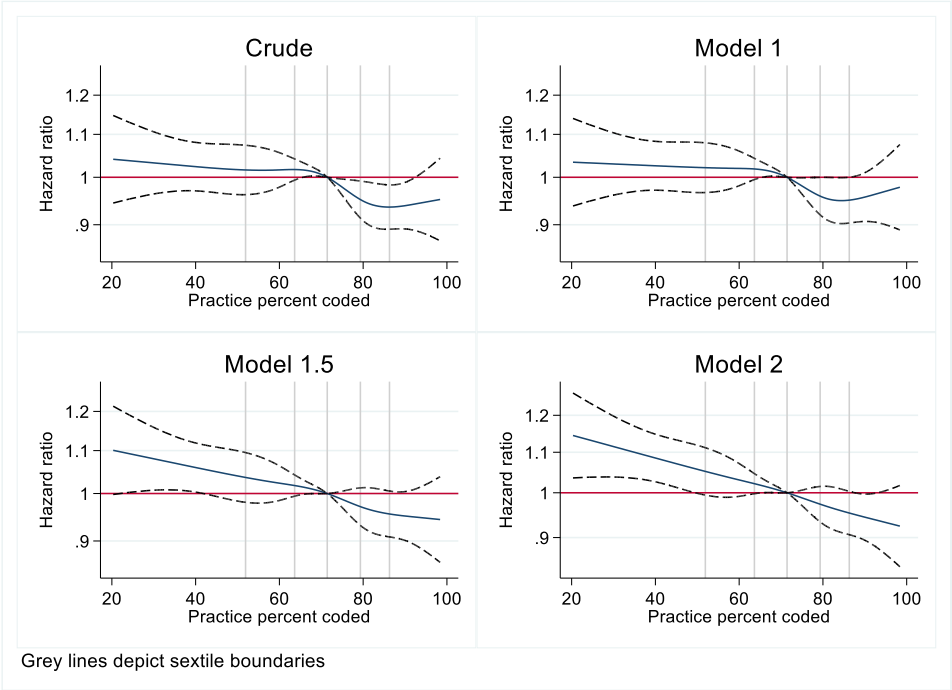


BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)

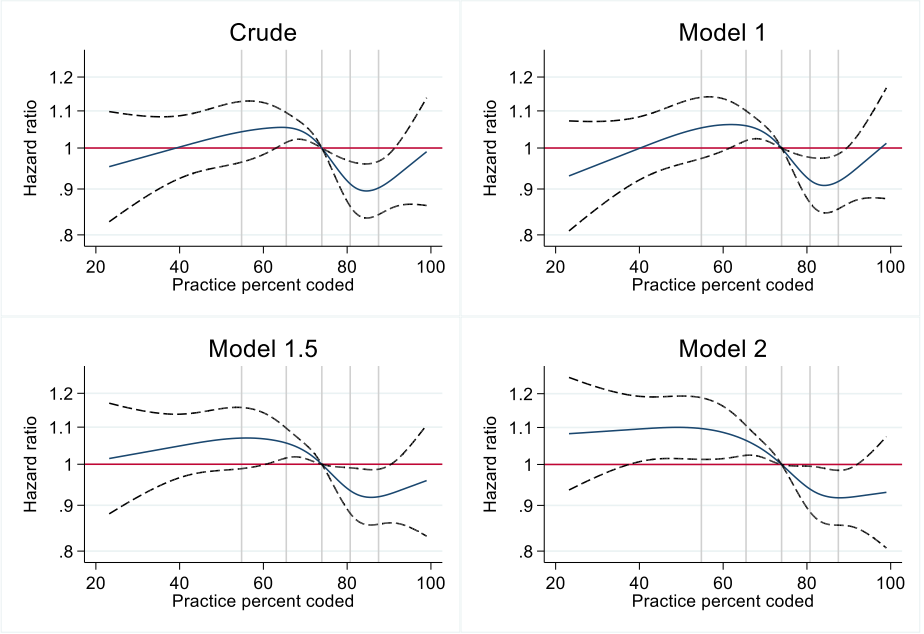


Supplementary Figure 8. Hazard ratio splines for time to first **CV event**, compared to median (71.4%) practice percent coded, in **no diabetes only** (sequential adjustments)



BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

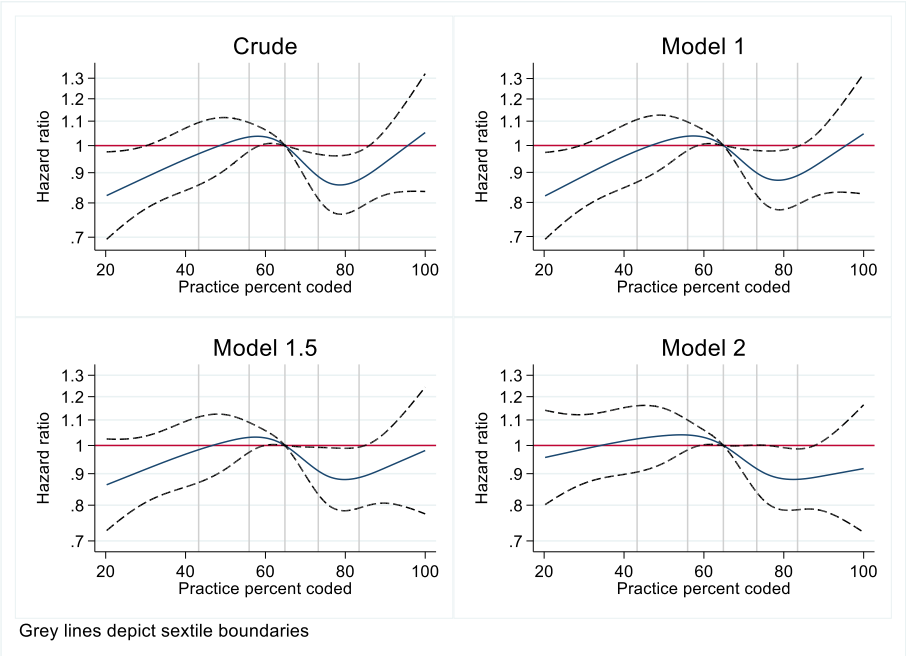
Supplementary Figure 9. Hazard ratio splines for time to first **HF** event, compared to median (73.9%) practice percent coded (sequential adjustments)



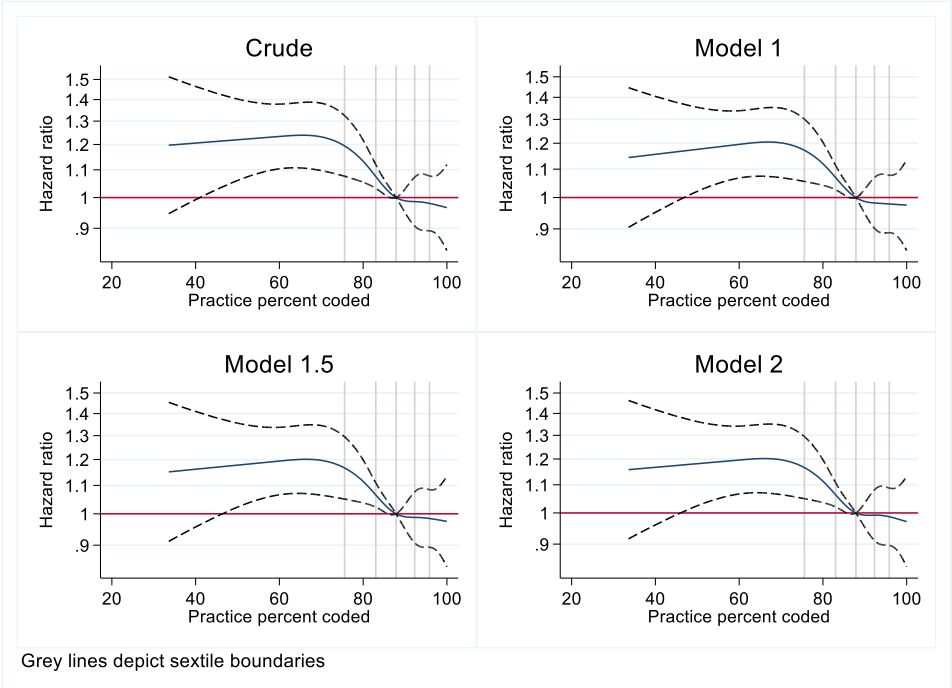
Grey lines depict sextile boundaries

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)

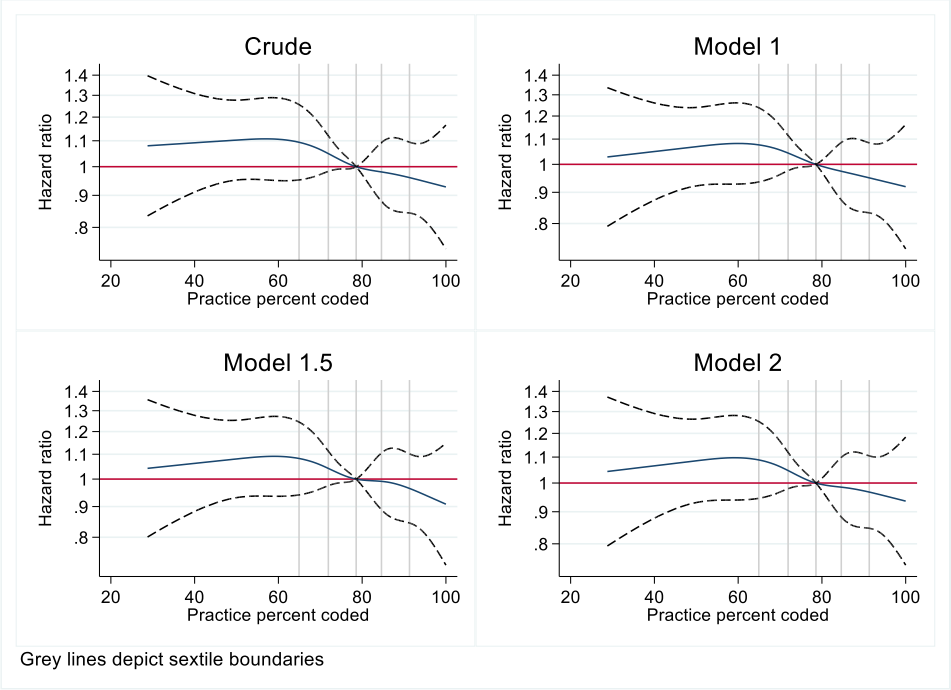


Supplementary Figure 11. Hazard ratio splines for time to first **HF** event, compared to median (87.9%) practice percent coded, in **CKD stages 3b-5** only (sequential adjustments)

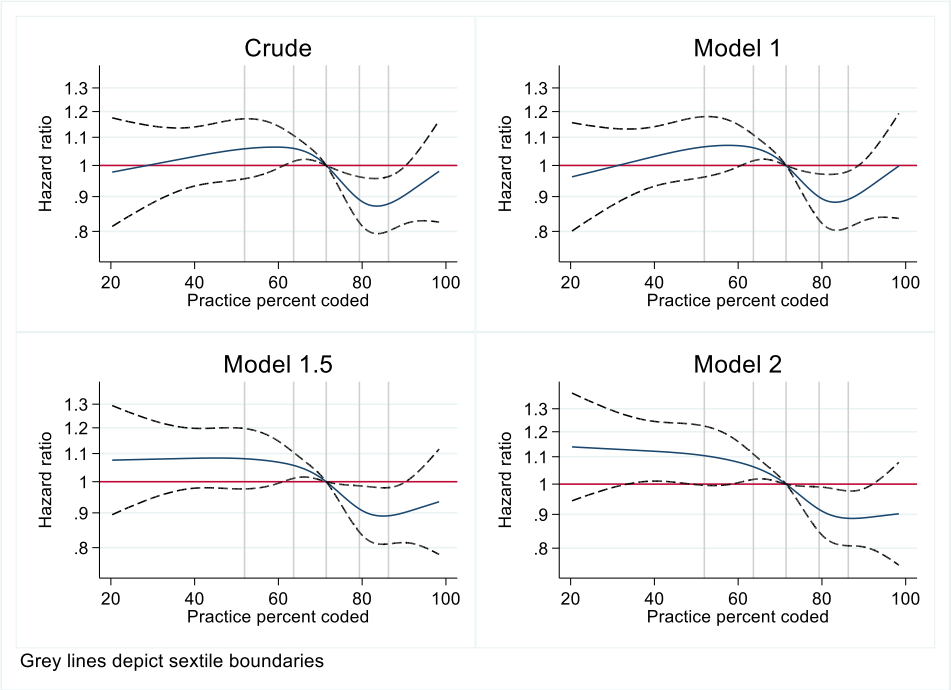


BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)

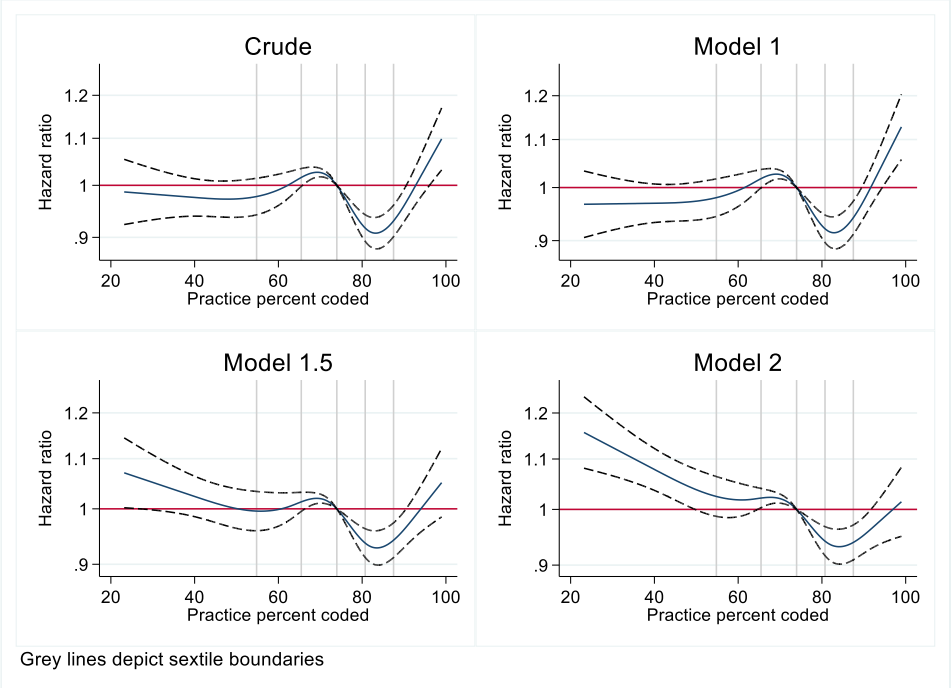


Supplementary Figure 13. Hazard ratio splines for time to first **HF** event, compared to median (71.4%) practice percent coded, in **no diabetes only** (sequential adjustments)



BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

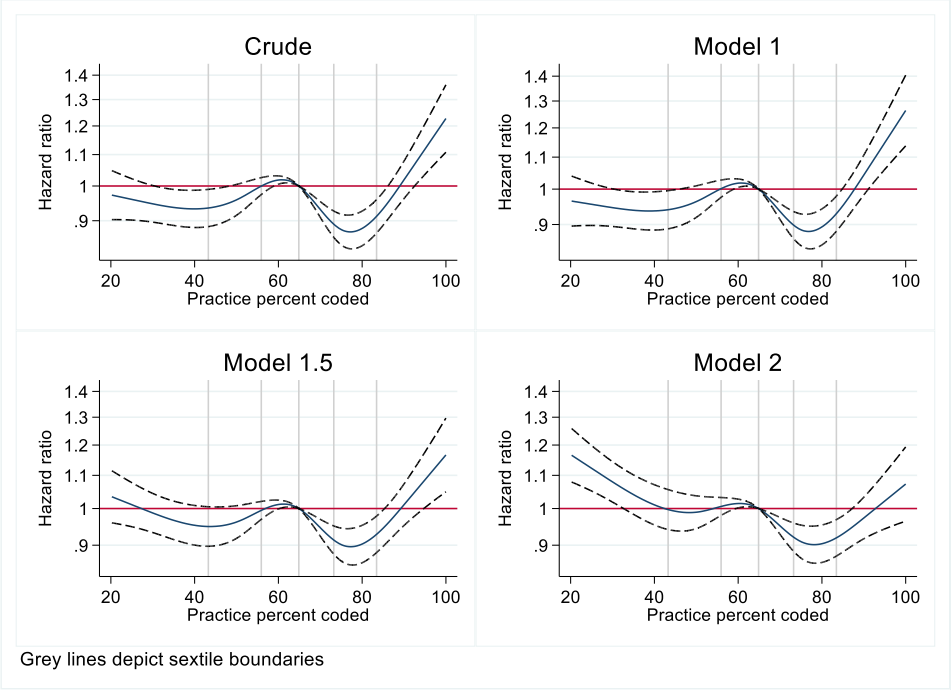
Supplementary Figure 14. Hazard ratio splines for time to first **AKI** event, compared to median (73.9%) practice percent coded (sequential adjustments)



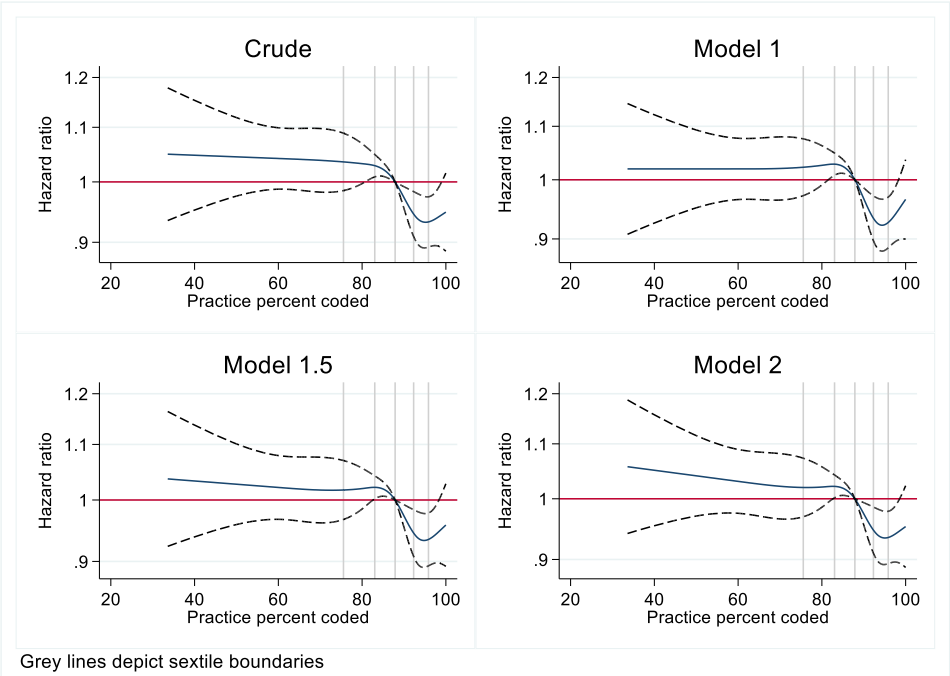
Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)



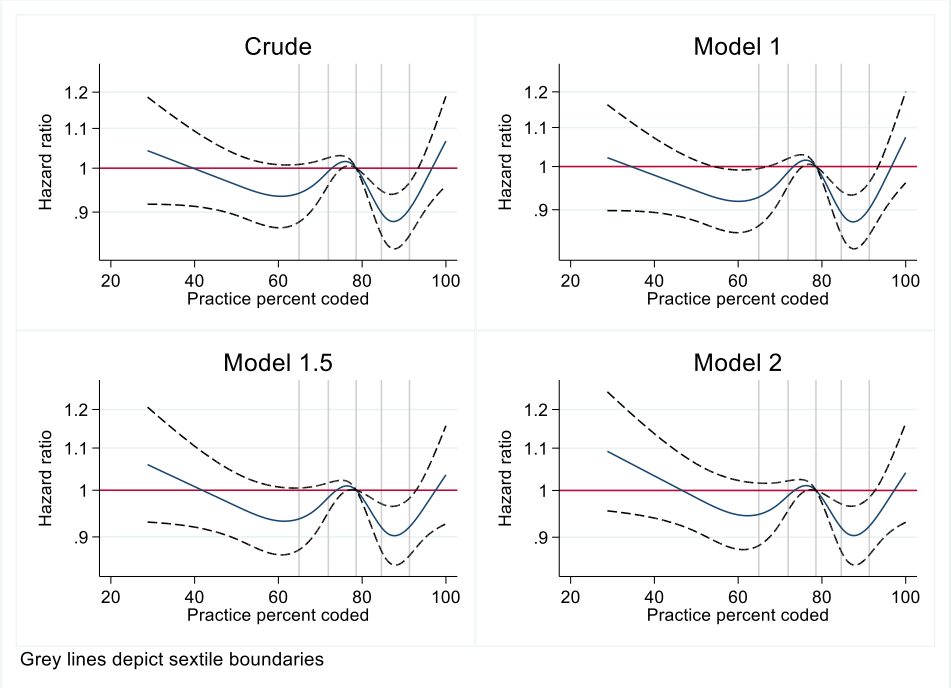
Supplementary Figure 16. Hazard ratio splines for time to first **AKI** event, compared to median (87.9%) practice percent coded, in **CKD stages 3b-5 only** (sequential adjustments)



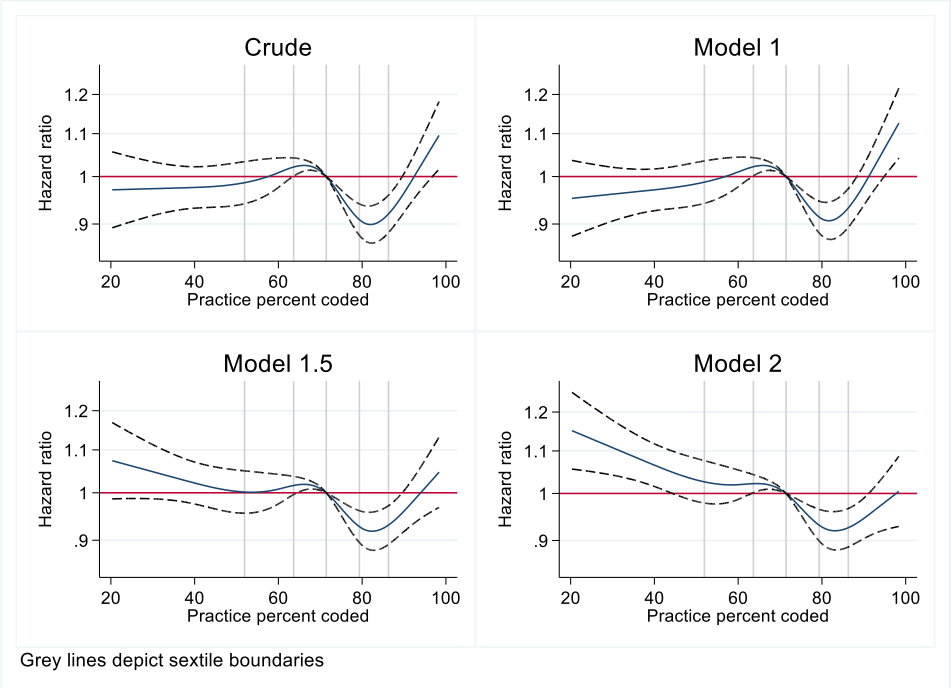
Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)



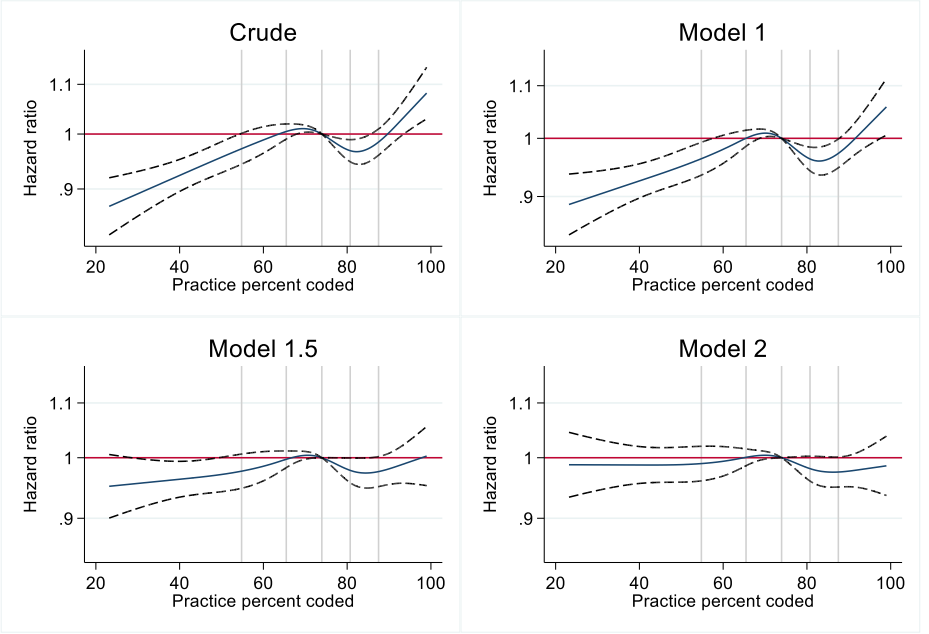
Supplementary Figure 18. Hazard ratio splines for time to first **AKI** event, compared to median (71.4%) practice percent coded, in **no diabetes only** (sequential adjustments)



Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Supplementary Figure 19. Hazard ratio splines for time to **death**, compared to median (73.9%) practice percent coded (sequential adjustments)

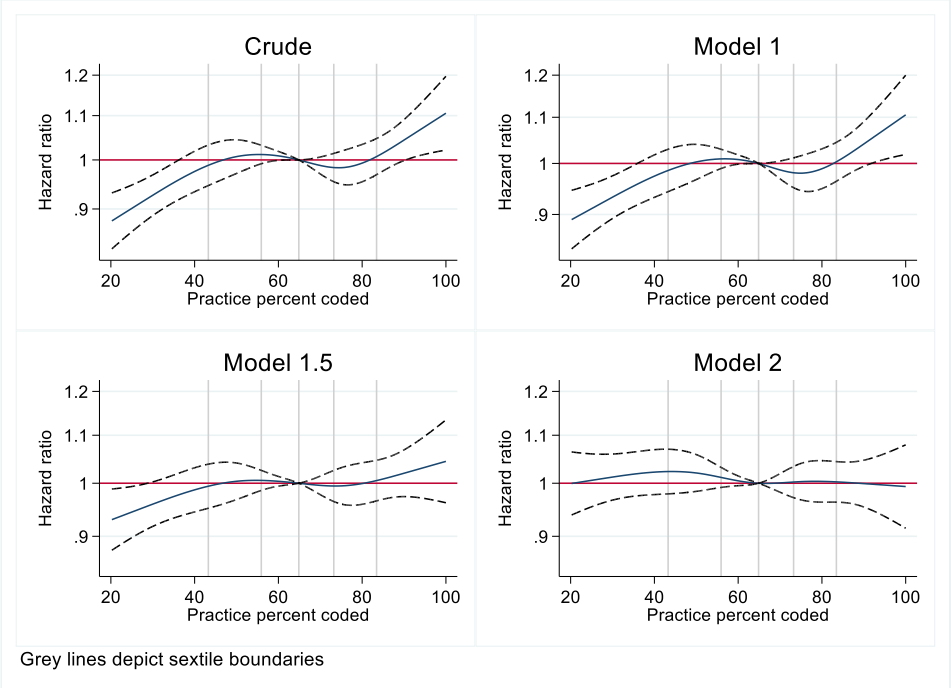


Grey lines depict sextile boundaries

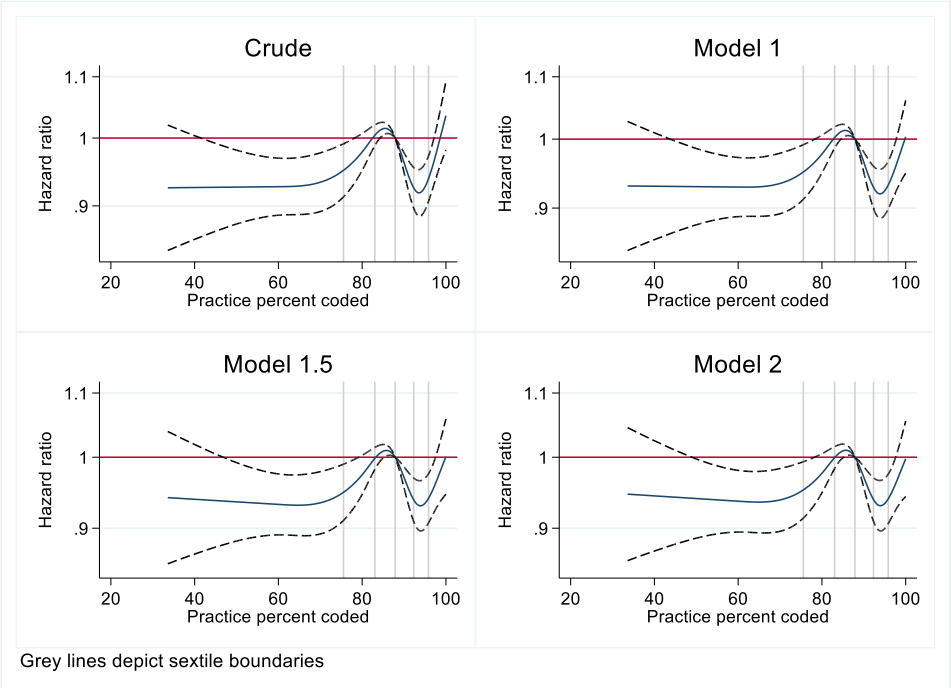
Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)

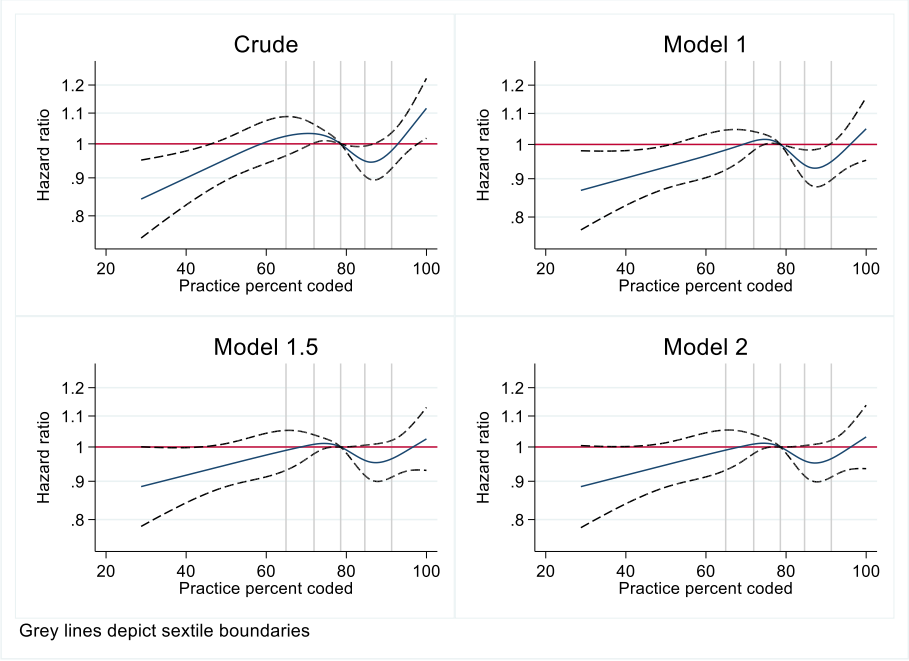


Supplementary Figure 21. Hazard ratio splines for time to **death**, compared to median (87.9%) practice percent coded, in **CKD stages 3b-5 only** (sequential adjustments)

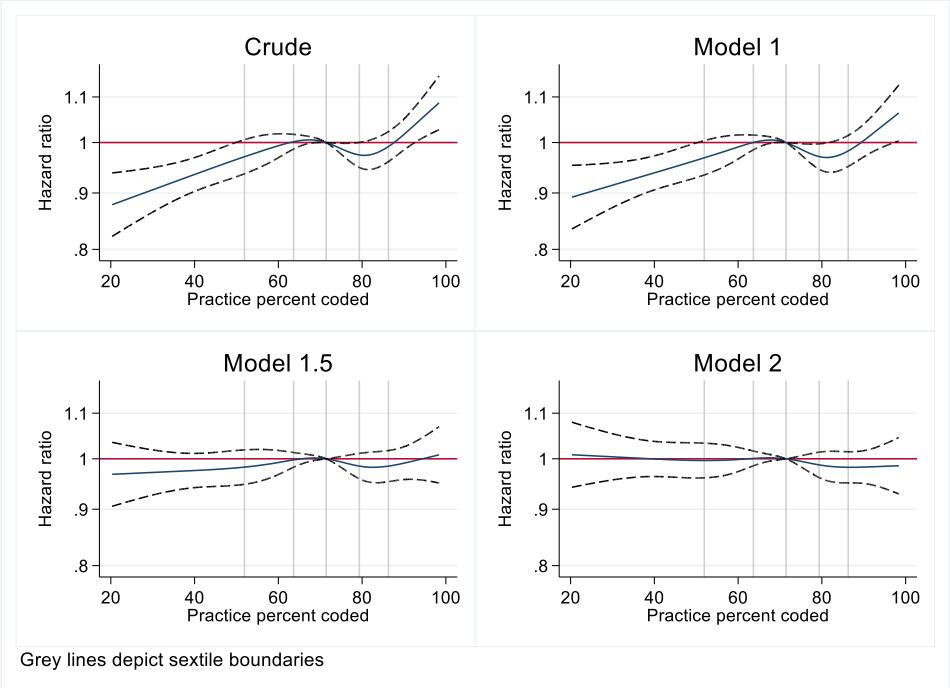


BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)



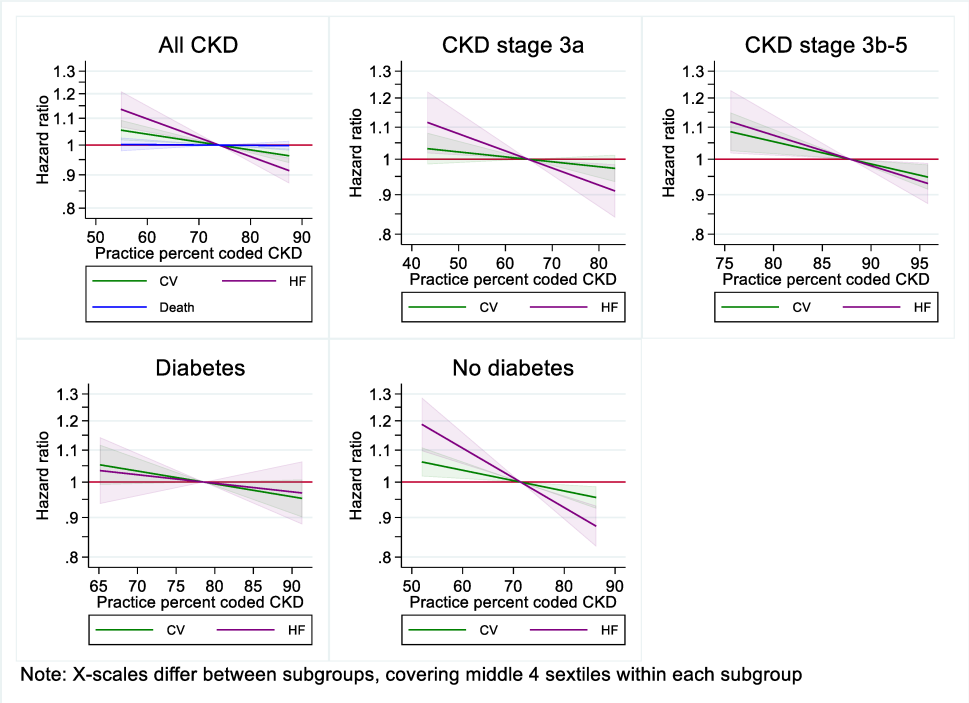
BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)*



BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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

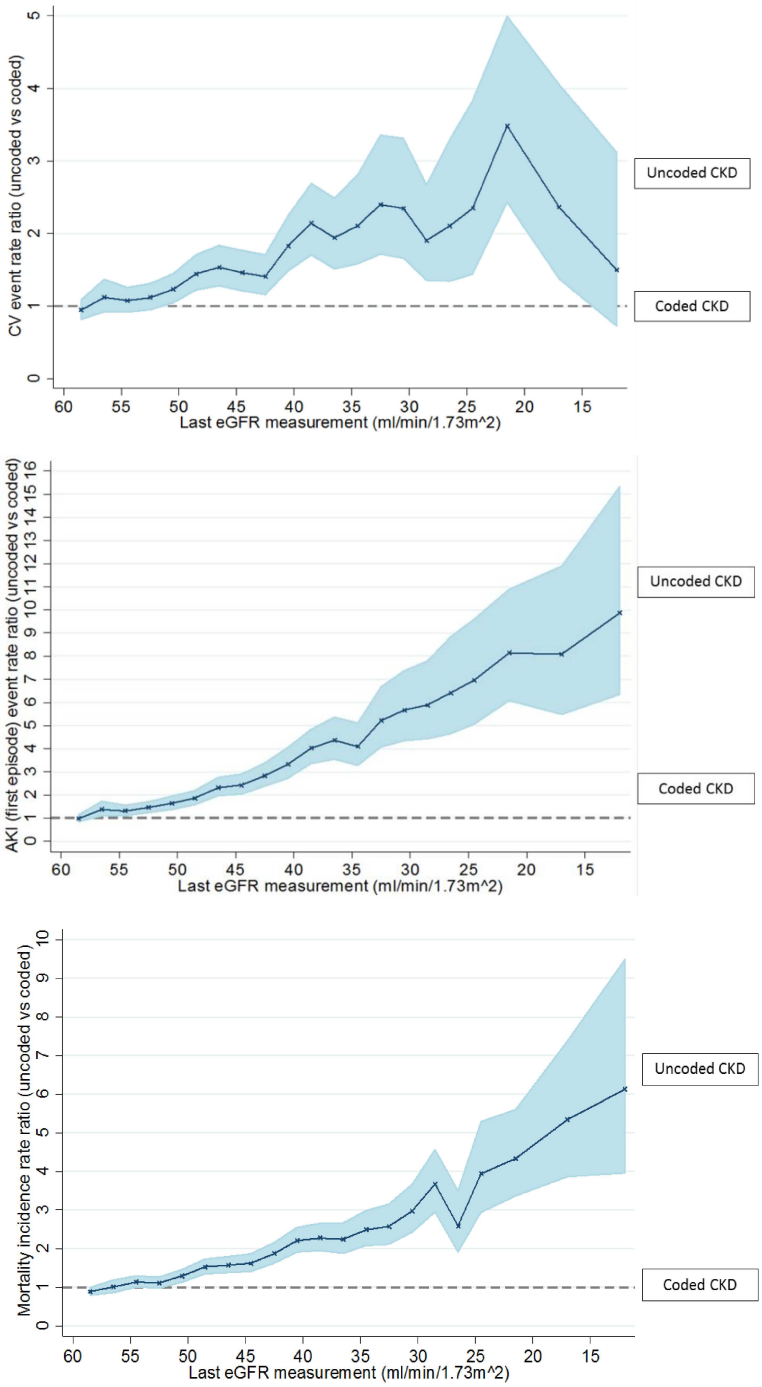
Practice behaviour	**HR (95% CI)**
Percent usage of ACEi/ARBs in hypertension (>76.6%)	0.935 (0.888, 0.984)*
Percent ACR/PCR test in last year in CKD (>58.7%)	0.951 (0.901, 1.004)
Percent flu vaccination in last year in CKD (>78.8%)	0.959 (0.908, 1.013)
Percent meeting blood pressure target in last year 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)

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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



BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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)

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

		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			

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

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

Cleary F, *et al. BMJ Open* 2022; 12:e064513. doi: 10.1136/bmjopen-2022-064513