

Diagnostic Validity of Chronic Kidney Disease in Health Claims Data Over Time: Results from a Cohort of Community-Dwelling Older Adults in Germany

Tim Bothe¹, Anne-Katrin Fietz², Elke Schaeffner¹, Antonios Douros³⁻⁵, Anna Pöhlmann², Nina Mielke¹, Cédric Villain^{1,6}, Muhammad Helmi Barghouth¹, Volker Wenning⁷, Natalie Ebert¹

¹Institute of Public Health, Charité – Universitätsmedizin Berlin, Berlin, Germany; ²Institute of Biometry and Clinical Epidemiology, Charité – Universitätsmedizin Berlin, Berlin, Germany; ³Departments of Medicine and of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada; ⁴Institute of Clinical Pharmacology and Toxicology, Charité – Universitätsmedizin Berlin, Berlin, Germany; ⁵Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, QC, Canada; ⁶Normandie Univ UNICAEN, INSERM U1075 COMETE, service de Gériatrie, CHU de Caen, Caen, France; ⁷AOK Nordost – Die Gesundheitskasse, Potsdam, Germany

Correspondence: Tim Bothe, Charité – Universitätsmedizin Berlin, Institute of Public Health, Charitéplatz 1, Berlin, DE-10117, Germany, Tel +49 30 450 570054, Email tim.bothe@charite.de

Purpose: The validity of ICD-10 diagnostic codes for chronic kidney disease (CKD) in health claims data has not been sufficiently studied in the general population and over time.

Patients and Methods: We used data from the Berlin Initiative Study (BIS), a prospective longitudinal cohort of community-dwelling individuals aged ≥ 70 years in Berlin, Germany. With estimated glomerular filtration rate (eGFR) as reference, we assessed the diagnostic validity (sensitivity, specificity, positive [PPV], and negative predictive values [NPV]) of different claims-based ICD-10 codes for CKD stages G3-5 (eGFR $< 60 \text{ mL/min/1.73m}^2$: ICD-10 N18.x-N19), G3 (eGFR $30 - < 60 \text{ mL/min/1.73m}^2$: N18.3), and G4-5 (eGFR $< 30 \text{ mL/min/1.73m}^2$: N18.4-5). We analysed trends over five study visits (2009–2019).

Results: We included data of 2068 participants at baseline (2009–2011) and 870 at follow-up 4 (2018–2019), of whom 784 (38.9%) and 440 (50.6%) had CKD G3-5, respectively. At baseline, sensitivity for CKD in claims data ranged from 0.25 (95%-confidence interval [CI] 0.22–0.28) to 0.51 (95%-CI 0.48–0.55) for G3-5, depending on the included ICD-10 codes, 0.20 (95%-CI 0.18–0.24) for G3, and 0.36 (95%-CI 0.25–0.49) for G4-5. Over the course of 10 years, sensitivity increased by 0.17 to 0.29 in all groups. Specificity, PPVs, and NPVs remained mostly stable over time and ranged from 0.82–0.99, 0.47–0.89, and 0.66–0.98 across all study visits, respectively.

Conclusion: German claims data showed overall agreeable performance in identifying older adults with CKD, while differentiation between stages was limited. Our results suggest increasing sensitivity over time possibly attributable to improved CKD diagnosis and awareness.

Keywords: CKD, diagnostic validity, health claims data, sensitivity, specificity

Introduction

Health claims data have been recognized as a relevant data source and used increasingly in clinical and epidemiological studies over the past years as they reflect real-world clinical practice and health care for a large number of individuals.¹⁻³ Claims data have also been used for estimating the prevalence of chronic kidney disease (CKD) and associated adverse outcomes.⁴⁻⁹ CKD is globally recognized as a relevant public health burden due to increased morbidity,^{10,11} risk of hospitalizations,^{12,13} vulnerability to infection,¹⁴ high costs,^{7,8} and mortality.^{11,15,16} Its prevalence is estimated to range from 23% to 36% in persons aged ≥ 65 years,¹⁷ or even higher in persons ≥ 80 years,¹⁸ and expected to further increase due to ageing populations.^{10,19}

Despite their potential, claims data should be used cautiously when aiming to reliably identify patients with CKD given their lack of laboratory values. Several studies, including two systematic reviews from 2011,^{20,21} evaluated the validity of CKD diagnostic codes and reported mostly poor sensitivity and limited generalisability of their findings. Sensitivity ranged from 3% to 88%, whereas specificity was overall high.^{20–28} However, broadly diverging sensitivity estimates as well as differences among countries and health care systems limit generalisability.²⁰ Furthermore, many of the existing studies were based on samples with an indication for CKD assessment (eg, after hospitalization or in high risk populations) and used cross-sectional older inpatient data. Up to date, the diagnostic validity for CKD has not been sufficiently studied in the general population (ie, populations not at high risk for CKD), particularly in older adults in Europe, and in more recent longitudinal data including outpatient claims. Whereas the diagnostic validity of acute kidney injury (AKI) in claims data seems to have improved over the years,²⁰ trends in diagnostic validity for CKD over time, assuming possible improvements after the implementation of the Kidney Disease: Improving Global Outcomes (KDIGO) guidelines in 2012,²⁹ have, to the best of our knowledge, not been investigated yet.

To address this knowledge gap, we used data from a longitudinal cohort study of community-dwelling older individuals in Germany with detailed laboratory data on kidney function and complementary health claims data to estimate the validity of claims-based identification of CKD. We compared different claims-based CKD definitions using inpatient and outpatient diagnosis codes for different CKD stages and investigated whether there was a change in diagnostic validity between 2009 and 2019.

Patients and Methods

Study Design and Data Source

We conducted an analysis investigating the diagnostic validity of CKD based on German health claims data over time. As the reference standard, we used data of the Berlin Initiative Study (BIS), a longitudinal cohort with detailed phenotyping of kidney measures in 2069 community-dwelling persons aged ≥ 70 years in Berlin, Germany, all insured at the statutory health insurance fund “AOK Nordost – Die Gesundheitskasse” (AOK). Exclusion criteria for participation in the BIS were age < 70 years, dependency on kidney replacement therapy, or kidney transplant recipients, and more details are described elsewhere.³⁰ Data were collected in five biennial face-to-face study visits from 2009 to 2019. The study was approved by the local ethics committee at the Charité – Universitätsmedizin Berlin (EA2/009/08). The BIS adheres to the Declaration of Helsinki and rules of “Good Scientific Practice”. All study participants gave written informed consent.

Information on sociodemographic characteristics, comorbidities, medications, and anthropometric measurements was routinely assessed using standardized computer-based questionnaires as well as blood and urine samples at biennial study visits. Kidney function was assessed from blood samples using serum creatinine which was measured using an isotope-dilution mass spectrometry-traceable enzymatic method. Detailed information on the BIS study is described elsewhere.³⁰ For this study, we included only participants with non-missing serum creatinine measurements at a respective study visit.

For all BIS participants, complementary AOK health claims data were available. Claims data included information on inpatient and outpatient diagnoses which were available in accordance with the German modification of the 10th revision of the International Classification of Diseases (ICD-10-GM). Additional information on the diagnosis type (ie, if an inpatient diagnosis was the main or secondary discharge diagnosis) was available for all diagnoses. The reporting of this study follows the STARD guidelines for diagnostic accuracy studies ([Table S1](#)).³¹

Definition of CKD Based on Kidney Function

As the reference standard, CKD was defined based on estimated glomerular filtration rate (eGFR) using the creatinine-based CKD Epidemiology Collaboration 2009 equation³² as recommended by the KDIGO guidelines from 2012. We included CKD stages G3–5 (eGFR < 60 mL/min/1.73m²) and further stratified in stage G3 (eGFR 30– < 60 mL/min/1.73m²) and stages G4–5 (eGFR < 30 mL/min/1.73m²). CKD was assessed at baseline (conducted in 2009–2011) and followed-up biennially (FU1–4; conducted in 2011–2013, 2014–2015, 2016–2017, and 2018–2019, respectively).

Definition of CKD Based on Claims Data

As the index test, we defined CKD based on claims data using inpatient and outpatient ICD-10-GM codes in correspondence with respective CKD stages. We used ICD-10-GM codes N18.3–5 for CKD G3–5, N18.3 for G3, and N18.4–5 for G4–5. For CKD G3–5, we extended the included codes from the primary analysis (N18.3–5, definition #1) and explored broader claims-based definitions: definition #2 also included unspecific diagnostic codes (N18.3–5, N18.8x, N18.9, N19), whereas definition #3 additionally included CKD stages G1–2 diagnoses (N18.x, N19; see [Table S2](#) for a detailed description). We only included inpatient diagnoses coded as “main” or “secondary” and outpatient diagnoses with diagnosis type “secure”. We used an observation period for claims data diagnoses of one year, spanning from six months preceding to six months after a study visit date (index date), or until death.

Statistical Analysis

Sociodemographic and clinical characteristics of study participants were assessed as absolute and relative frequencies for binary and mean (SD) for metric variables. As the primary outcome, diagnostic validity of claims-based CKD identification with eGFR as the reference was evaluated by assessing sensitivity, specificity, and positive (PPV) and negative predictive values (NPV). Diagnostic validity indices with 95% confidence intervals (CI) were estimated for all study visits. We assessed differences in diagnostic validity indices from baseline to FU4 (Δ) and trends over time with the Cochran-Armitage test in an exploratory manner.^{33,34} Additionally, we assessed the relative frequency of persons who received single N18.x and N19 diagnoses for a better understanding of diagnostic codes among CKD definitions.

Sensitivity was calculated as the proportion of persons with a respective claims-based diagnosis among persons with a respective eGFR (true-positive rate). *Specificity* was calculated as the proportion of persons without a respective claims-based diagnosis among persons without a respective eGFR (true-negative rate). *PPV* was calculated as the proportion of persons with a respective eGFR among all persons with a respective claims-based diagnosis. *NPV* was calculated as the proportion of persons without a respective eGFR among all persons without a respective claims-based diagnosis.

We performed pre-defined subgroup analyses stratified by age, sex, diabetes mellitus, and arterial hypertension. Diabetes mellitus was defined as HbA1c $\geq 6.5\%$ or the intake of antidiabetic medication (oral antidiabetics and insulin) and arterial hypertension as the intake of antihypertensive medications (angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium channel blockers, beta-blockers, and diuretics [excluding loop-diuretics]), assessed during the study visit based on the participants' medication lists and using the respective ATC (anatomical therapeutic chemical) codes. Further detail was published elsewhere.^{35,36}

As sensitivity analyses, we included the “chronicity” criterion for CKD as defined in the KDIGO guidelines (two eGFR values at least three months apart).²⁹ We applied the chronicity criterion to identify individuals with a respective CKD G3–5, G3, or G4–5 based on eGFR in two consecutive study visits. The latter study visit was used as the index date to search for claims-based diagnoses in both six months preceding and following. As studies suggest that sensitivity improves when choosing longer observation periods in claims data,^{20,23,28} we assessed the analytical flexibility of observation time for claims-based diagnoses and estimated indices of diagnostic validity of CKD from one up to 52 weeks preceding and following a study visit, extending the observation period to two years in total.

Results

Sociodemographic and Clinical Characteristics

We included data of 2068 participants at baseline, 1670 at FU1, 1421 at FU2, 1130 at FU3, and 870 at FU4 ([Table 1](#) and [Figure S1](#)). At all visits, more than half were female (53–56%) and mean age was 79.9 years (SD 6.7) at baseline and 85.1 years (SD 5.3) at FU4. Prevalence of diabetes and hypertension were between 23–26% and 79–83% across all study visits, respectively. Mean eGFR decreased slightly over time from 64.7 (SD 17.5) at baseline to 58.9 (SD 17.7) mL/min/1.73m² at FU4. At baseline, CKD prevalence based on eGFR was 38% for CKD G3–5, 35% for CKD G3, and 3% for CKD G4–5. CKD prevalence increased over time and was 51% for CKD G3–5, 45% for CKD G3, and 6% for CKD G4–5 at FU4.

Table I Sociodemographic and Clinical Characteristics for Each Study Visit Based on BIS Data

	Baseline (2009–2011)	FU1 (2011–2013)	FU2 (2014–2015)	FU3 (2016–2017)	FU4 (2018–2019)
Number of subjects, N	2068	1670	1421	1130	870
Demographics					
Females, n (%)	1088 (52.6)	889 (53.2)	766 (53.9)	618 (54.7)	487 (56.0)
Age (years), mean (SD)	79.9 (±6.7)	81.3 (±6.4)	82.6 (±6.1)	83.8 (±5.6)	85.1 (±5.3)
Age groups (years), n (%)					
70–74	573 (27.7)	273 (16.3)	66 (4.6)	0 (0.0)	0 (0.0)
75–79	475 (23.0)	502 (30.1)	472 (33.2)	321 (28.4)	115 (13.2)
80–84	429 (20.7)	375 (22.5)	377 (26.5)	356 (31.5)	363 (41.7)
85–89	385 (18.6)	309 (18.5)	270 (19.0)	248 (21.9)	208 (23.9)
≥90	206 (10.0)	211 (12.6)	236 (16.6)	205 (18.1)	184 (21.1)
Comorbidities					
Diabetes mellitus, n (%)	539 (26.1)	379 (22.7)	346 (24.3)	269 (23.8)	200 (23.0)
Missing, n (%)	3 (0.1)	6 (0.4)	9 (0.6)	6 (0.5)	7 (0.8)
Arterial hypertension, n (%)	1633 (79.0)	1342 (80.4)	1167 (82.1)	931 (82.4)	719 (82.6)
Missing, n (%)	6 (0.3)	2 (0.1)	3 (0.2)	1 (0.1)	2 (0.2)
eGFR (mL/min/1.73m²), mean (SD)					
CKD-EPI _{Crea}	64.7 (±17.5)	62.8 (±17.9)	61.8 (±17.9)	60.1 (±17.5)	58.9 (±17.7)
CKD prevalence (based on CKD-EPI_{Crea}), n (%)					
CKD G3-5 (eGFR <60 mL/min/1.73m ²)	784 (37.9)	696 (41.7)	618 (43.5)	529 (46.8)	440 (50.6)
CKD G3 (eGFR 30–<60 mL/min/1.73m ²)	723 (35.0)	621 (37.2)	544 (38.3)	467 (41.3)	388 (44.6)
CKD G4-5 (eGFR <30 mL/min/1.73m ²)	61 (2.9)	75 (4.5)	74 (5.2)	62 (5.5)	52 (6.0)

Notes: The table only shows data of BIS participants with non-missing serum creatinine measurements at a respective study visit. Excluded due to missing serum creatinine measurements were: n = 1, 29, 19, 36, and 64 at baseline, FU1, FU2, FU3, and FU4, respectively.

Abbreviations: FU, Follow-up; eGFR, estimated glomerular filtration rate; CKD, chronic kidney disease; CKD-EPI_{Crea}, creatinine-based CKD Epidemiology Collaboration 2009 equation; CKD G3-5, G3, and G4-5, CKD stages G3-5, G3, and G4-5.

Sensitivity

Indicators of diagnostic validity for different eGFR- and claims-based definitions of CKD are displayed in Table 2. Sensitivity at baseline was 0.25 (95% CI 0.22–0.28) for CKD G3-5 using definition #1 (ICD-10-GM N18.3–5). When additionally incorporating non-stage-corresponding diagnoses, sensitivity increased to 0.46 (95% CI 0.43–0.50) using definition #2 (ICD-10-GM N18.3–19), and to 0.51 (95% CI 0.48–0.55) using definition #3 (ICD-10-GM N18.x-19). Stratified by CKD stages, sensitivity was 0.20 (95% CI 0.18–0.24) for CKD G3 and 0.36 (95% CI 0.25–0.49) for CKD G4-5 at baseline. Sensitivity strongly increased from baseline to FU4 among all groups and definitions by 0.17 to 0.29 and was highest for CKD G3-5 using definition #3 (0.69, 95% CI 0.64–0.73) and lowest for CKD G3 (0.47, 95% CI 0.43–0.52) at FU4.

Specificity, PPV, and NPV

Specificity values ranged overall from 0.82 to 0.99 with slight downward trends from baseline to FU4 in all CKD definitions ($\Delta = -0.09$ to -0.02). Specificity was highest in CKD G4-5 (range: 0.97 to 0.99) and lowest in CKD G3-5 definition #3 (0.82 to 0.91). PPVs varied broadly between 0.47 and 0.89 over all groups and definitions. They were highest in CKD G3-5 definition #1 (0.86 to 0.89) and lowest in CKD G4-5 (0.47 to 0.57). Over time, PPVs remained stable in all groups ($\Delta = -0.01$ to 0.03) except for slight increases in CKD G4-5 ($\Delta = 0.08$). NPVs ranged from 0.66 to 0.98 and remained stable over time ($\Delta = -0.04$ to 0.00). NPVs were highest in CKD G4-5 (0.97 to 0.98) and lowest in CKD 3–5 definition #1 (0.66 to 0.69).

In CKD G3-5, incorporating more diagnoses was accompanied by slightly higher NPVs (0.66–0.69, 0.71–0.75, and 0.72–0.76 for definitions #1, #2, and #3, respectively) but slightly lower specificity (0.92–0.98, 0.87–0.93, and 0.82–

Table 2 Indicators of Diagnostic Validity (with 95% Confidence Intervals) for Different CKD Definitions for Each Study Visit with Trends Over Time

eGFR-Based CKD Definition	Claims-Based CKD Definition (ICD-10-GM Codes ^a)	Indicator of Diagnostic Validity	Baseline (2009–2011)	FU1 (2011–2013)	FU2 (2014–2015)	FU3 (2016–2017)	FU4 (2018–2019)	Trend
								Δ^b
G3-5 (eGFR <60 mL/min/1.73m ²)	#1 N18.3, N18.4, N18.5	Sensitivity	0.25 (0.22–0.28)	0.33 (0.30–0.37)	0.43 (0.39–0.47)	0.52 (0.48–0.56)	0.54 (0.49–0.58)	0.29*
		Specificity	0.98 (0.97–0.99)	0.96 (0.95–0.97)	0.95 (0.94–0.97)	0.93 (0.91–0.95)	0.92 (0.89–0.94)	–0.06*
		PPV	0.89 (0.84–0.93)	0.86 (0.81–0.89)	0.87 (0.83–0.91)	0.87 (0.83–0.90)	0.88 (0.83–0.91)	–0.01
		NPV	0.68 (0.66–0.70)	0.67 (0.64–0.69)	0.68 (0.66–0.71)	0.69 (0.65–0.72)	0.66 (0.62–0.70)	–0.02
	#2 N18.3, N18.4, N18.5, N18.8x, N18.9, N19	Sensitivity	0.46 (0.43–0.50)	0.51 (0.48–0.55)	0.61 (0.57–0.64)	0.64 (0.60–0.68)	0.65 (0.60–0.69)	0.19*
		Specificity	0.93 (0.92–0.94)	0.92 (0.90–0.93)	0.91 (0.89–0.93)	0.87 (0.84–0.90)	0.87 (0.83–0.90)	–0.06*
		PPV	0.80 (0.76–0.84)	0.81 (0.77–0.85)	0.84 (0.81–0.87)	0.81 (0.77–0.85)	0.84 (0.79–0.87)	0.03
		NPV	0.74 (0.72–0.76)	0.73 (0.70–0.75)	0.75 (0.72–0.78)	0.73 (0.70–0.76)	0.71 (0.67–0.75)	–0.03
	#3 N18.x, N19	Sensitivity	0.51 (0.48–0.55)	0.56 (0.52–0.60)	0.65 (0.61–0.68)	0.67 (0.63–0.71)	0.69 (0.64–0.73)	0.17*
		Specificity	0.91 (0.89–0.93)	0.89 (0.87–0.91)	0.87 (0.85–0.89)	0.82 (0.78–0.84)	0.82 (0.78–0.85)	–0.09*
		PPV	0.78 (0.74–0.81)	0.78 (0.75–0.82)	0.79 (0.76–0.83)	0.76 (0.72–0.80)	0.79 (0.75–0.83)	0.01
		NPV	0.75 (0.73–0.78)	0.74 (0.71–0.76)	0.76 (0.73–0.79)	0.74 (0.71–0.77)	0.72 (0.68–0.76)	–0.04
G3 (eGFR 30–<60 mL/min/1.73m ²)	N18.3	Sensitivity	0.20 (0.18–0.24)	0.27 (0.24–0.31)	0.36 (0.32–0.40)	0.45 (0.41–0.50)	0.47 (0.43–0.52)	0.27*
		Specificity	0.96 (0.95–0.97)	0.93 (0.91–0.94)	0.91 (0.89–0.93)	0.87 (0.84–0.90)	0.87 (0.84–0.90)	–0.09*
		PPV	0.72 (0.66–0.78)	0.69 (0.63–0.75)	0.72 (0.66–0.77)	0.71 (0.66–0.76)	0.75 (0.69–0.80)	0.03
		NPV	0.69 (0.67–0.71)	0.68 (0.66–0.71)	0.70 (0.67–0.72)	0.69 (0.66–0.72)	0.67 (0.64–0.71)	–0.02
G4-5 (eGFR <30 mL/min/1.73m ²)	N18.4, N18.5	Sensitivity	0.36 (0.25–0.49)	0.41 (0.31–0.53)	0.49 (0.38–0.60)	0.69 (0.57–0.79)	0.63 (0.50–0.75)	0.27*
		Specificity	0.99 (0.98–0.99)	0.98 (0.97–0.98)	0.98 (0.97–0.98)	0.97 (0.96–0.98)	0.97 (0.95–0.98)	–0.02*
		PPV	0.48 (0.34–0.62)	0.47 (0.35–0.59)	0.54 (0.42–0.65)	0.57 (0.45–0.67)	0.56 (0.43–0.68)	0.08
		NPV	0.98 (0.97–0.99)	0.97 (0.96–0.98)	0.97 (0.96–0.98)	0.98 (0.97–0.99)	0.98 (0.96–0.98)	0.00

Note: ^aOlder ICD-9-GM Version codes (before 2010) for CKD were recoded to their ICD-10-GM equivalent as follows: N18.81 to N18.1, N18.82 to N18.2, N18.83 to N18.3, N18.84 to N18.4, N18.0 to N18.5. ^bTrends were tested with the Cochran-Armitage test with *indicating $p < 0.05$.

Abbreviations: CKD, Chronic kidney disease; eGFR, Estimated glomerular filtration rate; ICD-10-GM, International Statistical Classification of Diseases and Related Health Problems, 10th Revision, German modification; Δ , difference from baseline to FU4.

0.91) and PPVs (0.86–0.89, 0.80–0.84, and 0.76–0.79). Comparing CKD G3 and G4–5, CKD G3 showed lower NPVs (0.67–0.70 for G3 vs 0.97–0.98 for G4–5) and lower specificity (0.87–0.96 vs 0.97–0.99) but higher PPVs (0.69–0.75 vs 0.48–0.57) over all study visits.

Subgroup and Sensitivity Analyses

Sensitivity was higher and specificity slightly lower in males compared to females but showed no difference between age strata (Figures S2 and S3). PPVs were slightly higher and NPVs slightly lower in older ages, while both measures differed inconsistently to a small degree between males and females (Figures S4 and S5). The subgroup analyses for diabetes and hypertension showed no differences between persons with either or both comorbidities. When neither was prevalent, sensitivity was lower, while NPVs were higher (Figure S6). Using the chronicity criterion for eGFR-based CKD definition, sensitivity and NPVs were higher, while specificity and PPVs were lower compared to the single definition (Figure S7). Trends over time remained mostly stable in all strata, subgroups, and the chronic CKD definition.

Figure 1 shows the results on the analytical flexibility of observation time for claims-based diagnoses using weekly thresholds from one year preceding to one year following the index date as observation time, respectively (see Figure S8 for time preceding and following index combined). Within an observation time of ± 3 months, sensitivity increased sharply, followed by a constantly increasing level with every additional week. Exceeding the observation time beyond ± 6 months resulted in overall higher sensitivity. NPVs increased slightly in extended observation periods, while specificity and PPVs showed a small decrease.

Frequencies of Coded Diagnoses

With regard to single N18.x and N19 diagnoses, overall coding of unspecific (N18.8x, N18.9, N19) diagnoses was high and remained constant over time (Figure 2). In persons with CKD G3–5, 37–43% received at least one unspecific diagnosis, whereas this increased to 69–79% in persons with CKD G4–5. Some persons received ICD-10-GM diagnoses not corresponding with their respective eGFR-based CKD stage. I.e., of persons with CKD G4–5, 14–21% were diagnosed with N18.1–2 (corresponding with CKD stages G1–2) and 52–74% with N18.3 (corresponding with CKD stage G3). Over time, the coding frequencies of N18.3 and N18.4–5 diagnoses in persons with CKD G3–5 increased from 23% and 6% at baseline to 49% and 13% at FU4, respectively, while the other CKD diagnoses remained mostly stable.

Discussion

In this study, we investigated the diagnostic validity of in- and outpatient health claims data for CKD over time using creatinine-based eGFR as the reference in 2068 community-dwelling persons aged ≥ 70 years. Diagnostic validity varied between different CKD stages and claims-based definitions. For CKD G3–5, sensitivity was lowest when restricting to stage-corresponding diagnosis codes (definition #1), was considerably higher when including unspecific diagnoses (definition #2), and further increased slightly when additionally using non-stage-corresponding diagnoses (definition #3). Sensitivity was higher in CKD G4–5 compared to G3, but still lower than in definition #2 and #3 for CKD G3–5, indicating limited identification of specific CKD stages. Specificity was overall high, while PPVs and NPVs varied broadly. From 2009 until 2019, sensitivity for CKD increased in all definitions and CKD stages, regardless of age, gender, and comorbidities, while specificity showed small decreases.

Several studies that investigated the diagnostic validity for CKD in claims data reported diverging but mostly poor sensitivity (range 0.03 to 0.88), high specificity (all >0.93), diverging PPVs (range 0.29 to 1.00), and agreeable NPVs (most >0.72).^{20–28} Comparing diagnostic validity across studies, our results showed higher sensitivity and mostly comparable specificity, PPVs, and NPVs. For example, a study on Dutch inpatient claims data of persons aged ≥ 75 years²⁴ reported a sensitivity of 0.27 for CKD G3–5 for the year 2014 compared to 0.43–0.65 in the BIS data assessed in 2014 and 2015. Differences may occur due to the overall higher CKD prevalence based on eGFR in the BIS (0.44 compared to 0.19 in the Dutch data). When comparing data from the REGARDS study including persons aged ≥ 65 years between 2004 and 2007²⁶ with the BIS baseline visit (2009–2011), prevalence of CKD G3–5, specificity, PPVs, and NPVs were comparable, while sensitivity was higher in the BIS (0.16 vs 0.25–0.51 for REGARDS vs BIS). In stratified analyses, we found a higher sensitivity in males compared to females and in individuals with diabetes or hypertension compared to those without which is in accordance with results on higher disease awareness for

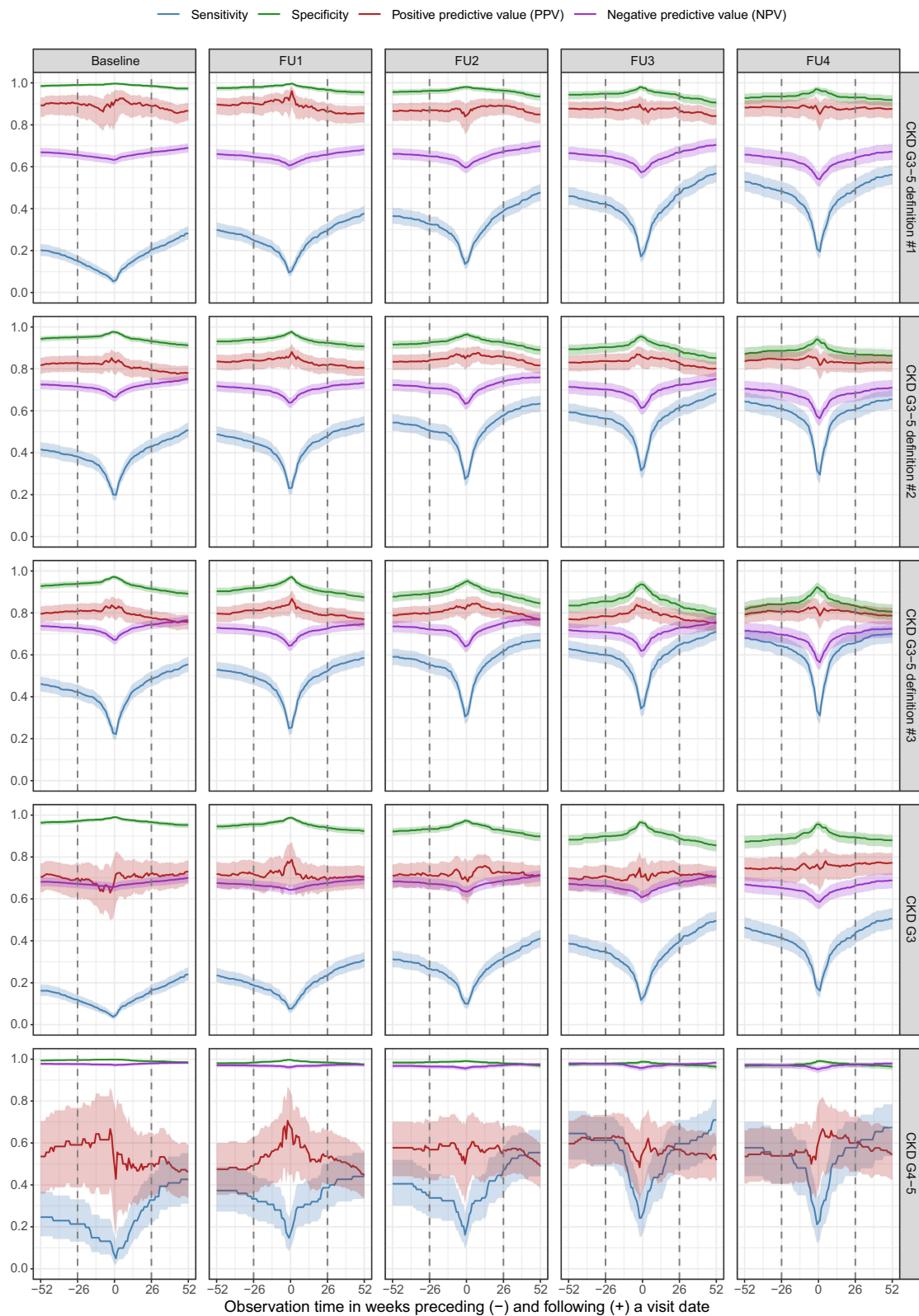


Figure 1 Indicators of diagnostic validity (sensitivity, specificity, PPV, and NPV) of different CKD stages and claims-based definitions using weekly thresholds from one year preceding to one year following a study visit as observation time for claims data diagnoses. Dashed lines represent the observation time window for the main analysis of ± 6 months. Shades represent 95% confidence intervals and are interpolated for graphical display. Included ICD-10-GM diagnoses for each stage and definition are presented in Table 2.

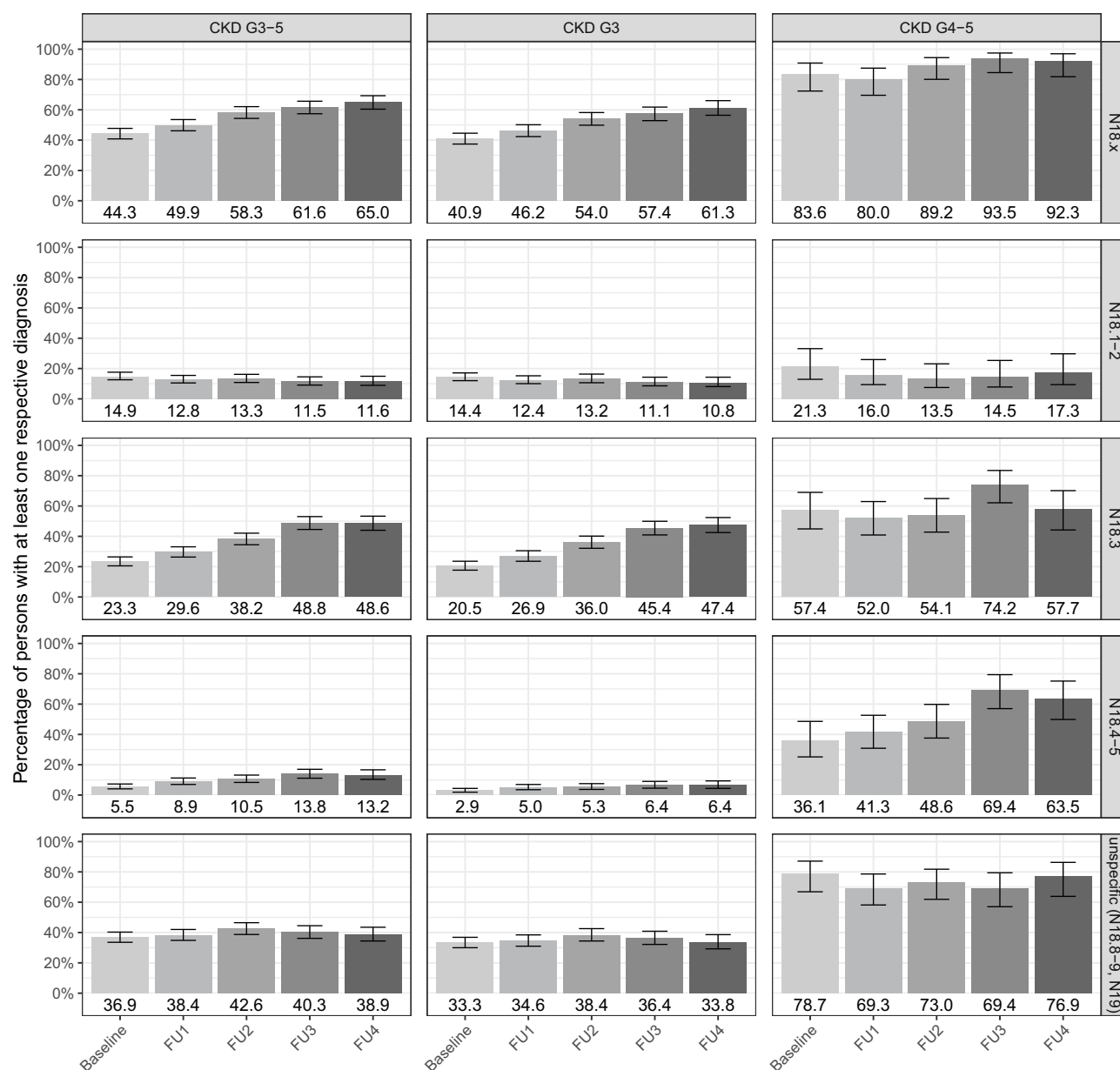


Figure 2 Coding frequencies of single ICD-10-GM N18.x and N19 diagnosis codes. Bars represent the relative frequencies of persons (with values below) with a respective diagnosis in +/- 6 months around a study visit date (index date) for CKD G3-5, G3, and G4-5 from baseline to follow-up (FU) 4. Error bars represent 95% confidence intervals.

CKD in males^{37,38} and previous results.^{20-23,25,26} Our results indicate that the detection of older CKD patients may perform better and may yield a higher validity in more recent German claims data compared to international data. Still, claims-based research on CKD may be further improved by the application of machine-learning algorithms incorporating more information than diagnostic codes only as done in other clinical parameters.³⁹⁻⁴¹ In addition, an electronic linkage to individual laboratory data as available in other countries would greatly enhance the value of German claims data.

A correct identification of CKD is of great clinical importance considering its associated risks of multimorbidity,⁴²⁻⁴⁴ cardiovascular events,⁴⁴⁻⁴⁶ and implications for drug dose adjustments.⁴⁷⁻⁴⁹ For CKD G3-5, incorporating more diagnostic codes within a claims-based definition resulted in increased sensitivity and was accompanied by slightly higher NPVs but slightly lower specificity and PPVs. These results indicate that using the ICD-10-GM codes N18.3-19 may provide the most preferable and balanced results for a claims-based definition of CKD, while ICD-10-GM codes

N18.3–5 may be more suitable if higher PPVs are required, eg, in drug safety studies. In the selection of the period of inclusion, extending the observation time in claims-based research for CKD from one (± 6 months) to two years (± 12 months) may be preferable as it results in a higher sensitivity and negligible decreases in other indicators of diagnostic validity. Given the increasing importance and potential use of claims-based research on chronic diseases, validation studies as well as the implementation of a CKD registry in Germany complemented by claims data could fundamentally improve the understanding of CKD prevalence and incidence as well as the quality of epidemiological studies on CKD.

We found that the validity of CKD detection improved between 2009 and 2019 as sensitivity steadily increased, in total by 0.17 to 0.29. This trend may be due to several reasons: BIS participants' ageing related eGFR declines resulting in higher CKD prevalence;^{18,50} standardized eGFR monitoring across all study visits and reporting to the treating general practitioners may having an impact on CKD coding; increased CKD awareness in some BIS-participants due to study participation; or improvements in CKD diagnosis as a consequence of effective KDIGO guideline implementation. Due to the observational design of the BIS, we cannot draw direct conclusions on differential effects and their impact. However, the trends are in accordance with increased sensitivity for claims-based identification of AKI²⁰ and may be supported by steep upward trends in the occurrence of stage-specific N18.3–5 diagnoses over the study visits, while unspecific diagnoses remained stable over time. The improved sensitivity and detection of specific CKD stages may indicate a better awareness amongst general practitioners for kidney disease and result in improved quality of care, more adequate drug dosing, and earlier referral to specialist care. In the absence of a German national CKD registry, a more unified use of the specific ICD-10-GM codes for CKD stages in primary care may improve the validity of CKD diagnosis in claims data which would also be of interest for German health authorities. It could serve as a basis to better navigate patients and avoid under-, over-, and misuse in health care, and thus guarantee an adequate distribution of resources in an increasingly expensive health care system.

Compared to CKD G3–5, the diagnostic validity of claims-based detection of CKD stages G3 and G4–5 performed worse, except for higher NPVs in CKD G4–5. This indicates an overall underdetection in claims-based research of specific CKD stages and limitations in differentiating CKD severity based on claims data, most likely due to the lack of laboratory values within claims data and the necessity of using narrower ICD-10-GM diagnosis-based definitions. Of note, PPVs were overall lowest in CKD G4–5 (0.47–0.57), although CKD G4–5 prevalence was very low in this cohort which may limit generalizability of the results. Still, this indicates many false-positive cases in this subgroup, with more than half CKD G4–5 patients showing N18.3 diagnoses (52–74%). Considering the continuous decrease of eGFR with age,⁵⁰ a strict cut-off based definition for CKD stages may also result in overestimation of CKD severity, especially in older age. Surprisingly, we detected a high share of persons with unspecific diagnoses of 37–43% in CKD G3–5 and 69–79% in G4–5 which calls for a more accurate CKD coding in clinical practice, especially in patients with advanced CKD.

Strengths and Limitations

Strengths of this study are the comprehensive use of a well-phenotyped cohort with detailed longitudinal measures of kidney function. We described the diagnostic validity of CKD within a community-dwelling cohort with an age range that becomes increasingly relevant against the background of demographic ageing. Due to longitudinal follow-ups, we were able to investigate trends over time from before and after the KDIGO guideline implementation. We compared different CKD definitions across various strata and subgroups and acknowledged the issue of analytical flexibility for observation time in claims data analyses.

Several limitations must be taken into account. First, the main analysis was based on only one eGFR measurement as no timely second confirmatory measures were available by design of the BIS. We conducted additional analyses including the eGFR of two consecutive study visits focusing on the chronicity of impaired kidney function and were able to confirm the results on trends over time. Second, low prevalence of CKD stages G4–5 in the BIS cohort may limit the generalisability of the results for this subgroup, and further studies including patients with more severe kidney disease are needed. Third, we only used data from one cohort study which might inhibit interpretability of general trends over time due to possible cohort effects. Selection biases due to the study region in Berlin, inclusion of only one health insurance fund, willingness for study participation, and missing serum creatinine values cannot be ruled out. Fourth, we

based the reference standard definition of CKD solely on eGFR values. Including data for albuminuria and applying a combined CKD definition in accordance with KDIGO guidelines may be subject to further research.

Conclusions

This study provides novel insights into the diagnostic validity of CKD in German health claims data for older community-dwelling individuals over time. The results suggest that German claims data show agreeable performance in CKD detection in older ages compared to other countries but may be limited in differentiating between CKD stages. An overall increase in diagnostic validity over time may also be attributable to the implementation of the KDIGO guidelines and improved diagnostic validity in clinical practice. Still, CKD prevalence may be underestimated when using only claims data due to limited sensitivity. In addition to a future linkage to patient's laboratory data, further studies to improve the understanding and use of claims-based research for CKD and other diseases using machine-learning algorithms to enhance diagnostic validity should be fostered.

Data Sharing Statement

The data used in this study cannot be made available in the manuscript, the [Supplementary Files](#), or in a public repository due to German data protection laws (Bundesdatenschutzgesetz). To facilitate the replication of results, the used data will be stored on a secure drive at Charité – Universitätsmedizin Berlin. Access to the raw data used in this study can only be provided to external parties under the conditions of a cooperation contract and can be accessed upon request, after written approval (bis@charite.de), if required.

Acknowledgments

We greatly thank the *AOK Nordost – Die Gesundheitskasse* for providing the health claims data of the BIS participants and supporting this study.

Author Contributions

TB, NE, ES, and AD conceptualised and designed the study and study protocol. Data were provided by ES, NE, NM, and VW. TB carried out the data analysis supported by AKF and AP. TB, NE, AKF, and AP interpreted the data. The manuscript was drafted by TB, substantially revised by NE, ES, and AD, and critically reviewed by AKF, AP, NM, CV, MHB, and VW. TB and AKF have accessed and verified the study data. All authors gave final approval of the version to be published and agreed on the journal to which the article was submitted. All authors agree to take responsibility and be held accountable for the contents of the article.

Funding

This research was partly funded by the Federal Joint Committee's innovation fund (project GUIDAGE-CKD; funding code 01VSF20020). The funding institution was not involved in planning, conducting, interpreting, or publishing of this study.

Disclosure

NE received grants by Bayer AG. ES receives honorarium from AstraZeneca for counseling and from the National Kidney Foundation for editorial work for the American Journal of Kidney Diseases. All other authors have nothing to declare for this work.

References

1. Mues KE, Liede A, Liu J, et al. Use of the Medicare database in epidemiologic and health services research: a valuable source of real-world evidence on the older and disabled populations in the US. *Clin Epidemiol*. 2017;9:267–277. doi:10.2147/CLEP.S105613
2. Thompson AM, Southworth MR. Real world data and evidence: support for drug approval: applications to kidney diseases. *Clin J Am Soc Nephrol*. 2019;14(10):1531–1532. doi:10.2215/CJN.02790319
3. Corrigan-Curay J, Sacks L, Woodcock J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. *JAMA*. 2018;320(9):867–868. doi:10.1001/jama.2018.10136

4. Yang C, Yang Z, Wang J, et al. Estimation of prevalence of kidney disease treated with dialysis in China: a study of insurance claims data. *Am J Kidney Dis.* 2021;77(6):889–897. doi:10.1053/j.ajkd.2020.11.021
5. Krishnamurthy S, Ks K, Dovgan E, et al. Machine learning prediction models for chronic kidney disease using national health insurance claim data in Taiwan. *Healthcare.* 2021;9(5):546. doi:10.3390/healthcare9050546
6. Krishnan M, Weinhandl ED, Jackson S, Gilbertson DT, Lacson E. Comorbidity ascertainment from the ESRD Medical Evidence Report and Medicare claims around dialysis initiation: a comparison using US Renal Data System data. *Am J Kidney Dis.* 2015;66(5):802–812. doi:10.1053/j.ajkd.2015.04.015
7. Gandjour A, Armsen W, Wehmeyer W, Multmeier J, Tschulena U, Alam K. Costs of patients with chronic kidney disease in Germany. *PLoS One.* 2020;15(4):e0231375. doi:10.1371/journal.pone.0231375
8. van Oosten MJM, Logtenberg SJJ, Leegte MJH, et al. Age-related difference in health care use and costs of patients with chronic kidney disease and matched controls: analysis of Dutch health care claims data. *Nephrol Dial Transplant.* 2020;35(12):2138–2146. doi:10.1093/ndt/gfz146
9. Kolbrink B, Schussel K, von Samson-Himmelstjerna FA, et al. Patient-focused outcomes after initiation of dialysis for ESRD: mortality, hospitalization, and functional impairment. *Nephrol Dial Transplant.* 2023;38(11):2528–2536. doi:10.1093/ndt/gfad099
10. Levey AS, Coresh J. Chronic kidney disease. *Lancet.* 2012;379(9811):165–180. doi:10.1016/S0140-6736(11)60178-5
11. Tonelli M, Wiebe N, Guthrie B, et al. Comorbidity as a driver of adverse outcomes in people with chronic kidney disease. *Kidney Int.* 2015;88(4):859–866. doi:10.1038/ki.2015.228
12. Ishigami J, Grams ME, Chang AR, et al. CKD and risk for hospitalization with infection: the Atherosclerosis Risk in Communities (ARIC) study. *Am J Kidney Dis.* 2017;69(6):752–761. doi:10.1053/j.ajkd.2016.09.018
13. Wong E, Ballew SH, Daya N, et al. Hospitalization risk among older adults with chronic kidney disease. *Am J Nephrol.* 2019;50(3):212–220. doi:10.1159/000501539
14. Su G, Iwagami M, Qin X, et al. Kidney disease and mortality in patients with respiratory tract infections: a systematic review and meta-analysis. *Clin Kidney J.* 2021;14(2):602–611. doi:10.1093/ckj/sfz188
15. Bowe B, Xie Y, Li T, et al. Changes in the US burden of chronic kidney disease from 2002 to 2016: an analysis of the Global Burden of Disease study. *JAMA Network Open.* 2018;1(7):e184412. doi:10.1001/jamanetworkopen.2018.4412
16. Chen TK, Knicely DH, Grams ME. Chronic kidney disease diagnosis and management: a review. *JAMA.* 2019;322(13):1294–1304. doi:10.1001/jama.2019.14745
17. Zhang QL, Rothenbacher D. Prevalence of chronic kidney disease in population-based studies: systematic review. *BMC Public Health.* 2008;8(1):117. doi:10.1186/1471-2458-8-117
18. Bowling CB, Sharma P, Fox CS, O'Hare AM, Muntner P. Prevalence of reduced estimated glomerular filtration rate among the oldest old from 1988-1994 through 2005-2010. *JAMA.* 2013;310(12):1284–1286. doi:10.1001/jama.2013.252441
19. Grams ME, Chow EK, Segev DL, Coresh J. Lifetime incidence of CKD stages 3-5 in the United States. *Am J Kidney Dis.* 2013;62(2):245–252. doi:10.1053/j.ajkd.2013.03.009
20. Vlasschaert ME, Bejimal SA, Hackam DG, et al. Validity of administrative database coding for kidney disease: a systematic review. *Am J Kidney Dis.* 2011;57(1):29–43. doi:10.1053/j.ajkd.2010.08.031
21. Grams ME, Plantinga LC, Hedegeman E, et al. Validation of CKD and related conditions in existing data sets: a systematic review. *Am J Kidney Dis.* 2011;57(1):44–54. doi:10.1053/j.ajkd.2010.05.013
22. Fleet JL, Dixon SN, Shariff SZ, et al. Detecting chronic kidney disease in population-based administrative databases using an algorithm of hospital encounter and physician claim codes. *BMC Nephrol.* 2013;14(1):81. doi:10.1186/1471-2369-14-81
23. Ronksley PE, Tonelli M, Quan H, et al. Validating a case definition for chronic kidney disease using administrative data. *Nephrol Dial Transplant.* 2012;27(5):1826–1831. doi:10.1093/ndt/gfr598
24. 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.* 2021;14(6):1586–1593. doi:10.1093/ckj/sfaa167
25. Weckmann G, Wirkner J, Kasbohm E, et al. Monitoring and management of chronic kidney disease in ambulatory care - analysis of clinical and claims data from a population-based study. *BMC Health Serv Res.* 2022;22(1):1330. doi:10.1186/s12913-022-08691-y
26. Muntner P, Gutierrez OM, Zhao H, et al. Validation study of medicare claims to identify older US adults with CKD using the Reasons for Geographic and Racial Differences in Stroke (REGARDS) Study. *Am J Kidney Dis.* 2015;65(2):249–258. doi:10.1053/j.ajkd.2014.07.012
27. Paik JM, Paterno 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(4):467–475. doi:10.1002/pds.5398
28. 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;7:1–13. doi:10.1177/2054358120959908
29. Eknayan G, Lameire N, Eckardt K, et al. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl.* 2013;3(1):5–14. doi:10.1038/kisup.2012.77
30. Schaeffner ES, van der Giet M, Gaedeke J, et al. The Berlin Initiative Study: the methodology of exploring kidney function in the elderly by combining a longitudinal and cross-sectional approach. *Eur J Epidemiol.* 2010;25(3):203–210. doi:10.1007/s10654-010-9424-x
31. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ.* 2015;351:h5527. doi:10.1136/bmj.h5527
32. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med.* 2009;150(9):604–612. doi:10.7326/0003-4819-150-9-200905050-00006
33. Cochran WG. Some methods for strengthening the common χ^2 tests. *Biometrics.* 1954;10(4):417–451. doi:10.2307/3001616
34. Armitage P. Tests for linear trends in proportions and frequencies. *Biometrics.* 1955;11(3):375–386. doi:10.2307/3001775
35. Douros A, Schneider A, Ebert N, et al. Control of blood pressure in older patients with heart failure and the risk of mortality: a population-based prospective cohort study. *Age Ageing.* 2021;50(4):1173–1181. doi:10.1093/ageing/afaa261
36. Douros A, Tolle M, Ebert N, et al. Control of blood pressure and risk of mortality in a cohort of older adults: the Berlin Initiative Study. *Eur Heart J.* 2019;40(25):2021–2028. doi:10.1093/eurheartj/ehz071

37. Stolpe S, Scholz C, Stang A, et al. Eine chronische Niereninsuffizienz, auch in höherem Stadium, ist Patienten häufig unbekannt - aber warum wissen Frauen noch seltener von ihrer Erkrankung als Männer? [High patient unawareness for chronic kidney disease even in later stages - but why is it more frequent in women than in men?] *Dtsch Med Wochenschr.* **2022**;147(17):e70–e81. German. doi:10.1055/a-1819-0870
38. Hodlmoser S, Winkelmayer WC, Zee J, et al. Sex differences in chronic kidney disease awareness among US adults, 1999 to 2018. *PLoS One.* **2020**;15(12):e0243431. doi:10.1371/journal.pone.0243431
39. Kim DH, Schneeweiss S, Glynn RJ, et al. Measuring frailty in Medicare data: development and validation of a claims-based frailty index. *J Gerontol a Biol Sci Med Sci.* **2018**;73(7):980–987. doi:10.1093/gerona/glx229
40. Desai RJ, Lin KJ, Paterno E, et al. Development and preliminary validation of a Medicare claims-based model to predict left ventricular ejection fraction class in patients with heart failure. *Circ Cardiovasc Qual Outcomes.* **2018**;11(12):e004700. doi:10.1161/CIRCOUTCOMES.118.004700
41. Mansour O, Paik JM, Wyss R, et al. A novel chronic kidney disease phenotyping algorithm using combined electronic health record and claims data. *Clin Epidemiol.* **2023**;15:299–307. doi:10.2147/CLEP.S397020
42. Warnock DG, Muntner P, McCullough PA, et al. Kidney function, albuminuria, and all-cause mortality in the REGARDS (Reasons for Geographic and Racial Differences in Stroke) study. *Am J Kidney Dis.* **2010**;56(5):861–871. doi:10.1053/j.ajkd.2010.05.017
43. Gansevoort RT, Matsushita K, van der Velde M, et al. Lower estimated GFR and higher albuminuria are associated with adverse kidney outcomes. A collaborative meta-analysis of general and high-risk population cohorts. *Kidney Int.* **2011**;80(1):93–104. doi:10.1038/ki.2010.531
44. Kühn A, van der Giet M, Kuhlmann MK, et al. Kidney function as risk factor and predictor of cardiovascular outcomes and mortality among older adults. *Am J Kidney Dis.* **2021**;77(3):386–396. doi:10.1053/j.ajkd.2020.09.015
45. Matsushita K, van der Velde M, Astor BC, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *Lancet.* **2010**;375(9731):2073–2081. doi:10.1016/S0140-6736(10)60674-5
46. Lee M, Saver JL, Chang KH, et al. Low glomerular filtration rate and risk of stroke: meta-analysis. *BMJ.* **2010**;341:c4249. doi:10.1136/bmj.c4249
47. D'Angelo K, Paul S, Ranjeeta F, et al. Awareness and knowledge among internal medicine house-staff for dose adjustment of cardiovascular drugs in chronic kidney disease. *High Blood Press Cardiovasc Prev.* **2021**;28(2):177–184. doi:10.1007/s40292-021-00438-w
48. Karsch-Volk M, Schmid E, Wagenpfeil S, et al. Kidney function and clinical recommendations of drug dose adjustment in geriatric patients. *BMC Geriatr.* **2013**;13(1):92. doi:10.1186/1471-2318-13-92
49. Matzke GR, Aronoff GR, Atkinson AJ, et al. Drug dosing consideration in patients with acute and chronic kidney disease-A clinical update from Kidney Disease: improving Global Outcomes (KDIGO). *Kidney Int.* **2011**;80(11):1122–1137. doi:10.1038/ki.2011.322
50. Schaeffner ES, Ebert N, Kuhlmann MK, et al. Age and the course of GFR in persons aged 70 and above. *Clin J Am Soc Nephrol.* **2022**;17(8):1119–1128. doi:10.2215/CJN.16631221

Clinical Epidemiology

Dovepress

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>