Original Clinical Research Quantitative

Agreement Between Administrative Database and Medical Chart Review for the Prediction of Chronic Kidney Disease G category

Louise Roy[1], Michael Zappitelli[2], Brian White-Guay[3] (iD),
Jean-Philippe Lafrance[4], Marc Dorais[5], and Sylvie Perreault[6] (iD)

Abstract

Background: Chronic kidney disease (CKD) is a major health issue and cardiovascular risk factor. Validity assessment of administrative data for the detection of CKD in research for drug benefit and risk using real-world data is important. Existing algorithms have limitations and we need to develop new algorithms using administrative data, giving the importance of drug benefit/risk ratio in real world.

Objective: The aim of this study was to validate a predictive algorithm for CKD GFR category 4-5 (eGFR < 30 mL/min/1.73 m^2 but not receiving dialysis or CKD G4-5ND) using the administrative databases of the province of Quebec relative to estimated glomerular filtration rate (eGFR) as a reference standard.

Design: This is a retrospective cohort study using chart collection and administrative databases.

Setting: The study was conducted in a community outpatient medical clinic and pre-dialysis outpatient clinic in downtown Montreal and rural area.

Patients: Patient medical files with at least 2 serum creatinine measures (up to 1 year apart) between September 1, 2013, and June 30, 2015, were reviewed consecutively (going back in time from the day we started the study). We excluded patients with end-stage renal disease on dialysis. The study was started in September 2013.

Measurement: Glomerular filtration rate was estimated using the CKD Epidemiological Collaboration (CKD-EPI) from each patient's file. Several algorithms were developed using 3 administrative databases with different combinations of physician claims (diagnostics and number of visits) and hospital discharge data in the 5 years prior to the cohort entry, as well as specific drug use and medical intervention in preparation for dialysis in the 2 years prior to the cohort entry.

Methods: Chart data were used to assess eGFR. The validity of various algorithms for detection of CKD groups was assessed with sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

Results: A total of 434 medical files were reviewed; mean age of patients was 74.2 ± 10.6 years, and 83% were older than 65 years. Sensitivity of algorithm #3 (diagnosis within 2-5 years and/or specific drug use within 2 years and nephrologist visit ≥ 4 within 2-5 years) in identification of CKD G4-5ND ranged from 82.5% to 89.0%, specificity from 97.1% to 98.9% with PPV and NPV ranging from 94.5% to 97.7% and 91.1% to 94.2%, respectively. The subsequent subgroup analysis (diabetes, hypertension, and <65 and ≥ 65 years) and also the comparisons of predicted prevalence in a cohort of older adults relative to published data emphasized the accuracy of our algorithm for patients with severe CKD (CKD G4-5ND).

Limitations: Our cohort comprised mostly older adults, and results may not be generalizable to all adults. Participants with CKD without 2 serum creatinine measurements up to 1 year apart were excluded.

Conclusions: The case definition of severe CKD G4-5ND derived from an algorithm using diagnosis code, drug use, and nephrologist visits from administrative databases is a valid algorithm compared with medical chart reviews in older adults.

Abrégé

Contexte: L'insuffisance rénale chronique (IRC) est un problème de santé majeur et un facteur de risque cardiovasculaire. La validité de la détection de l'IRC à partir des bases de données administratives est importante pour les études évaluant en situation réelle les bénéfices et les risques des médicaments. Les algorithmes existants comportent des limites et, compte tenu de l'importance revêtue par ce rapport bénéfices/risques, le développement de nouveaux algorithmes utilisant les bases de données administratives s'avère essentiel.

Objectif: Valider le pouvoir prédictif d'un algorithme pour détecter l'insuffisance rénale chronique sévère (DFGe $<$30 mL/min/1.73 m^2, patient non-dialysé ou CKD G4-5ND) à partir des banques de données administratives de la province de Québec, avec le débit de filtration glomérulaire estimé (DFGe) comme point de référence.

Type d'étude: Étude de cohorte rétrospective réalisée à partir des dossiers médicaux et de données administratives.

Cadre: Des cliniques médicales communautaires et de protection rénale de Montréal et des régions rurales périphériques.

Sujets: Les dossiers médicaux de patients avec au moins deux mesures de la créatinine sérique (en moins d'un an) entre le 1er septembre 2013 et le 30 juin 2015 ont été revus consécutivement, en reculant dans le temps. Les patients avec insuffisance rénale terminale et dialysés ont été exclus. L'étude a débuté en septembre 2013.

Mesures: Le DFG a été estimé à l'aide de la formule CKD Epidemiological Collaboration (CKD-EPI) à partir du dossier médical de chaque patient. Nous avons développé différents algorithmes en utilisant trois banques de données administratives avec différentes combinaisons de facturations médicales (diagnostics et nombre de visites en néphrologie) et de données colligées au congé de l'hôpital dans les cinq ans précédant l'entrée dans la cohorte, de même qu'avec la consommation de certains médicaments et les interventions médicales subies en préparation à la dialyse dans les deux ans précédant l'entrée dans la cohorte.

Méthodologie: Les données des dossiers médicaux ont été utilisées pour définir le DFGe. La validité des algorithmes développés a été évaluée en utilisant la sensibilité, la spécificité, la valeur prédictive positive (VPP) et la valeur prédictive négative (VPN).

Résultats: En tout, 434 dossiers médicaux ont été revus; l'âge moyen des patients était de 74.2 $\pm$ 10.6 ans et 83% avaient plus de 65 ans. La sensibilité de l'algorithme no.3 (diagnostic dans un délai de 2 à 5 ans et/ou l'usage de médicaments spécifiques dans un délai de 2 ans, et au moins quatre visites médicales en néphrologie dans les 2 à 5 ans précédant la date d'entrée dans la cohorte) dans l'identification d'une insuffisance rénale sévère (CKD G4-5ND) variait de 82.5% à 89.0%. La spécificité de ce même algorithme variait de 97.1% à 98.9% avec une PPV et une NPV allant respectivement de 94.5.% à 97.7% et de 91.1% à 94.2%. L'analyse de sous-groupes (patients diabétiques, hypertendus, âgés de moins de 65 ans ou âgés de 65 ans et plus) ainsi que la comparaison de la prévalence prédite dans une cohorte de patients âgés par rapport aux données de la littérature font valoir la précision de notre algorithme pour les patients avec insuffisance rénale sévère (CKD G4-5ND).

Limites: Notre cohorte était composée essentiellement de sujets âgées, les résultats pourraient ne pas s'appliquer à tous les adultes. Les patients n'ayant pas eu deux mesures de la créatinine sérique à l'intérieur d'un an ont été exclus.

Conclusion: Chez les personnes âgées, la définition de cas pour une insuffisance chronique rénale sévère (CKD G4-5ND) estimée par un algorithme utilisant les codes diagnostic, la consommation de médicaments spécifiques et les services médicaux de néphrologie tirés des données administratives s'avère un algorithme valide comparativement à l'examen du dossier médical.

What was known before

Chronic kidney disease (CKD) is a major health issue and cardiovascular risk factor, and its detection in an administrative databank is important in research on drug benefit and risk using real-world data. Algorithms using physician claim for CKD have low predictive positive value.

What this adds

We designed several algorithms to define adult CKD with administrative health data, using physician claims, number of visits to a nephrologist, and specific drugs. We then compared these algorithms with patients' medical chart and the reference standard of estimated glomerular filtration rate

[1]Faculty of Medicine, University of Montreal, University of Montreal Hospital Center, QC, Canada

[2]Faculty of Medicine, Department of Pediatrics, Pediatric Nephrology, Toronto Hospital for Sick Children, University of Toronto, ON, Canada

[3]Faculty of Medicine, University of Montreal, QC, Canada

[4]Faculty of Medicine, Department of Pharmacology and Physiology, University of Montreal, QC, Canada

[5]StatSciences Inc., Notre-Dame-de-l'Île-Perrot, QC, Canada

[6]Faculty of Pharmacy, University of Montreal, QC, Canada

Corresponding Author:

Sylvie Perreault, Chaire Sanofi sur l'utilisation des médicaments, Faculté de Pharmacie, Université de Montréal, Case Postale 6128, Succursale Centre-Ville, Montréal, QC, Canada H3C 3J7.

Email: sylvie.perreault@umontreal.ca

(eGFR). Our algorithm (#3) has an excellent specificity and positive predictive value to detect severe kidney failure (CKD G4-5ND), which is the subgroup of CKD patients more at risk and hence more of interest.

Introduction

Chronic kidney disease (CKD) is an important public health burden associated with increased morbidity, mortality, and substantial health care costs worldwide.[1-5] Approximately 11% of the adult population and 25% of individuals >70 years of age have CKD G3-5ND[6] in North America[7-9]; CKD is an important clinical endpoint in various medical conditions such as diabetes, hypertension, cardiovascular disease, and use of certain drugs; and it is also a risk factor for cardiovascular disease and death[10,11] and larger use of health care resources.[12-14] Detection and management of CKD have a significant impact by reducing the incidence of cardiovascular disease,[1,15-18] the rate of progression of kidney function,[19-24] as well as the rate of adverse events by optimizing drug management and health care costs.

Measuring serum creatinine and estimating GFR are recommended in all patients with any risk factor for CKD (Canadian Society of Nephrology guidelines and predictive model of CKD).[25] Following initial evaluation, if CKD is detected, routine evaluation of GFR is the standard of care. In pharmacoepidemiologic studies at the population level, databanks are a central tool but they often miss specific clinical data (eg, BP measurements) and lab results (eg, creatinine). Whether it is for cardiovascular assessment or mortality risk factors, as a clinical endpoint in specific diseases or conditions, or simply as justification for drug use and dose, identifying CKD is a very valuable addition to any pharmacoepidemiologic study regarding cardiovascular morbidity and mortality, hypertension, diabetes, or drug use.[26]

Two systematic reviews have recently assessed the validity of existing data sources to identify CKD[27,28] and showed major discrepancies in sensitivity values ranging from 3% to 88%. In addition, most of the studies included in those reviews had some transferability flags, such as a lack of a valid reference standard or the development of algorithms without consideration for the period of time, number of codes or medical services, and/or specific drug uses to define disease. Studies of CKD validation have reported that administrative databases are not recommended for CKD surveillance but may be a useful tool when an algorithm with high specificity is required, such as in pharmacoepidemiologic research.[29,30] We aimed to determine the validity of a more accurate algorithm derived from administrative data (Quebec, Canada) for identifying severe CKD (G4-5) compared with the reference standard of estimated glomerular filtration rate (eGFR).

Materials and Methods

Design, Setting, and Patients

This is a retrospective diagnostic accuracy study of administrative data using a cohort of patients followed in 2 community outpatient medical clinics in Montreal (CMFU-Notre-Dame in downtown Montreal) and Valleyfield (Group of Familial Medicine Medival) and 2 pre-dialysis clinics in downtown Montreal and Valleyfield, Quebec, Canada. Medical files of patients 23 years and older receiving follow-up care in one of these clinics, with at least 2 serum creatinine measures (up to 1 year apart) between September 1, 2013, and June 31, 2015, were studied consecutively (going back in time from the day we started the study). The date of cohort entry was the date of the first eGFR during the period of 2013 to 2015. Patients had to be insured by the *Régie de l'Assurance Maladie du Québec* (RAMQ) drug plan for at least 2 years prior to the cohort entry. We collected the administrative data from RAMQ medical services and Med-Echo for data on hospitalizations for the last 5 years prior to the cohort entry.

Patients treated with peritoneal dialysis or hemodialysis in the 3 months prior to the date of cohort entry were excluded.[29] In addition, to reduce the impact of possible episodes of acute kidney injury, laboratory measurements associated with hospital admission were also excluded. The selection of the study population is shown in Figure 1. We obtained approvals from institutional research ethics boards of the *Centre Hospitalier de l'Université de Montréal* (CHUM) and the *Commission d'Accès à l'Information du Québec* (CAI, provincial ethics body), as well as approval to waive requirement for patient consent.

Data Collection and Sources

Baseline patient characteristics and treatment were collected by retrospective chart review. These data were de-identified and merged with the administrative health databases (RAMQ and Med-Echo) from September 1, 2013, to June 30, 2015. The administrative records of hospitalization and medical services were provided in the 5-year period prior to the cohort entry, and the pharmaceutical files in the 2-year period prior to cohort entry.

The administrative health databases contain information about patient demographics, inpatient and outpatient International Classification of Disease (ICD-9 and ICD-10) diagnostic codes, and the physician claim database; however, no lab results are available, and were therefore retrieved from individual chart review. The acute care hospitalization data include admission and discharge dates, primary diagnosis, physician information, procedures, up to 18 secondary diagnosis (ICD-9/10) codes, and the length of stay. The physician database contains information on physician services such as dates and location of the visits, diagnostic code
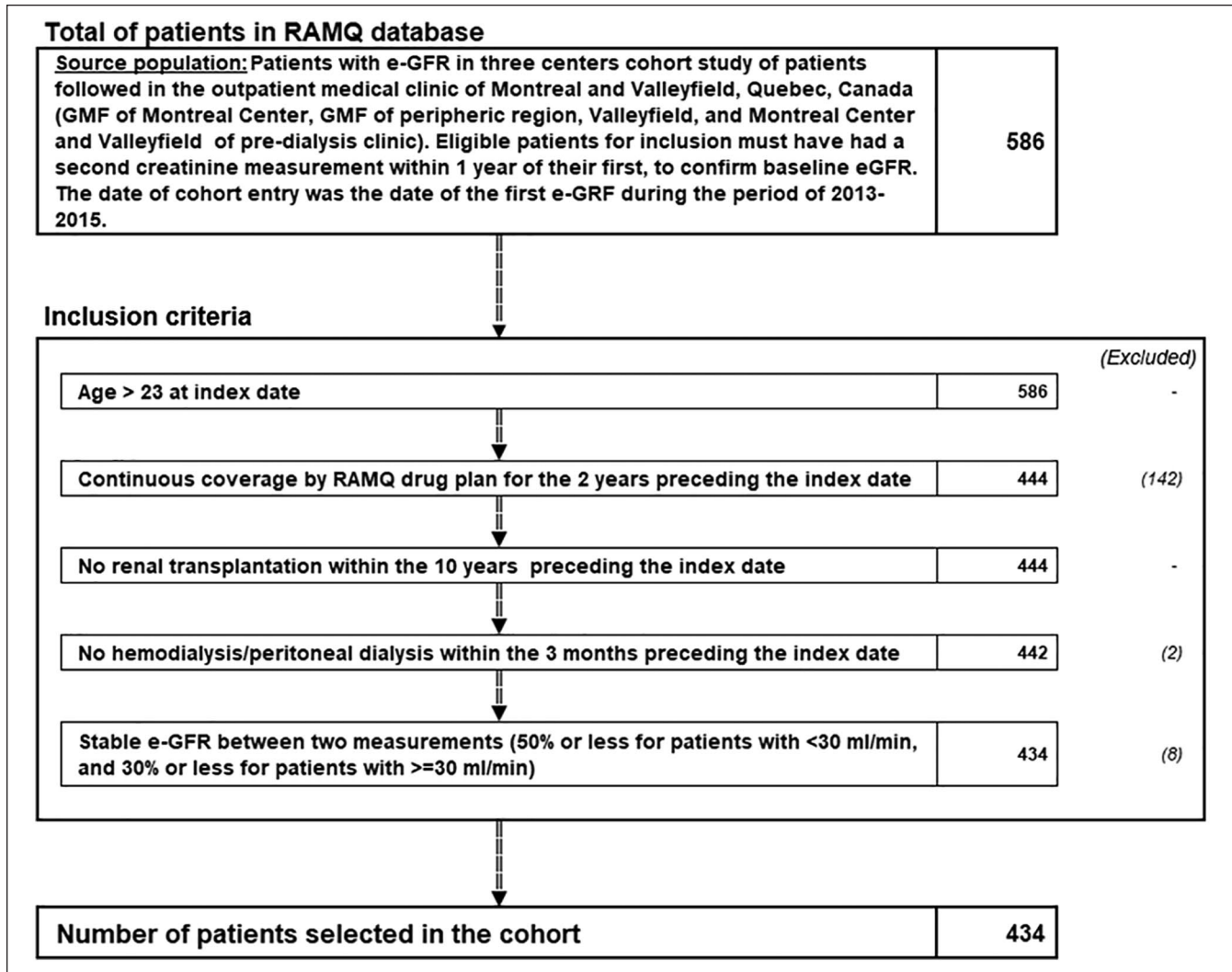
Figure 1. Flow chart of the study population.

(ICD-9), and provider specialty. The pharmaceutical database contains outpatient prescription information on patients with provincial medication insurance, representing more than 95% of the older adult population.

The RAMQ and Med-Echo databases have been used extensively to perform pharmacoepidemiologic studies.[31-36] Data recorded in RAMQ prescription files (outpatient only) have been evaluated and found to be comprehensive and valid,[37] as were medical diagnoses in the Med-Echo database.

Assessment of Kidney Function and Defining CKD

We estimated eGFR using the CKD Epidemiological Collaboration (CKD-EPI).[38] Based on the eGFR, CKD was classified as CKD G3 (<60 mL/min/1.73 m^2) and CKD G4-5 (<30 mL/min/1.73 m^2). The date of cohort entry was the date of the first eGFR for classification of CKD. The

CKD-EPI creatinine equation is the accurate method for estimating GFR for diverse populations.[38-41]

Administrative Data to Define CKD

Using the unique provincial health insurance identifier, all patient files were linked to the administrative databases. To identify relevant ICD-9 and ICD-10 codes to define CKD, a detailed review of the literature was performed. Eight articles, all in adult populations, and one systematic review on the validity of administrative database coding for kidney disease were found.[21,29,35,42-47] Based on these studies and expert opinion, we selected diagnostic codes and assessed the frequency at which these codes appeared within the physician claims database and Med-Echo database among patients with CKD (Table S3). We identified the codes with the highest frequencies to be ICD-9 585, 403, or 404 and ICD-10 N18, I12, or I13, which we then used to define the algorithms

for CKD (Table S4). The resulting algorithms thus defined CKD for each patient using administrative data with different combinations of physician claims and hospital discharge data within the 5 years prior to the cohort entry. We first defined CKD G 3-5ND with algorithm 1 and 2 (with specific medications); then CKD G 4-5ND with algorithm 1 and 2 with the inclusion of ≥ 4 outpatient medical visits to a nephrologist for algorithm 3; then CKD G5ND with algorithm 4. By elimination, patients who were not classified in the G 3-5ND group using the algorithms above were automatically classified in the CKD G 1-2.

Algorithm 1 (diagnostic codes only): (1) one physician claim or one hospital discharge as primary or secondary diagnosis within 2, 3, and 5 years; (2) two physician claims or one hospital discharge within 2, 3, and 5 years; and (3) three physician claims or one hospital discharge within 2, 3, and 5 years.

Algorithm 2 (diagnostic codes and/or use of a specific drug for CKD): algorithm 1 with the addition of specific medications and doses used in CKD. The outpatient medications included in the definition were selected based on previous research and expert opinion.[35] As medications can be used for indications other than CKD, we included strict parameters in the algorithm to maximize specificity. Specifically, we included users of medications including carbonate calcium (≥1500 mg daily), and/or furosemide (≥20 mg daily), and/or specific dosage of calcitriol, alfacalcidol, doxercalciferol, and/or any dosage of sevelamer, lanthanum, cinacalcet, darbepoetine, or erythropoietin in the 2 years prior to the index date (Table S5).

Algorithm 3 (diagnostic and/or use of a specific drug for CKD and ≥4 nephrology visits): algorithm 2 with the inclusion of ≥4 outpatient medical visits to a nephrologist within 2, 3, and 5 years prior to the cohort entry.

Algorithm 4 (diagnostic codes and/or use of a specific drug for CKD and nephrology visit ≥4 or medical procedures): algorithm 3 for CKD G4-5ND with the addition of either nephrology visits within 2, 3, and 5 years prior to the cohort entry or the presence of medical procedures in preparation for peritoneal dialysis or hemodialysis, or duplex ultrasound of forearms in the 2 years prior to the index date.

Subgroups Analysis

We assessed the case definitions across different subgroups in our cohort, defined by administrative databases, as either older or younger than 65 years, gender, presence or absence of diabetes, and presence or absence of hypertension among patients with G4-5ND. These 3 cohorts represent subgroups of patients particularly at risk for CKD.

Comparison With a Cohort of Older Adults

Using algorithm 3 for a 5-year period case definition, we proceeded to assess the predicted prevalence of CKD G4-5ND among a cohort of older adults based on a 40% random sample of individuals in the province of Quebec for the period of January 2010 to December 2015, compared with literature data. We evaluated the prevalence of CKD G4-5ND among age groups including 66-69, 70-74, 75-79, and ≥80 years for men and women of the total cohort, with 2 further subgroups of patients having a diagnosis of diabetes, and those with a diagnosis of chronic heart failure, both groups of patients at higher risk of CKD. We selected 2 subcohorts among the total cohort to assess the ascertainment of CKD G4-5ND among patient with diabetes using ICD-9 codes (ICD-9: 250.xx, 357.2x, 362.0x, 366.41/ICD-10: E8, E9, E10, E11, E13) and chronic heart failure (ICD-9 code "428.0, 428.1, 428.9" or ICD-10 code I50.0, I50.1, I50.9) in the 5-year period prior to the cohort entry (Table S4).[48,49]

Statistical Analyses

Descriptive statistics of the population were stratified by eGFR using 2 algorithms. We reviewed the medical records of patients selected in a community setting in 3 different Quebec centers over 2013 to 2015, to estimate the sensitivity, specificity, positive predictive values, and negative predictive values of the diagnostic CKD codes, medications used, and medical visits for nephrology using the Quebec administrative databases. Validity indices were estimated for each case definition combination using laboratory data as the reference standard.

Sensitivity was defined as the proportion of patients classified by the algorithm as having a given eGFR among all patients within this eGFR category in clinical charts. *Specificity* was defined as the proportion of patients classified by the algorithm as not having a given eGFR among all patients within this eGFR category as defined by medical charts. We defined *positive predictive value* (PPV) as the proportion of patients who were assigned a given eGFR in medical charts among all patients classified by the algorithm as being in the selected eGFR category. We defined *negative predictive value* (NPV) as the proportion of patients who were not assigned a given eGFR in medical charts among all patients classified by the algorithm as not being in the eGFR category. All analyses were planned a priori and conducted using SAS statistical software, release 9.4 (SAS Institute Inc., Cary, NC).

Sample Size Calculation

In this retrospective study, sample size was determined by considering the number of elements included in the algorithm, 10 patients per element, for each level of eGFR, and with an alpha error of 0.05.[50] The elements considered comprised medical visits, diagnosis of CKD, use of drugs, and medical interventions in preparation for dialysis. Patients' charts were reviewed consecutively up to the required number. A total of 434 medical files were studied, of which 154 participants had G4-5ND, including 41 from the pre-dialysis

clinics, 122 participants had G3, and 158 participants had G1-2.

Results

Patient Characteristics

A total of 434 patients met the inclusion criteria (Figure 1). Demographic and main clinical characteristics of these patients stratified by disease stage are shown in Table 1 (supplementary data are presented in Table S1). Mean age varied between 71.8 and 75.2 (74.2 ± 10.6) years, 83% of them being older than 65, and 48.1% to 63.4% of patients were female. When measured with administrative data in the 5 years prior to the index date, >90% of patients with G4-5ND had a diagnosis of CKD, while approximately 65% of patients with G3-5ND had such a diagnosis. Chronic kidney disease patients had a higher number of medical services compared with those without CKD (defined as G1-2). They also presented a higher prevalence of diabetes and hypertension compared with patients without CKD. In addition, the identification of diabetes and hypertension using administrative data closely resembled to the retrospective chart review.

Estimated Glomerular Filtration Rate Validation

The validity of administrative data in determining the presence of CKD compared with the reference standard (G4-5ND) varied across case definitions and length of administrative data observation (Table 2). Across the different algorithms tested, sensitivity ranged from 82.5% to 99.4%, specificity ranged from 76.1% to 98.9%, PPV ranged from 69.5% to 97.7%, and NPV ranged from 91.1% to 99.5%. Algorithm 2, using diagnosis and/or specific drug use, presented the least favorable validity results with respect to algorithms 1 and 3. On the contrary, algorithm 1, using diagnosis only, led to high estimates of sensitivity, specificity, and NPV (around >90%), where the PPV was a little lower (estimates ranging around >80%); and algorithm 3 led to similar estimates, but sensitivity estimates were lower. These results suggest that algorithm 1 favors the identification of true cases but increases chances of identifying false positives, while the addition of drug marker and nephrologist visits (algorithm 3) favors the identification of true positives by the algorithm, but also false negatives. To prioritize a higher specificity, algorithm 3 with its minimum 3-year observation period led to the most stable and optimal results.

Results of the validity of administrative data in determining the presence of severe CKD (G5ND) compared with the reference standard are shown in Table 3. The use of algorithm 4 led to estimates of sensitivity ranging from 85.4% to 87.8%, specificity ranging from 72.0% to 75.1%, PPV ranging from 24.7% to 26.3%, and NPV ranging from 97.7% to 98.3%. Although it is possible that the algorithm identifies many false positives, the low PPV estimates could be a result of the small sample size (n = 41).

The validity of administrative data in determining the presence of CKD compared with the reference standard group CKD (G3-5ND) is shown in Table S6. The use of algorithms 1 and 2 led to sensitivity estimates ranging from 57.2% to 72.1%, specificity ranging from 86.7% to 98.7%, PPV ranging from 90.5% to 98.8%, and NPV ranging from 59.1% to 64.0%. Algorithm 2 allowed a better identification of true cases (higher sensitivity estimates) compared with algorithm 1 at the detriment of specificity estimates, although they remained above 80%. Patients who were not classified as CKD G3-5ND using the algorithm were automatically classified in the CKD G1-2 group, with sensitivity of 86.8% (95% CI = 81.1-91.2), specificity of 70.7% (95% CI = 67.7-73.0), PPV of 61.2% (95% CI = 57.2-64.3), and NPV of 90.9% (95% CI = 87.0-93.9).

Subgroup Analysis of Predicted CKD G4-5NDn

As shown in Table 4, we compared case definitions to the reference standard G4-5ND, in subgroups according to age, sex, diabetes, or hypertension, and found high sensitivity, specificity, PPV, and NPV estimates.

Comparison of Predicted Prevalence of CKD 4-5NDn Among an Older Adult Cohort

Demographic and clinical characteristics according to gender of the selected cohort of older adults can be found in Table S2. The results of our case definition for CKD G4-5ND were stratified by age group and gender among older adult patients with additional distinctions for those with diabetes, and those with chronic heart failure, as shown in Table 5. Among men aged 66-69, 70-74, 75-79, and ≥80 years old, the predicted prevalence of CKD G4-5ND was 1.7%, 2.3%, 3.1%, and 4.3%, respectively; the corresponding values were 1.0%, 1.3%, 2.0%, and 2.5% for women of the same age groups, respectively.

The predicted prevalence of CKD G4-5ND among older men with diabetes was 4.2%, 5.0%, 5.9%, and 7.2% for age groups 66-69, 70-74, 75-79, and ≥80 years, respectively; and those estimates were 2.7%, 3.4%, 4.4%, and 4.5% for women of the same age groups, respectively.

The predicted prevalence of G4-5ND among older men with chronic heart failure was 9.6%, 11.7%, 11.2%, and 11.3% for age groups 66-69, 70-74, 75-79, and ≥80 years, respectively; and those estimates for similarly grouped women were 8.5%, 8.7%, 10.0%, and 7.5%, respectively.

Discussion

We assessed the validity of an algorithm in the Quebec (Canada) administrative databank (RAMQ) to detect severe

Table 1. Demographic and Clinical Characteristics of CKD Patients and Non-CKD Patients as Reference According to Chart Review and Administrative Databases.

	CKD G5ND		CKD G4-5ND		CKD G3-5ND		CKD G1-2	
CKD	Clinical data (n = 41) N (%)	Databases[a] (n = 41) N (%)	Clinical data (n = 154) N (%)	Databases[a] (n = 154) N (%)	Clinical data (n = 276) N (%)	Databases[a] (n = 276) N (%)	Clinical data (n = 158) N (%)	Databases[a] (n = 158) N (%)
Age (mean ± SD)	71.8 (13.0)		72.8 (11.8)		75.2 (11.1)		72.5 (9.4)	
Female (%)	26 (63.4)		74 (48.1)		143 (51.8)		93 (58.9)	
eGFR (first value) (mean ± SD)	11.6 ± 1.7		18.4 ± 5.4		30.9 ± 15.6		74.7 ± 11.2	
eGFR (second value) (mean ± SD)	13.6 ± 2.7		19.3 ± 5.7		32.2 ± 16.6		74.5 ± 12.4	
Chronic kidney disease 5-year prior index date (%)		38 (92.7)		149 (96.8)		181 (65.6)		6 (3.8)
Health care use 5 years prior index date								
Nephrology community visit (mean ± SD)	—	11.1 ± 8.5	—	11.3 ± 7.6	—	6.6 ± 7.9	—	0.04 ± 0.3
Nephrology community visits (median)	—	10	—	11	—	3	—	0
Peritoneal dialysis or hemodialysis procedures in the last 2 years (%)	—	9 (22.0)	—	15 (9.7)	—	16 (5.8)	—	0
Comorbidities 5 years prior index date (%)[b]								
Diabetes	27 (65.9)	—	99 (65.6)	—	148 (54.2)	—	72 (45.9)	—
DX in the last 5 years	—	29 (70.7)	—	108 (70.1)	—	151 (54.7)	—	63 (39.9)
Procedure in the last 5 years	—	20 (48.8)	—	75 (48.7)	—	100 (36.2)	—	27 (17.1)
RX in the last 2 years	—	26 (63.4)	—	95 (61.7)	—	137 (49.6)	—	68 (43.0)
DX or procedure or RX	—	29 (70.7)	—	108 (70.1)	—	157 (56.9)	—	72 (45.6)
Hypertension	40 (97.6)	—	144 (94.1)	—	248 (90.2)	—	122 (77.2)	—
DX in the last 5 years	—	38 (92.7)	—	127 (82.5)	—	206 (74.6)	—	86 (54.4)
RX in the last 2 years	—	38 (92.7)	—	145 (94.2)	—	249 (90.2)	—	124 (78.5)
DX or RX	—	40 (97.6)	—	150 (97.4)	—	256 (92.8)	—	128 (81.0)
Renal medication 2-year prior index date (%)[c]	—	32 (26.2)	—	116 (75.3)	—	53 (19.7)	—	17 (10.8)
Calcium carbonate (≥1500 mg/day)		5 (12.2)		18 (11.7)		18 (6.5)		0.04 ± 0.3
Calcitriol (yes vs no)		4 (9.8)		5 (3.3)		5 (1.8)		0
Sevelamer (yes vs no)		5 (12.2)		8 (5.2)		8 (2.9)		0
Doxercalciferol (yes vs no)		0		0		0		0
Alfacalcidol (yes vs no)		8 (19.5)		29 (18.8)		29 (10.5)		0
Cinacalcet (yes vs no)		0		0		0		0
Lanthanum (yes vs no)		1 (2.4)		1 (0.7)		1 (0.4)		0
Erythropoietin (yes vs no)[d]		0		0		0		0
Darbepoietin (yes vs no)[d]		18 (43.9)		48 (31.2)		48 (17.4)		0
Furosemide (mg/day)								
≥20 mg		29 (70.7)		101 (65.6)		13 (48.2)		17 (10.8)
≥40 mg		24 (58.5)		78 (50.7)		102 (37.0)		5 (3.2)
≥80 mg		17 (41.5)		47 (30.5)		60 (21.7)		0

Note. CKD = chronic kidney disease; eGFR = estimated glomerular filtration rate; DX = diagnostic; RX = medication; SD = standard deviation.

[a] Administrative databases are RAMQ/Med-Echo.

[b] Diagnosis definition with ICD-9/10 codes (ICD-9 585, 403, 404, ICD-10 N18, I12, I13) found in Table S4, drug markers found in Table S5.

[c] Specific dosages found in Table S5.

[d] No multiple myeloma.

Table 2. Validity of Case Definitions for CKD Compared with the Reference Standard of CKD G4-5ND.

Case definition	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Algorithm 1: diagnosis only within 2 to 5 years				
1 claim or 1 hospitalization in 2 years	95.5 (91.4-97.9)	90.7 (88.5-92.1)	85.0 (81.4-87.1)	97.3 (94.9-98.8)
1 claim or 1 hospitalization in 3 years	95.5 (91.3-97.9)	88.9 (86.7-90.3)	82.6 (79.0-84.7)	97.3 (94.8-98.7)
1 claim or 1 hospitalization in 5 years	96.8 (92.8-98.8)	86.4 (84.3-87.5)	79.7 (76.4-81.3)	98.0 (95.5-99.2)
2 claims or 1 hospitalization in 2 years	92.2 (87.8-95.3)	91.8 (89.4-93.5)	86.1 (82.0-89.0)	95.5 (93.0-97.3)
2 claim or 1 hospitalization in 3 years	93.5 (89.2-96.4)	90.4 (88.0-92.0)	84.2 (80.3-86.9)	96.2 (93.7-97.9)
2 claim or 1 hospitalization in 5 years	94.8 (90.6-97.4)	88.9 (86.6-90.4)	82.5 (78.8-84.8)	96.9 (94.4-98.5)
3 claims or 1 hospitalization in 2 years	89.0 (84.3-92.5)	91.8 (89.2-93.7)	85.6 (81.1-89.1)	93.8 (91.2-95.8)
3 claim or 1 hospitalization in 3 years	92.9 (88.5-95.9)	90.4 (87.9-92.0)	84.1 (80.1-86.9)	95.8 (93.3-97.6)
3 claim or 1 hospitalization in 5 years	94.2 (89.9-97.0)	89.3 (86.9-98.0)	82.9 (79.1-85.3)	96.5 (94.0-98.2)
Algorithm 2: diagnosis within 2 to 5 years and/or specific drug use within 2 years				
1 claim or 1 hospitalization in 2 years OR 2-year selected drugs	98.7 (95.2-99.8)	78.6 (76.7-79.2)	71.7 (69.2-72.5)	99.1 (96.7-99.8)
1 claim or 1 hospitalization in 3 years OR 2-year selected drugs	98.7 (95.2-99.8)	78.2 (76.3-78.8)	71.4 (68.8-72.1)	99.1 (96.7-99.8)
1 claim or 1 hospitalization in 5 years OR 2-year selected drugs	99.4 (96.1-99.9)	76.1 (74.3-76.4)	69.5 (67.3-70.0)	99.5 (97.2-99.9)
2 claims or 1 hospitalization in 2 years OR 2-year selected drugs	97.4 (93.5-99.2)	79.3 (77.1-80.2)	72.1 (69.2-73.4)	98.2 (95.6-99.4)
2 claim or 1 hospitalization in 3 years OR 2-year selected drugs	98.1 (94.3-99.5)	78.9 (76.9-79.7)	71.9 (69.2-73.0)	98.7 (96.1-99.7)
2 claim or 1 hospitalization in 5 years OR 2-year selected drugs	98.1 (94.3-99.5)	77.9 (75.8-78.6)	70.9 (68.2-71.9)	98.6 (96.1-99.6)
3 claims or 1 hospitalization in 2 years OR 2-year selected drugs	96.8 (92.7-98.8)	79.3 (77.1-80.4)	72.0 (69.0-73.5)	97.8 (95.0-99.2)
3 claim or 1 hospitalization in 3 years OR 2-year selected drugs	98.1 (94.3-99.5)	78.9 (76.9-79.7)	71.9 (69.2-73.0)	98.7 (96.1-99.7)
3 claim or 1 hospitalization in 5 years OR 2-year selected drugs	98.1 (94.3-99.5)	78.2 (76.2-79.0)	71.2 (68.5-72.3)	98.6 (96.1-99.6)
Algorithm 3: diagnosis within 2 to 5 years and/or specific drug use within 2 years and nephrologist visit ≥4 within 2 to 5 years				
1 claim or 1 hospitalization in 2 years OR 2-year selected drugs and visit in 2 years	83.1 (79.6-84.6)	98.9 (97.0-99.7)	97.7 (93.6-99.4)	91.4 (89.7-92.1)
1 claim or 1 hospitalization in 3 years OR 2-year selected drugs and visit in 3 years	88.3 (84.8-90.0)	98.6 (96.6-99.5)	97.1 (93.3-99.0)	93.9 (92.0-94.8)
1 claim or 1 hospitalization in 5 years OR 2-year selected drugs and visit in 5 years	89.0 (85.1-91.5)	97.1 (95.0-98.5)	94.5 (90.3-97.2)	94.1 (92.0-95.5)
2 claims or 1 hospitalization in 2 years OR 2-year selected drugs and visit in 2 years	82.5 (79.0-83.9)	98.9 (97.0-99.7)	97.7 (93.6-99.4)	91.1 (89.4-91.8)
2 claims or 1 hospitalization in 3 years OR 2-year selected drugs and visit in 3 years	88.3 (84.8-90.0)	98.6 (96.6-99.5)	97.1 (93.3-99.0)	93.9 (92.0-94.8)
2 claims or 1 hospitalization in 5 years OR 2-year selected drugs and visit in 5 years	89.0 (85.2-91.2)	97.9 (95.8-99.1)	95.8 (91.8-98.2)	94.2 (92.2-95.3)
3 claims or 1 hospitalization in 2 years OR 2-year selected drugs and visit in 2 years	83.1 (79.6-84.6)	98.9 (97.0-99.7)	97.7 (93.6-99.4)	91.4 (89.7-92.1)
3 claims or 1 hospitalization in 3 years OR 2-year selected drugs and visit in 3 years	88.3 (84.8-90.0)	98.6 (96.6-99.5)	97.1 (93.3-99.0)	93.9 (92.0-94.8)
3 claims or 1 hospitalization in 5 years OR 2-year selected drugs and visit in 5 years	89.0 (85.1-91.5)	97.1 (95.0-98.5)	94.5 (90.3-97.2)	94.1 (92.0-95.5)

Note. CI = confidence interval; CKD = chronic kidney disease; NPV = negative predictive value; OR = odds ratio; PPV = positive predictive value.

CKD compared with the reference standard of eGFR. The results show that algorithm 3 has a sensitivity ranging from 82.5% to 89.0%, specificity from 97.1% to 98.9%, PPV from 94.5% to 97.7%, and NPV from 91.4% to 94.2% for detection of CKD G4-5ND. The increasing validity measurement was highly dependent on the number of variables of administrative

Table 3. Validity of Case Definitions Compared to the Reference Standard of CKD G5ND.

Case definition	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Algorithm 4: diagnosis within 2 to 5 years and/or use of specific drug within 2 years and nephrologist visit ≥4 within 2 to 5 years or medical procedure within 2 years				
1 claim or 1 hospitalization in 2 years OR 2-year selected drugs and visits ≥4 in 2- or 2-year procedure	85.4 (71.0-93.8)	75.1 (73.6-75.9)	26.3 (21.9-28.9)	98.0 (96.0-99.2)
1 claim or 1 hospitalization in 3 years OR 2-year selected drugs and visits ≥4 in 3- or 2-year procedure	87.8 (73.7-95.4)	73.3 (71.8-74.1)	25.5 (21.4-27.7)	98.3 (96.3-99.4)
1 claim or 1 hospitalization in 5 years OR 2-year selected drugs and visits ≥4 in 5- or 2-year procedure	87.8 (73.7-95.4)	72.0 (70.5-72.8)	24.7 (20.7-26.8)	98.3 (96.3-99.3)
2 claims or 1 hospitalization in 2 years OR 2-year selected drugs and visits ≥4 in 2- or 2-year procedure	82.9 (68.2-92.2)	75.1 (73.5-76.0)	25.8 (21.2-26.8)	97.7 (95.7-98.9)
2 claims or 1 hospitalization in 3 years OR 2-year selected drugs and visits ≥4 in 3- or 2-year procedure	87.8 (73.7-95.4)	73.3 (71.8-74.1)	25.5 (21.4-27.7)	98.3 (96.3-99.4)
2 claims or 1 hospitalization in 5 years OR 2-year selected drugs and visits ≥4 in 5- or 2-year procedure	87.8 (73.7-95.4)	72.5 (71.1-73.3)	25.0 (21.0-27.2)	98.3 (96.3-99.3)
3 claims or 1 hospitalization in 2 years OR 2-year selected drugs and visits ≥4 in 2- or 2-year procedure	85.4 (71.0-93.8)	75.1 (73.6-75.9)	26.3 (21.9-28.9)	98.0 (96.0-99.2)
3 claims or 1 hospitalization in 3 years OR 2-year selected drugs and visits ≥4 in 3- or 2-year procedure	87.8 (73.7-95.4)	73.3 (71.8-74.1)	25.5 (21.4-27.7)	98.3 (96.3-99.4)
3 claims or 1 hospitalization in 5 years OR 2-year selected drugs and visits ≥4 in 5- or 2-year procedure	87.8 (73.7-95.4)	72.0 (70.5-72.8)	24.7 (20.7-26.2)	98.3 (96.3-99.3)

Note. CI = confidence interval; NPV = negative predictive value; OR = odds ratio; PPV = positive predictive value.

Table 4. Validity of Selected Case Definition (Using Algorithm 3 Within 5 Years), Compared to Reference Standard G4-5ND, Stratified by Subgroups Defined in Administrative Database.

	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Gender				
Female (n = 236)	85.1 (78.1-89.5)	96.3 (93.1-98.3)	91.3 (83.8-96.0)	93.4 (90.3-95.4)
Male (n = 198)	92.5 (87.3-94.5)	98.3 (94.8-99.7)	97.4 (91.9-99.5)	95.1 (91.7-96.4)
Age				
<65 (n = 68)	84.4 (74.1-84.4)	100.0 (90.9-100.0)	100.0 (87.9-100.0)	87.8 (79.8-87.8)
≥65 (n = 366)	90.2 (85.5-93.2)	96.7 (94.4-98.3)	93.2 (88.4-96.4)	95.2 (92.9-96.7)
Diabetes				
Yes (n = 229)	89.8 (85.2-92.2)	96.7 (92.6-98.9)	96.0 (91.1-98.6)	91.4 (87.5-93.4)
No (n = 205)	87.0 (77.3-92.4)	97.5 (94.7-99.0)	90.9 (80.9-96.6)	96.3 (93.5-97.8)
Hypertension				
Yes (n = 384)	88.7 (84.7-91.3)	96.6 (94.0-98.2)	94.3 (90.1-97.1)	93.0 (90.5-94.6)
No (n = 50)	100.0 (47.5-100.0)	100.0 (95.4-100.0)	100.0 (47.5-100.0)	100.0 (95.4-100.0)

Note. CI = confidence interval; NPV = negative predictive value; PPV = positive predictive value.

data used. A final case definition employing 3 physician claims or 1 hospitalization within a 5-year period and/or specific use of drug in the last 2-year period and at least 4 nephrologist visits in the last 5-year period offered the best results with a sensitivity of 89.0%, specificity of 97.1%, PPV of 94.5%, and NPV of 94.1%. Regarding the validity of the administrative case definitions of CKD G3-5ND and CKD G5ND, there was low variation across case definitions and length of administrative data observation, but estimates were not as accurate as those of G4-5ND compared with the reference standard.

However, the accuracy of these case definitions would still make them useful for research purposes; for instance, NPV value for the patients with CKD G1-2 was at 90.9% and 98.3% for CKD G5 (eGFR < 15 mL/min/1.73 m^2).

In the subgroup analysis (diabetes, hypertension, and different age groups) with our final case definition (algorithm 3) for CKD G4-5ND compared with the reference standard, we observed similar estimates specificity as reported in the study by Ronksley et al,[29] but much better sensitivity, PPV and NPV.

Table 5. Prevalence of Predicted G4-5ND in a Quebec Adult Cohort ≥66 Years Old During the Period of 2010-2015, Using Algorithm 3 Within 5 Years.

Prevalence of eGFR in the whole Quebec cohort of older adults		
Men	n = 68 209	<30 mL/min/1.73 m²
Age group		
66-69	n = 22 209	368 (1.7%)
70-74	n = 20 906	488 (2.3%)
75-79	n = 12 829	402 (3.1%)
≥80	n = 12 265	525 (4.3%)
Total	n = 68 209	1783 (2.6%)
Women	n = 88 620	<30 mL/min/1.73 m²
Age group		
66-69	n = 26 545	253 (1.0%)
70-74	n = 24 540	319 (1.3%)
75-79	n = 16 243	323 (2.0%)
≥80	n = 21 292	522 (2.5%)
Total	n = 88 620	1417 (1.6%)
Overall	n = 156 829	3200 (2.0)

Prevalence of eGFR in Quebec older adults with diabetes		
Men	n = 19 549	<30 mL/min/1.73 m²
Age group		
66-69	n = 5716	238 (4.2%)
70-74	n = 6201	307 (5.0%)
75-79	n = 3971	235 (5.9%)
≥80	n = 3661	265 (7.2%)
Total	n = 19549	1045 (5.4%)
Women	n = 19604	<30 mL/min/1.73 m²
Age group		
66-69	n = 4992	133 (2.7%)
70-74	n = 5247	177 (3.4%)
75-79	n = 4127	183 (4.4%)
≥80	n = 5238	237 (4.5%)
Total	n = 19 604	730 (3.7%)
Overall	N = 39 153	1775 (4.5%)

Prevalence of eGFR in Quebec older adults with chronic heart failure		
Men	n = 6165	<30 mL/min/1.73 m²
Age group		
66-69	n = 1246	120 (9.6%)
70-74	n = 1558	182 (11.7%)
75-79	n = 1342	150 (11.2%)
≥80	n = 2019	229 (11.3%)
Total	n = 6165	681 (11.1%)
Women	n = 6289	<30 mL/min/1.73 m²
Age group		
66-69	n = 904	77 (8.5%)
70-74	n = 1185	103 (8.7%)
75-79	n = 1186	119 (10.0%)
≥80	n = 3014	225 (7.5%)
Total	n = 6289	524 (8.3%)
Overall	n = 12454	1205 (9.7%)

Note. eGFR = estimated glomerular filtration rate.

Age was not a study entry criteria, but we nonetheless observed that the mean age was 74±10 and that 83% of patients were >65. We therefore proceeded to compare with other cohorts of patients in the same age group. The comparison of case definitions using the prediction of CKD G 4-5ND among a cohort of older adults based on a 40% random sample of individuals in the province of Quebec for the period of January 2010 to December 2015 stratified by age group of 66-69, 70-74, 75-79, and ≥ 80 years of age predicted a prevalence of 1.7 and 1.0%, 2.3 and 1.3%, 3.1 and 2.0%, and 4.3 and 2.5%, for males and females, respectively. Those estimates are similar to other reports[51-53] although lower than others,[15] however all below 10%. Moreover, the predicted prevalence of CKD G4-5ND among the older adult cohort with diabetes ranged from 2.7% to 7.2% for ages 66 years and above. Those estimates agree with other published estimates in patients with diabetes, in which CKD G4-5ND ranged from 1.75% to 7.84% for a similar cohort.[54-58] Again, the predicted prevalence of CKD G4-5ND among older adults with chronic heart failure ranged from 9.6% to 11.7%, which is quite close to other published estimates ranging from 6% to 12%.[59-61] The prevalence of CKD was expected to be higher in these subgroups of patients (diabetes, heart failure).

The use of algorithms based on the 3 linked administrative databases allowed us to reduce the amount of missing data and improve the validity of the variables under study. As for limitations, first, the characteristics obtained from this population may be not generalizable to the entire population of adults, as our cohort was of only older patients. Second, participants with CKD without serum creatinine measurements were excluded. It is possible that patients with CKD who did not visit a physician or were not admitted to the hospital during the study period were missed. Nevertheless, this measurement ensured that we selected true cases of normal and abnormal eGFR. Third, the comparison with other cohorts may be limited, since a case definition including a management by nephrologists could be linked to more moderate-to-severe CKD; it is possible that in certain remote areas, CKD patients are treated and followed by general practitioners or internal medicine specialists. Finally, as with all algorithms, defining disease based on administrative data may not provide results generalizable to other countries where health care registration is organized in a different way and/or with differences in coding for claims and hospitalizations; this would have to be confirmed. Nevertheless, apart from the fact that in certain provinces (eg, Alberta), serum creatinine results may be available in the databank, we believe that our results apply across Canada, as we share guidelines (Canadian Society of Nephrology) and have very similar health care systems in different provinces. Although other studies of CKD validation reported that administrative databases are not recommended for CKD surveillance, they still are useful tools when algorithms with high specificity are required, such as in pharmacoepidemiologic research.[29,30]

Conclusions

We suggest that our case definition of CKD G4-5ND derived from a composite of diagnosis code, drug use, and nephrologist visits using administrative databases is a valid algorithm when compared with medical chart reviews for older adults with CKD.

Disclosure Statement

Drs Perreault, Roy, Zappitelli, White-Guay, Lafrance, and Mr Dorais report no disclosures.

Ethics Approval and Consent to Participate

We obtained approvals from institutional research ethics boards of the *Centre Hospitalier de l'Université de Montréa*l (CHUM) and the *Commission d'Accès à l'Information du Québe*c (CAI, provincial ethics body), as well as approval to waive requirement for patient consent.

Consent for Publication

All authors consent for publication.

Availability of Data and Materials

Deidentified patient data from medical information, RAMQ and Med-Echo administrative databases are not available according to the rules of *Commission d'Accès à l'Information du Québec* (CAI, provincial ethics body).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Trial Registration

Trial registration number is not applicable because this is a retrospective study.

ORCID iDs

Brian White-Guay [iD] https://orcid.org/0000-0003-0500-5452
Sylvie Perreault [iD] https://orcid.org/0000-0002-0066-0127

Supplemental Material

Supplemental material for this article is available online.

References

1. Culleton BF, Larson MG, Wilson PW, Evans JC, Parfrey PS, Levy D. Cardiovascular disease and mortality in a community-based cohort with mild renal insufficiency. *Kidney Int*. 1999; 56(6):2214-2219.

2. Drey N, Roderick P, Mullee M, Rogerson M. A population-based study of the incidence and outcomes of diagnosed chronic kidney disease. *Am J Kidney Dis*. 2003;42(4):677-684.

3. Eknoyan G, Lameire N, Barsoum R, et al. The burden of kidney disease: improving global outcomes. *Kidney Int*. 2004; 66(4):1310-1314.

4. Garella S. The costs of dialysis in the USA. *Nephrol Dial Transplant*. 1997;12(suppl 1):10-21.

5. Meguid El NA, Bello AK. Chronic kidney disease: the global challenge. *Lancet*. 2005;365(9456):331-340.

6. Andrassy KM. Comments on "KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease." *Kidney Int*. 2013;84(3):622-623.

7. National Kidney F. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis*. 2002;39(2 suppl 1):S1-S266.

8. Coresh J, Byrd-Holt D, Astor BC, et al. Chronic kidney disease awareness, prevalence, and trends among U.S. adults, 1999 to 2000. *J Am Soc Nephrol*. 2005;16(1):180-188.

9. Zhang QL, Rothenbacher D. Prevalence of chronic kidney disease in population-based studies: systematic review. *BMC Public Health*. 2008;8:117.

10. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N Engl J Med*. 2004;351(13):1296-1305.

11. Keith DS, Nichols GA, Gullion CM, Brown JB, Smith DH. Longitudinal follow-up and outcomes among a population with chronic kidney disease in a large managed care organization. *Arch Intern Med*. 2004;164(6):659-663.

12. Kent S, Schlackow I, Lozano-Kuhne J, et al. What is the impact of chronic kidney disease stage and cardiovascular disease on the annual cost of hospital care in moderate-to-severe kidney disease? *BMC Nephrol*. 2015;16:65.

13. Eriksson JK, Neovius M, Jacobson SH, Elinder CG, Hylander B. Healthcare costs in chronic kidney disease and renal replacement therapy: a population-based cohort study in Sweden. *BMJ Open*. 2016;6(10):e012062.

14. Welch JL, Meek J, Bartlett Ellis RJ, Ambuehl R, Decker BS. Patterns of healthcare encounters experienced by patients with chronic kidney disease. *J Ren Care*. 2017;43(4):209-218.

15. Bello AK, Ronksley PE, Tangri N, et al. Prevalence and demographics of CKD in Canadian primary care practices: a cross-sectional study. *Kidney Int Rep*. 2019;4(4):561-570.

16. Hill NR, Fatoba ST, Oke JL, et al. Global prevalence of chronic kidney disease—a systematic review and meta-analysis. *PLoS One*. 2016;11(7):e0158765.

17. Arora P, Vasa P, Brenner D, et al. Prevalence estimates of chronic kidney disease in Canada: results of a nationally representative survey. *CMAJ*. 2013;185(9):E417-E423.

18. Levey AS, Coresh J. Chronic kidney disease. *Lancet*. 2012;379(9811):165-180.

19. Brenner BM, Cooper ME, de Zeeuw D, et al. Effects of losartan on renal and cardiovascular outcomes in patients with type 2 diabetes and nephropathy. *N Engl J Med*. 2001;345(12): 861-869.

20. Hsu CY, Bates DW, Kuperman GJ, Curhan GC. Diabetes, hemoglobin A(1c), cholesterol, and the risk of moderate chronic renal insufficiency in an ambulatory population. *Am J Kidney Dis*. 2000;36(2):272-281.

21. Lewis EJ, Hunsicker LG, Clarke WR, et al. Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. *N Engl J Med*. 2001;345(12):851-860.

22. Mann JF, Gerstein HC, Pogue J, Bosch J, Yusuf S. Renal insufficiency as a predictor of cardiovascular outcomes and the impact of ramipril: the HOPE randomized trial. *Ann Intern Med*. 2001;134(8):629-636.

23. Peterson JC, Adler S, Burkart JM, et al. Blood pressure control, proteinuria, and the progression of renal disease. The Modification of Diet in Renal Disease Study. *Ann Intern Med*. 1995;123(10):754-762.

24. Wright JT, Jr, Bakris G, Greene T, et al. Effect of blood pressure lowering and antihypertensive drug class on progression of hypertensive kidney disease: results from the AASK trial. *JAMA*. 2002;288(19):2421-2431.

25. Bang H, Vupputuri S, Shoham DA, et al. SCreening for Occult REnal Disease (SCORED): a simple prediction model for chronic kidney disease. *Arch Intern Med*. 2007;167(4): 374-381.

26. Foley RN, Murray AM, Li S, et al. Chronic kidney disease and the risk for cardiovascular disease, renal replacement, and death in the United States Medicare population, 1998 to 1999. *J Am Soc Nephrol*. 2005;16(2):489-495.

27. Grams ME, Plantinga LC, Hedgeman 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.

28. Vlasschaert ME, Bejaimal 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.

29. 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.

30. Friberg L, Gasparini A, Carrero JJ. A scheme based on ICD-10 diagnoses and drug prescriptions to stage chronic kidney disease severity in healthcare administrative records. *Clin Kidney J*. 2018;11(2):254-258.

31. Perreault S, Yu AY, Cote R, Dragomir A, White-Guay B, Dumas S. Adherence to antihypertensive agents after ischemic stroke and risk of cardiovascular outcomes. *Neurology*. 2012;79(20):2037-2043.

32. Perreault S, Nuevo J, Baumgartner S, Morlock R. Any link of gout disease control among hypertensive patients and onset of end-stage renal disease? Results from a population-based study. *World J Nephrol*. 2017;6(3):132-142.

33. D'Arienzo D, Hessey E, Ali R, et al. A validation study of administrative health care data to detect acute kidney injury in the pediatric intensive care unit. *Can J Kidney Health Dis*. 2019;6.doi:10.1177/2054358119827525.

34. Perreault S, de Denus S, White M, et al. Older adults with heart failure treated with carvedilol, bisoprolol, or metoprolol tartrate: risk of mortality. *Pharmacoepidemiol Drug Saf*. 2017;26(1):81-90.

35. Roy L, White-Guay B, Dorais M, Dragomir A, Lessard M, Perreault S. Adherence to antihypertensive agents improves risk reduction of end-stage renal disease. *Kidney Int*. 2013;84(3):570-577.

36. Simard P, Presse N, Roy L, et al. Association between metformin adherence and all-cause mortality among new users of metformin: a nested case-control study. *Ann Pharmacother*. 2018;52(4):305-313.

37. Tamblyn R, Lavoie G, Petrella L, Monette J. The use of prescription claims databases in pharmacoepidemiological research: the accuracy and comprehensiveness of the prescription claims database in Quebec. *J Clin Epidemiol*. 1995;48(8): 999-1009.

38. Levey AS, Stevens LA. Estimating GFR using the CKD Epidemiology Collaboration (CKD-EPI) creatinine equation: more accurate GFR estimates, lower CKD prevalence estimates, and better risk predictions. *Am J Kidney Dis*. 2010;55(4): 622-627.

39. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150(9):604-612.

40. Ix JH, Wassel CL, Stevens LA, et al. Equations to estimate creatinine excretion rate: the CKD epidemiology collaboration. *Clin J Am Soc Nephrol*. 2011;6(1):184-191.

41. Stevens LA, Schmid CH, Greene T, et al. Comparative performance of the CKD Epidemiology Collaboration (CKD-EPI) and the Modification of Diet in Renal Disease (MDRD) Study equations for estimating GFR levels above 60 mL/min/1.73 m2. *Am J Kidney Dis*. 2010;56(3):486-495.

42. Muhajarine N, Mustard C, Roos LL, Young TK, Gelskey DE. Comparison of survey and physician claims data for detecting hypertension. *J Clin Epidemiol*. 1997;50(6):711-718.

43. Siddiqui NF, Coca SG, Devereaux PJ, et al. Secular trends in acute dialysis after elective major surgery—1995 to 2009. *CMAJ*. 2012;184(11):1237-1245.

44. Tu K, Campbell NR, Chen ZL, Cauch-Dudek KJ, McAlister FA. Accuracy of administrative databases in identifying patients with hypertension. *Open Med*. 2007;1(1):e18-e26.

45. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *J Clin Epidemiol*. 2004;57(2):131-141.

46. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43(11):1130-1139.

47. Navaneethan SD, Jolly SE, Schold JD, et al. Development and validation of an electronic health record-based chronic kidney disease registry. *Clin J Am Soc Nephrol*. 2011;6(1):40-49.

48. Saczynski JS, Andrade SE, Harrold LR, et al. A systematic review of validated methods for identifying heart failure using administrative data. *Pharmacoepidemiol Drug Saf*. 2012;21(suppl 1):129-140.

49. Kadhim-Saleh A, Green M, Williamson T, Hunter D, Birtwhistle R. Validation of the diagnostic algorithms for 5 chronic conditions in the Canadian Primary Care Sentinel Surveillance Network (CPCSSN): a Kingston Practice-based Research Network (PBRN) report. *J Am Board Fam Med*. 2013;26(2):159-167.

50. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. 2009;338:b375.

51. Sim JJ, Batech M, Danforth KN, Rutkowski MP, Jacobsen SJ, Kanter MH. End-stage renal disease outcomes among the Kaiser Permanente Southern California Creatinine Safety Program (Creatinine SureNet): opportunities to reflect and improve. *Perm J*. 2017;21:16-143.

52. Coresh J, Selvin E, Stevens LA, et al. Prevalence of chronic kidney disease in the United States. *JAMA*. 2007;298(17): 2038-2047.

53. Armstrong MWR, Pannu N. *Pannu N. Prevalence and quality of care in chronic kidney disease*. Alberta Kidney Care Report; 2019, Alberta Health Services, Kidney Health Strategic Clinical Network. February 2019.

54. Zelnick LR, Weiss NS, Kestenbaum BR, et al. Diabetes and CKD in the United States Population, 2009-2014. *Clin J Am Soc Nephrol*. 2017;12(12):1984-1990.

55. Rodriguez-Poncelas A, Garre-Olmo J, Franch-Nadal J, et al. Prevalence of chronic kidney disease in patients with type 2 diabetes in Spain: PERCEDIME2 study. *BMC Nephrol*. 2013; 14:46.

56. Boehme MW, Buechele G, Frankenhauser-Mannuss J, et al. Prevalence, incidence and concomitant co-morbidities of type 2 diabetes mellitus in South Western Germany—a retrospective cohort and case control study in claims data of a large statutory health insurance. *BMC Public Health*. 2015;15:855.

57. Bramlage P, Lanzinger S, van Mark G, et al. Patient and disease characteristics of type-2 diabetes patients with or without chronic kidney disease: an analysis of the German DPV and DIVE databases. *Cardiovasc Diabetol*. 2019;18(1):33.

58. Metsärinne K, Broijersen A, Kantola I, et al. High prevalence of chronic kidney disease in Finnish patients with type 2 diabetes treated in primary care. *Prim Care Diabetes*. 2015;9(1): 31-38.

59. Molnar AO, Petrcich W, Weir MA, Garg AX, Walsh M, Sood MM. The association of beta-blocker use with mortality in elderly patients with congestive heart failure and advanced chronic kidney disease. *Nephrol Dial Transplant*. 2020;35:782–789.

60. Martinez-Milla J, Garcia MC, Urquia MT, et al. Blockade of renin-angiotensin-aldosterone system in elderly patients with heart failure and chronic kidney disease: results of a single-center, observational cohort study. *Drugs Aging*. 2019; 36(12):1123-1131.

61. Hakopian NN, Gharibian D, Nashed MM. Prognostic impact of chronic kidney disease in patients with heart failure. *Perm J*. 2019;23:18-273.