Original Research Article

Validation of Self-Reported Race in a Canadian Provincial Renal Administrative Database

Aiza Waheed[1] (iD), Ognjenka Djurdjev[2], Jianghu Dong[2],
Jagbir Gill[1,2], and Sean Barbour[1,2]

Abstract

Background: Administrative data are commonly used to study clinical outcomes in renal disease. Race is an important determinant of renal health delivery and outcomes in Canada but is not validated in most administrative data, and the correlation with census-based definitions of race is unknown.

Objectives: Validation of self-reported race (SRR) in a Canadian provincial renal administrative database (Patient Records and Outcome Management Information System [PROMIS]) and comparison with the Canadian census categories of race.

Design: Prospective patient survey study to validate SRR in PROMIS.

Setting: British Columbia, Canada.

Patients: Adult patients registered in PROMIS.

Measurements: Survey SRR was used as gold standard to validate SRR in PROMIS. Self-reported race in PROMIS was compared with census race categories.

Methods: This is a cross-sectional telephone survey of a random sample of all adults in PROMIS conducted between February 2016 and November 2016. Responders selected a race category from PROMIS and from the Canadian census. Sensitivity (Sn) and specificity (Sp) were calculated with 95% confidence intervals (CIs).

Results: A total of 21 039 patients met inclusion criteria, 1677 were selected for the survey and 637 participated (38% response rate). There were no differences between the PROMIS, sampled, and responder populations. PROMIS SRR had an accuracy of 95.3% (95% CI: 94.2%-97.0%) when validated against the survey SRR with Sn and Sp ≥90% in all race groups except in Aboriginals (Sn 87.5%). The positive and negative predictive values were ≥95%, except in very low and high–prevalence groups, respectively. The Canadian census had an accuracy of 95.7% (95% CI: 94.4%-97.6%) when validated against PROMIS SRR with Sn and Sp ≥90%. The results did not differ in subgroups based on age, sex, birth outside Canada, or renal group (glomerulonephritis, chronic kidney disease, hemodialysis, peritoneal dialysis, transplant recipients, or live donors).

Limitations: Analysis of minority groups and lower prevalence groups is limited by sample size. Results may not be generalizable to other administrative databases.

Conclusions: We have shown high accuracy of PROMIS SRR that validates its use in the secondary analysis of administrative data for research. There is high correlation between PROMIS and census race categories which allows linkage with other data sources that use census-based definitions of race.

Abrégé

Contexte: Les données administratives sont fréquemment utilisées pour étudier les issues cliniques en néphrologie. L'origine ethnique (OE) du patient est un déterminant important de la prestation de soins et des résultats en santé rénale au Canada, mais n'est pas validée dans la plupart des données administratives et la corrélation avec les définitions d'ethnies fondées sur le recensement demeure inconnue.

Objectifs: L'étude visait à valider l'origine ethnique autodéclarée (OEAD) dans une base de données administrative provinciale relative à la santé rénale (PROMIS), et à la comparer à l'origine ethnique inscrite au recensement canadien.

Type d'étude: Une étude prospective menée sous forme de sondage auprès de patients pour valider l'OEAD dans PROMIS.

Cadre: Colombie-Britannique, Canada

Sujets: Des patients adultes inscrits dans PROMIS

Mesures: L'OE mentionnée dans le sondage a servi d'étalon-or pour valider l'OEAD dans PROMIS, et cette dernière a été comparée à l'OE rapportée par le recensement.

Méthodologie: Une enquête transversale conduite par téléphone entre février et novembre 2016 auprès d'un échantillon aléatoire d'adultes inscrits dans PROMIS. Les répondants devaient choisir une OE dans PROMIS et dans les catégories du recensement canadien. La sensibilité (Sn) et la spécificité (Sp) ont été calculées avec un intervalle de confiance à 95 % (IC 95 %).
Résultats: Des 21 039 patients qui satisfaisaient les critères d'inclusion, 1 677 ont été sélectionnés pour le sondage et 637 ont participé (taux de réponse: 38 %). Aucune différence n'a été observée entre les populations de PROMIS, de l'échantillon et de répondants. L'OEAD dans PROMIS était exacte dans 95,3 % des cas (IC 95 %: 94,2-97,0 %), lorsque validée contre l'OEAD dans le sondage, avec une Sn et une Sp d'au moins 90 % pour tous les groupes ethniques, à l'exception des Autochtones (Sn: 87,5 %). Les valeurs prédictives positive et négative étaient d'au moins 95 %, sauf dans les groupes à très faible et à forte prévalence, respectivement. Le recensement canadien a montré une précision de 95,7 % (IC 95 %: 94,4-97,6 %) lorsque validé contre l'OEAD dans PROMIS avec une Sn et une Sp d'au moins 90 %. Les résultats n'ont pas varié dans les sous-groupes selon l'âge, le sexe, la naissance hors Canada ou le groupe de néphrologie (glomérulonéphrite, insuffisance rénale chronique, hémodialyse, dialyse péritonéale, receveurs d'une greffe ou donneurs vivants).
Limites: L'analyse des groupes minoritaires et des groupes à faible prévalence est limitée par la taille de l'échantillon. Les résultats pourraient ne pas être généralisables à d'autres bases de données administratives.
Conclusion: Nous avons montré la grande précision de l'OEAD dans PROMIS, ce qui valide son utilisation pour l'analyse secondaire de données administratives à des fins de recherche. Une forte corrélation existe entre les définitions de l'OE dans PROMIS et le recensement, ce qui permet d'établir des liens avec d'autres sources de données qui utilisent les mêmes définitions que le recensement.

What was known before

Administrative data are being used for research purposes, but use is limited by the accuracy of this data. Race is an important determinant of renal health delivery and outcomes and is studied extensively in Canadian health research.

What this adds

The race variable in our renal administrative data set is validated, allowing its use in clinical research. We compare the renal administrative dataset race variable with the Canadian census definitions of race with good accuracy, allowing linkage to other data sets.

Introduction

Race is an important determinant of patient outcome and health care utilization at all categories of kidney disease, including chronic kidney disease (CKD), end-stage kidney disease (ESKD), and transplantation.[1,2] In recent years, the role of race has been studied extensively in these populations in Canada, including rates of transplantation and allograft failure,[3-8] kidney donation,[9,10] outcomes on dialysis,[11-13] and progression of CKD and glomerulonephritis (GN).[14-19] Studying the impact of race in kidney disease has identified disparities in access to care that have been targeted by health policies resulting in improved patient outcomes.[20]

There is increasing use of administrative data to study clinical outcomes in kidney disease across large geographically and racially diverse populations.[21-23] However, administrative databases often do not collect race or use methods that have not been previously validated, unlike other important variables that have undergone extensive validation of case definitions.[24-30] In British Columbia, the Patient Records and Outcome Management Information System (PROMIS) is the administrative database that captures all patients with CKD and dialysis patients, and transplant recipients and donors in the province. Because PROMIS was designed for administrative purposes, there is not a standardized mechanism to capture self-reported race (SRR). As such, the accuracy of race in PROMIS or in other renal administrative databases remains unknown. Furthermore, broad race categories often include a mixed demographic of patients of

[1]The University of British Columbia, Vancouver, Canada
[2]BC Provincial Renal Agency, Vancouver, Canada

Corresponding Author:
Sean Barbour, The University of British Columbia, 2775 Laurel Street, 5th Floor, Vancouver, BC, Canada V5Z 1M9.
Email: Sean.Barbour@vch.ca

Table 1. Race Categories for PROMIS and the 2011 Canadian Census.

PROMIS categories of race	2011 census categories of race	
Select one	Select all that apply	2011 census capture of aboriginal status
Caucasian	White	Are you Aboriginal: yes/no
South Asian	South Asian	
East Asian	Chinese	
Black	Black	If Aboriginal, to which group do you belong:
Filipino	Filipino	First Nations
Latin American	Latin American	Metis
Middle Eastern	Arab	Inuit
Aboriginal	Southeast Asian	
Other/Multiracial	West Asian	
	Korean	
	Japanese	
	Other	

Note. PROMIS = Patient Records and Outcome Management Information System.

varying age and country of birth, making it more difficult to draw meaningful conclusions from health outcomes associations in these groups.

Therefore, we sought to characterize patients within race groups captured in PROMIS, validate the capture of race in PROMIS against the gold standard of patient SRR, and to compare it with the Canadian Census definitions of race.[31]

Methods

Study Design

This is a prospective validation study that used a cross-sectional survey of a randomly sampled population of patients from PROMIS between February 2016 and November 2016. PROMIS is the provincial administrative database in British Columbia (BC) for patients with kidney disease and is managed by the BC Provincial Renal Agency. Registration in PROMIS is mandatory when patients have advanced all-cause CKD needing renal-specific medications or multidisciplinary clinics, at the time of renal biopsy diagnosis of GN, or at the time of kidney transplantation, live kidney donation, or commencing dialysis (including hemodialysis [HD] and peritoneal dialysis [PD]). As such, PROMIS captures all living kidney donors in BC and all patients with GN, ESKD, and advanced CKD. At the time of registration in PROMIS, patient demographics are captured during the usual processes of clinical care, including SRR (PROMIS SRR).

Sampling Strategy

The source population consisted of all patients registered in PROMIS who were alive and ≥18 years old at the time of survey completion. Children were not sampled due to the sensitive nature of the survey questions. Eligible patients were randomly selected until the desired number of patients had

responded to the survey. The sample size was chosen to allow 90% power to validate a 2-level categorical race variable with a prevalence of 15% at a sensitivity of 0.95 using an alpha of .05. The prevalence rate was based on the prevalence of different race groups in PROMIS (11% East Asian, 9% South Asian, etc).[15] The prevalence of patients with GN in PROMIS is 7%. To ensure sufficient power to validate race in a priori–defined renal subgroups, patients with GN were oversampled from 7% to 15%. The sample size was calculated to be 487 and was increased from 487 to 600 to broaden sampling and ensure an adequate mix of race and renal groups.

Survey Details and Definitions

The survey was developed by the investigators and included questions on demographics, including age, sex, and country of birth as these have been associated with differential accuracy of race measurements.[32-34] Patients were asked to report their race using 2 different categorizations. First, they were asked to report their race based on the usual categorization available in PROMIS (see Table 1). Second, they were asked to report their race based on the categorization from the 2011 Canadian Census (see Table 1). Only one selection was allowed for the PROMIS categorization of race—this was defined as the Survey SRR. The Census race categories, on the contrary, allowed multiple entries. When patients chose multiple answers for these questions, the first response was defined as the Census SRR. Mapping of Census to PROMIS SRR categories is shown in Supplementary Table 1. The survey was pilot tested on healthy volunteers of different racial backgrounds to ensure content validity prior to its administration.

Eligible patients were contacted by multilingual trained research coordinators via telephone and consenting patients completed the questionnaire over the telephone. Surveys

Table 2. Characteristics of the Patients Included in the Survey Based on PROMIS Self-Reported Race Categories.

	All (N = 637)	Caucasian (N = 404)	East Asian (N = 65)	South Asian (N = 62)	Filipino (N = 20)	Aboriginal (N = 19)	P value
Mean age (SD)	65 (14)	65 (13)	66 (14)	63 (15)	56 (11)	63 (13)	<.001
Female (%)	47	47	37	56	65	53	.180
Household income (%)							
Less than $25,000	21	18	28	26	5	32	.036
$25,000- $49,000	24	24	20	16	40	16	
$50,000- $74,000	16	19	14	15	10	5	
$75,000- $99,000	12	11	14	13	20	5	
More than $100,000	16	17	12	15	20	21	
Prefer not to say	11	11	12	16	5	21	
Renal group (%)							
GN	14	15	17	11	5	16	<.001
CKD	19	18	15	13	5	16	
HD	23	19	22	44	30	46	
PD	11	11	19	10	25	5	
TX	20	22	19	16	15	16	
Live donors	13	15	9	6	20	0	
Area of primary residence (%)							
Vancouver Metropolitan Area	61	48	90	81	95	40	<.001
Born outside of Canada (%)	36	19	89	89	100	23	<.001
Most Common Country of Birth (%)	Canada (64) China (6) India (6)	Canada (81), England (5) Scotland (2)	China (55), Canada (11), Malaysia(8)	India (55), Fiji (22), Canada (11)	Philippines (100)	Canada (77), El Salvador (12), Mexico (4)	

Note. Other and unknown categories have been omitted for clarity. GN = glomerulonephritis; CKD = chronic kidney disease; HD = hemodialysis; PD = peritoneal dialysis; TX = transplant.

were completed in English, Cantonese, Mandarin, and Punjabi by research coordinators. For patients speaking other languages, interpreters were used for consent and to complete the survey.

Statistical Analysis

Continuous variables were summarized as mean (SD) and compared across groups using the t test, and categorical variables were summarized as count (frequency) and compared across groups using the χ^2 test. When validating the capture of race in PROMIS, survey SRR was considered the gold standard and was compared with PROMIS SRR. When validating the Canadian Census capture of race, PROMIS SRR was considered the gold standard and compared with the Census SRR. This approach was specifically designed so that the results could inform future research that merges PROMIS with other administrative data sets that only capture Census definitions of race. Sensitivity and specificity were calculated based on the presence or absence of each category of race, with 95% confidence intervals generated by the simple asymptotic method.

The analyses were repeated in a priori–determined subgroups based on age, sex, country of birth other than Canada, and type of renal program (CKD, GN, transplant recipient, PD, HD, and living donors) to investigate consistency of results.

All analyses were performed using SAS version 9.4. *P* values <.05 were considered statistically significant. The study was approved by the University of British Columbia research ethics board. The Standards of Reporting of Diagnostic Accuracy Studies (STARD) guidelines and checklists were followed for this validation study.[35]

Results

Description of the Cohort

There were 21 039 patients in PROMIS who met criteria for inclusion. A total of 1677 patients were approached to participate and 640 consented to participate (survey response rate 38.0%). Three individuals did not answer key questions and were therefore excluded (see Supplementary Figure 1). Within each renal program, there were no qualitative differences in age, sex, and PROMIS SRR between those in the source population, those selected for the survey, and those who responded (see Supplementary Table 2).

A description of the cohort is provided in Table 2. According to the PROMIS categorization of SRR, of the 637 patients included in the analysis, 20 were Filipino, 62 were South Asian, 65 were East Asian, 404 were Caucasian, 19

Table 3. Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value for Each Race Category in PROMIS (PROMIS SRR) Compared With the Survey SRR as the Gold Standard.

PROMIS SRR	Sensitivity (%)	95% CI	Specificity (%)	95% CI	PPV (%)	NPV (%)
Filipino	90.5	75.1-99.9	99.8	99.5-100.0	95.0	99.7
South Asian	95.3	91.3-100.0	99.8	98.5-100.0	98.4	99.5
East Asian	96.9	87.1-99.0	99.5	98.7-100.0	95.4	99.7
Caucasian	91.5	94.2-98.1	93.4	89.3-98.8	96.5	84.5
Aboriginal	87.5	48.8-90.9	99.2	97.9-99.6	73.7	99.0
Other	54.8	46.3-65.8	99.3	95.2-99.8	85.2	96.9

Note. Other PROMIS SRR category includes Black, Middle Eastern, and Latin American. PROMIS = Patient Records and Outcome Management Information System; SRR = self-reported race; CI = confidence interval; PPV = positive predictive value; NPV = negative predictive value.

were Aboriginal, 27 were multiracial or other (including Black, Latin American, and Middle Eastern), and 40 were listed as having unknown race.

The mean age was 65 years and was significantly younger in the Filipino group (56 years, $P = <.001$). Overall, 36% of patients were born outside Canada, which was more common in the East Asian, South Asian, and Filipino groups (87%-100%, $P = <.001$.). There were also significant differences in renal groups, with Filipinos being more commonly on PD (25%) or HD (30%), Aboriginal and South Asians more commonly on HD (46 and 44%, respectively), and fewer South Asian, East Asian, and Aboriginal living donors (6%, 9% and 0%, respectively). The frequency of patients in the lowest income bracket was 21% and was significantly higher in the East Asian, South Asian, and Aboriginal groups (28%, 26%, and 32% respectively), when compared with the Caucasian and Filipino groups (18% and 5%, respectively). It is important to note, however, that the sample sizes in some race groups were quite small (n = 20 in Filipino group and n = 19 in Aboriginal group). Most (61%) of the cohort lived in the Vancouver Metropolitan Area, but this was less common in the Caucasian and Aboriginal groups (48% and 40% respectively, $P = <.001$).

Validation of PROMIS SRR

Compared with the gold standard of SRR reported in the survey, SRR captured in PROMIS classified 95.3% (95% confidence interval [CI]: 94.2%-97.0%) of patients into the correct race group. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for each race category are shown in Table 3. The sensitivity and specificity were both >90% in all race groups, except in the Aboriginal group where the sensitivity was 87.5% (95% CI: 48.8%-90.9%). The sensitivity for the Other category was 54.8%, lower than the other race categories.

The PPV for each group was >95% except for the Aboriginal and Other group. The NPV for each group was >95% except for the Caucasian group. In the Caucasian group, the PPV was 96.5%, whereas the NPV was 84.5%,

and in the Aboriginal group, the PPV was 73.7%, whereas the NPV was 99.0%, likely as a result of high and low prevalence, respectively. Similarly, the Other group had a lower PPV 85.2% due to low prevalence.

Supplementary Figure 2 outlines the accuracy for SRR captured in PROMIS across a priori–defined subgroups. There were no major differences in the accuracy according to subgroups based on age, sex, birth outside of Canada, or renal program.

Comparing the Canadian Census to PROMIS Categorization of Race

When validated against SRR captured in PROMIS, the Canadian Census categorization of SRR had an overall accuracy of 95.7% (95% CI: 94.4%-97.6%). Of the individuals who identified as Caucasian in PROMIS, 99% selected Caucasian on the Census and 1% selected South Asian. Most (97%) of South Asians in PROMIS selected South Asian on the Census, whereas 1.5% selected each of Caucasian and Middle Eastern. Among East Asians from PROMIS, 82% reported Chinese origin on the Canadian Census, compared with 4.7% who reported each of Japanese, Korean, and South East Asian (Vietnamese, Cambodian, Laotian, Thai, etc) ethnic origins. Among the Middle Eastern/Arabian population in PROMIS, 60% reported West Asian and 20% reported each of Caucasian and Arab origins on the Canadian Census. Most (95%) of Filipinos in PROMIS selected Filipino on the Census and 5% selected Caucasian. Only 11 patients reported multiple responses to the ethnic origins Census question.

The sensitivity, specificity, PPV, and NPV of the Census SRR compared with PROMIS SRR as the gold standard are shown in Table 4. The specificity for most groups was very good (>90%), except for Caucasians (specificity 89.1%). The sensitivity was also >90% in most groups, except in the Aboriginal, Latin American, and West Asian groups (sensitivity: 66.7%-89.5%). The accuracy was not different in subgroups based on age, sex, birth outside of Canada, or renal program, as shown in Supplementary Figure 2.

Table 4. Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value of SRR Captured Using the Canadian Census Categorization (Census SRR) Compared With PROMIS SRR as the Gold Standard.

Census SRR	Sensitivity (%)	95% CI	Specificity (%)	95% CI	PPV (%)	NPV (%)
Filipino	95.0	93.3-96.7	99.8	99.5-100.0	95.0	99.8
South Asian	98.0	97.4-99.4	99.3	98.6-99.9	93.8	99.8
East Asian	92.3	90.2-94.3	99.5	98.9-100	95.2	99.1
Caucasian	99.8	99.3-100.0	89.1	86.7-91.5	94.9	99.5
Aboriginal	89.5	87.1-91.8	97.7	96.5-98.8	54.8	99.7
Black	100.0	100.0-100.0	99.8	99.5-100.0	80.0	100.0
Middle Eastern/ Arabian	66.7	63.0-70.3	99.8	99.5-100.0	80.0	99.7
Latin American	66.7	63.0-70.3	99.7	99.2-100.0	67.0	100.0
Other	9.1	6.8-11.3	99.8	99.5-100.0	50.0	98.3

Note. SRR = self-reported race; PROMIS = Patient Records and Outcome Management Information System; CI = confidence interval; PPV = positive predictive value; NPV = negative predictive value.

Discussion

We surveyed 637 patients from a large provincial renal administrative database in British Columbia (PROMIS) and demonstrated 95% accuracy for the capture of race. The PPV and NPV for individual categories of race were ≥90% except in very low and very high–prevalence populations, respectively. We additionally compared the categorization of race in PROMIS with that taken from the 2011 Canadian Census and demonstrated an accuracy of 96%. Our results were similar in subgroups based on renal program, age, sex, and non-Canadian country of birth.

There are several characteristics of our cohort that may have affected the accuracy of SRR capture in PROMIS. Immigration patterns in BC have led to large populations of certain minority groups, such as the South Asian, East Asian, and Filipino populations. The survey results identified these as relatively homogeneous populations, the majority of which were born outside of Canada and are predominantly from India, China, Hong Kong, and the Philippines. For example, 89% of East Asians, 87% of South Asians, and 100% of Filipinos were born outside of Canada. The proportion of patients within these groups that are born outside of Canada is much higher than what is reported in the 2011 Canadian census, in which comparatively fewer South Asians (69.3%) and Chinese (73.3%) individuals were born outside Canada.[36] This suggests that outcomes for individuals with renal disease may be associated with immigration status and other social determinants of health.[37] In addition, the capture of race may be less accurate in administrative data sets servicing more heterogeneous populations with a larger admixture of immigrant, Canadian-born, and multiethnic groups.

Additional factors may also influence the accuracy of PROMIS SRR. Race entry in PROMIS occurs on registration and may be done by clerks or other health care workers. While the standard is to confirm race with the patient, assumptions are often made which can lead to inaccuracies in the data. This has been shown to be especially true for minority groups.[38] Barriers also exist in obtaining race information, for example, asking patients about their racial background may be perceived as race playing a role in their care, leading to breakdown of trust in the patient—health care worker relationship, which may be particularly relevant to underserviced ethnic groups such as the Aboriginal population.[38] Furthermore, the "other" SRR category may have been affected by errors related to multiracial or multiethnic backgrounds or in individuals in whom the categories presented do not accurately reflect their race.[38-41]

The secondary analysis of health administration data frequently requires linkage of multiple data sets to improve the capture of variables relevant to health research.[42,43] Canada has been a leader in this regard, with an established infrastructure in many provinces for linking regional, provincial, and national administrative databases to support health research that could not have been addressed otherwise.[23,43-45] One commonly used source of data on race in linked administrative data sets is the Canadian Census, which may not be interchangeable with the categorization of race in other administrative data sets.[42] As such, we sought to compare the Census capture of race with that from PROMIS. Our results show good overall accuracy with sensitivity and specificity greater than 90% for most categories of race. There were no differences noted between subgroups based on age, sex, renal program, and birth outside of Canada. These results justify using linkages between PROMIS and the Canadian Census databases for future studies investigating race in kidney disease.

There are several limitations to consider when interpreting our results. The analysis of certain minority groups, such as Aboriginals, may be limited by small sample size that increases error. Additional research is required to further validate other racial groups that also have low prevalence in PROMIS, including Black, South American, and West Asian populations. Our results may not generalize to other administrative databases, but might nonetheless apply to systems with similar

methods of primary capture of race. Finally, the purpose of this study was to validate the capture of SRR and was not intended to address whether the measurement of race in administrative data captures true differences in biologic, ethnic, cultural, or environmental factors between racial groups.

In conclusion, this study validated the capture of race in PROMIS as a large provincial renal administrative database and demonstrated excellent correlation with the Canadian Census capture of race. Our results justify the secondary use of PROMIS data alone or linked to other administrative databases in future health research exploring race in kidney disease.

Ethics Approval and Consent to Participate

The study was approved by the University of British Columbia ethics board. Participants consented to participate prior to administration of the survey.

Consent for Publication

Consent for publication was obtained from all authors.

Availability of Data and Materials

Data in PROMIS can be accessed through the BC Renal Agency. Results of the survey are stored on UBC research servers.

Declaration of Conflicting Interests

Funding

ORCID iD

Aiza Waheed [iD] https://orcid.org/0000-0002-3134-4880

Supplemental Material

Supplemental material for this article is available online.

References

1. Risch N, Burchard E, Ziv E, Tang H. Categorization of humans in biomedical research: genes, race and disease. *Genome Biol*. 2002;3(7):comment2007.
2. Lin SS, Kelsey JL. Use of race and ethnicity in epidemiologic research: concepts, methodological issues, and suggestions for research. *Epidemiol Rev*. 2000;22(2):187-202.
3. Promislow S, Hemmelgarn B, Rigatto C, et al. Young aboriginals are less likely to receive a renal transplant: a Canadian national study. *BMC Nephrol*. 2013;14:11.
4. Yeates K, Wiebe N, Gill J, et al. Similar outcomes among black and white renal allograft recipients. *J Am Soc Nephrol*. 2009;20(1):172-179.
5. Tonelli M, Hemmelgarn B, Manns B, et al. Death and renal transplantation among Aboriginal people undergoing dialysis. *CMAJ*. 2004;171(6):577-582.
6. Hemmelgarn BR, Chou S, Wiebe N, et al. Differences in use of peritoneal dialysis and survival among East Asian, Indo Asian, and white ESRD patients in Canada. *Am J Kidney Dis*. 2006;48(6):964-971.
7. Tonelli M, Hemmelgarn B, Gill JS, et al. Patient and allograft survival of Indo Asian and East Asian dialysis patients treated in Canada. *Kidney Int*. 2007;72(4):499-504.
8. Tonelli M, Hemmelgarn B, Kim AK, et al. Association between residence location and likelihood of kidney transplantation in Aboriginal patients treated with dialysis in Canada. *Kidney Int*. 2006;70(5):924-930.
9. Gill J, Dong J, Rose C, Johnston O, Landsberg D. The effect of race and income on living kidney donation in the United States. *J Am Soc Nephrol*. 2013;24(11):1872-1879.
10. Lunsford SL, Simpson KS, Chavin KD, et al. Racial disparities in living kidney donation: is there a lack of willing donors or an excess of medically unsuitable candidates? *Transplantation*. 2006;82(7):876-881.
11. Sood MM, Hemmelgarn B, Rigatto C, et al. Association of modality with mortality among Canadian Aboriginals. *Clin J Am Soc Nephrol*. 2012;7(12):1988-1995.
12. Sood MM, Komenda P, Sood AR, et al. Adverse outcomes among Aboriginal patients receiving peritoneal dialysis. *CMAJ*. 2010;182(13):1433-1439.
13. Tonelli M, Hemmelgarn B, Manns B, et al. Use and outcomes of peritoneal dialysis among Aboriginal people in Canada. *J Am Soc Nephrol*. 2005;16(2):482-488.
14. Barbour SJ, Cattran DC, Kim SJ, et al. Individuals of Pacific Asian origin with IgA nephropathy have an increased risk of progression to end-stage renal disease. *Kidney Int*. 2013;84(5):1017-1024.
15. Barbour SJ, Er L, Djurdjev O, Karim M, Levin A. Differences in progression of CKD and mortality amongst Caucasian, Oriental Asian and South Asian CKD patients. *Nephrol Dial Transplant*. 2010;25(11):3663-3672.
16. Barbour SJ, Er L, Djurdjev O, Karim MA, Levin A. The prevalence of hematologic and metabolic abnormalities during chronic kidney disease stages in different ethnic groups. *Kidney Int*. 2008;74(1):108-114.
17. Gao S, Manns BJ, Culleton BF, et al. Prevalence of chronic kidney disease and survival among aboriginal people. *J Am Soc Nephrol*. 2007;18(11):2953-2959.
18. Conley J, Tonelli M, Quan H, et al. Association between GFR, proteinuria, and adverse outcomes among White, Chinese, and South Asian individuals in Canada. *Am J Kidney Dis*. 2012;59(3):390-399.
19. Gao S, Manns BJ, Culleton BF, et al. Access to health care among status Aboriginal people with chronic kidney disease. *CMAJ*. 2008;179(10):1007-1012.
20. Zacharias JM, Young TK, Riediger ND, Roulette J, Bruce SG. Prevalence, risk factors and awareness of albuminuria on a Canadian First Nation: a community-based screening study. *BMC Public Health*. 2012;12:290.
21. Chowdhury TT, Hemmelgarn B. Evidence-based decision-making 6: utilization of administrative databases for health services research. *Methods Mol Biol*. 2015;1281:469-484.
22. Bello A, Hemmelgarn B, Manns B, Tonelli M, Network AKD. Use of administrative databases for health-care planning in CKD. *Nephrol Dial Transplant*. 2012;27(suppl 3): iii12-iii18.

23. Hemmelgarn BR, Clement F, Manns BJ, et al. Overview of the Alberta Kidney Disease Network. *BMC Nephrol*. 2009; 10:30.

24. Moist LM, Richards HA, Miskulin D, et al. A validation study of the Canadian Organ Replacement Register. *Clin J Am Soc Nephrol*. 2011;6(4):813-818.

25. Lam NN, McArthur E, Kim SJ, Knoll GA. Validation of kidney transplantation using administrative data. *Can J Kidney Health Dis*. 2015;2:20.

26. Quinn RR, Laupacis A, Austin PC, et al. Using administrative datasets to study outcomes in dialysis patients: a validation study. *Med Care*. 2010;48(8):745-750.

27. 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.

28. Clement FM, James MT, Chin R, et al. Validation of a case definition to define chronic dialysis using outpatient administrative data. *BMC Med Res Methodol*. 2011;11:25.

29. 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.

30. Li AH, Kim SJ, Rangrej J, et al. Validity of physician billing claims to identify deceased organ donors in large healthcare databases. *PLoS ONE*. 2013;8(8):e70825.

31. Statistics Canada. National Household Survey 2011. http://www23.statcan.gc.ca/imdb/p3Instr.pl?Function=assembleInstr&lang=en&Item_Id=114641#qb115335 (accessed June 19, 2019).

32. Shah BR, Chiu M, Amin S, Ramani M, Sadry S, Tu JV. Surname lists to identify South Asian and Chinese ethnicity from secondary data in Ontario, Canada: a validation study. *BMC Med Res Methodol*. 2010;10:42.

33. Macfarlane GJ, Lunt M, Palmer B, Afzal C, Silman AJ, Esmail A. Determining aspects of ethnicity amongst persons of South Asian origin: the use of a surname-classification programme (Nam Pehchan). *Public Health*. 2007;121(3):231-236.

34. Quan H, Wang F, Schopflocher D, et al. Development and validation of a surname list to define Chinese ethnicity. *Med Care*. 2006;44(4):328-333.

35. 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.

36. Statistics Canada. *Immigration and Ethnocultural Diversity in Canada*. Ottawa, ON: Statistics Canada; 2013.

37. Kuile S, Rousseau C, Munoz M, Nadeau L, Ouimet M. The universality of the Canadian Health Care System in question: barriers to services for immigrants and refugees. *Int J Migrat Health Soc Care*. 2007; 3(1): 15-26.

38. Hasnain-Wynia R, Baker DW. Obtaining data on patient race, ethnicity, and primary language in health care organizations: current challenges and proposed solutions. *Health Serv Res*. 2006;41(4Pt1)1501-18.

39. Arday SL, Arday DR, Monroe S, Zhang J. HCFA's racial and ethnic data: current accuracy and recent improvements. *Health Care Financ Rev*. 2000;21(4):107-116.

40. Eicheldinger C, Bonito A. More accurate racial and ethnic codes for Medicare administrative data. *Health Care Financ Rev*. 2008;29(3):27-42.

41. Ford ME, Kelly PA. Conceptualizing and categorizing race and ethnicity in health services research. *Health Serv Res*. 2005;40(5, pt 2):1658-1675.

42. Gliklich RE, Dreyer NA, Leavy MB. *Registries for Evaluating Patient Outcomes: A User's Guide*. Rockville, MD: Agency for Healthcare Research and Quality; 2014. http://www.ncbi.nlm.nih.gov/books/NBK208616/. Accessed June 12. 2019.

43. Roos LL, Wajda A, Nicol JP. The art and science of record linkage: methods that work with few identifiers. *Comput Biol Med*. 1986;16(1):45-57.

44. Chamberlayne R, Green B, Barer ML, Hertzman C, Lawrence WJ, Sheps SB. Creating a population-based linked health database: a new resource for health services research. *Can J Public Health*. 1998;89(4):270-273.

45. Chiu M, Lebenbaum M, Lam K, et al. Describing the linkages of the immigration, refugees and citizenship Canada permanent resident data and vital statistics death registry to Ontario's administrative health database. *BMC Med Inform Decis Mak*. 2016;16(1):135.