

# UCSF

## UC San Francisco Previously Published Works

### Title

Estimating Prevalence of CKD Stages 3-5 Using Health System Data

### Permalink

<https://escholarship.org/uc/item/2wq3q9q9>

### Journal

American Journal of Kidney Diseases, 61(6)

### ISSN

0272-6386

### Authors

Shahinian, Vahakn B  
Hedgeman, Elizabeth  
Gillespie, Brenda W  
[et al.](#)

### Publication Date

2013-06-01

### DOI

10.1053/j.ajkd.2013.01.018

Peer reviewed



Published in final edited form as:

*Am J Kidney Dis.* 2013 June ; 61(6): 930–938. doi:10.1053/j.ajkd.2013.01.018.

## Estimating Prevalence of CKD Stages 3–5 Using Health System Data

Vahakn B. Shahinian, MD, MS<sup>1,2</sup>, Elizabeth Hedgeman, MS, MPH<sup>1,2</sup>, Brenda W. Gillespie, PhD<sup>3,4</sup>, Eric W. Young, MD<sup>1,5</sup>, Bruce Robinson, MD, MS<sup>6</sup>, Chi-yuan Hsu, MD, MS<sup>7</sup>, Laura C. Plantinga, MS<sup>8</sup>, Nilka Ríos Burrows, MT, MPH<sup>9</sup>, Paul Eggers, PhD<sup>10</sup>, Sharon Saydah, PhD<sup>9</sup>, Neil R. Powe, MD, MPH, MBA<sup>8</sup>, and Rajiv Saran, MBBS, MD, MRCP, MS<sup>1,2,\*</sup> on behalf of the CDC CKD Surveillance System

<sup>1</sup>Department of Internal Medicine, Ann Arbor, MI

<sup>2</sup>Kidney Epidemiology and Cost Center, Ann Arbor, MI

<sup>3</sup>Center for Statistical Consultation and Research, Ann Arbor, MI

<sup>4</sup>Department of Biostatistics, University of Michigan, Ann Arbor, MI

<sup>5</sup>Veterans Administration Ann Arbor Healthcare System, Ann Arbor, MI

<sup>6</sup>Arbor Research Collaborative for Health, Ann Arbor, MI

<sup>7</sup>Division of Nephrology, University of California, San Francisco

<sup>8</sup>Department of Medicine, San Francisco General Hospital and University of California, San Francisco, CA

<sup>9</sup>Centers for Disease Control and Prevention, Atlanta, GA

<sup>10</sup>National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD

### Abstract

**Background**—The feasibility of using health system data to estimate prevalence of chronic kidney disease (CKD) stages 3–5 was explored.

**Study Design**—Cohort study.

**Setting & Participants**—A 5% national random sample of patients from the Veterans Affairs (VA) health care system, enrollees in a managed care plan in Michigan (M-CARE), and participants from the 2005–2006 National Health and Nutrition Examination Survey (NHANES).

---

Address correspondence to Rajiv Saran, MBBS, MD, MRCP, MS, Division of Nephrology, Department of Internal Medicine, University of Michigan-Kidney Epidemiology and Cost Center, 1415 Washington Heights, Ste 3645, Ann Arbor, MI 48109-2029. rsaran@umich.edu.

\*A list of the CDC CKD Surveillance System team appears in the Acknowledgements.

*Financial Disclosure:* The authors declare that they have no other relevant financial interests.

### SUPPLEMENTARY MATERIAL

Table S1: Disease definitions.

Table S2: Observed prevalence of individuals with CKD stages 3–5 as estimated using the MDRD Study equation.

Table S3: Comparison of observed versus model-predicted prevalence of CKD stages 3–5.

Note: The supplementary material accompanying this article (<http://dx.doi.org/10.1053/j.ajkd.2013.01.018>) is available at [www.ajkd.org](http://www.ajkd.org).

**Predictor**—Observed CKD prevalence estimates in the health system population were calculated as patients with an available outpatient serum creatinine measurement with estimated glomerular filtration rate  $<60$  mL/min/1.73 m<sup>2</sup>, among those with at least one outpatient visit during the year.

**Outcomes & Measurements**—A logistic regression model was fitted using data from the 2005–2006 NHANES to predict CKD prevalence in those untested for serum creatinine in the health system population, adjusted for demographics and comorbid conditions. Model results then were combined with the observed prevalence in tested patients to derive an overall predicted prevalence of CKD within the health systems.

**Results**—Patients in the VA system were older, had more comorbid conditions, and were more likely to be tested for serum creatinine than those in the M-CARE system. Observed prevalences of CKD stages 3–5 were 15.6% and 0.9% in the VA and M-CARE systems, respectively. Using data from NHANES, the overall predicted prevalences of CKD were 20.4% and 1.6% in the VA and M-CARE systems, respectively.

**Limitations**—Health system data quality was limited by missing data for laboratory results and race. A single estimated glomerular filtration rate value was used to define CKD, rather than persistence over 3 months.

**Conclusions**—Estimation of CKD prevalence within health care systems is feasible, but discrepancies between observed and predicted prevalences suggest that this approach is dependent on data availability and quality of information for comorbid conditions, as well as the frequency of testing for CKD in the health care system.

## INDEX WORDS

Chronic kidney disease; surveillance; prevalence; epidemiology

---

Chronic kidney disease (CKD) is now recognized as a major public health problem.<sup>2,3</sup> It is associated with substantial morbidity and mortality and affects more than 20 million adults in the United States.<sup>4,5</sup> Whereas national registries exist for patients requiring renal replacement therapy, no surveillance system has been available until recently for the much larger population with earlier forms of CKD.<sup>1</sup> A CKD surveillance system ideally should be able to identify individuals with CKD and track the burden, or prevalence, of CKD in the nation as a whole, as well as in specific populations or geographic regions.

Although the National Health and Nutrition Examination Survey (NHANES) has been used effectively to estimate the national prevalence of CKD, lack of geographic granularity limits its utility for assessing prevalence within populations served by individual health care systems. Such information could be of great use for directing and monitoring quality improvement efforts in those populations. With the expansion of electronic medical records across the country, use of data from health care systems may represent an efficient approach to estimating the prevalence of CKD in those populations.<sup>6</sup>

As part of a US Centers for Disease Control and Prevention–sponsored project charged with developing a national surveillance system for CKD,<sup>1</sup> we explored the feasibility of using health care system data to identify persons with CKD and produce estimates of CKD prevalence. We used 2 examples of health care system data, from a 5% national sample of

patients from the US Department of Veterans Affairs (VA) health care system and a regional managed care plan, and compared the results obtained with NHANES-derived estimates of prevalence in those populations.

## METHODS

### Data Sources/Study Participants

**National Health and Nutrition Examination Survey**—The NHANES program has been well described previously.<sup>7</sup> Data used for this report were from the 2005–2006 survey and included information from participants 20 years or older. By design, all participants received testing for serum creatinine. Participants with missing data for serum creatinine (n = 270), who had received dialysis within 12 months of the survey (n = 16), or who were pregnant or menstruating at the time of examination (n = 440) were excluded. All analyses incorporated examination weights, strata, and pseudostrata to reflect the survey sample design.

**VA Health System**—The Veterans Health Administration is the largest integrated health care system in the United States. National data for VA patients are processed at the Austin Automation Center, Austin, TX, which abstracts information from VA facilities such as patient demographics, medical diagnoses and procedures, outpatient and inpatient visits, and vital status.<sup>8</sup> The data files include the Medical SAS Datasets, which contain information for diagnoses and procedures using the *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* and *Current Procedural Terminology* coding systems. In addition, the VA Decision Support System files include pharmacy information and selected laboratory results. For this study, a 5% random sample of national VA data from fiscal year 2005 was extracted. Although a linkage with VA data and Medicare claims is available, this was not used for this study to avoid limiting the population to older patients. For comparability with NHANES data, patients 20 years and older were included in the analyses, and patients with diagnosis or procedure codes indicating pregnancy, dialysis, or kidney transplantation were excluded. To ensure that patients were using the VA health care system, the cohort was limited further to individuals who had at least one outpatient visit to a VA facility in 2005.

**M-CARE**—M-CARE was a managed care health plan in the Southeast Michigan region that was owned and operated by the University of Michigan Health System until December 31, 2006. M-CARE created and maintained a high-quality claims data repository that describes health care utilization since 1997 for its covered population. The database includes billing claims with diagnosis and procedure codes using the *ICD-9-CM* and *Current Procedural Terminology* systems, pharmacy data files, and laboratory results in a subset of patients. For comparability to the NHANES and VA health care system data sets, data from calendar year 2005 were used, patients 20 years and older with at least one claim for an outpatient visit were selected, and those with diagnosis codes indicating dialysis, kidney transplantation, or pregnancy were excluded. A substantial proportion (49.7%) of patients had missing laboratory data (ie, a claim was present for serum creatinine testing, but no result was available). In the subset of patients who had a University of Michigan primary care

physician (UM-PCP), laboratory data were available more consistently (75% of cases with a claim for serum creatinine testing had a laboratory result). Results for M-CARE data were stratified accordingly by whether patients had a UM-PCP; characteristics of the groups were compared.

### Definitions of Study Variables

Detailed descriptions for selected study variables are listed in Table S1 (provided as online supplementary material). The CKD Epidemiology Collaboration (CKD-EPI) creatinine equation was used for calculation of estimated glomerular filtration rate (eGFR; for comparison, results using the 4-variable Modification of Diet in Renal Disease Study equation are listed in Table S2).<sup>9</sup> For this study, CKD then was defined as eGFR <60 mL/min/1.73 m<sup>2</sup> (ie, National Kidney Foundation–Kidney Disease Outcomes Quality Initiative [NKF-KDOQI]<sup>10</sup> stages 3–5), excluding patients on dialysis therapy or after kidney transplantation. No attempt was made to identify CKD stages 1 or 2 because urine protein assessments were available infrequently in the VA and M-CARE data sets. For the VA and M-CARE data sets, only outpatient serum creatinine values were used to minimize misclassification of episodes of acute kidney injury as CKD. If multiple serum creatinine values were available in a year, a single value was randomly chosen. In the VA or M-CARE data only, CKD also was identified separately based on the presence of at least one inpatient or outpatient *ICD-9-CM* diagnosis code (see Table S1). The CKD diagnosis codes were selected based on previous literature and study team consensus.<sup>11,12</sup>

Relevant comorbid conditions identified in the data sets included diabetes mellitus, hypertension, and anemia. Race was classified in NHANES as non-Hispanic white, non-Hispanic black, Mexican American, other Hispanic, and other. Although the National Center for Health Statistics recommends keeping the Hispanic categories separate, the other Hispanic and Mexican American categories were grouped together as Hispanic solely for purposes of comparisons with the VA data set. In the VA data set, race was classified as non-Hispanic white, non-Hispanic black, Hispanic, Asian, Pacific Islander, Native American, and other or unknown, with the latter 5 categories grouped under the heading other/unknown. Race was unavailable in the M-CARE database.

### Statistical Analyses

Demographic information for each of the data sets was tabulated. The proportion of patients undergoing testing for serum creatinine within a year in the VA and M-CARE health care systems was estimated based on the availability of an actual laboratory result and/or the presence of an outpatient claim indicating serum creatinine testing.

The prevalence of CKD stages 3–5 was calculated in NHANES data with adjustment for sampling weights and the survey design. *Prevalence of CKD From Health System Data* In the case of VA and M-CARE data, observed prevalence was calculated as follows: the number of persons with CKD stages 3–5 divided by the number of persons within the health care system 20 years and older who had at least one outpatient visit during the 1-year study period. For the VA and M-CARE data sets, the prevalence of CKD based on claims alone or a combination of claims and eGFR also was calculated.

When race information was missing for the VA and M-CARE data sets, eGFR was calculated using non-Hispanic white race as the default. To improve prevalence estimates when race information was missing, eGFRs also were calculated using a single imputation for black race based on the proportion of blacks in the zip code of residence of the patient (from the US Census<sup>7</sup>) as follows: the race coefficient in the CKD-EPI equation was calculated as  $1 + (0.159 \times \text{proportion of blacks in zip code})$ . To assess the validity of the imputation, a comparison was made between true CKD status (ie, eGFR  $\geq 60$  or  $<60$  mL/min/1.73 m<sup>2</sup>) and CKD status imputed by zip code–based race coefficient in the group of patients with known race. The  $\kappa$  statistic for the comparison was 0.99, suggesting excellent agreement between classifications of CKD based on actual race versus race from the zip code imputation.

Because the observed prevalence estimates in the health care systems were expected to be underestimates due to incomplete testing for serum creatinine, a logistic regression model was fitted using the NHANES data set to predict the presence of CKD stages 3–5 based on age, race (non-Hispanic black vs nonblack), diabetes mellitus, hypertension, anemia, and sex. The coefficients from this model then were applied to the subgroups of patients who did not undergo testing for serum creatinine within the VA and M-CARE data sets to produce estimates of the predicted prevalence of CKD stages 3–5 within the untested populations. These were combined with the observed prevalence estimates based on the tested populations to produce overall estimates of the predicted prevalence of CKD stages 3–5 within the VA and M-CARE health care systems. To assess how well the model predicted CKD stages 3–5 within these health care systems, it also was applied to the subgroup that had undergone testing for serum creatinine in order to compare the predicted prevalence of CKD with the actual observed prevalence of CKD. Within NHANES data, the C statistic for the model was 0.92, with C statistics of 0.81 and 0.87 in the VA and M-CARE populations, respectively. Observed versus predicted prevalence in the tested VA and M-CARE populations across the full strata of patient clinical and demographic characteristics is listed in Table S3.

Because the VA data set contained a substantial proportion of patients with multiple serum creatinine values within a year, the stability of the CKD stage classification based on eGFR estimation was examined. Patients in the VA system with 2 available serum creatinine measurements 3–6 months apart were selected. CKD stage classification based on eGFR at the first measurement was cross-tabulated with CKD stage classification at the second measurement to determine the proportion of patients for whom CKD stage classification remained stable.

All analyses were performed with SAS, version 9.2 (SAS Institute Inc). The study protocol was approved by the institutional review boards of the University of Michigan and VA Ann Arbor Healthcare System.

## RESULTS

### Comparison of Population Characteristics

Table 1 presents a comparison of population characteristics across the NHANES (ie, general United States), VA, and M-CARE data sets. Compared to the NHANES population, the VA system included a predominantly male, older population with a higher prevalence of comorbid conditions, such as diabetes mellitus and hypertension. The M-CARE population included a generally younger, healthier population.

### Proportion of Patients Tested

The proportion of patients tested for serum creatinine was significantly higher in the VA system (75.8%) than in the M-CARE UM-PCP (36.6%) or non-UM-PCP (41.3%) populations (Table 2). In both the VA and M-CARE populations, testing increased with age and was higher in patients with diabetes mellitus or hypertension, with persisting differences between the VA and M-CARE systems. The proportion tested for albuminuria was very low regardless of health care system; 10.3% had measured albuminuria in the VA system (35.6% with diabetes and 14.1% with hypertension) and 6.0% had albuminuria measurement in the overall M-CARE population (45.2% in patients with diabetes and 17.1% in patients with hypertension).

### Estimated CKD Stages 3–5 Prevalence in VA and M-CARE Systems

Table 3 lists observed estimates of the prevalence of CKD stages 3–5 in the NHANES, VA, and M-CARE (including only the subset with a UM-PCP) data sets. Observed prevalence was highest in the VA, intermediate in NHANES, and lowest in the M-CARE system. In all 3 data sets, prevalence was higher with increasing age, presence of diabetes mellitus, and presence of hypertension. Imputation of race in the M-CARE and VA data sets led to a small reduction in the estimated prevalence of CKD stages 3–5 (ie, an increase in population mean eGFR) in the VA system only compared to the assumption that all patients with unknown race were white. CKD prevalence based on *ICD-9-CM* diagnosis claims alone produced substantially lower estimates than those based on eGFR calculation. A definition of CKD based on eGFR and/or diagnosis claims produced prevalence estimates slightly higher than for eGFR alone for both the VA and M-CARE data sets.

Figure 1 shows results for the VA and M-CARE health systems comparing the observed prevalence of CKD stages 3–5 versus the prevalence predicted based on a model using NHANES data. To evaluate how well the model predicted CKD stages 3–5 in the health care systems, the predicted and observed prevalences of CKD stages 3–5 initially were compared using only patients who had undergone testing for serum creatinine (Fig 1, left panel). The 2 estimates were virtually identical for the M-CARE data, but the predicted prevalence of CKD stages 3–5 was slightly higher than the observed prevalence within the VA data. There were substantial gaps in both the VA and M-CARE systems between observed and predicted prevalences in the overall population (Fig 1, right panel), with the observed prevalence being roughly two-thirds of the predicted prevalence.

### Stability of CKD Staging

Table 4 presents the stability of CKD staging by eGFR in the VA system among the subset of individuals who had 2 serum creatinine results available 3–6 months apart ( $n = 26,080$ ; mean age, 66.1 years; diabetes, 39.6%; and hypertension, 83.0%). In those with an initial eGFR  $>60$  mL/min/1.73 m<sup>2</sup>,  $>90\%$  remained stable. For those with CKD stages 3–5 at initial assessment, two-thirds to three-quarters remained within the same classification at assessment 3–6 months later. In most cases in which stage classification was different, individuals were classified into a less severe stage (ie, with a higher eGFR). For example, in those initially classified with stage 3 CKD, 20% were classified as not having CKD on the basis of subsequent serum creatinine–based eGFR calculation.

## DISCUSSION

This study assessed the feasibility of producing estimates of CKD prevalence using available data from a large national and a large regional health care system. A number of potential issues and challenges were identified. One major issue was the choice of denominator for the calculation of prevalence. In NHANES, all participants undergo testing for serum creatinine by design, allowing estimation of the “true prevalence” of CKD within the population. The numerator of the prevalence calculation includes persons with CKD and the denominator is simply the entire survey population (using appropriate survey weights when estimating national prevalence). When using health care system data, estimating prevalence is more complex because testing for serum creatinine generally is performed only for clinical indications. Using a denominator that includes only those tested, as some studies have done,<sup>13,14</sup> therefore would tend to overestimate the true prevalence of CKD. We instead chose to use a denominator that included the entire population available (with noted exclusions) in the health care data set. However, the extent to which this prevalence approximates the “true prevalence” will depend on testing practices. In an idealized scenario, if untested persons were uniformly free of CKD, the prevalence of CKD as observed in the health care system data would equal the true prevalence of CKD in the population under consideration. To the extent that some persons with CKD are not tested, the prevalence based on health care system data would be an underestimate of the true prevalence. We used modeling based on NHANES data to assess the degree to which this underestimation occurred by examining the predicted prevalence of CKD in the untested populations of the health care systems. In the context of this study, the observed prevalence of CKD stages 3–5 in the total population was a third lower than the predicted prevalence (Fig 1), even within the VA system with its robust electronic health record and implementation of clinical reminders.<sup>15</sup> This highlights the ongoing need to optimize strategies for appropriate testing and identification of persons with CKD in both the general population and high-risk groups.

Another important issue is the availability of laboratory data within the health care system for calculation of eGFR. Previous studies have shown that diagnosis claims for CKD, although relatively specific, have poor sensitivity for the presence of CKD, potentially leading to substantial underestimation of prevalence.<sup>12,13</sup> In this study, prevalence estimates for CKD based on diagnosis claims alone were 60% lower than those based on eGFR in the



VA and M-CARE systems. Even when laboratory results generally are available in the health care system, issues with missing values can exist, as we noted in this study within the M-CARE system. Although we were able to identify a subset of patients (with UM-PCPs) who more consistently had laboratory results available, the data were still incomplete, and this likely also contributed to the large gap in observed versus predicted prevalence within the M-CARE system. Finally, error also could be introduced from variations in assays and lack of standardization for measurement of serum creatinine across laboratories used by health care systems.<sup>16</sup>

The quality of the data for patient characteristics relevant to CKD also is important. Missing information for race is a frequent issue in health care system data because it often is not routinely or accurately collected. We used a method to impute race from census data on the proportion of blacks within the zip code of the patient's residence, similar to what others have done previously.<sup>17</sup> This method produces results that are reasonably consistent with those based on actual patient-level race information. Comparisons across health care systems or with NHANES data also require accurate information for comorbid conditions to adequately adjust for differences in population characteristics that may influence the prevalence of CKD. Even when some information for comorbid conditions is available, it may not be easily comparable across health care systems or compared with NHANES data. In the context of this study, for example, the presence of diabetes mellitus was necessarily defined differently in NHANES versus M-CARE or VA data (see Table S1). As a result, inadequate adjustment for comorbid conditions could have led to errors in estimation of the predicted prevalence of CKD.

A final important point relates to the stability of eGFR assessments and the definition of CKD. The formal definition of CKD requires a reduced GFR for at least 3 months, although this often is not strictly followed, such as with NHANES data, which includes only one measurement of serum creatinine.<sup>18</sup> When we examined this issue in the VA data, we found that for patients with at least 2 serum creatinine measurements 3–6 months apart, nearly 20% initially identified as stage 3 CKD by eGFR would no longer be classified as having CKD based on their second measurement. However, the absolute change in eGFR between the 2 assessments was relatively modest (median difference, 6.5 mL/min/1.73 m<sup>2</sup>) and likely of minimal clinical significance, with the change in classification occurring predominantly in patients with eGFRs close to the top of the range of their initial CKD stage. Furthermore, restricting the sample of patients to those with at least 2 serum creatinine measurements introduces a substantial selection bias, tending to include a sicker population of patients. As such, it probably is reasonable to base estimates of CKD prevalence in the health care system on a single outpatient measurement of serum creatinine. However, users of the data should be made aware that the estimates do not meet the strict requirement of chronicity.

In summary, estimation of the prevalence of CKD within health care systems, although feasible in principle, remains a challenge due to issues related to the definition of the denominator and, in particular, the quality of the data. With the recent legislated emphasis on electronic health records, the availability of data relevant to these efforts is likely to expand in coming years. Our findings suggest that important areas for improvement for electronic health care data include: ensuring the availability and completeness of laboratory

results, collecting detailed information for race, and developing consistent definitions for recording comorbid conditions. Nevertheless, health care system data in its present state can be useful for surveillance relevant to the care of CKD. Within health care systems with relatively complete laboratory data, such as in the VA system, we have shown that it is possible to derive estimates of prevalence, supplemented with data from NHANES. With appropriate analysis and inference, this could be used by health care systems to track the burden of CKD within the covered populations over time, to implement disease management strategies, or to help assess the potential impact of interventions directed at preventing or treating CKD. It also is possible to estimate the magnitude of and trends in potential gaps in testing practices for CKD by comparing observed prevalence to predicted estimates based on adjusted NHANES data, which can be used by health care systems to direct and monitor quality improvement efforts.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

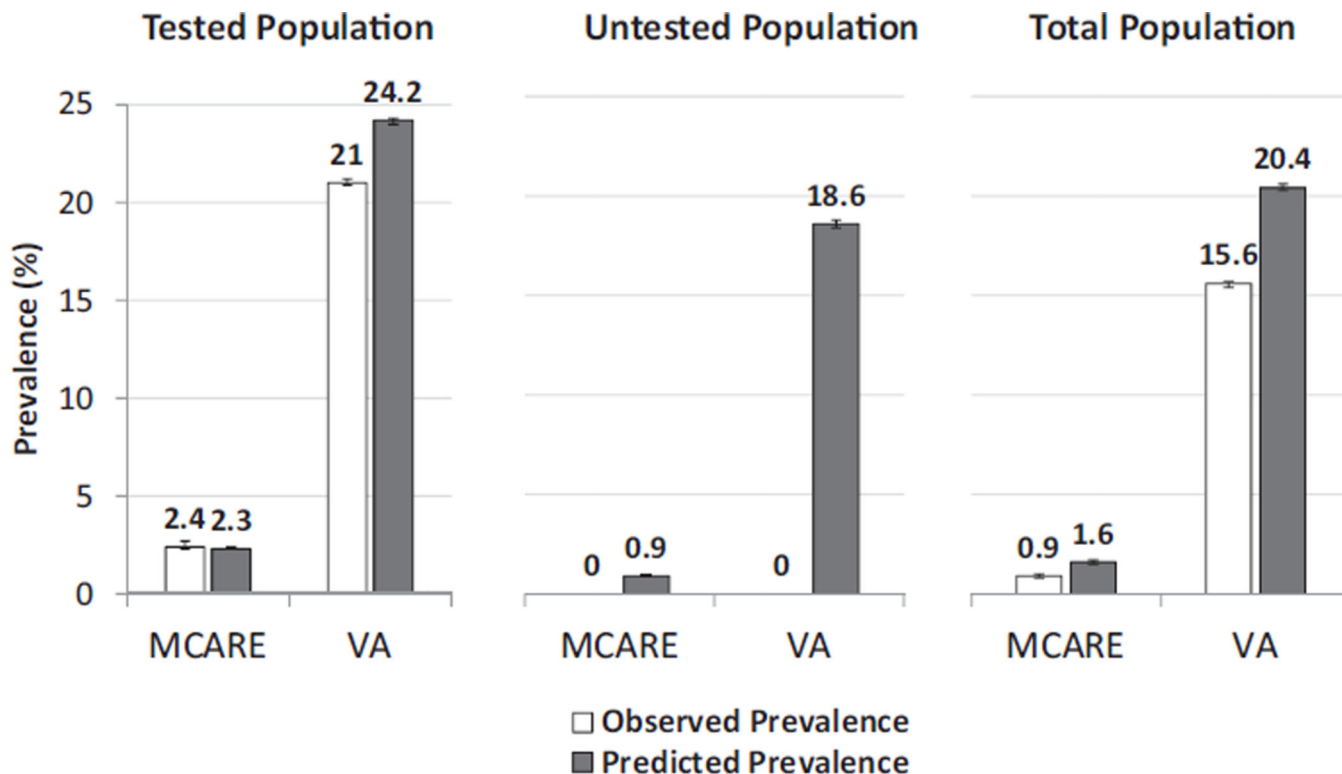
The CDC CKD Surveillance System group consists of the following members. Centers for Disease Control and Prevention: Desmond Williams, Mark Eberhardt, Kristina Ernst, Nicole Flowers, Linda Geiss, Regina Jordan, Nilka Ríos Burrows, Juanita Mondesire, Bernice Moore, Gary Myers, Meda Pavkov, Deborah Rolka, Susan Hailpern, Sharon Saydah, Anton Schoolwerth, Rodolfo Valdez, and Paul Eggers (National Institutes of Health); University of California San Francisco: Neil Powe (Principal Investigator), Tanushree Banerjee, Kirsten Bibbins-Domingo, L. Ebony Boulware, Josef Coresh, Chi-Yuan Hsu; Andrew Levey (Tufts Medical Center and Tufts School of Medicine), Charles McCulloch, Edgar R. Miller, III, Laura Plantinga, Annie Rein-Weston, Lesley Stevens (Tufts Medical Center and Tufts School of Medicine); University of Michigan: Rajiv Saran (Principal Investigator), Brenda Gillespie, Elizabeth Hedgeman, William Herman, Michael Heung, Yi Li, Hal Morgenstern, Bruce Robinson (Arbor Research Collaborative for Health), Vahakn Shahinian, Diane Steffick, Anca Tilea, Jerry Yee (Henry Ford Health System), and Eric Young (Department of Veterans Affairs).

*Support:* This work is funded by the Centers for Disease Control and Prevention (CDC), grant 1U58DP003836 for the project titled “Establishing a Surveillance System for Chronic Kidney Disease.” This material is the result of work partially supported with resources and the use of facilities at the VA Ann Arbor Healthcare System. Publication and report contents are solely the responsibility of the authors and do not necessarily represent the official views of the CDC, the VA, or the US Government.

## REFERENCES

1. CKD Surveillance Team. Establishing a national chronic kidney disease surveillance system for the United States. *Clin J Am Soc Nephrol.* 2010; 5:152–161. [PubMed: 19965534]
2. Schieppati A, Remuzzi G. Chronic renal diseases as a public health problem: epidemiology, social, and economic implications. *Kidney Int Suppl.* 2005; 98:S7–S10. [PubMed: 16108976]
3. Schoolwerth AC, Engelgau MM, Hostetter TH, et al. Chronic kidney disease: a public health problem that needs a public health action plan [abstract]. *Prev Chronic Dis.* 2006; 3:A57. [PubMed: 16539798]
4. Coresh J, Selvin E, Stevens LA, et al. Prevalence of chronic kidney disease in the United States. *JAMA.* 2007; 298:2038–2047. [PubMed: 17986697]
5. 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; 351:1296–1305. [PubMed: 15385656]
6. Hillestad R, Bigelow J, Bower A, et al. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Aff (Millwood).* 2005; 24:1103–1117. [PubMed: 16162551]

7. United States Census Bureau. [Accessed February 21, 2013] 2010 Census. [http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid\\_DEC\\_10\\_SF2\\_PCT1&prodType\\_table](http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid_DEC_10_SF2_PCT1&prodType_table).
8. Maynard C, Chapko MK. Data resources in the Department of Veterans Affairs. *Diabetes Care*. 2004; 27(suppl 2):B22–B26. [PubMed: 15113778]
9. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009; 150:604–612. [PubMed: 19414839]
10. National Kidney Foundation. K/DOQI Clinical Practice Guidelines for Chronic Kidney Disease: evaluation, classification, and stratification. *Am J Kidney Dis*. 2002; 39 suppl 1(2):S1–S266. [PubMed: 11904577]
11. 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; 16:489–495. [PubMed: 15590763]
12. Winkelmayer WC, Schneeweiss S, Mogun H, et al. Identification of individuals with CKD from Medicare claims data: a validation study. *Am J Kidney Dis*. 2005; 46:225–232. [PubMed: 16112040]
13. Kern EF, Maney M, Miller DR, et al. Failure of *ICD-9-CM* codes to identify patients with comorbid chronic kidney disease in diabetes. *Health Serv Res*. 2006; 41:564–580. [PubMed: 16584465]
14. O'Hare AM, Bertenthal D, Covinsky KE, et al. Mortality risk stratification in chronic kidney disease: one size for all ages? *J Am Soc Nephrol*. 2006; 17:846–853. [PubMed: 16452492]
15. Patel TG, Pogach LM, Barth RH. CKD screening and management in the Veterans Health Administration: the impact of system organization and an innovative electronic record. *Am J Kidney Dis*. 2009; 53(suppl 3):S78–S85. [PubMed: 19231765]
16. Miller WG. Estimating glomerular filtration rate. *Clin Chem Lab Med*. 2009; 47:1017–1019. [PubMed: 19728841]
17. Guessous I, McClellan W, Vupputuri S, Wasse H. Low documentation of chronic kidney disease among high-risk patients in a managed care population: a retrospective cohort study. *BMC Nephrol*. 2009; 10:25. [PubMed: 19758452]
18. Levey AS, Eckardt KU, Tsukamoto Y, et al. Definition and classification of chronic kidney disease: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO). *Kidney Int*. 2005; 67:2089–2100. [PubMed: 15882252]



**Figure 1.** Observed versus predicted prevalence of chronic kidney disease (CKD) stages 3–5 in the M-CARE and Veterans Affairs (VA) health care systems. Predicted prevalence was derived from a logistic regression model fitted using the National Health and Nutrition Examination Survey (NHANES) 2005–2006 data set to predict the presence of CKD stages 3–5 based on age, race (non-Hispanic black vs nonblack), diabetes mellitus, hypertension, anemia, and sex. The coefficients from this model were used to produce estimates of the prevalence of CKD within the VA and M-CARE cohorts based on the respective characteristics of those populations. Observed and predicted prevalence are shown for both the population tested for serum creatinine (Tested Population) and those not tested (Untested Population). The predicted prevalence presented for the total population is a combination of the observed prevalence in the tested population and the prevalence in the untested population based on the NHANES model. The vertical thin bars represent 95% confidence intervals.

**Table 1**

Characteristics of the VA Health System, M-CARE, and NHANES 2005–2006 Populations

	NHANES 2005–2006 (N = 4,107)	VA (N = 203,820)	M-CARE	
			Non-UM-PCP (n = 76,508)	UM-PCP (n = 41,957)
Age (y)	47.3 ± 0.7	64.0 ± 14.9	41.2 ± 11.8	41.7 ± 11.9
Age category				
20–29 y	16.8	2.5	19.4	19.1
30–39 y	18.7	4.2	26.1	25.4
40–49 y	22.2	9.8	28.4	27.0
50–59 y	18.6	23.8	19.6	21.2
60–69 y	11.4	20.3	6.2	6.8
70 y	12.3	39.4	0.5	0.4
Male sex	49.5	95.0	45.5	42.1
Race				
Non-Hispanic white	72.9	41.1	NA	NA
Non-Hispanic black	10.9	8.1	NA	NA
Hispanic <sup>a</sup>	11.1 <sup>a</sup>	3.3	NA	NA
Other/unknown	5.1	47.6	NA	NA
Diabetes mellitus	9.6	24.1	9.4	5.4
Hypertension	40.0	65.3	20.8	12.3

Note: Values for categorical variables are given as percentages; values for age, as mean ± standard deviation.

Abbreviations and definitions: M-CARE, a managed care health plan in the Southeast Michigan region owned and operated by UM Health System until 2006; NA, not available; NHANES, National Health and Nutrition Examination Survey; PCP, primary care physician; VA, Veterans Affairs health care system; UM, University of Michigan.

<sup>a</sup>The National Center for Health Statistics recommends keeping Mexican American Hispanic and other Hispanic categories separate for reporting purposes for NHANES surveys prior to 2007, but these categories were collapsed for comparability to the VA database.

**Table 2**

Proportion of Population Who Underwent Testing for Serum Creatinine

	VA	M-CARE	
		Non-UM-PCP	UM-PCP Only
Overall	75.8 (75.6–76.0)	41.3 (41.0–41.7)	36.6 (36.2–37.1)
Age category			
20–29 y	51.3 (49.9–52.7)	25.2 (24.5–25.9)	21.6 (20.7–22.5)
30–39 y	60.8 (59.7–61.8)	34.1 (33.4–34.8)	27.1 (26.3–28.0)
40–49 y	68.7 (68.0–69.3)	43.6 (42.9–44.3)	38.0 (37.1–38.9)
50–59 y	78.9 (78.5–79.2)	55.7 (54.9–56.5)	51.3 (50.3–52.3)
60–69 y	81.8 (81.4–82.2)	64.7 (63.3–66.1)	60.6 (58.8–62.4)
70 y	75.8 (75.5–76.1)	61.8 (56.7–66.9)	74.1 (67.6–80.7)
Sex			
Male	76.1 (75.9–76.3)	42.1 (41.6–42.6)	37.6 (36.9–38.3)
Female	70.3 (69.4–71.2)	40.6 (40.2–41.1)	35.9 (35.3–36.5)
Race			
Non-Hispanic white	82.0 (81.8–82.3)	NA	NA
Non-Hispanic black	79.4 (78.7–80.0)	NA	NA
Hispanic <sup>a</sup>	83.1 (82.2–84.0)	NA	NA
Other/unknown	69.4 (69.1–69.6)	NA	NA
Diabetes mellitus			
Yes	88.7 (88.4–89.0)	78.0 (77.1–79.0)	79.1 (77.4–80.8)
No	71.7 (71.5–71.9)	37.5 (37.1–37.9)	34.2 (33.7–34.7)
Hypertension			
Yes	84.6 (84.4–84.8)	72.6 (71.9–73.3)	77.5 (76.3–78.6)
No	59.3 (58.9–59.6)	33.1 (32.7–33.4)	30.9 (30.4–31.4)

*Note:* Based on available laboratory results or claims in M-CARE and the VA health system. Values are given as percentage tested (95% confidence interval).

Abbreviations and definitions: M-CARE, a managed care health plan in the Southeast Michigan region owned and operated by UM Health System until 2006; NA, not available; PCP, primary care physician; VA, Veterans Affairs health care system; UM, University of Michigan.

<sup>a</sup>The National Center for Health Statistics recommends keeping Mexican American Hispanic and other Hispanic categories separate for reporting purposes for NHANES surveys prior to 2007, but these categories were collapsed for comparability to the VA database.

Observed Prevalence of Individuals With CKD Stages 3–5 in the VA Health System and M-CARE and Estimated US Prevalence From NHANES 2005–2006

Table 3

Characteristic	VA			M-CARE <sup>a</sup>		NHANES 2005–2006
	CKD 3–5 by eGFR	CKD by Diagnosis Claim	CKD 3–5 by eGFR or Diagnosis Claim	CKD 3–5 by eGFR	CKD by Diagnosis Claim	CKD 3–5 by eGFR
Overall <sup>b</sup>	15.6 (15.4–15.7)	6.1 (6.0–6.2)	17.6 (17.4–17.7)	0.9 (0.8–1.0)	0.5 (0.4–0.5)	7.0 (5.4–8.6)
CKD stage						
3	14.1 (14.0–14.3)	—	—	0.8 (0.8–0.9)	—	6.5 (5.0–8.0)
4	1.2 (1.1–1.2)	—	—	0 (0–0.1)	—	0.4 (0.2–0.6)
5	0.3 (0.3–0.3)	—	—	0 (0–0)	—	0.1 (0–0.1)
Overall <sup>b</sup> (missing race imputed)	15.3 (15.1–15.4)	—	17.3 (17.1–17.5)	0.9 (0.8–0.9) <sup>c</sup>	—	—
Age category						
20–29 y	0.1 (0–0.1)	0.3 (0.2–0.5)	0.3 (0.2–0.5)	0 (0–0.1)	0.2 (0.1–0.3)	— <sup>d</sup>
30–39 y	0.5 (0.4–0.7)	1.0 (0.8–1.2)	1.3 (1.0–1.5)	0.2 (0.1–0.2)	0.4 (0.3–0.5)	— <sup>d</sup>
40–49 y	1.6 (1.4–1.7)	1.6 (1.4–1.8)	2.7 (2.4–2.9)	0.4 (0.3–0.5)	0.4 (0.2–0.5)	— <sup>d</sup>
50–59 y	5.1 (4.9–5.3)	3.8 (3.6–4.0)	7.1 (6.9–7.3)	1.7 (1.5–2.0)	0.8 (0.6–1.0)	4.3 (1.8–6.9)
60–69 y	11.8 (11.5–12.1)	5.9 (5.6–6.1)	14.2 (13.9–14.5)	4.4 (3.7–5.2)	0.9 (0.6–1.3)	10.4 (7.8–13.0)
70 y	29.9 (29.6–30.2)	9.6 (9.4–9.8)	32.2 (31.9–32.5)	13.2 (8.1–18.3)	2.3 (0–4.5)	38.9 (34.7–43.1)
Sex						
Male	16 (15.8–16.1)	6.3 (6.2–6.4)	18 (17.9–18.2)	0.9 (0.8–1.1)	0.7 (0.5–0.8)	6.0 (4.4–7.6)
Female	8.1 (7.6–8.6)	1.8 (1.6–2.1)	9.0 (8.4–9.5)	0.9 (0.8–1.0)	0.3 (0.3–0.4)	7.9 (5.7–10.2)
Race						
Non-Hispanic white	17.7 (17.4–17.9)	7.1 (6.9–7.2)	19.9 (19.6–20.2)	NA	NA	8.3 (6.5–10.1)
Non-Hispanic black	9.8 (9.4–10.3)	8.3 (7.9–8.7)	13.3 (12.8–13.8)	NA	NA	5.7 (3.7–7.7)
Hispanic <sup>e</sup>	13.5 (12.7–14.3)	6.9 (6.3–7.5)	15.6 (14.7–16.4)	NA	NA	1.7 (0.8–2.7)
Other/unknown	14.9 (14.7–15.1)	4.8 (4.7–5.0)	16.4 (16.2–16.7)	NA	NA	—
Diabetes mellitus						
Yes	24.5 (24.1–24.9)	12.3 (12.0–12.6)	28.6 (28.2–29)	5.3 (4.4–6.3)	2.5 (1.9–3.2)	18.9 (15.1–22.7)

Characteristic	VA		M-CARE <sup>a</sup>		NHANES 2005–2006
	CKD 3–5 by eGFR	CKD by Diagnosis Claim	CKD 3–5 by eGFR	CKD by Diagnosis Claim	CKD 3–5 by eGFR
No	12.7 (12.6–12.9)	4.1 (4.0–4.2)	14.1 (13.9–14.3)	0.6 (0.6–0.7)	5.7 (4.1–7.3)
Hypertension					
Yes	21.7 (21.4–21.9)	8.6 (8.5–8.8)	24.4 (24.1–24.6)	4.9 (4.3–5.5)	15.0 (11.7–18.4)
No	4.1 (4.0–4.3)	1.3 (1.2–1.4)	4.8 (4.7–5.0)	0.3 (0.3–0.4)	1.6 (1.0–2.2)

*Note:* Values given as percentage (95% confidence interval).

Abbreviations and definitions: CKD, chronic kidney disease; CKD 3–5, CKD stages 3–5; eGFR, estimated glomerular filtration rate; M-CARE, a managed care health plan in the Southeast Michigan region owned and operated by University of Michigan Health System until 2006; NA, not available; NHANES, National Health and Nutrition Examination Survey; VA, Veterans Affairs health care system.

<sup>a</sup>Patients with non–University of Michigan primary care physicians were excluded due to a high proportion of missing laboratory values in those who had been tested.

<sup>b</sup>For the overall prevalence calculation, all patients with missing race were assumed to be non-Hispanic white. The overall (missing race imputed) prevalence was calculated using race imputation based on zip code, as described in the Methods section.

<sup>c</sup>There were 6,146 patients with missing zip code information and who therefore were assumed to be white for the purposes of the eGFR calculation.

<sup>d</sup>Estimate omitted due to low sample size.

<sup>e</sup>The National Center for Health Statistics recommends keeping Mexican American Hispanic and other Hispanic categories separate for reporting purposes for surveys prior to 2007, but these categories were collapsed for comparability to the VA data set.



**Table 4**

Stage of CKD Based on Second Versus First eGFR Measurement

CKD Stage at First eGFR	CKD Stage at Second eGFR			
	eGFR >60 mL/min/1.73 m <sup>2</sup>	CKD Stage 3	CKD Stage 4	CKD Stage 5
eGFR >60 mL/min/1.73 m <sup>2</sup>	92.64	7.23	0.11	0.01
CKD stage 3	19.68	76.97	3.31	0.04
CKD stage 4	3.10	26.83	65.63	4.43
CKD stage 5	0	6.06	18.18	75.76

*Note:* Among patients with multiple serum creatinine measurements in the Veterans Affairs health care system, the second measurement occurred 3–6 months after first measurement. Values are given as percentages.

Abbreviations: CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript