

Under-documentation of chronic kidney disease in the electronic health record in outpatients

Herbert S Chase,^{1,2} Jai Radhakrishnan,² Shayan Shirazian,² Maya K Rao,² David K Vawdrey¹

¹Department of Biomedical Informatics, Columbia University, New York, New York USA

²Department of Medicine, Columbia University, New York, New York USA

Correspondence to

Herbert Chase, Department of Biomedical Informatics, CUMC, VC-5, 622 West 168th Street, New York, NY 10032, USA; herbert.chase@dbmi.columbia.edu

Received 15 October 2009

Accepted 25 June 2010

ABSTRACT

Objective To ascertain if outpatients with moderate chronic kidney disease (CKD) had their condition documented in their notes in the electronic health record (EHR).

Design Outpatients with CKD were selected based on a reduced estimated glomerular filtration rate and their notes extracted from the Columbia University data warehouse. Two lexical-based classification tools (classifier and word-counter) were developed to identify documentation of CKD in electronic notes.

Measurements The tools categorized patients' individual notes on the basis of the presence of CKD-related terms. Patients were categorized as appropriately documented if their notes contained reference to CKD when CKD was present.

Results The sensitivities of the classifier and word-count methods were 95.4% and 99.8%, respectively. The specificity of both was 99.8%. Categorization of individual patients as appropriately documented was 96.9% accurate. Of 107 patients with manually verified moderate CKD, 32 (22%) lacked appropriate documentation. Patients whose CKD had not been appropriately documented were significantly less likely to be on renin-angiotensin system inhibitors or have urine protein quantified, and had the illness for half as long (15.1 vs 30.7 months; $p < 0.01$) compared to patients with documentation.

Conclusion Our studies show that lexical-based classification tools can accurately ascertain if appropriate documentation of CKD is present in a EHR. Using this method, we demonstrated under-documentation of patients with moderate CKD. Under-documented patients were less likely to receive CKD guideline recommended care. A tool that prompts providers to document CKD might shorten the time to implementing guideline-based recommendations.

INTRODUCTION

Early recognition is key to preventing the progression of chronic kidney disease (CKD) by allowing the implementation of recommended treatments. Multiple studies, conducted in the primary care setting, have shown poor detection of CKD as well as suboptimal adherence to guideline-recommended care.^{1–3} One mechanism to prompt early recognition of CKD might be a clinical decision support system (CDSS), which automatically determines if providers caring for the patients with CKD have mentioned the illness in the patients' electronic health record (EHR). If not, the CDSS could notify the provider and suggest guideline-based recommendations. The

purpose of this study was to develop methods to electronically ascertain if CKD was appropriately documented in the notes of an individual patient's EHR, to test and validate the tool's ability to perform this task, and to use the tool to assess appropriate documentation of CKD in a population of patients with known moderate disease.

BACKGROUND

There is a mounting epidemic of CKD and end-stage renal disease (ESRD) in the US.^{4–5} As of 2002, between 4 million and 20 million Americans were affected with CKD, and about 300 000 were defined as having ESRD or requiring renal replacement therapy.^{6–9} It is estimated that by 2015 the number of patients with ESRD will be 712 000.¹⁰ The total number of expected patients receiving dialysis by 2010 will reach 560 000 resulting in an annual Medicare spending of \$28.3 billion by 2010.¹¹ As of 2007, the total Medicare cost for CKD reached \$57.5 billion.¹²

Patients with CKD are at risk for not only progression to ESRD but also increased cardiovascular morbidity and mortality.^{13–14} The key to preventing either of these two outcomes is recognition of the earliest stages of kidney disease and initiation of a targeted and aggressive management plan. The National Kidney Foundation provides evidence-based clinical practice guidelines for all stages of CKD and related complications,¹⁵ which include a recommendation for referral to a nephrologist if CKD is sufficiently advanced. The importance of a timely referral to a nephrologist is evident in multiple studies that have shown an association with late nephrology referral and poor outcomes when starting hemodialysis.^{16–18} Patients with unrecognized CKD may be referred by their provider at a later stage than a patient with recognized CKD.

Only if providers recognize that their patients have CKD will the appropriate targeted management be initiated. Several investigators have demonstrated considerable under-recognition by primary care practitioners. De Lusignan and colleagues demonstrated that less than 4% of patients with CKD had been coded as having renal disease.¹⁹ Studies conducted by manual chart review (bypassing the known *International Classification of Diseases* (ICD)-9 coding sensitivity issues²⁰) demonstrated that over three-quarters of patients with CKD were not recognized as having CKD.^{1–2–21–23}

A first step in creating a tool to prompt early recognition of CKD is to determine if the provider



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://jamia.bmj.com/site/about/unlocked.xhtml>

has recognized the patient's CKD. The tool could search for appropriate documentation of CKD in the patient's notes as a proxy for recognition. If documentation is lacking, the tool could prompt the provider to re-examine the patient's record thereby potentially increasing awareness of the patient's condition. Because manual review of notes for documentation is not feasible on a large scale, we considered that natural language processing (NLP) based methods would be useful in ascertaining whether patients with CKD had the diagnosis of CKD documented in their electronic outpatient visit. Several groups have successfully used NLP methods to find documentation of specific diseases or conditions (diabetes mellitus, hypertension, obesity, and pneumonia).^{24–31} We reasoned that we could use a similar strategy to assess whether disease documentation was present in the notes of patients with CKD.

The purpose of this study was to develop, validate and use a CKD-documentation-verification tool to determine whether CKD had been appropriately documented in individual outpatient notes in the EHR.

METHODS

Patient records

The data source for this study was the Clinical Data Warehouse (CDW)—the research database of the Columbia University Medical Center (CUMC) of the New York Presbyterian Hospital (NYP) system. The CDW contains a broad range of clinical information such as diagnoses, procedures, discharge summaries, laboratory tests, and pharmacy data for the past 20 years. Our study sample was selected from all the patient records of the years 2003–2006.

Associates of Internal Medicine (AIM) clinic outpatients

The patients whose records were examined in this study were regularly cared for in the AIM primary care clinic on the CUMC campus (see figure 1 for an overview of patient and note selection). Approximately 92% of the patients in the AIM clinic are on Medicaid or Medicare, and some of them lack insurance. During a 4-year period (2003–2006) over 10 000 patients were seen regularly in the AIM clinic; the average age was 61 years and 69% were female. Approximately 20% were Hispanic, 10% were African American, and the remainder patients were White, Asian, or 'Other'.

Patients of the AIM clinic are cared for by approximately 150 CUMC residents in internal medicine, each of whom have individual clinic assignments and carry their own panel of patients during 3 years of training. The AIM clinic is overseen by clinic chiefs who have leadership roles in the department of medicine. In addition, there are approximately 10 senior faculty members who practice in the AIM clinic seeing their own patients and serving as mentors to the residents. During the time period of this study, providers used an EHR (WebCIS³²) developed at CUMC. On a computer in each examination room, providers entered notes using one of four titles to capture the clinical encounter: initial visit; follow-up; summary; or other clinical. These notes were entered as free-text by the provider using one of the four titles described above without the use of templates, pre-populated content, or automatic import of information from the EHR.

Identifying patients with reduced kidney function

Patients in whom CKD may be present were identified by searching the research database for patients with at least two episodes of an elevated serum creatinine of greater than 1.6 mg/dl, during 2003–2006. In addition, all subjects in our cohort had

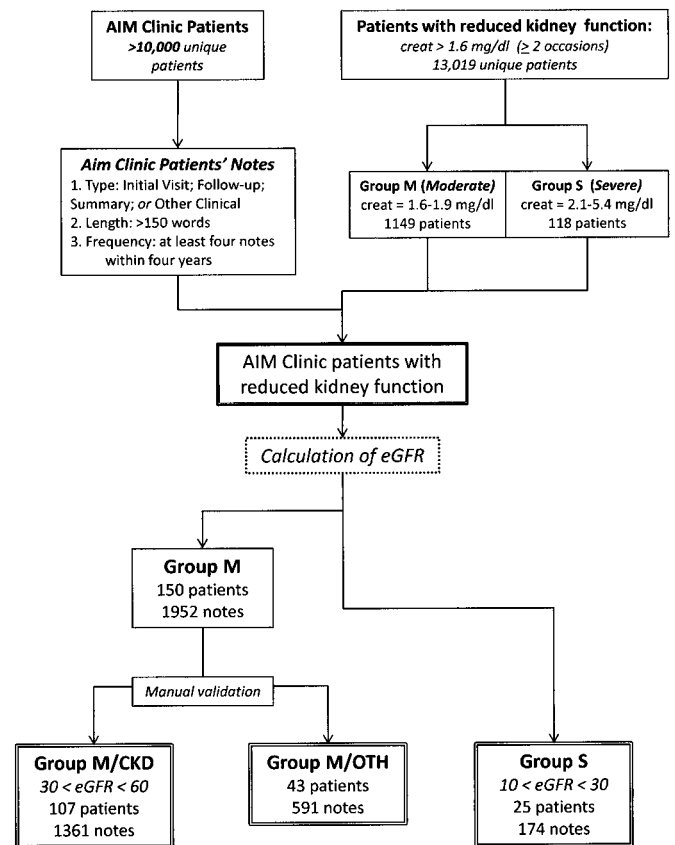


Figure 1 Patient selection criteria study period (patient enrollment and data extractions) was between 1 January 2003 and 31 December 2006.

hypertension as indicated by appropriate ICD-9-CM codes. Patients were stratified on the basis of their level of renal function, reflected in the magnitude of the creatinine: group M (moderately reduced function) had at least two creatinine values between 1.6–1.9 mg/dl; group S (significantly reduced) had at least two creatinine values between 2.6–5.4 mg/ml. A set of control patients, without structural or functional kidney damage or hypertension, was selected by identifying medical record numbers of patients with creatinine values between 0.9 and 1.0 mg/dl reflecting estimated glomerular filtration rate (eGFR) values of greater than 60 ml/min/1.73 m² for the majority of patients.

Extraction of notes

Using patient identification numbers of the group M and S patients, clinic notes of AIM patients were extracted from the CDW (see center box of figure 1) and eGFR for each patient was estimated using the four-variable Modification of Diet in Renal Disease (MDRD) formula³³ based on the patient's creatinine values, gender, race (African-American or not), and age. The group M patients were subdivided into two groups: those with manually verified stage 3 CKD (group M/CKD), defined in the National Kidney Foundation Kidney Disease Outcomes Quality Initiative as an eGFR of less than 60 ml/min/1.73 m² for greater than or equal to 3 months,¹⁵ and those with transient or inconsistently reduced renal function (group M/OTH). Group S patients were defined as having eGFR <30 ml/min/1.73 m². The time course of CKD (how long a patient had CKD) was computed by calculating the eGFR based on the most recent creatinine and on successively older creatinine values until the eGFR was no longer less than 60 ml/min/1.73 m².

We selected only those patients who had been cared for ‘consistently’ as an outpatient, which we defined as having a minimum of four notes in the EHR written over the 4-year period. In a preliminary study we found that approximately 30% of patients had fewer than four notes written over the 4-year period and that many of these patients had only one recorded outpatient note; these patients were not included in the study.

Validation of CKD status

Patients with a reduced eGFR could have either CKD or acute kidney injury (AKI). Validation of disease status (CKD or AKI) was carried out by three experts in nephrology (the first three authors) who manually reviewed the medical records of the group M patients. Agreement in the experts’ classifications was analyzed using the κ statistic.²⁸ In the initial rounds of pair-wise validation there was moderate agreement ($\kappa=0.56$). The disputed patients were those with transplants (cardiac or renal) with exceedingly complex medical histories, often on a variety of nephrotoxic or potentially nephrotoxic agents. By consensus, these patients were omitted from the study population. A final list of 151 unique patients with either CKD or AKI, as determined unanimously by the three nephrologists, was used to explore the CKD-documentation-verification studies.

CKD-documentation-verification tools

Two methods were used to identify reference to CKD in a note: a lexical-based classification method and a word-count method. Both methods consulted a library of words, terms, and abbreviations that providers might use when documenting an encounter with a patient with CKD. The library was created by first identifying terms used by kidney specialists. Starting with approximately 1000 notes written by kidney specialists, we used a word frequency analysis to identify 39 medical terms commonly associated with kidney disease. To identify which of these terms would provide the most specificity for CKD mentions, we ran the IbK and Naïve Bayes classifiers (Weka workbench³⁴) using these terms as attributes, on notes of patients with severe CKD (group S) and controls. Repeat runs of classification identified terms, with low attribute strength, which were removed from the list leaving a shortened list of terms with high attribute strength (table 1). For each term on the final list we used regular expressions to capture variations in the choice of words for documentation. Manual review of notes of patients with known CKD identified predictable terms, such

Table 1 CKD-related terms used by classification algorithm (Naïve Bayes) or word-count methods

Category	CKD terms (attribute strength)
CKD syndrome	CKD (202.4)
	CRI (687.0)
	CRF (142.6)
	Chronic (kidney or renal) (211.4)
Reduced renal function	(Renal or kidney) insufficiency (164.2)
	eGFR (27.2)
	GFR (252.3)
Damage to parenchyma	Glomerulopathy (38.2)
	Nephrotic (104.8)
	Proteinuria (296.6)

CKD, chronic kidney disease; CRI, chronic renal insufficiency; CRF, chronic renal failure; GFR, glomerular filtration rate; eGFR, estimated GFR. Attribute strength was calculated by the χ^2 AttributeEval method, which evaluates predictive value of attribute in classification by computing the χ^2 statistic with respect to the class.

as ‘chronic renal insufficiency’, as well as unpredictable ones, such as ‘chronic renal insuffic’.

Categorization of notes

Group M (and control) notes were categorized as either containing documentation of CKD (CKD+) or not (CKD-). The classifier (Naïve Bayes) uses the library of CKD terms as attributes, trains on a gold standard of 174 notes of patients with severe disease (group S), in whom CKD was documented in each and every note, and categorizes individual group M notes (and controls). The word-count method looks for the presence of any of the terms in the library and categorizes as CKD+ if the note contains at least one instance. To score each attribute, the number of times each term appeared in a single chart is counted and normalized to the total number of words in the chart. Notes of fewer than 150 words were not processed because they were usually brief communications between various members of the healthcare team and not intended to document a patient’s visit.

RESULTS

Categorization of individual notes

The CKD classifier categorized 1361 notes belonging to 107 group M patients with verified CKD and 1133 notes of 154 control patients. Of the control patients’ notes, the tool categorized all but 2 of the 1133 as CKD-. However, of the notes of patients with verified CKD, only 708 (52%) of 1361 notes were categorized as CKD+ and the remainder (653) were categorized as CKD- (table 2, lines 1 and 2, respectively).

Categorization of notes of patients with CKD as CKD- could be due to the absence of documentation of CKD in the note or to erroneous categorization where documentation of CKD was actually present in the note but unrecognized by the tool. Of the 653 CKD- notes of patients with verified CKD (table 2, line 2), 59 CKD- notes contained at least one CKD term from table 1 and were, thus, considered to have been incorrectly categorized (table 2, line 3). Examination of these notes indicated that most of the 59 notes contained references to either ‘CRI’ (chronic renal insufficiency), which is not the preferred term of renal physicians (CKD), or ‘proteinuria’. The remaining 594 CKD- notes (653-59) did not contain a single instance of any CKD terms. While this suggested that these remaining notes had been correctly categorized as CKD- due to lacking CKD terms, it was possible that some of these notes actually contained documentation of CKD but expressed in manner not included in the CKD

Table 2 Categorization of notes of patients with verified chronic kidney disease (CKD)

Line	Categorization (CORRECT or INCORRECT)	Classifier	Word count
1	CKD+ notes CORRECT	708	767
2	CKD- notes, some CORRECT, some INCORRECT	653	594
3	CKD- notes containing CKD terms from library INCORRECT	59	0
4	CKD- notes potentially containing CKD terms not in library PROJECTED INCORRECT	3	3
5	CKD- notes INCORRECT (line 3+line 4)	62	3
6	CKD- notes CORRECT (line 2-line 5)	591	591
7	Total correct (line 1+line 6)	1299	1358

Breakdown of notes of patients with verified CKD that were correctly or incorrectly categorized.

Notes were categorized using the classifier or word-count methods as either containing CKD documentation (CKD+) or not (CKD-).

regular expression library (table 1). Manual review of 370 CKD–notes in patients with CKD identified two additional instances such as ‘...worsened renal function’ or ‘...DM (retinopathy, nephropathy...)’ for an overall rate of 0.54%. Based on this rate, it was projected that three additional notes (0.54% of 594 notes belonging to the group M patients originally categorized as CKD–) might have been incorrectly categorized (table 2, line 4). Therefore, the total rate of actual or potential incorrect categorization of CKD–notes was 62 (table 2, line 5, the sum of lines 3 and 4). The remaining 591 CKD–notes were classified as correctly categorized (the difference between line 2 and line 5 in table 1).

Sensitivity and specificity of the classifier were calculated based on the number of correct and incorrect categorization rates in CKD patients and controls (table 3, second column from the right under ‘Classifier’). Totals for correct/incorrect were taken from table 2. Sensitivity and specificity were 95.4% (94.2%–96.5) and 99.8% (99.3–100.0), respectively (table 3).

The word-count method of categorization yielded similar results to the classification method with a significantly higher sensitivity and negative predictive value (tables 2 and 3). However, there were additional false-positives in the control group: 14 notes of control patients with no reduction of renal function were categorized as CKD+. Manual review discovered that seven of these patients were noted to have proteinuria with no mention of CKD (thus explaining why they had been originally categorized as CKD– by the classifier). The remaining seven were considered false-positives (‘the patient does not have CRI...’). Totals of correctly/incorrectly categorized notes using the word-count method are listed in table 2 and sensitivity and specificity of the word-count method, 99.8% (99.3–99.9) and 98.8% (97.9–99.3), respectively, are listed in table 3.

Categorization of group M patients

The notes of individual patients were collected and then categorized as a unit. Appropriate documentation was defined as CKD documentation in the notes at the time the patient had CKD. Of 107 group M patients with verified CKD (group M/CKD, figure 1), 75 (70%, 95% CI 60% to 78%) were appropriately documented. In 20 patients, each and every one of their notes documented CKD. In an additional 55 patients, documentation appeared appropriately at the time the patient developed CKD (as determined by manual review). The notes progressed from no documentation to documentation during the time period that the patient’s eGFR was declining to below the CKD cut-off of 60 ml/min/1.73 m².

The remaining 32 group M/CKD patients lacked appropriate documentation (30%, 95% CI 22% to 39%) despite having developed CKD. Given the small but finite false-negative rate observed in the categorization of individual notes, we manually reviewed all the notes of these 32 patients lacking documentation to determine if any had been misclassified (as a result of the false-negative categorization). Of the 32 patients lacking CKD documentation in their notes, one patient had to be reclassified because several individual notes in the set had been falsely categorized as CKD– by the classifier (but correctly categorized by the word-count method). The remaining 31 (97%) patients were correctly classified as appropriate documentation lacking (95% CI 84.3% to 99.5%). The observed false-negative rate of patient categorization (as opposed to note categorization) was 3% (95% CI 0.6% to 15.8%).

Identification of the provider

One explanation for the lack of appropriate documentation in patients with CKD was that the notes were not written by the primary provider. We manually reviewed all of the notes of the 32 patients lacking appropriate documentation to determine if any or all of the notes were written by someone other than the primary provider. In 24 of the 32 patients, the majority and sometimes all of the notes lacking CKD documentation were written by the primary provider. In the remaining eight patients, notes had been written exclusively by physicians in specialty clinics (mainly cardiology). Thus, 22% of the group M/CKD patients lacked appropriate documentation by their primary providers.

Differences between documented and undocumented patients

We explored the possibility that patients whose CKD had been documented in every note had worse renal function and were cared for differently than those whose CKD had not been documented in any notes. Results in table 4 demonstrate that documented patients had significantly lower eGFR values (and higher average creatinine values) over the 4-year period of observation. However, the proportion of patients with significant proteinuria, was similar in the two groups. Patients with documentation in all their notes had CKD for over twice as long as the cohort with no notes documenting CKD (time-course, table 4). Targets of evidence-based management also seemed to differ between the two groups. Significantly fewer patients lacking documentation were on inhibitors of the renin-angiotensin system (RAS-blockade) or had their urine protein quantified at least once, both of which are Kidney Disease Outcomes

Table 3 Accuracy of categorization of chronic kidney disease (CKD) patients’ notes using classifier or word-count methods

Patient status (number of notes)	Categorization of notes	Categorization method	
		Classifier	Word count
Verified CKD (1361)	CKD+ (correct)	1299	1358
	CKD– (incorrect)	62	3
Control (1133)	CKD+ (incorrect)	2	14
	CKD– (correct)	1131	1119
Sensitivity (95% CI)		95.4% (94.2 to 99.3%)	99.8% (99.3 to 99.9%)
Specificity (95% CI)		99.8% (99.3 to 100)	98.8% (97.9 to 99.3)
Positive predictive value		99.8% (99.4 to 100%)	99.0% (98.2 to 99.4%)
Negative predictive value		94.8% (93.3 to 96.0%)	99.7% (99.7 to 99.9)

The sensitivity, specificity, positive predictive value and negative predictive value of categorization of notes from patients with CKD using the lexical-based classifier or word-count methods. Values for correct or incorrect are from table 2.

Quality Initiative recommendations.¹⁵ Although not reaching statistical significance, fewer had low-density lipoprotein (LDL) levels below the recommended target range of 100 mg/dl. Demographical analysis suggested no significant differences between the two groups.

Documentation of CKD in notes of group M patients without CKD

Documentation of CKD was explored in those patients who did not have CKD but rather AKI (group M/OTH). Of the 43 patients without CKD, 3 (7%) had CKD documentation in all of their notes and 19 (44%) had both CKD+ and CKD- notes. Exhaustive manual review of CKD+ notes belonging to 19 patients with AKI identified two instances of false-positives out of total 136 individual notes (such as the provider recorded ‘...family history: CRI’). However, reclassification of these two notes from CKD+ to CKD- did not alter the overall proportion of patients who were inappropriately documented.

Automated identification of patients with CKD

We explored the possibility that the MDRD-based eGFR calculator, used to determine the time-course of CKD, could be used to automatically identify patients with CKD and distinguish them from patients with AKI. On a subset of group M patients with either CKD or AKI, validated by the experts, eGFR was calculated for every value of creatinine measured over the 4-year time period 2003–2006. Those patients with two measurements of eGFR below the cut-off for stage 3 CKD (<60 ml/min/1.73/m²) of at least 90 days were classified as CKD, the rest AKI. Sensitivity and specificity of this CKD-identification algorithm was 90.7% (76.9–96.9) and 83.3% (61.8–94.5), respectively, and accuracy was 87.0%. We modified the algorithm to determine if the patient had a single instance of a eGFR value of greater than 60 ml/min/1.73 m² during the 90 day period. Fluctuating eGFR (above and below 60 ml/min/1.73 m²) is a hallmark of pre-renal azotemia, common in our patient population, which is indicative of AKI not CKD. The modified algorithm had a lower

Table 4 Comparison of documentation of group M/CKD: patients with CKD documentation in all notes had decreased kidney function and were seen longer by providers

Parameter	All notes documented	No notes documented	Significance
<i>Renal function</i>			
(creatinine), mg/dl	1.54 (0.11)	1.41 (0.14)	p<0.005
eGFR	47 (5.9)	52 (9.4)	p<0.05
Time-course (months)	30.7 (15.4)	15.1 (14.1)	p<0.005
Significant proteinuria*	64%	42%	NS
<i>CKD targets</i>			
RAS blockade lacking	12.5%	36.4%	p<0.05
Urine protein quantification lacking	29%	59%	p<0.05
LDL > 100 mg/dl	35%	55%	p<0.10
<i>Demographics</i>			
Age (y)	70 (13.7)	68 (12.9)	NS
Gender (female)	25%	26.3%	NS
African-American	33.3%	22.7%	NS
Diabetes (ICD-9)	60%	50%	NS

Patients were categorized using classifier (see Methods). Qualitatively similar results were obtained categorizing patients using the word-count method. *Significant proteinuria is defined as ≥2+ protein on routine urine analysis or an albumin/creatinine ratio of 30 µg/mg. CKD, chronic kidney disease; eGFR, estimated GFR; ICD, International Classification of Diseases; LDL, low-density lipoprotein; RAS, renin-angiotensin system.

sensitivity of 83.3% (74.7%–89.6%) but a higher specificity of 92.9% (79.4%–98.1%), and accuracy was 86%. The positive predictive value was 96.8% (90.2–99.2%).

DISCUSSION

Early recognition and prompt implementation of recommended management guidelines are essential to prevent worsening kidney function and cardiovascular morbidity in patients with early stage CKD.⁴ A major hurdle in achieving these goals is what may be the lack of recognition by the primary care physicians that their patients have early stage CKD.^{1 2 19 35} One approach to improving recognition might be through a CDSS, which ascertains if the provider recognizes the patient’s condition. Recognition could be indirectly assessed by the presence or absence of documentation of CKD consisting of words or concepts that communicate the presence of CKD. If providers caring for patients with CKD had not appropriately documented CKD in the patients’ notes, then the CDSS system could notify the provider as well as suggest guideline-based recommendations.

A key element to such a notification system is a tool that determines if CKD has been appropriately documented in the patient’s chart. The purpose of this study was to develop, validate, and use lexical-based methods to ascertain electronically if CKD was appropriately documented in an individual patient’s EHR.

CKD-documentation-verification tools

Our approach to determine if patients were appropriately documented for CKD included two steps: categorizing individual notes as containing CKD documentation or not, and then categorizing patients as appropriately documented or not based on whether CKD documentation corresponded to when the patient had CKD. Categorization of individual notes by the classification or the word-search methods achieved high accuracy, sensitivity, and specificity (table 3). The false-negative rate was higher using the classifier because the classifier weighed and favored certain terms (such as ‘CKD’). A note containing a less definitive reference to CKD (such as ‘proteinuria’ or even ‘chronic renal insufficiency (CRI)’) was not scored as high as CKD resulting in some of the notes being categorized as CKD-. The low scoring for ‘CRI’ was due to the fact that the gold standard notes were of patients with severe CKD nearly all of whom were being taken care of by nephrology attending or fellows who used the preferred term ‘CKD’. Categorization using the simple word-count engine eliminated these false negatives by crediting ‘CRI’ but in turn increased the false positive rate. Notes of controls with an occasional ‘the patient does not have CRI’ were not categorized by the classifier as CKD+.

Categorization of individual patients achieved a similar level of accuracy. Manual review of the set of patients’ notes of patients lacking appropriate documentation identified 1 of 32 which had been incorrectly categorized by the classifier (for an overall rate of 3%). This patient was correctly categorized as appropriate using the word-search engine.

Appropriate documentation in patients with moderate CKD

We found that 32 of 107 patients (30%) with verified moderate CKD lacked appropriate documentation of CKD. Manual review determined that of these 32 patients, 24 (22%) had been followed closely and regularly by their primary care provider. This proportion of patients lacking documentation is substantially lower than the 78% rate observed in a study of primary providers employing manual chart review.^{1 2} One explanation of

the discrepancy is that in these prior studies providers were members of an outpatient family medicine practice whereas the majority of our providers were internal medicine residents under the watchful eye of senior attending physicians. The residents are in a demanding training program and are routinely lectured to about standards of care and following guidelines. Another explanation for our higher rates of documentation is that all of our patients had hypertension; providers may have been more aware of CKD in this higher risk group of patients.

The discrepancy might also be due to the fact that prior studies measured recognition whereas we measured only documentation. The prior studies required demonstration of follow-up or referral of the patient, which would only take place if the provider recognized the patient's CKD. Our classifier searched only for the presence of words and gave credit for merely noting 'CKD' with no additional requirement of acting on this observation. However, our studies suggest that documentation and action might be associated (table 4). Patients whose CKD was completely undocumented in any of their notes, despite having CKD for, on average, 15 months, were significantly less likely to be on an inhibitor of the RAS or have had the protein in their urine quantified, key recommendations promulgated by the National Kidney Foundation¹⁵ and the Canadian Society of Nephrology.³⁶

Our studies also demonstrated evidence of inappropriate documentation. Of 43 patients without CKD (AKI, instead), 22 (51%) had some CKD documentation in their notes. Three of the forty-three (7%) patients had documentation of CKD in every single one of their notes. Misdiagnosis of patients with AKI as CKD obviously sets in motion an inappropriate diagnostic and treatment plan, delays the identification of the cause of AKI, and, in this era of 'cut and paste' electronic record keeping, has the potential to propagate an erroneous diagnosis.

Automating CKD-documentation-verification

Our studies suggest that it may be feasible to ascertain automatically in real time if patients have CKD and whether or not their condition is being appropriately documented in the EHR. Creatinine values and demographical information could be extracted enabling calculation of eGFR, as is already being done at many institutions.³⁷ A CKD-identification tool, such as the one we used to calculate time-course (see Methods), could automatically identify patients with early CKD. The individual notes of these patients could then be processed and categorized using the simple word-count method, in consultation with an ever expanding library of suitable ways to communicate CKD. The patient's set of notes could then be categorized as appropriate or lacking CKD documentation by ascertaining if documentation and eGFR less than 60 ml/min/1.73 m² were contemporaneous. The providers of those patients lacking appropriate documentation could then be prompted to do so.

However, there are several potential sources of error in the proposed automated CKD-verification-assessment tool, which would have to be addressed before implementing. First, the CKD-identification tool had a positive predictive value of 96.8% which means that of patients identified as having CKD based on automated calculation of eGFR, as many as 4% might be incorrectly identified as CKD when they actually had AKI. Second, some individual notes of patients with bona fide CKD would be incorrectly categorized as lacking documentation (CKD-). Although both the classifier and word-search tool had a high negative predictive values (94.8 and 99.7%, respectively), a small but finite number of notes would, nevertheless, be incorrectly categorized. These misclassifications of individual

notes would in turn lead to patients being incorrectly categorized as lacking appropriate documentation. Our manual review suggested that 3% (95% CI 2% to 15%) of patients might be incorrectly categorized using the classification tool.

Given the well-known overrides of medication alerts,^{38–40} each of these steps would have to be modified to achieve higher accuracy before implementation as a documentation-verification and notification tool. However, it is not clear that the positive predictive value of the CKD-recognition tool could be improved. There are patients with various conditions, such as heart failure, who sustain long periods of reduced renal function only to improve months later. Perhaps a semi-automated approach, where kidney experts review possible CKD candidates, would eliminate false-positive identifications of CKD. The negative predictive value of the two categorization steps could be improved by continually updating the CKD library to include unconventional terms that are infrequently used to communicate CKD.

LIMITATIONS

There are several major limitations of this study. First, the study size of both patients and providers is small. Conclusions regarding the prevalence of inappropriate documentation or the association of documentation and action must be considered specific to our institution where the study population was comprised of clinic patients cared for at a tertiary medical center by providers who were enrolled in a rigorous internal medicine residency program. Confirmatory studies with different populations of patients and providers would be required to assess the degree to which appropriate documentation is present in the EHR.

Second, any conclusions regarding a possible association between provider documentation and action (table 4) must formally account for possibly confounding factors, such as the presence of diabetes (for which RAS inhibition is recommended irrespective of the degree of proteinuria), hypertension, or age. While there were no significant differences between fully documented and fully undocumented CKD patients regarding these factors, multivariate analysis on a larger sample would be required to enable a more definitive conclusion.

A third limitation is that the false-negative rate cannot be known given the infinite number of ways (including misspellings) that a provider could refer to CKD in an electronic note. We attempted to project an upper limit to the false-negative rate by taking into account the false negatives we discovered on manual review of hundreds of notes (table 2). Nevertheless, providers document CKD (and other illnesses) in unconventional ways, which cannot be anticipated. False positives (negation such as 'patient has no CKD') are not an issue in a CKD-CDSS notification system because the patients have already been determined to have CKD by the CKD-identification tool. However, if the tool was used to confirm the presence of CKD, the false-positive rate would have to be reduced. There are several successful approaches to identifying negation^{27 31 41} or locating where terms are within a note (family history)^{42 43} that can be integrated into our current CKD-identification algorithm.

CONCLUSION

Our studies show that it is possible to ascertain electronically if patients' CKD has been appropriately documented in their clinic notes using lexical-based classification tools. We found that more than a fifth of patients with verified CKD lack appropriate documentation of their illness in the EHR. A tool that prompts

providers to document CKD might shorten the time to implementing guideline-based recommendations such as RAS inhibition or urine protein quantification, which in turn might slow the progression of CKD in an individual patient.

Acknowledgments The authors thank Gilad Kuperman, and Noemie Elhadad for their thorough reviews of this manuscript; Alla Babina, Senior Programmer and Analyst, for assisting with data extraction, and the reviewers for their many helpful suggestions.

Funding This work was supported by NLM Grant T15LM007079-16.

Competing interests None.

Ethics approval This study was conducted with the approval of the Columbia University Medical Center Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Akbari A, Swedko PJ, Clark HD, *et al*. Detection of chronic kidney disease with laboratory reporting of estimated glomerular filtration rate and an educational program. *Arch Intern Med* 2004;**164**:1788–92.
2. Fox CH, Swanson A, Kahn LS, *et al*. Improving chronic kidney disease care in primary care practices: an Upstate New York Practice-based Research Network (UNYNET) Study. *J Am Board Fam Med* 2008;**26**:522–30.
3. Patwardhan MB, Samsa GP, Matchar DB, *et al*. Advanced chronic kidney disease practice patterns among nephrologists and non-nephrologists: a database analysis. *Clin J Am Soc Nephrol* 2007;**2**:277–83.
4. McClellan WM, Powe NR. Introduction to the Proceedings of a Centers for Disease Control and Prevention Expert Panel Workshop: developing a comprehensive public health strategy for preventing the development, progression, and complications of CKD. *Am J Kid Dis* 2009;**53**(3 Suppl 3):S1–3.
5. Schoolwerth AC, Engलगau MM, Hostetter TH, *et al*. Chronic kidney disease: a public health problem that needs a public health action plan. *Prev Chronic Dis* 2006;**3**:A57.
6. Nissenon A, Pereira B, Collins A, *et al*. Prevalence and characteristics of individuals with chronic kidney disease in a large health maintenance organization. *Am J Kid Dis* 2001;**37**:1177–83.
7. Xue JL, Ma JZ, Louis TA, *et al*. Forecast of the number of patients with end-stage renal disease in the United States to the year 2010. *J Am Soc Nephrol* 2001;**12**:2753–8.
8. Stevens LA, Coresh J, Greene T, *et al*. Assessing kidney function—measured and estimated glomerular filtration rate. *NEJM* 2006;**354**:2473–83.
9. Coresh J, Selvin E, Stevens LA, *et al*. Prevalence of chronic kidney disease in the United States. *JAMA* 2007;**298**:2038–47.
10. Gilbertson DT, Liu J, Xue JL, *et al*. Projecting the number of patients with end-stage renal disease in the United States to the year 2015. *J Am Soc Nephrol* 2005;**16**:3736–41.
11. Collins AJ, Li S, Gilbertson DT, *et al*. Chronic kidney disease and cardiovascular disease in the medicare population. *Kidney Int Suppl* 2003;**64**(Suppl 87):s24–31.
12. Unites States Renal Data System. http://www.usrds.org/2009/pdf/V1_09_09.PDF (accessed 1 Apr 2010).
13. 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**:489–95.
14. Fried LF, Shlipak MG, Crump C, *et al*. Renal insufficiency as a predictor of cardiovascular outcomes and mortality in elderly individuals. *J Am Coll Cardiol* 2003;**41**:1364–72.
15. National Kidney Foundation Kidney Disease Outcomes Quality Initiative (NKF KDOQI) GUIDELINES 2009. http://www.kidney.org/professionals/KDOQI/guidelines_ckd/toc.htm (accessed 23 May 2009).
16. Levinsky NG. Specialist evaluation in chronic kidney disease: too little, too late. *Ann Int Med* 2002;**137**:542–3.
17. Kinchen KS, Sadler J, Fink N, *et al*. The timing of specialist evaluation in chronic kidney disease and mortality. *Ann Int Med* 2002;**137**:479–86.
18. Navaneethan SD, Aloudat S, Singh S. A systematic review of patient and health system characteristics associated with late referral in chronic kidney disease. *BMC Nephrology* 2008;**9**:3.
19. de Lusignan S, Chan T, Stevens P, *et al*. Identifying patients with chronic kidney disease from general practice computer records. *Fam Pract* 2005;**22**:234–41.
20. Wilcheskya M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *J Clin Epidemiol* 2004;**57**:131–41.
21. Quartarolo JM, Thoeke M, Schafers SJ. Reporting of estimated glomerular filtration rate: effect on physician recognition of chronic kidney disease and prescribing practices for elderly hospitalized patients. *J Hosp Med* 2007;**2**:74–8.
22. Stevens LA, Fares G, Fleming J, *et al*. Low rates of testing and diagnostic codes usage in a commercial clinical laboratory: evidence for lack of physician awareness of chronic kidney disease. *J Am Soc Nephrol* 2005;**16**:2439–48.
23. Anandarajah S, Tai T, de Lusignan S, *et al*. The validity of searching routinely collected general practice computer data to identify patients with chronic kidney disease (CKD): a manual review of 500 medical records. *Nephrol Dial Transplant* 2005;**20**:2089–96.
24. Turchin A, Pendergrass ML, Kohane IS. DITTO—a tool for identification of patient cohorts from the text of physician notes in the electronic medical record. *Proc AMIA Symp* 2005:744–8.
25. Pakhomov S, Weston SA, Jacobsen SJ, *et al*. Electronic medical records for clinical research: application to the identification of heart failure. *Am J Manag Care* 2007;**13**:281–8.
26. Pakhomov S, Jacobsen SJ, Chute CG, *et al*. Agreement between patient-reported symptoms and their documentation in the medical record. *Am J Manag Care* 2008;**14**:530–9.
27. Pakhomov S, Hemingway H, Weston SA, *et al*. Epidemiology of angina pectoris: role of natural language processing of the medical record. *Am H J* 2007;**153**:666–73.
28. Fizman M, Chapman WW, Aronsky D, *et al*. Automatic detection of acute bacterial pneumonia from chest X-ray reports. *J Am Med Inform Assoc* 2000;**7**:593–604.
29. Fizman M, Roseblat G, Ahlers CB, *et al*. Identifying risk factors for metabolic syndrome in biomedical text. *AMIA Annu Symp Proc* 2007:249–53.
30. Voorham J, Denig P. Computerized extraction of information on the quality of diabetes care from free text in electronic patient records of general practitioners. *J Am Med Inform Assoc* 2007;**14**:349–54.
31. Friedman C, Shagina L, Lussier Y, *et al*. Automated encoding of clinical documents based on natural language processing. *J Am Med Inform Assoc* 2004;**11**:392–402.
32. Hripcsak G, Cimino JJ, Sengupta S. WebCIS: large scale deployment of a web-based clinical information system. *Proc AMIA Symp* 1999:804–8.
33. Levey AS, Bosch JP, Lewis JB, *et al*. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. *Ann Int Med* 1999;**130**:461–70.
34. Witten IH, Eibe F. *Data mining: practical machine learning tools*. San Francisco, CA: Morgan Kaufman Publishers, an imprint of Elsevier, 2005.
35. Martínez-Ramírez HR, Jalomo-Martínez B, Cortés-Sanabria L, *et al*. Renal function preservation in type 2 diabetes mellitus patients with early nephropathy: a comparative prospective cohort study between primary health care doctors and nephrologist. *Am J Kid Dis* 2006;**47**:78–87.
36. Levin A, Hemmelgarn B, Culleton B, *et al*. Guidelines for the management of chronic kidney disease. *CMAJ* 2008;**179**:1154–62.
37. Hemmelgarn BR, Zhang J, Manns BJ, *et al*. Nephrology visits and health care resource use before and after reporting estimated glomerular filtration rate. *JAMA* 2010;**303**:1151–8.
38. Shah NR, Seger AC, Seger DL, *et al*. Improving acceptance of computerized prescribing alerts in ambulatory care. *J Am Med Inform Assoc* 2006;**13**:5–11.
39. Sellier E, Colombet I, Sabatier B, *et al*. Effect of alerts for drug dosage adjustment in inpatients with renal insufficiency. *J Am Med Inform Assoc* 2009;**16**:203–10.
40. Chused AE, Kuperman GJ, Stetson PD. Alert override reasons: a failure to communicate. *AMIA Annu Symp Proc* 2008:111–15.
41. Chapman WW, Dowling JN, Wagner MM. Classification of emergency department chief complaints into 7 syndromes: a retrospective analysis of 527,228 patients. *Ann Emerg Med* 2005;**46**:445–55.
42. Friedman C, Hripcsak G, Shagina L, *et al*. Representing information in patient reports using natural language processing and the extensible markup language. *J Am Med Inform Assoc* 1999;**6**:76–87.
43. Soman S, Zasuwa G, Yee J. Automation, decision support, and expert systems in nephrology. *Adv Chronic Kidney Dis* 2008;**15**:42–55.