

Original Article

Accounting for overdispersion when determining primary care outliers for the identification of chronic kidney disease: learning from the National Chronic Kidney Disease Audit

Lois G. Kim¹, Ben Caplin², Faye Cleary¹, Sally A. Hull³, Kathryn Griffith⁴, David C. Wheeler² and Dorothea Nitsch¹ on behalf of the UK National Chronic Kidney Disease Audit

¹Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK, ²Centre for Nephrology, UCL Medical School, London, UK, ³Clinical Effectiveness Group, Centre for Primary Care and Public Health, Queen Mary University of London, London, UK and ⁴Dr Price & Partners, University Health Centre, Heslington, York, UK

Correspondence and offprint requests to: Lois G. Kim; E-mail: lois.kim@lshtm.ac.uk

ABSTRACT

Background. Early diagnosis of chronic kidney disease (CKD) facilitates best management in primary care. Testing coverage of those at risk and translation into subsequent diagnostic coding will impact on observed CKD prevalence. Using initial data from 915 general practitioner (GP) practices taking part in a UK national audit, we seek to apply appropriate methods to identify outlying practices in terms of CKD stages 3–5 prevalence and diagnostic coding.

Methods. We estimate expected numbers of CKD stages 3–5 cases in each practice, adjusted for key practice characteristics, and further inflate the control limits to account for overdispersion related to unobserved factors (including unobserved risk factors for CKD, and between-practice differences in coding and testing).

Results. GP practice prevalence of coded CKD stages 3–5 ranges from 0.04 to 7.8%. Practices differ considerably in coding of CKD in individuals where CKD is indicated following testing (ranging from 0 to 97% of those with and glomerular filtration rate <60 mL/min/1.73 m²). After adjusting for risk factors and overdispersion, the number of ‘extreme’ practices is reduced from 29 to 2.6% for the low-coded CKD prevalence outcome, from 21 to 1% for high-uncoded CKD stage and from 22 to 2.4% for low total (coded and uncoded) CKD prevalence. Thirty-one practices are identified as outliers for at least one of these outcomes. These can then be categorized into practices needing to address testing, coding or data storage/transfer issues.

Conclusions. GP practice prevalence of coded CKD shows wide variation. Accounting for overdispersion is crucial in providing useful information about outlying practices for CKD prevalence.

Keywords: audit, chronic kidney disease, outliers, overdispersion, prevalence

INTRODUCTION

Chronic kidney disease (CKD) is a precursor to subsequent deterioration in kidney function and adverse outcomes associated with this decline [1]. It is therefore of growing importance to characterize and understand differences in diagnosis of this condition in primary care in order to improve management and outcomes for those with CKD.

Factors known to be associated with developing CKD include increasing age, diabetes, hypertension and previous cardiovascular events [2–6]. Since 2006, general practitioners (GPs) in England and Wales have been incentivized to maintain a register of those with CKD stages 3–5 through the NHS Quality and Outcomes Framework (QOF). GPs can add patients to their register by assigning an electronic ‘read code’. Most prescribing software systems used by GPs rely on the presence of coded CKD, and not on serum creatinine or estimated glomerular filtration rate (eGFR) values. If patients are not coded for CKD then, depending on the software used, the

Table 1. Summary of practice-level characteristics

Practice-level characteristics	No. of practices with data (/915) ^a	Median (IQR)
List size (including under 18s)	756	7456 (4568, 10 310)
Median age (years) ^b	756	40 (40, 45)
Female (%)	756	49.4 (48.4, 50.6)
Median index of multiple deprivation ^c	601	17 179 (10 575, 22 866)
Black ethnicity (%)	756	0.27 (0, 1.27)
Diabetes (%)	756	6.1 (5.2, 7.2)
Hypertension (%)	756	17.7 (15.0, 20.4)
CVD (%)	756	6.2 (4.9, 7.4)
Testing in at-risk population (%)	915	92.7 (90.6, 94.2)
		Mean (SD)
Age–sex standardized CKD prevalence (%)	756	3.1 (1.2)
Proportion with uncoded CKD if GFR evidence of CKD (%)	915	35.8 (17.6)
Combined coded and uncoded age–sex standardized CKD prevalence (%)	756	4.3 (1.2)

^aOnly 756 practices had list size data available.

^bUsing all ages list size data to identify median age in practice, using mid-points of 5-year age bands.

^cAmongst patients with CKD or at high risk from CKD only. Estimated only for those practices that also have list size data (i.e. excludes 15 practices with IMD data, but no list size data).

inbuilt safety alerts related to adapting drug choice or dosing to level of renal function may not work. QOF data suggest that there is substantial practice variation in the proportion of patients with read codes for CKD stages 3–5, with prevalences ranging from 0.01 to 27% [7]. The next steps are to investigate the sources of this heterogeneity and to identify outliers for CKD prevalence, with a view to contacting individual practices to improve testing for and/or coding of CKD (and thus care and management of these patients) [8, 9].

We use initial data from a UK national CKD audit to illustrate the importance of accounting for large between-practice variability when seeking to identify the most extreme outlying practices for CKD prevalence.

MATERIALS AND METHODS

Subjects

GP practices submitting baseline data to the National Chronic Kidney Disease Audit (NCKDA) between March and 12 November 2015 are included in the analysis. More information on the audit can be found on the NCKDA website [10]. Data were collected for 915 practices in England and Wales using Informatica software. The software is compatible with the range of GP clinical computer systems and can directly extract anonymized-coded clinical data, including diagnostic and laboratory test results, from the patient record.

Baseline data were extracted for all individuals within a practice aged 18+ years with either a QOF code for CKD (any stage; for list of codes see [Supplementary Table S1](#) or with a risk factor/renal disease diagnosis at least 1 year prior to data extraction. This latter group includes those with diagnosed genetic renal conditions or any of the following: hypertension, diabetes, gout, ischaemic heart disease, congestive cardiac failure, cerebrovascular disease, peripheral arterial disease, kidney stones, prostatic hypertrophy, prescription of lithium/tacrolimus/cyclosporin in past year, systemic lupus erythematosus and other connective tissue disorders. Retrospective creatinine

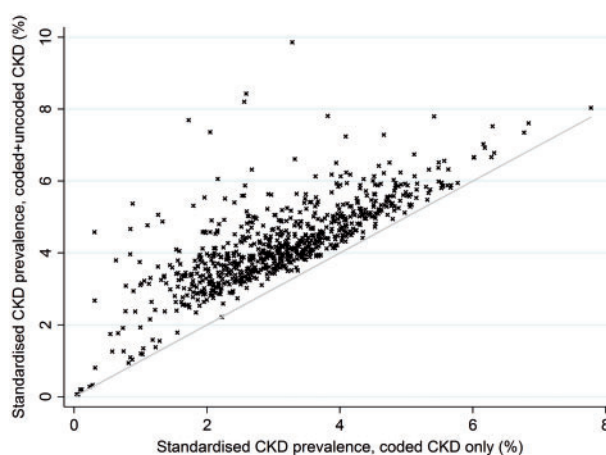


FIGURE 1: Scatter plot showing increase in standardized CKD prevalence after inclusion of uncoded CKD, for 756 practices for whom prevalence can be calculated from list size data.

eGFR measurements were also obtained for all these individuals, together with baseline characteristics including age, sex and index of multiple deprivation (IMD), a validated measure of area socio-economic deprivation [11]. Where available, complete practice age–sex distribution data were also obtained. Practices were given an email summary of their data after the initial extraction (round 1) to allow the practices to check the data for errors and amend these. After 3 months, a second and final extraction of data on CKD stages 3–5 prevalence were taken (round 2). It is the initial round 1 data that are used for the analyses in this article, and results shown here use anonymized practice data. The full results of the audit, using round 2 data, will be published in a report at the end of 2016.

This work aims to identify and compare outlying practices in terms of three key outcomes of interest: (i) prevalence of read-coded CKD (stages 3–5), (ii) prevalence of uncoded CKD amongst those with eGFR evidence of CKD stages 3–5 and (iii) total CKD prevalence (combining those with coded and

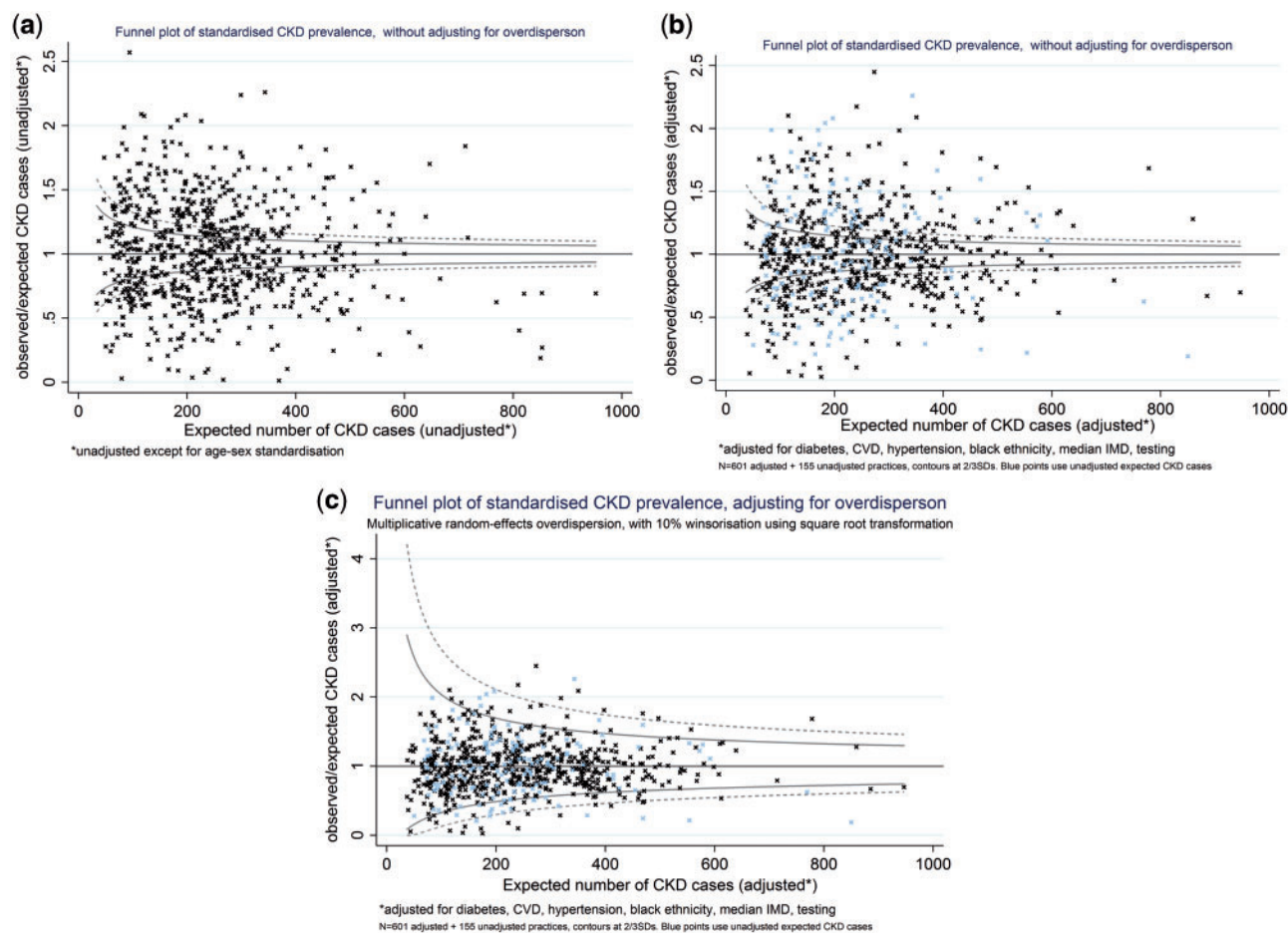


FIGURE 2: (a) Crude funnel plot of standardized coded CKD prevalence. (b) Funnel plot of standardized coded CKD prevalence. Expected cases adjusted for practice median IMD and proportion of practice tested for CKD and proportion with black ethnicity, diabetes, hypertension and CVD. (c) Funnel plot of standardized coded CKD prevalence, with 10% winsorization and adjustment for overdispersion using multiplicative random effects. Expected cases adjusted for practice median IMD and proportion of practice tested for CKD and proportion with black ethnicity, diabetes, hypertension and CVD.

uncoded CKD stages 3–5). The recording of urinary protein tests is generally poor; we therefore do not report results on proteinuria in this article.

(i) Read-coded CKD prevalence

Practice-coded CKD prevalence is calculated using read codes for CKD stages 3–5 and list size data for all ages. Direct standardization is carried out using the GP practice database of the age–sex distribution (using 5-year age bands) for GP practices in England [12]. Note that this calculation assumes no CKD in under 18s (as data are not collected for this group in the audit); however, this is not unreasonable as the number of such cases within a practice is likely to be negligible.

(ii) Uncoded CKD amongst those with eGFR evidence of CKD

Patients with eGFR evidence of CKD are defined as those for whom either: (a) the two most recent eGFR measurements are both $<60 \text{ mL/min/1.73 m}^2$ and where at least 3 months have elapsed between measurements (the most recent measurement must be in the last 2 years and both measurements

since 1 January 2008), or (b) the most recent eGFR measurement (since 1 January 2008, in last 2 years) is $<60 \text{ mL/min/1.73 m}^2$, and this is the only eGFR measurement ever taken. Individuals meeting one of these criteria are then defined as having uncoded CKD if they do not have a QOF code for stages 3–5 CKD.

(iii) Total CKD prevalence

The number of individuals with (a) coded and (b) uncoded CKD are combined to obtain a practice total CKD prevalence. This is then age–sex standardized in the same way as measure (i).

Statistical methods

Funnel plots are produced for each of the three outcomes of interest (i)–(iii), with a view to identifying outlying practices. However, when seeking to identify outlying practices, it would be naïve to take the view that there is a single common target CKD prevalence that all practices should be compared to. Rather, there are a range of factors, both observed and unobserved, that must be taken into account.

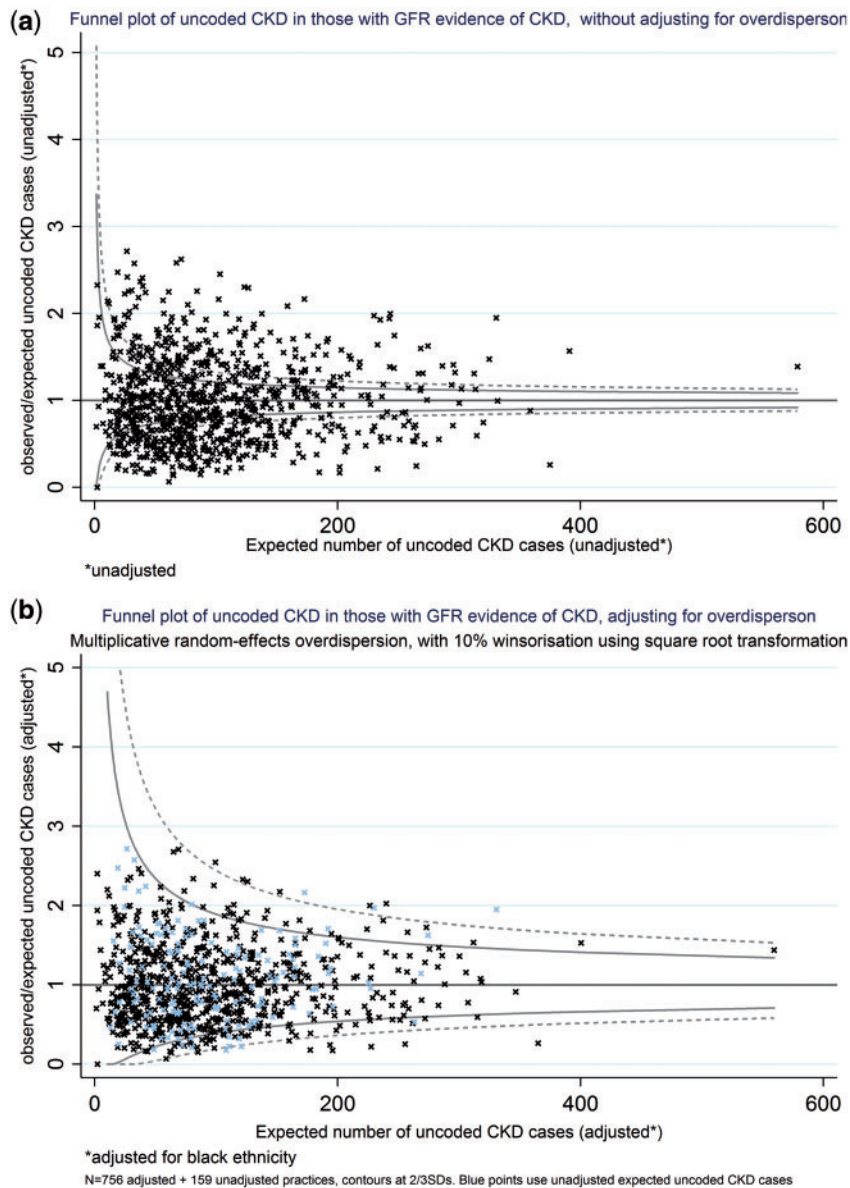


FIGURE 3: (a) Crude funnel plot of proportion of uncoded CKD in those with GFR evidence of CKD, without adjusting for overdispersion. (b) Funnel plot of proportion of uncoded CKD in those with GFR evidence of CKD, with 10% winsorization and adjustment for overdispersion using multiplicative random effects. Expected cases adjusted for proportion of practice with black ethnicity.

Observed factors

Target prevalences for each practice should ideally reflect underlying differences in the practice populations, and the acceptable range of values should also reflect the population size from which the observed value is derived. This can be achieved using a funnel plot based on observed–expected cases against expected cases. Expected cases for outcomes (i) and (iii) for each practice were back-calculated from a logistic regression model for the practice proportion of cases, adjusted for practice-level proportions of individuals with diabetes, hypertension, cardiovascular disease (CVD), black ethnicity and median IMD, and the proportion of at-risk patients undergoing testing for CKD. For outcome (ii), expected cases were taken from a logistic regression model adjusted only for black ethnicity, since there was no *a priori* reason to expect coding given eGFR measurement to be associated with any other factor.

Where data are not available on all explanatory variables used in calculating the expected cases for some practices, unadjusted cases (crude mean) are plotted instead—these practices are identified on the funnel plots using red points.

Unobserved factors. Wide variation between practices in terms of both testing and coding of CKD results in considerable overdispersion of all of these outcomes. This phenomenon has been described previously in between-practice comparisons of similar types of outcomes [13, 14]. Some of this overdispersion can be accounted for by adjusting the contours for factors related to the between-practice heterogeneity [13]. We have opted to apply multiplicative random effects methods [15] here as this approach is more conservative with respect to practices with very small numbers of expected cases. Contours were additionally winsorized by 10% in order to further reduce the

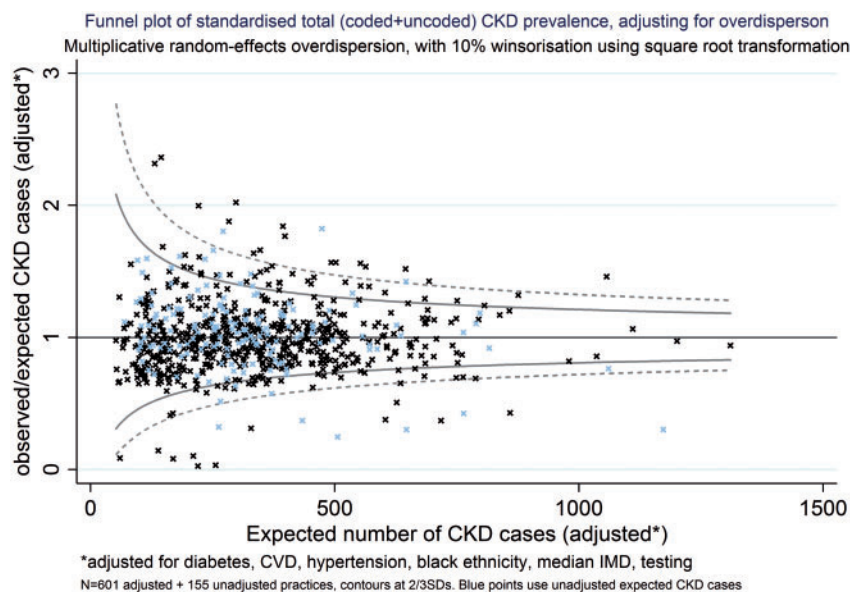


FIGURE 4: Funnel plot of combined coded and uncoded standardized CKD prevalence, with 10% winsorization and adjustment for overdispersion using multiplicative random effects. Expected cases adjusted for practice median IMD and proportion of practice tested for CKD and proportion with black ethnicity, diabetes, hypertension and CVD.

Table 2. Number of practices identified as outliers for CKD prevalence

	n	No. of outlying practices identified at 3 SDs (% of practices analysed)	
		Crude contours	Adjusted contours ^a
Low-coded CKD	756	223 (29)	20 (2.6)
High-uncoded CKD	915	194 (21)	9 (1.0)
Low combined coded and uncoded CKD	756	163 (22)	18 (2.4)

^aContours adjusted for 10% winsorization and overdispersion; expected cases adjusted for practice-level diabetes, CVD, hypertension, black ethnicity, IMD and testing (black ethnicity only for uncoded CKD) where all data are available, otherwise unadjusted cases are used.

potential influence of extreme values on the contours whilst still retaining the same number of z-scores. This is achieved by ranking the z-values calculated above, and replacing the bottom 10% with the value of the 10th centile and top 10% with the value of the (100–10)th centile.

RESULTS

In total, 915 GP practices recruited at round 1 are included in the analysis of uncoded CKD [outcome (ii)] and 756 of these practices with list size data available are included in the analysis of prevalence [outcomes (i) and (iii)]. Basic practice characteristics are summarized in Table 1. The large heterogeneity in age-sex standardized CKD prevalence is illustrated in Supplementary Figure S1; it is also clear from these plots that key practice-level characteristics such as the proportion of patients with diabetes only explain around 20% of the between-practice differences in prevalence ($R^2=21.6\%$, Appendix Table S2).

Figure 1 shows the increase in CKD prevalence after inclusion of uncoded cases in whom there is biochemical evidence of CKD, by practice. Whilst the majority of practices increase by

an additional 0–1% after this adjustment, uncoded cases increase the prevalence by over 5% in a small number of practices.

Funnel plots for each of the outcomes of interest using crude and adjusted contours are shown in Figures 2–4 and Supplementary Figure S2, with accompanying numbers of outliers given in Table 2. It is clear that neither the crude contours nor those with expected cases adjusted for known risk factors are useful in identifying outliers for CKD prevalence. This reflects the large underlying between-practice heterogeneity in a range of factors that is not accounted for in the calculation of expected cases. This will include both identification of the at-risk population and coding of CKD as well as unmeasured practice population characteristics (the proportion of the identified at-risk population who are tested is adjusted for here). This heterogeneity must be taken into account when seeking to identify practices that are really extreme in terms of these outcomes. This is achieved by increasing the contours by a factor ($\sqrt{\hat{\phi}}$) that is related to the percentage of overdispersion around the target [15] (see Supplementary Table S3).

These methods have enabled the identification of a small number of practices that warrant further investigation (Table 3

Table 3. Summary of outlier practices identified using control limits accounting for overdispersion

Practice ID	Mean age (years)	Observed cases	List size	Outlier low prevalence	Outlier high-uncoded CKD	Outlier low total CKD	Low testing ^a	High IMD ^a	Low diabetes ^a	Low ht ^a	Low CVD ^a	Group
142	42.5	99	11 769	✓	✓							A
448	45.8	46	5425	✓	✓							A
454	46.9	124	8140	✓	✓							A
581	45.1	79	6230	✓	✓							A
92	36.9	174	15 170	✓		BL						B
525	39.3	14	5477	✓	BL							B
788	39.1	23	7874	✓	BL	BL						B
827	43.5	192	8900	✓		BL	✓					B
100	44.9	280	9455		✓		✓					C
164	37.7	240	11 990	BL	✓			✓				C
796	-	355	-		✓							C
892	-	261	-		✓							C
909	-	132	-		✓							C
57	43.2	23	6060	BL		✓		✓	✓	✓	✓	D
155	47.6	39	12 700	✓		✓			✓	✓	✓	D
266	42.7	136	10 390	BL		✓			✓	✓	✓	D
388	45.9	257	28 200	✓		✓		✓	✓	✓	✓	D
408	41.9	9	6916	✓		✓	✓		✓	✓	✓	D
459	44.4	6	12 190	✓		✓	✓		✓	✓	✓	D
477	49.3	91	6280	BL		✓			✓	✓	✓	D
568	42.1	112	12 120	✓		✓			✓	✓	✓	D
577	37.9	4	8800	✓		✓			✓	✓	✓	D
611	51.1	165	15 500	✓		✓			✓	✓	✓	D
634	44.5	4	2622	BL		✓	✓		✓	✓	✓	D
640	54.3	263	28 100	✓		✓			✓	✓	✓	D
651	47.7	197	20 800	✓		✓			✓	✓	✓	D
738	50.5	166	18 300	✓		✓			✓	✓	✓	D
772	38.0	15	7716	✓		✓			✓	✓	✓	D
804	45.4	338	26 800	✓		✓			✓	✓	✓	D
822	42.0	240	20 100	BL		✓			✓	✓	✓	D
332	39.1	256	16 340	BL		✓		BL			BL	D

^aLow diabetes: <10th centile = <4%; low hypertension (ht): <10th centile = <11%; low CVD: <10th centile = <3.5%; high IMD: >90th centile = 26 000; low testing: <10th centile = <88% of at-risk.

BL = borderline (outlier at 2 SD).

and Figure 5). Such practices can generally be categorized in order to tailor quality improvement:

- (i) Outlier for low-coded/combined CKD prevalence and for high-uncoded CKD
Suggests that coding could be improved.
- (ii) Outlier for low-coded/combined CKD prevalence but not for high-uncoded CKD
 - (a) This may still be a coding issue, or be due to particular practice characteristics: coding issue, but not an outlier for this due to low numbers tested. Low testing may indicate that testing of those at risk could be improved, or may be due to low numbers at risk related to low practice mean age/small list size (practice 525, 788).
 - (b) expected cases do not adequately reflect practice characteristics, either due to unobserved practice characteristics or the model not fully capturing the reduction in expected cases for practices with extreme characteristics [practice 92 (low age, high black ethnicity, low IMD), 827 (low testing)].
- (iii) Not an outlier for low prevalence but outlier for high-uncoded CKD

Suggests some improvements can still be made to coding. Some of these practices may be unidentified outliers for CKD prevalence (due to missing list size data), which would place them in group (i).

- (iv) Either (i) or (ii) together with very low diabetes, hypertension and/or CVD

Suggests more a general issue with coding and/or data extraction for diabetes and vascular disease as well as CKD.

For these data, 4 practices are identified in group A, 4 in group B, 5 in group C and 18 practices are in group D. This means that a sizeable majority of outliers identified (18/31 practices) are likely to be due to poor data quality, with smaller groups identified as needing to address issues with coding.

DISCUSSION

It is good practice to seek to identify outliers that may be able to improve their CKD coding compared with the rest of the practices in the analysis. In reality, this means we want to identify a small number of extreme outlying practices in terms of

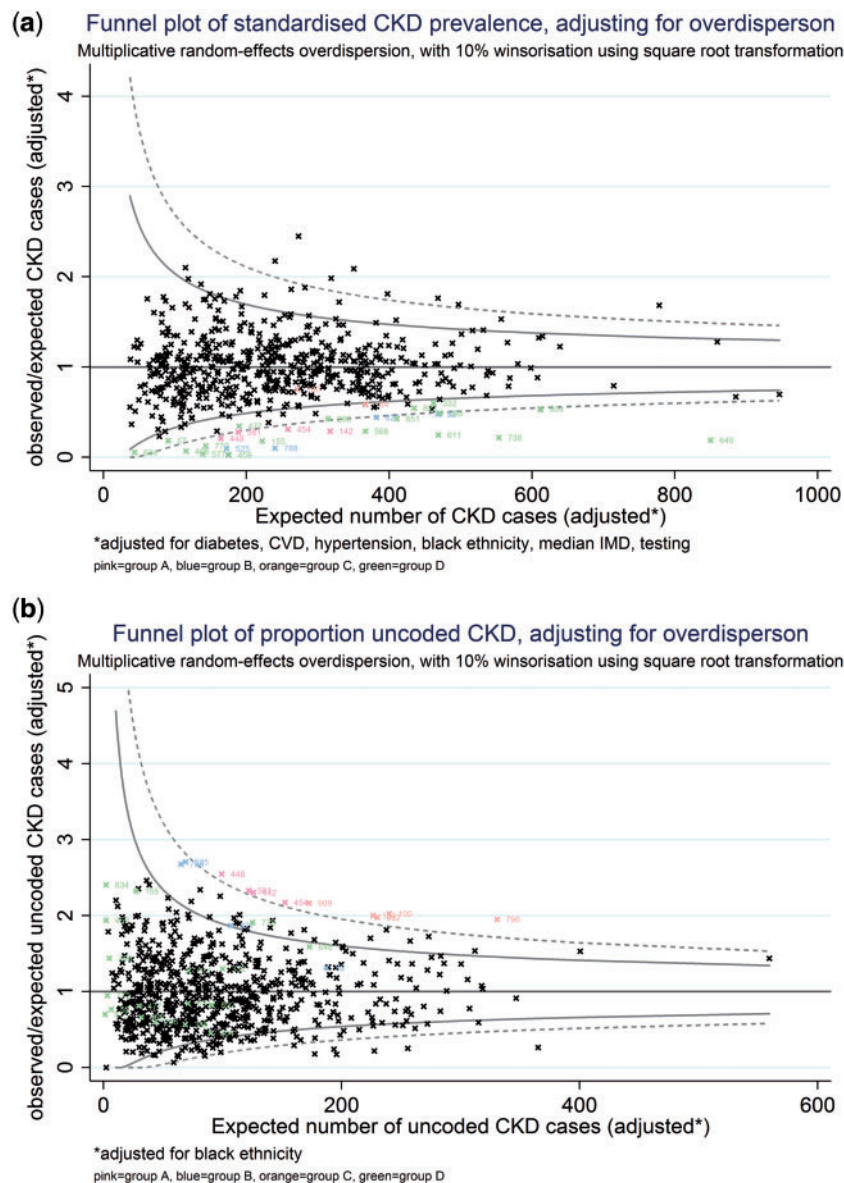


FIGURE 5: (a) Funnel plot of standardized CKD prevalence, with outlying practices highlighted by group. Contours with 10% winsorization and adjustment for overdispersion using multiplicative random effects. Expected cases adjusted for practice median IMD and proportion of practice tested for CKD and proportion with black ethnicity, diabetes, hypertension and CVD. (b) Funnel plot of uncoded CKD, with outlying practices highlighted by group. Contours with 10% winsorization and adjustment for overdispersion using multiplicative random effects. Expected cases adjusted for black ethnicity.

performance in order to focus efforts for improvement. However, there is generally large between-practice heterogeneity in outcomes in primary care that rely on a range of processes (here, identification of at-risk patients, testing and coding) as well as underlying risk factors for the outcome itself. This heterogeneity results in substantial overdispersion of prevalences, hampering efforts to identify outlying practices. It is therefore imperative that an appropriate strategy is applied in order to account for this when carrying out such analyses.

Identification of outlying practices requires comparison of practice outcomes with a defined target. However, this target should reflect both observed underlying risk factors and unobserved influences resulting in overdispersion; such influences may include unobserved risk factors and differences in the

processes relating to estimating the outcome. This practice-specific target estimation can be carried out by direct adjustment for known risk factors at practice level, and by additionally accounting for overdispersion in the outcome. For CKD prevalence, this has resulted in a substantial reduction of the number of outlying practices identified, which in turn focusses attention on those practices where there may be errors in data acquisition or deficiencies in identifying and coding CKD cases.

There are a number of limitations to our work. Our data are limited to practices using the Informatica Audit+ software, and as such these practices may represent a self-selecting sub-population of practices who are most interested in quality improvement. In light of this, it is reasonable to interpret the data presented here as potentially the best possible case

scenario for the UK. The observed mean practice age- and sex-standardized prevalence of coded CKD in this sample of 915 practices in England and Wales (3.1%) is lower than prevalences previously reported for studies in other countries [16]. That this was observed despite the incentives of QOF coding may be due to a number of reasons, including the increased complexity of CKD coding and the presence of other codes and flags for medication alerts. However, estimates of total CKD prevalence based on eGFR measurements from UK primary care data do not differ greatly from those based on nationally representative samples [17]. Our study defined CKD prevalence as having two recorded measurements of eGFR <60 mL/min/1.73 m² more than 3 months apart, and included all individuals in each practice with, therefore, a lower mean age than samples used for other prevalence estimates [16]. Differences in underlying risk factors, identifying at-risk patients, and in testing and diagnosing of CKD are also likely to vary significantly between countries. Although generalizability to other health systems is somewhat limited due to differing practice software systems and funding incentives, the dominant issue is likely to be the high degree of variability of CKD coding and care especially amongst those without diabetes.

The clear strength of this study lies in the large sample of practices, representing an underlying population of around 6 million patients. This audit provides a good snap-shot of the quality of routine CKD data in primary care. We have used these initial pilot data to demonstrate the importance of accounting for between-practice variability in epidemiological studies investigating CKD. Without appropriate handling of this heterogeneity, findings from large primary care data on CKD will be confounded by individual doctor's management of CKD.

SUPPLEMENTARY DATA

Supplementary data are available online at <http://ndt.oxfordjournals.org>.

ACKNOWLEDGEMENTS

The National Chronic Kidney Disease Audit is commissioned by the Healthcare Quality Improvement Partnership (HQIP), and funded by NHS England, as part of the National Clinical Audit and Patient Outcomes Programme (NCAPOP) and the Welsh Government.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Levey AS, Coresh J. Chronic kidney disease. *Lancet* 2012; 379: 165–180
2. Matsushita K, van der Velde M, Astor BC *et al*. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *Lancet* 2010; 375: 2073–2081
3. Roderick PJ, Atkins RJ, Smeeth L *et al*. Detecting chronic kidney disease in older people; what are the implications? *Age Ageing* 2008; 37: 179–186
4. Zdrojewski L, Zdrojewski T, Rutkowski M *et al*. Prevalence of chronic kidney disease in a representative sample of the Polish population: results of the NATPOL 2011 survey. *Nephrol Dial Transplant* 2016; 31: 433–439
5. Bruck K, Stel VS, Gambaro G *et al*. CKD prevalence varies across the European general population. *J Am Soc Nephrol* 2016; 27: 2135–2147
6. Fraser SD, Aitken G, Taal MW *et al*. Exploration of chronic kidney disease prevalence estimates using new measures of kidney function in the health survey for England. *PLoS One* 2015; 10: e0118676
7. (HSCIC) HaSCIC. *Quality and Outcomes Framework (QOF), 2014-15*. 2015. www.hscic.gov.uk/catalogue/PUB18887 (31 May 2015, date last accessed)
8. Lv J, Ehteshami P, Sarnak MJ *et al*. Effects of intensive blood pressure lowering on the progression of chronic kidney disease: a systematic review and meta-analysis. *CMAJ* 2013; 185: 949–957
9. Kurella Tamura M, Li S, Chen SC *et al*. Educational programs improve the preparation for dialysis and survival of patients with chronic kidney disease. *Kidney Int* 2014; 85: 686–692
10. BMJ. National Chronic Kidney Disease Audit [19/05/2016]. www.ckdaudit.org.uk
11. Statistics On. *English Indices of Deprivation* 2015. <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015> (31 May 2015, date last accessed)
12. (HSCIC) HaSCIC. *Number of Patients Registered at a GP Practice - January 2015*. 2015. www.hscic.gov.uk/catalogue/PUB16357 (31 May 2015, date last accessed)
13. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005; 24: 1185–1202
14. Marshall C, Best N, Bottle A *et al*. Statistical issues in the prospective monitoring of health outcomes across multiple units. *J R Stat Soc A* 2004; 167: 541–560
15. Spiegelhalter DJ, Sherlaw-Johnson C, Bardsley M *et al*. Statistical methods for healthcare regulation: rating, screening and surveillance. *J R Stat Soc A* 2012; 175: 1–47
16. De Nicola L, Zoccali C. Chronic kidney disease prevalence in the general population: heterogeneity and concerns. *Nephrol Dial Transplant* 2016; 31: 331–335
17. Iwagami M, Tomlinson LA, Mansfield KE *et al*. Validity of estimated prevalence of decreased kidney function and renal replacement therapy from primary care electronic health records compared to national survey and registry data in the United Kingdom. *Nephrol Dial Transplant* 2017; 32 (Suppl 2): ii142–ii150

Received for publication: 31.5.2016; Accepted in revised form: 15.10.2016