Original article

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 UK

Masao Iwagami,[1] Laurie Tomlinson,[1] Kathryn Mansfield,[1] Anna Casula,[2] Fergus Caskey,[2] Grant Aitken,[3] Simon Fraser[4], Paul Roderick[4], Dorothea Nitsch[1]

[1]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK; [2]Renal Registry, UK Renal Association, Bristol, UK; [3]Geography & Environment, Faculty of Social and Human Sciences, University of Southampton, Southampton, UK; [4]Academic Unit of Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton, UK

Correspondence:

Masao Iwagami, MD, MPH, MSc, Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK.

Address: Keppel Street, London, WC1E 7HT, UK

E-mail: masao.iwagami@lshtm.ac.uk

For Peer Review

Abstract

Background: Anonymous primary care records are an important resource for observational studies. However, their external validity is unknown in identifying the prevalence of decreased kidney function and renal replacement therapy (RRT). We thus compared the prevalence of decreased kidney function and RRT in the Clinical Practice Research Datalink (CPRD) with a nationally representative survey and national registry.

Methods: Among all people aged ≥25 registered in CPRD for ≥1 year on 31st March 2014, we identified patients with estimated glomerular filtration rate (eGFR) <60 mL/min/1.73m^2, according to their most recent serum creatinine in the past five years using the Chronic Kidney Disease Epidemiology Collaboration equation, and patients with recorded diagnoses of RRT. Denominators were the entire population in each age-sex band irrespective of creatinine measurement. The prevalence of eGFR <60 mL/min/1.73m^2 was compared with that in the Health Survey for England (HSE) 2009/2010, and the prevalence of RRT was compared with that in UK Renal Registry (UKRR) 2014.

Results: We analysed 2,761,755 people in CPRD (mean age 53 [SD 17], men 49%), of whom 189,581 (6.86%) had eGFR <60 mL/min/1.73m^2 and 3,293 (0.12%) were on RRT. The prevalence of eGFR <60 mL/min/1.73m^2 in CPRD was similar to that in HSE and the prevalence of RRT was close to that in UKRR across all age groups in men and women, although the small number of younger patients with eGFR <60 mL/min/1.73m^2 in HSE might

For Peer Review

have hampered precise comparison.

Conclusions: UK primary care data have good external validity for the prevalence of decreased kidney function and RRT.

Keywords

Epidemiology; Chronic kidney disease; Renal replacement therapy; Validity; Primary care

Short Summary

- We examined the external validity of the prevalence of decreased kidney function (eGFR <60 mL/min/1.73m^2) and RRT in a UK primary care database (CPRD), by comparing them with results from two nationally representative surveys (Health Survey for England and UK Renal Registry).
- Of 2,761,755 eligible adults in CPRD, 189,581 (6.86%) patients had eGFR <60 mL/min/1.73m^2 and 3,293 (0.12%) had recorded diagnoses of RRT.
- The prevalence of eGFR <60 mL/min/1.73m^2 in CPRD was similar to that in Health Survey for England, and the prevalence of RRT in CPRD was close to that in UK Renal Registry across all age groups (every 10 years), both in men and women.
- This study is the first to indicate that using recent UK primary care data for cross-sectional research on CKD and RRT is valid.

For Peer Review

Background

Chronic kidney disease (CKD) is a major public health problem, which increases in prevalence with age and is associated with increased morbidity and mortality [1-3]. The number of people with end-stage renal disease (ESRD) requiring renal replacement therapy (RRT) has been increasing worldwide and is predicted to double by 2030 [4]. Appropriate identification of CKD is thus important for early intervention, including prevention of both CKD progression and cardiovascular diseases [5]. At the population level, accurate estimation of CKD prevalence is essential to assess the burden of CKD in the community and to evaluate the effectiveness of population approaches for CKD [6]. However, potential methodological difficulties may make it problematic to determine the community prevalence of CKD [7, 8]. For example, people who have kidney function measured routinely by serum creatinine may not represent the general population and serum creatinine assays may not be uniformly standardised.

Data derived from routine patient care, such as the anonymous primary care records held in the UK Clinical Practice Research Datalink (CPRD) [9], are an important resource for observational studies [10]. Because CRPD broadly represents the UK population in terms of demographics [11], it can be a useful source to estimate a disease prevalence in the UK. However, using routine electronic records to investigate renal disease is only possible if the general practitioners (GPs) who contribute data to CPRD appropriately test, identify, and

For Peer Review

record all those with kidney disease in the population. Reliable measures of renal disease in electronic health records would allow a more robust use of primary care data to investigate renal disease epidemiology, for example, researchers would be able to investigate the association between kidney diseases and other comorbidities or medications recorded in primary care data. To date, a number of definitions for diseases or specific conditions have been validated in CPRD at the individual or population level [12, 13]. However, to our knowledge, there has been no external validation study for the prevalence of decreased kidney function and RRT in CPRD. The best available methods to identify CKD and RRT in CPRD are to use serum creatinine records measured by GPs and recorded diagnoses of RRT in CPRD, respectively, yet the validity or appropriateness of these methods are not known.

The Health Survey for England (HSE), a nationally representative survey of health condition, included measurement of kidney function in 2009 and 2010 [14]. Every consenting participant had kidney function measured, giving representative statistics for the prevalence of decreased kidney function in the general population. Meanwhile, the UK Renal Registry (UKRR), which records information regarding all people on RRT in the UK, provides annual reports of the prevalence of RRT [15]. Referring to these two nationally representative sources of data, we aimed to evaluate the external validity of the prevalence of decreased kidney function and RRT in CPRD.

For Peer Review

Methods

Details of CPRD and study population

In the UK, the primary care system acts as a gate-keeper to healthcare – patients need to be registered with a primary care doctor to access National Health Service (NHS) non-emergency care. Health care is free at the point of access. Primary care practices have used computerised electronic health records since the early 1990s. There are only a limited number of suppliers of GP electronic health record software. CPRD uses data from VISION software system (In Practice Systems Ltd., London, UK) and has evolved as an observational data and interventional research service provided by the NHS. Currently over 650 GP practices contribute data meeting quality control standards to CPRD, covering and representing nearly 7% of the UK population [11]. Previous studies have suggested that the distribution of age, sex, ethnicity, practice location deprivation, and other health indicators such as smoking and morbidities are similar to that of external UK-based sources [11, 16-19]. The database includes patient demographics, coded diagnoses, and outpatient laboratory test results. The Secretary of State waived informed consent for CPRD data because data are anonymised and there is an overall benefit for research. Ethical approval for this study was obtained from the Independent Scientific Advisory Committee, which oversees research on CPRD data (protocol No. 16_055), as well as the London School of Hygiene and Tropical Medicine Ethics Committee (reference: 9196).

For Peer Review

The study population was all people aged 25 or older, who were alive and registered in CPRD for at least one year on 31st March 2014. The choice of age 25 as a lower limit was made for the best comparability between CPRD and HSE or UKRR: HSE and UKRR collected data of people under 25 differently (HSE grouped people aged 16-24, while UKRR grouped people aged 18-24). One year registration was considered necessary for GPs to record a history of RRT for newly registered patients, or to test their kidney function if they had a key CKD risk factor such as diabetes [5].

Details of external data

For the prevalence of decreased kidney function, we compared the data from CPRD with those from HSE 2009 and 2010 (combined) [14]. Briefly, the HSE 2009/2010 included a cross-sectional study of kidney disease among people selected using a multistage stratified random probability sampling method. Blood samples were taken from nearly 6,000 consenting participants, accounting for 77% for men and 73% for women among all the HSE participants. Data were weighted for non-response to reduce response bias. Creatinine was measured by an internationally standardized enzymatic method, which is traceable to isotope dilution mass spectrometry (IDMS) [20]. Estimated glomerular filtration rate (eGFR) was calculated from serum creatinine value using the Modification of Diet in Renal Disease Study equation in the original HSE report [14], whereas a post-hoc analysis was conducted using

For Peer Review

the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation [21]. The prevalence of people with single eGFR <60 ml/min/1.73m^2 was reported according to age (every 10 years) and sex.

For RRT prevalence, we referred to the data from the UKRR 2014 [15]. The UKRR 2014 collects data from all 71 renal centres in the UK. The prevalence of RRT in 2013 was estimated by dividing the number of patients on RRT by the 2013 UK population, according to age (every 10 years), sex, and RRT modality: haemodialysis, peritoneal dialysis or kidney transplantation.

Definition of decreased kidney function and RRT in CPRD

We identified patients with eGFR <60 ml/min/1.73m^2, according to their most recent single serum creatinine measured by GP in the past five years (i.e. the period between 1st April 2009 and 31st March 2014), using the CKD-EPI equation [22]. We used a single eGFR to define decreased kidney function in the main analysis because HSE (reference data in this study), as well as previous large epidemiological studies [23, 24], have used this definition. For the main analysis, we made the following assumptions: (i) all the UK laboratories reported IDMS traceable creatinine; (ii) people with missing record of ethnicity in CPRD had non-black ethnicity; (iii) people without any creatinine measurement for the past five years did not have decreased kidney function.

For Peer Review

We identified patients on RRT based on the diagnoses recorded in CPRD anytime from the date of their registration to 31st March 2014. The list of diagnosis codes (Read codes) indicative of RRT was determined by using a recommended strategy [25], and agreed among the authors (**Supplementary data Table 1**). In addition, in order to examine the validity of diagnoses of different RRT modality in CPRD, we classified patients with RRT into those with haemodialysis, peritoneal dialysis, or kidney transplantation. We used the most recent recorded diagnosis as this is the best available approach to estimate the prevalence of the current RRT modality in CPRD.

Data analysis

We calculated the prevalence (95% confidence interval [CI]) of eGFR <60 ml/min/1.73m^2 according to age (every 10 years) and sex in CPRD and HSE, respectively, using the CKD-EPI equation. Denominators in CPRD were the entire population in each age-sex band irrespective of creatinine measurement in the past five years. Patients aged 75 or older were grouped in CPRD to be consistent with HSE. We calculated the difference (95% CI) in the prevalence of eGFR <60 ml/min/1.73m^2 between CPRD and HSE. We also reported the proportion of patients with at least one creatinine measurement for the past five years in CPRD.

Similarly, we calculated the prevalence of RRT in CPRD and UKRR, respectively,

For Peer Review

and then the difference between CPRD and UKRR, in 10-year age bands by sex. We also reported results by RRT modality.

All statistical analyses were conducted using Stata 14 software (Stata Corp, Texas).

Sensitivity analyses

We repeated our analyses using a number of alternative eGFR definitions and restricted study populations in order to determine the impact of the definition for decreased kidney function that we used. We defined decreased kidney function as follows: i) We assumed that all the UK laboratories reported non-IDMS traceable creatinine, and therefore multiplied the recorded creatinine value by 0.95 to use the CKD-EPI equation for IDMS-traceable creatinine [26]; ii) We conducted a complete case analysis for ethnicity (restricting the analysis to people with recorded ethnicity in CPRD); iii) We used the participants' most recent creatinine in the past two years, instead of five years; iv) We restricted the region to England, by excluding data from Scotland, Wales, and Northern Ireland; v) We additionally required a measure of chronicity to define decreased kidney function [27]: two or more eGFR results of <60 ml/min/1.73m^2 needed to be recorded consecutively ≥3 months apart in the past five years; and vi) We conducted a complete case analysis for creatinine by restricting the analysis to people with at least one creatinine measurement in the past five years.

We also compared the prevalence of eGFR <45 ml/min/1.73m^2 (calculated from the

For Peer Review

most recent creatinine in the past five years) between CPRD and HSE, which may be a more robust indicator of decreased kidney function with prognostic implications [28, 29].

For Peer Review

Results

From 685 GP practices, we identified 2,761,755 people (mean age 53 [SD17], men 49%) who were alive and registered in CPRD for ≥1 year on 31st March 2014. Their age-sex distribution was broadly similar to that of the UK Census 2013 (**Supplementary data Table 2**). Of those identified, 189,581 patients (6.86%) had eGFR <60 mL/min/1.73m^2 and 3,293 patients (0.12%) were on RRT.

The prevalence of eGFR <60 ml/min/1.73m^2 increased steeply with age (**Table 1** and **Figure 1**). There was no evidence that the prevalence of eGFR <60 ml/min/1.73m^2 in CPRD was different from that in HSE across age groups, both in men and women, except for the group of men aged 25-34 in which no one had eGFR <60 ml/min/1.73m^2 in HSE. The proportion of people who had recorded measurement of creatinine increased with age, with 26% of men and 46% of women aged 25-34 with tests in the past five years, up to 92% (both men and women) among people aged 75 years or older.

The prevalence of RRT gradually increased according to age (**Table 2** and **Figure 2**). The difference between CRPD and the UKRR was small across all age groups, both in men and women. **Table 3** shows the subgroup analysis by RRT modality. The prevalence of patients with haemodialysis in CPRD was slightly lower than that in UKRR across all age groups, whilst the prevalence of those with peritoneal dialysis and kidney transplantation in CPRD were similar to or slightly higher than those in UKRR.

For Peer Review

Table 4 shows the results of sensitivity analyses. By assuming all creatinine results were non-IDMS traceable, the prevalence of eGFR <60 ml/min/1.73m^2 in CPRD decreased predominantly among older people, and overall prevalence decreased from 6.86% to 5.35%., Restricting to people with recorded ethnicity in CPRD, using serum creatinine value in the past two years, and restricting to English data produced similar results to the main analysis. By defining decreased kidney function including a measure of chronicity, the prevalence decreased slightly in each age group, and overall prevalence decreased from 6.86% to 6.27%. Finally, in a complete case analysis using as the denominator only those with tests for serum creatinine, the prevalence of eGFR <60 ml/min/1.73m^2 became substantially larger than that in the main analysis.

The overall prevalence of eGFR <45 ml/min/1.73m^2 was 2.33% (64,425/2,761,755) in CPRD. The number of people with eGFR <45 ml/min/1.73m^2 was small and confidence intervals of the prevalence estimates were large in HSE (**Table 5**). The proportion of people with eGFR <45 ml/min/1.73m^2 in the age group 75 or older in CPRD was significantly higher than that of HSE, both in men and women.

For Peer Review

Discussion

In this study, we examined the external validity of the prevalence of decreased kidney function (based on serum creatinines measured by GPs) and RRT (based on recorded diagnoses) in CPRD, by comparing them with results from two nationally representative sources (HSE and UKRR). Across all ages for men and women the prevalence of eGFR <60 ml/min/1.73m^2 in CPRD was similar to that in HSE, although the small number of younger patients with eGFR <60 mL/min/1.73m^2 in HSE might have hampered precise comparison. The prevalence of RRT in CPRD was broadly similar to that obtained from UKRR, although there were differences in the RRT modality-specific prevalence between CPRD and UKRR.

Routinely collected primary care data can be a useful resource for epidemiological studies, particularly in the UK, where over 98% of citizens are registered with NHS GPs [11]. Although the prevalence or incidence of various diseases in CPRD have good comparability with other UK-based data sources [12, 13], the external validity of the prevalence of decreased kidney function and RRT has not been studied. Concerns specific to kidney diseases include that GPs do not test every registered patient's kidney function, which could lead to the underestimation of the true prevalence of decreased kidney function. In our study, the proportion of people with creatinine measurement was small among young and middle-aged people, especially men. However, using the entire practice population as a denominator, the prevalence of eGFR <60 ml/min/1.73m^2 in CPRD was close to that in HSE

For Peer Review

across all age groups, both in men and women. A possible explanation would be that, in line with the current National Institute for Health and Care Excellence (NICE) guidance for CKD [5], GPs are efficiently testing kidney function for people with CKD risk factors, including hypertension, diabetes, cardiovascular diseases, and hereditary kidney disease (e.g. autosomal dominant polycystic kidney disease). In addition, the Quality and Outcome Framework (QOF) incentivises GPs to register and manage patients with CKD [30]. Since the launch of QOF for CKD in 2006/7, the identification and management of patients with CKD have been improving in the UK [31], although there are delays in coding patients with CKD in the system [32]. In older age groups, very high proportions had undergone testing of kidney function, and it is likely that those not tested are healthier, with a lower risk of CKD.

In sensitivity analyses, we examined to what extent the prevalence estimates for decreased kidney function changed under different assumptions related to uncertainties in CPRD. Firstly, the estimation changed considerably with the assumption over whether the UK laboratories reported creatinines traceable to IDMS or not. We expect that most of the UK laboratories reported IDMS-traceable creatinines during the study period, yet if a few laboratories reported non-IDMS-traceable creatinines, the true prevalence of eGFR <60 ml/min/1.73m^2 in CPRD would become lower than our estimation in the main analysis. Standardisation of serum creatinine assays is thus important in studies regarding CKD epidemiology. Second, the assumption of non-black ethnicity for people with missing

For Peer Review

ethnicity data in CPRD affected the prevalence estimates only slightly. This is probably because the proportion of people with black ethnicity is small in the UK, at around 3% [18]. Third, using creatinine records for the past two, instead of five, years made little change to prevalence estimates for decreased kidney function. This may relate to recommendations for regular testing in line with QOF and the current NICE guidance for CKD [5]. Fourth, in CPRD the prevalence of eGFR <60 ml/min/1.73m^2 in England was similar to that in the whole UK, ensuring the comparability between HSE and CPRD in our study. Fifth, the prevalence estimates slightly decreased by using the CKD criteria including chronicity. This may suggest that some patients with single eGFR <60 ml/min/1.73m^2 had transient kidney dysfunction, probably because serum creatinine was measured at the time of acute illness when they may have developed acute kidney injury. Finally, the prevalence of decreased kidney function was likely to be overestimated by restricting the denominator to only people with creatinine measurement. This suggests that GPs selectively test people at high risk of CKD, especially among younger people.

The prevalence of RRT was also similar between CPRD and UKRR across all age groups in men and women. Patients receiving RRT are in frequent contact with kidney units, so GPs do not provide comprehensive routine care for these individuals. However, patients on RRT remain registered with their GPs and, therefore, we would anticipate that GPs update patient records to reflect commencement of RRT. Our results suggest that the estimated

For Peer Review

prevalence of RRT based on recorded diagnoses in CPRD was broadly valid when compared against comprehensive UK Registry statistics. However, using the most recent diagnosis indicating RRT modality, the prevalence of haemodialysis was underestimated in CPRD, while those of peritoneal dialysis and kidney transplantation were similar, or somewhat overestimated, especially among older people. This may be because patients with peritoneal dialysis and kidney transplantation are often healthier and have more regular contact with their GPs compared to those on haemodialysis. In addition, for patients with a change in their RRT modality (e.g. from peritoneal dialysis to haemodialysis) there may be a delay in updating the modality in the GP record. Therefore, some patients currently on haemodialysis might be misclassified into the group of peritoneal dialysis or kidney transplantation because their previous diagnoses (i.e. peritoneal dialysis or kidney transplantation) are not yet updated. Another possibility is that patients commencing haemodialysis died before this was recorded in CPRD, given the high early mortality rates of these patients [33].

There are several limitations to our study. First, this is a cross-sectional study examining the validity of prevalence of decreased kidney function and RRT. Our results do not ensure that UK primary care data are reliable for identifying the incidence of CKD and RRT. Unless we were prepared to assume that kidney function stays constant for people without creatinine measurements, we cannot estimate incident CKD for the general population. Second, our comparison of data between CPRD and HSE or UKRR was only at

For Peer Review

the population rather than the individual level. Our analyses did not allow us to calculate sensitivity or specificity of RRT diagnoses. In the absence of linked data, it is possible that there was a similar extent of misclassification between cases and non-cases resulting in an overall agreement of the prevalence estimates in CPRD with those in HSE and UKRR. Third, the prevalence of decreased kidney function in HSE was the best available estimate, but not a perfect reference standard. The survey did not include people who were temporarily hospitalised for acute illness or were in residential care [14]. In addition, people with poor health might be reluctant to give a blood sample, and the existing adjustment for non-response in HSE may not have fully dealt with this bias. This may explain the finding in our sensitivity analysis that the proportion of people with eGFR <45 ml/min/1.73m^2 in the oldest age group in CPRD was significantly higher than that of HSE. Blood sampling was conducted on only one occasion in HSE. Accordingly, we defined decreased kidney function in CPRD using one serum creatinine measurement in our main analysis. However, some patients might have had their kidney function checked as a result of acute illness, and therefore their decreased kidney function might have been transient. A previous research has shown that creatinine fluctuation can affect the CKD prevalence estimates in routinely collected data [34], although the influence was not large in our study. At around 6,000, the sample size in HSE was not small, yet the relatively wide 95% CIs for the prevalence estimates in each age-sex group hampered more precise comparisons. In particular, the

For Peer Review

number of patients with eGFR <60 ml/min/1.73m^2 was small among younger age groups. We could not compare the prevalence of more severe kidney dysfunction, because patients with eGFR <30 ml/min/1.73m^2 were rare even among older people in HSE [14]. Meanwhile, testing of albuminuria is known to be incomplete in UK primary care electronic health records [32], which prevented us from comparing the prevalence of albuminuria, or CKD stage 1 and 2, between CPRD and HSE. Because albuminuria is an important prognostic factor in people with and without low eGFR [35], the unknown validity of albuminuria in UK primary care remains an obstacle to study of CKD using CPRD. Finally, our findings may not be generalisable to other GP practices in the UK if GP practices contributing to CPRD were more likely to measure kidney function and record the diagnoses of RRT. Generalisability to primary care electronic health records in other European countries is also uncertain, because the frequency of practices such as blood testing, chronic disease monitoring, recording of diagnoses, and incentives, and access to public primary care clinics differ.

In the era of a rising global prevalence of ESRD [4], high-quality epidemiological research on kidney diseases is becoming more important. Routinely collected electronic health record data would play an important role for kidney research, because most patients with CKD are diagnosed and managed in primary care. Accurate identification of CKD and RRT in CPRD would allow investigation of the association between kidney diseases and other comorbidities or medications. It is also possible to investigate equity of care (e.g.

For Peer Review

referral to nephrologists), given that the database is less biased for ascertaining advanced CKD than population surveys and disease registry. In this study, we demonstrated that identifying the prevalence of CKD and RRT is valid at the population level in CPRD. Although further validation of individual level data is needed, our findings support the use of UK primary care data for research into kidney disease.

For Peer Review

Conclusions

We examined the external validity of the prevalence of decreased kidney function and RRT in CPRD. The prevalence of eGFR <60 ml/min/1.73m^2 in CPRD was similar to that in a national sampling survey (HSE 2009/2010), and the prevalence of RRT in CPRD was close to that obtained from a national disease registry (UKRR 2014) across all age groups, both in men and women. These findings suggest that UK primary care data can be used to identify the prevalence of decreased kidney function and RRT in future studies.

For Peer Review

Competing interests

The authors declare that they have no any competing interests (both financial and non-financial) related to this manuscript.

Authors' contributions

MI, LT, and DN planned the study. MI carried out the data extraction from Clinical Practice Research Datalink, cleaning, and, analysis, and drafted the manuscript. KM supported the data analysis and drafted the manuscript. AC and FC managed the data from UK Renal Registry 2014. GA, SF and PR managed the data from Health Survey for England 2009 and 2010. All authors contributed substantially to the interpretation of the results, and the writing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Dr Masao Iwagami is funded by Honjo International Scholarship Foundation. Dr Laurie Tomlinson is funded by a Wellcome Trust Intermediate Clinical Fellowship [WT101143MA]. The funders had no role in the execution of this study or interpretation of results.

For Peer Review

References

1. Jha V, Garcia-Garcia G, Iseki K et al. Chronic kidney disease: global dimension and perspectives. *Lancet* 2013; 382: 260-72
2. Castro AF, Coresh J. CKD surveillance using laboratory data from the population-based National Health and Nutrition Examination Survey (NHANES). *Am J Kidney Dis* 2009; 53: S46-55
3. Levey AS, Atkins R, Coresh J et al. Chronic kidney disease as a global public health problem: approaches and initiatives - a position statement from Kidney Disease Improving Global Outcomes. *Kidney Int* 2007; 72: 247-59
4. Liyanage T, Ninomiya T, Jha V et al. Worldwide access to treatment for end-stage kidney disease: a systematic review. *Lancet* 2015; 385: 1975-82
5. The National Institute for Health and Care Excellence: NICE guidelines [CG182] Chronic kidney disease in adults: assessment and management, [online] July 2014 [cited 2016 March 1]. Available from: https://www.nice.org.uk/guidance/cg182
6. Ackland P. Prevalence, detection, evaluation and management of chronic kidney disease. *BMJ* 2014; 348: f7688
7. Brück K, Jager KJ, Dounousi E et al. Methodology used in studies reporting chronic kidney disease prevalence: a systematic literature review. *Nephrol Dial Transplant* 2015; 30: 6-16

For Peer Review

8. McCullough K, Sharma P, Ali T et al. Measuring the population burden of chronic kidney disease: a systematic literature review of the estimated prevalence of impaired kidney function. *Nephrol Dial Transplant* 2012; 27: 1812-21

9. Clinical Practice Research Datalink. [online] [cited 2016 March 1]. Available from: http://www.cprd.com/intro.asp

10. de Lusignan S, van Weel C. The use of routinely collected computer data for research in primary care: opportunities and challenges. *Fam Pract* 2006; 23: 253-63

11. Herrett E, Gallagher AM, Bhaskaran K et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015; 44: 827-36

12. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010; 69: 4-14

13. Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. *Br J Gen Pract* 2010; 60: e128-36

14. Health Survey for England - 2010, Respiratory health: Chapter 8, Kidney disease and renal function. [online] December 15, 2011 [cited 2016 March 1]. Available from: http://www.hscic.gov.uk/pubs/hse10report

15. Rao A, Casula A, Castledine C. UK Renal Registry 17th Annual Report: Chapter 2 UK Renal Replacement Therapy Prevalence in 2013: National and Centre-specific Analyses.

For Peer Review

Nephron 2015; 129 (Suppl 1): 31-56

16. Hollowell J. The General Practice Research Database: quality of morbidity data. *Popul Trends* 1997; 87: 36-40
17. Booth HP, Prevost AT, Gulliford MC. Validity of smoking prevalence estimates from primary care electronic health records compared with national population survey data for England, 2007 to 2011. *Pharmacoepidemiol Drug Saf* 2013; 22: 1357-61
18. Mathur R, Bhaskaran K, Chaturvedi N et al. Completeness and usability of ethnicity data in UK-based primary care and hospital databases. *J Public Health (Oxf)* 2014; 36: 684-92
19. Kontopantelis E, Springate D, Reeves D, Ashcroft DM, Valderas JM, Doran T. Withdrawing performance indicators: retrospective analysis of general practice performance under UK Quality and Outcomes Framework. *BMJ* 2014; 348: g330
20. Health Survey for England - 2010, Respiratory health: Volume 2, Methods and documentation. [online] December 15, 2011 [cited 2016 March 1]. Available from: http://www.hscic.gov.uk/pubs/hse10report
21. 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
22. Levey AS, Stevens LA, Schmid CH et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; 150: 604-12

For Peer Review

23. Tonelli M, Wiebe N, Culleton B et al. Chronic kidney disease and mortality risk: a systematic review. *J Am Soc Nephrol* 2006; 17: 2034-47

24. Anderson J, Glynn LG. Definition of chronic kidney disease and measurement of kidney function in original research papers: a review of the literature. *Nephrol Dial Transplant* 2011; 26: 2793-8

25. Davé S, Petersen I. Creating medical and drug code lists to identify cases in primary care databases. *Pharmacoepidemiol Drug Saf* 2009; 18: 704-7

26. Levey AS, Coresh J, Greene T et al. Expressing the Modification of Diet in Renal Disease Study equation for estimating glomerular filtration rate with standardized serum creatinine values. *Clin Chem* 2007; 53: 766-72

27. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney Int* 2013; Suppl 3(1):1-150.

28. Roderick PJ, Atkins RJ, Smeeth L et al. Detecting chronic kidney disease in older people; what are the implications? *Age Ageing* 2008; 37: 179-86

29. Roderick PJ, Atkins RJ, Smeeth L et al. CKD and mortality risk in older people: a community-based population study in the United Kingdom. *Am J Kidney Dis* 2009; 53: 950-60

30. Quality and Outcomes Framework. [online] October 29, 2015 [cited 2016 March 1]. Available from: http://www.hscic.gov.uk/qof

For Peer Review

31. Stevens PE, de Lusignan S, Farmer CK, Tomson CR. Engaging primary care in CKD initiatives: the UK experience. *Nephrol Dial Transplant* 2012; 27 (Suppl 3): iii5-11

32. Fraser SD, Parkes J, Culliford D, Santer M, Roderick PJ. Timeliness in chronic kidney disease and albuminuria identification: a retrospective cohort study. *BMC Fam Pract* 2015; 16: 18

33. Goodkin DA, Young EW, Kurokawa K, Prutz KG, Levin NW. Mortality among hemodialysis patients in Europe, Japan, and the United States: case-mix effects. *Am J Kidney Dis* 2004; 44(5 Suppl 2): 16-21

34. de Lusignan S, Tomson C, Harris K, van Vlymen J, Gallagher H. Creatinine fluctuation has a greater effect than the formula to estimate glomerular filtration rate on the prevalence of chronic kidney disease. *Nephron Clin Pract* 2011; 117: c213-24

35. 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: 2073-81

For Peer Review

Table 1. Prevalence of estimated glomerular filtration rate <60 mL/min/1.73m^2 in Clinical Practice Research Datalink and Health Survey for England.

	Age group (years)					
	25-34	35-44	45-54	55-64	65-74	75+
Men:						
Prevalence of eGFR <60 mL/min/1.73m^2 in CPRD, % (95%CI)	0.11 (0.10 – 0.12)	0.27 (0.25 – 0.29)	0.76 (0.73 – 0.79)	2.59 (2.53 – 2.66)	10.16 (10.02 – 10.29)	35.32 (35.07 – 35.57)
Prevalence of eGFR <60 mL/min/1.73m^2 in HSE, % (95%CI)	0	0.19 (0 – 1.06)	1.22 (0.45 – 2.63)	1.94 (0.84 – 3.78)	14.04 (10.27 – 18.56)	31.39 (25.36 – 37.92)
Difference (prevalence in CPRD - that in HSE), % (95%CI)	0.11 (0.09 – 0.12)	0.08 (-0.30 – 0.45)	-0.46 (-1.43 – 0.51)	0.65 (-0.68 – 1.99)	-3.88 (-7.87 – 0.10)	3.93 (-2.17 – 10.03)
Proportion of patients with serum creatinine measurement in past five years in CPRD, % (numerator/denominator)	25.85 (61,339/237,284)	38.47 (98,759/256,739)	55.61 (163,001/293,104)	72.44 (167,841/231,695)	86.15 (163,933/190,292)	92.29 (132,103/143,144)
Women:						
Prevalence of eGFR <60 mL/min/1.73m^2 in CPRD, % (95%CI)	0.10 (0.09 – 0.12)	0.27 (0.25 – 0.29)	0.90 (0.87 – 0.94)	3.22 (3.15 – 3.29)	11.13 (10.99 – 11.27)	38.50 (38.29 – 38.72)
Prevalence of eGFR <60 mL/min/1.73m^2 in HSE, % (95%CI)	0.65 (0.13 – 1.89)	0.78 (0.21 – 1.97)	2.00 (0.96 – 3.64)	4.70 (2.93 – 7.09)	9.48 (6.53 – 13.19)	35.41 (30.04 – 41.06)
Difference (prevalence in CPRD - that in HSE), % (95%CI)	-0.55 (-1.29 – 0.19)	-0.51 (-1.27 – 0.25)	-1.10 (-2.32 – 0.13)	-1.48 (-3.44 – 0.48)	1.65 (-1.53 – 4.83)	3.09 (-2.28 – 8.47)
Proportion of patients with serum creatinine measurement in past five years in CPRD, % (numerator/denominator)	46.22 (108,767/235,341)	55.30 (139,977/253,145)	67.35 (192,872/286,386)	75.27 (174,268/231,517)	84.45 (171,620/203,227)	91.88 (183,655/199,881)

CI = confidence interval, CPRD = Clinical Practice Research Datalink, eGFR = estimated glomerular filtration rate, HSE = Health Survey for England

For Peer Review

Table 2. Prevalence of renal replacement therapy in Clinical Practice Research Datalink and UK Renal Registry.

	Age group (years)					
	25-34	35-44	45-54	55-64	65-74	75+
Men:						
Prevalence of RRT in CPRD, % (95%CI)	0.05 (0.04 – 0.06)	0.09 (0.08 – 0.10)	0.14 (0.12 – 0.15)	0.17 (0.16 – 0.19)	0.22 (0.20 – 0.24)	0.25 (0.23 – 0.28)
Prevalence of RRT in UKRR, % (95%CI)	0.05 (0.05 – 0.06)	0.10 (0.10 – 0.10)	0.17 (0.16 – 0.17)	0.22 (0.21 – 0.22)	0.25 (0.25 – 0.26)	0.29 (0.28 – 0.29)
Difference (prevalence in CPRD - that in UKRR), % (95%CI)	0 (-0.01 – 0.01)	-0.01 (-0.02 – 0)	-0.03 (-0.05 – -0.02)	-0.05 (-0.06 – -0.03)	-0.03 (-0.05 – -0.01)	-0.03 (-0.06 – 0)
Women:						
Prevalence of RRT in CPRD, % (95%CI)	0.04 (0.03 – 0.05)	0.07 (0.06 – 0.08)	0.09 (0.08 – 0.11)	0.13 (0.11 – 0.14)	0.15 (0.13 – 0.17)	0.12 (0.10 – 0.13)
Prevalence of RRT in UKRR, % (95%CI)	0.04 (0.03 – 0.04)	0.07 (0.06 – 0.07)	0.11 (0.10 – 0.11)	0.13 (0.13 – 0.14)	0.15 (0.15 – 0.16)	0.11 (0.11 – 0.12)
Difference (prevalence in CPRD - that in UKRR), % (95%CI)	0 (-0.01 – 0.01)	0 (-0.01 – 0.01)	-0.01 (-0.03 – 0)	-0.01 (-0.02 – 0.01)	0 (-0.02 – 0.02)	0 (-0.01 – 0.02)

CI = confidence interval, CPRD = Clinical Practice Research Datalink, RRT = renal replacement therapy, UKRR = UK renal registry

For Peer Review

Table 3. Prevalence of renal replacement therapy by modality in Clinical Practice Research Datalink and UK Renal Registry.

	Age group (years)					
	25-34	35-44	45-54	55-64	65-74	75+
Clinical Practice Research Datalink:						
Denominator, N	472,625	509,884	579,490	463,212	393,519	343,025
Number of patients with haemodialysis, n (‰)	39 (0.08)	84 (0.16)	144 (0.25)	202 (0.44)	257 (0.65)	378 (1.10)
Number of patients with peritoneal dialysis, n (‰)	27 (0.06)	15 (0.03)	48 (0.08)	37 (0.08)	67 (0.17)	67 (0.20)
Number of patients with kidney transplantation, n (‰)	141 (0.30)	299 (0.59)	480 (0.83)	455 (0.98)	399 (1.01)	154 (0.45)
UK Renal Registry:						
Denominator, N	8,676,837	8,463,148	9,030,893	7,297,460	6,030,602	5,101,203
Number of patients with haemodialysis, n (‰)	887 (0.10)	1,677 (0.20)	3,513 (0.39)	4560 (0.62)	5939 (0.98)	7324 (1.44)
Number of patients with peritoneal dialysis, n (‰)	180 (0.02)	321 (0.04)	585 (0.06)	740 (0.10)	918 (0.15)	830 (0.16)
Number of patients with kidney transplantation, n (‰)	2,836 (0.33)	5,047 (0.60)	8,361 (0.93)	7538 (1.03)	5224 (0.87)	1269 (0.25)

For Peer Review

Table 4. Prevalence of estimated glomerular filtration rate <60 mL/min/1.73m^2 in Clinical Practice Research Datalink: results of main analysis and sensitivity analyses.

	Age group (years)						Total
	25-34	35-44	45-54	55-64	65-74	75+	
Main analysis, % (numerator/denominator):	0.11 (497/472,625)	0.27 (1,360/509,884)	0.83 (4,800/579,490)	2.90 (13,455/463,212)	10.66 (41,949/393,519)	37.18 (127,520/343,025)	6.86 (189,581/2,761,755)
Sensitivity analyses, % (numerator/denominator):							
(i) Assuming all creatinine results non-IDMS traceable	0.09 (429/472,625)	0.21 (1,066/509,884)	0.56 (3,246/579,490)	1.86 (8,612/463,212)	7.45 (29,304/393,519)	30.66 (105,171/343,025)	5.35 (147,828/2,761,755)
(ii) Complete case analysis for ethnicity	0.11 (332/295,942)	0.27 (815/299,641)	0.93 (2,710/292,837)	3.24 (7,313/225,881)	11.26 (22,496/199,805)	38.25 (64,923/169,738)	6.64 (98,589/1,483,844)
(iii) Using creatinine records in past two years	0.09 (439/472,625)	0.24 (1,202/509,884)	0.73 (4,202/579,490)	2.59 (11,993/463,212)	9.93 (39,081/393,519)	35.08 (120,323/343,025)	6.42 (177,240/2,761,755)
(iv) Restricting region to England	0.11 (368/346,641)	0.26 (999/377,675)	0.85 (3,596/423,030)	3.00 (9,935/331,404)	10.71 (30,302/282,983)	37.22 (93,887/252,246)	6.91 (139,087/2,013,979)
(v) Using CKD criteria including chronicity[a]	0.07 (353/472,625)	0.19 (977/509,884)	0.56 (3,234/579,490)	2.19 (10,156/463,212)	9.34 (36,770/393,519)	35.44 (121,564/343,025)	6.27 (173,054/2,761,755)
(vi) Complete case analysis for creatinine	0.29 (497/170,148)	0.57 (1,360/238,786)	1.35 (4,800/355,929)	3.93 (13,455/342,151)	12.50 (41,949/335,581)	40.38 (127,520/315,777)	10.78 (189,581/1,758,372)

CKD = chronic kidney disease, IDMS = isotope dilution mass spectrometry

[a]estimated glomerular filtration rate <60 mL/min/1.73m^2 twice consecutively for ≥3 months in the past five years

For Peer Review

Table 5. Prevalence of estimated glomerular filtration rate <45 mL/min/1.73m^2 in Clinical Practice Research Datalink and Health Survey for England.

	Age group (years)					
	25-34	35-44	45-54	55-64	65-74	75+
Men:						
Prevalence of eGFR <45 mL/min/1.73m^2 in CPRD, % (95%CI)	0.08 (0.07 – 0.09)	0.16 (0.14 – 0.17)	0.30 (0.28 – 0.32)	0.66 (0.63 – 0.70)	2.36 (2.30 – 2.43)	13.49 (13.32 – 13.67)
Prevalence of eGFR <45 mL/min/1.73m^2 in HSE, % (95%CI)	0	0	0.41 (0.05 – 1.46)	0	4.11 (2.14 – 7.07)	8.97 (5.56 – 13.51)
Difference (prevalence in CPRD - that in HSE), % (95%CI)	0.08 (0.07 – 0.09)	0.16 (0.14 – 0.17)	-0.11 (-0.67 – 0.45)	0.66 (0.63 – 0.70)	-1.75 (-4.02 – 0.53)	4.53 (0.77 – 8.28)
Women:						
Prevalence of eGFR <45 mL/min/1.73m^2 in CPRD, % (95%CI)	0.06 (0.05 – 0.07)	0.12 (0.11 – 0.14)	0.24 (0.23 – 0.26)	0.60 (0.57 – 0.63)	2.36 (2.29 – 2.43)	15.14 (14.99 – 15.30)
Prevalence of eGFR <45 mL/min/1.73m^2 in HSE, % (95%CI)	0	0	0.40 (0.05 – 1.43)	0.89 (0.24 – 2.28)	2.75 (1.27 – 5.16)	10.82 (7.57 – 14.86)
Difference (prevalence in CPRD - that in HSE), % (95%CI)	0.06 (0.05 – 0.07)	0.12 (0.11 – 0.14)	-0.16 (-0.71 – 0.40)	-0.29 (-1.17 – 0.58)	-0.39 (-2.17 – 1.38)	4.32 (0.83 – 7.81)

CI = confidence interval, CPRD = Clinical Practice Research Datalink, eGFR = estimated glomerular filtration rate, HSE = Health Survey for England

For Peer Review

Figure legends

Figure 1. Prevalence of estimated glomerular filtration rate <60 mL/min/1.73m^2 in Clinical Practice Research Datalink and Health Survey for England

CPRD = Clinical Practice Research Datalink, HSE = Health Survey for England

Figure 2. Prevalence of renal replacement therapy in Clinical Practice Research Datalink and UK Renal Registry

CPRD = Clinical Practice Research Datalink, UKRR = UK Renal registry

For Peer Review

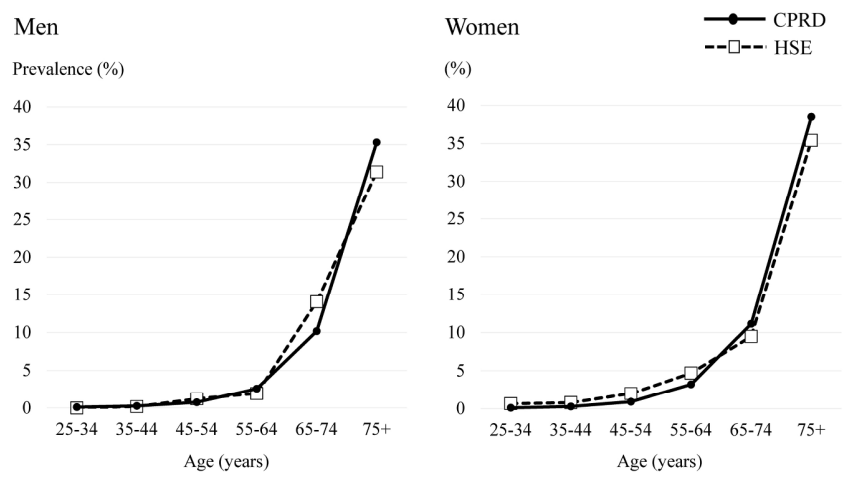

Figure 1. Prevalence of estimated glomerular filtration rate <60 mL/min/1.73m2 in Clinical Practice Research Datalink and Health Survey for England
CPRD = Clinical Practice Research Datalink, HSE = Health Survey for England
209x148mm (300 x 300 DPI)

For Review

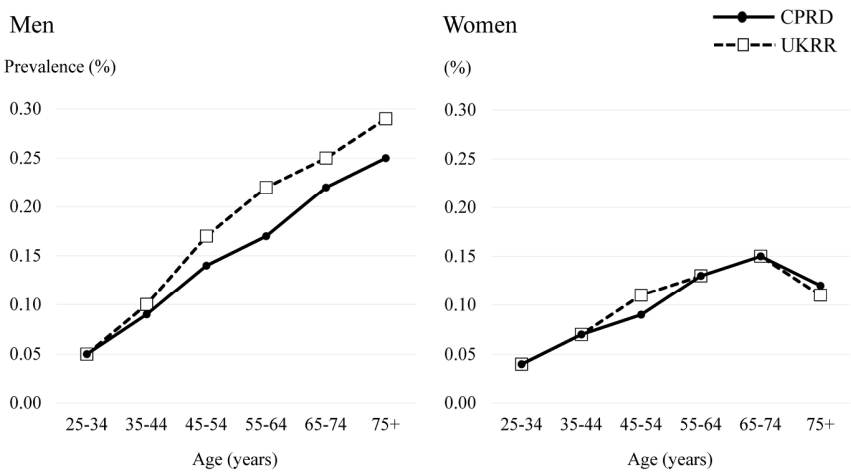

Figure 2. Prevalence of renal replacement therapy in Clinical Practice Research Datalink and UK Renal Registry
CPRD = Clinical Practice Research Datalink, UKRR = UK Renal registry
209x148mm (300 x 300 DPI)

For Review