Author's response to reviews

Title: Identification of people with autosomal dominant polycystic kidney disease using routine data: a cross sectional study

Authors:

Andrew P McGovern (andy@mcgov.co.uk)
Simon Jones (simonjones@surrey.ac.uk)
Jeremy van Vlymen (j.vanvlymen@surrey.ac.uk)
Anand K Saggar (a.saggar@sgul.ac.uk)
Richard Sandford (rns13@medschl.cam.ac.uk)
Simon de Lusignan (s.lusignan@surrey.ac.uk)

Version: 3
Date: 14 October 2014

Author's response to reviews: see over
Response to reviewers

Reviewer 1

Major Compulsory Revisions

McGovern et al have examined a large repository of primary care clinical data in order to determine whether it is possible to use routinely collected data in those under 60 in order to identify individuals at high risk of polycystic kidney disease, with a view to potential future demand for population screening for ADPKD. Major strengths of the study are the large sample size and the high quality of the statistical analysis used. In this real-world UK population the prevalence of a diagnosis of polycystic kidney disease in primary care records (1:2700) is well below the accepted prevalence of the disease (~1:800), implying that the diagnosis is not being made (or at least accurately recorded in GP records) in a significant number of people with the disease.

This is indeed a limitation of using currently available primary care datasets. The apparently low prevalence of PKD will also be partly because the condition is not yet apparent in younger members of our population. This limitation is now made clearer in the limitations section of the discussion.

We are also aware of recent debate over the true prevalence of ADPKD with many commentators suggesting that the true prevalence of apparent disease is lower than previously estimated – indeed a recent population based analysis demonstrates an estimate for clinically apparent disease of ~1 in 3000 (Neumann et al., 2013). We have also updated the manuscript to highlight this point.

However, there are a number of issues which make it difficult for this paper to achieve its stated aim of showing that routine data could be used to stratify patients’ risk of having undiagnosed ADPKD.

1. The PKD group are identified by their coding in the primary care records. Since their primary care provider must therefore be aware of the diagnosis, it follows that they are more likely to undergo blood pressure and renal function measurements than the non-PKD population, and that more aggressive blood pressure control is likely to be employed, so statistical differences in these parameters may be caused by the diagnosis rather than the disease. This hypothesis is supported by their finding that undocumented renal function has an odds ratio for PKD of 0.28 (p<0.001). This ascertainment bias implies that extrapolation to the undiagnosed PKD population of the identified risk factors may not be completely justified.

We concur with the reviewers comments here. A longitudinal analysis would help to determine whether renal function, blood pressure, or other abnormalities occurred before the diagnosis of ADPKD was made. However these changes and the disease progression have been established elsewhere. The dataset we have used for this analysis only has six years’ worth of high quality longitudinal data and is therefore not fully suitable for this analysis given the rarity of ADPKD and the small number of individuals with the condition. We have now added a discussion of this limitation to the discussion section of the manuscript. We again feel that this highlights the need for a larger integrated primary and secondary care dataset.
2. The authors build a regression model using a number of parameters including CKD stages 3B (OR 42.35), 4 (OR 101.27) and 5 (OR 75.08) and haematuria (OR 1.85). According to the National Institute of Health and Clinical Excellence Guidelines for the Identification and Management of Chronic Kidney Disease, these are all indications for a renal ultrasound scan. This implies that individuals with these clinical features should all be either diagnosed with PKD or have a diagnosis of PKD excluded, so incorporating these factors in a risk score for undiagnosed PKD is unlikely to improve the clinical management of these patients. In addition, proteinuria and hypertension (particularly hypertension requiring multiple agents) in a young (<60 year old) patient is likely to warrant clinical investigation that would include a renal ultrasound scan in many cases.

Our findings suggest that a combination of early stage renal impairment (CKD stage 3a) with other risk factors (proteinuria, diastolic BP over 90mmHg) are associated with a high likelihood of ADPKD and may merit ultrasound scanning (in addition to the indications for ultrasound scan already recommended by NICE). Use of a risk stratification tool may prompt earlier screening in individuals with these features before they would otherwise be referred as recommended in current guidelines. We have also added a discussion of this limitation to the manuscript.

3. There is no mention of the role of family history in stratification of patients. Since ADPKD is a Mendelian condition with high penetrance, a majority (probably over 80%) of affected individuals have a family history of kidney disease. According to the National Institute of Health and Clinical Excellence Guidelines for the Identification and Management of Chronic Kidney Disease a family history is also an indication for offering screening for this condition. Unfortunately this data is also not available from our dataset, however, as stated people with a family history of ADPKD should be screened according to existing guidelines. We have again added this as a limitation of our analysis.

I do not think the aim of this paper has been completely met. In order to address the question posed, the authors would need to re-analyse the data including only those patients in whom ultrasound scanning is not indicated by clinical data and family history under current national guidelines.

As the reviewer notes, the objective for us in this work is to attempt to ascertain whether or not routine data can be used in risk stratification of rare diseases using ADPKD as a primary case. Should this approach demonstrate promise then additional analyses using merged primary and secondary care datasets (e.g. UK Renal Registry data) could be undertaken to refine/confirm the identification model. A merged primary and secondary care dataset fully suitable for this purpose is not currently available in the UK, to the best of our knowledge, and would require additional financial resources to generate.

The limitations of all existing UK based primary care big datasets, that we are aware of, preclude this analysis at present. We are therefore unable to comply with the reviewers request for incorporation of these additional clinical features. We feel that publication of these preliminary analyses will help to facilitate the collection and amalgamation of an appropriate dataset for such an analysis. In this regard we have downscaled the conclusion of the paper as detailed above. However we have the limitations of the available data preclude our undertaking of the additional analyses suggested by the reviewer.
We still feel our analysis demonstrates that routine data has the potential to be used as an adjunctive screening tool in the identification of ADPKD and one which merits further investigation using a linked primary and secondary care dataset. We hope that publication of these preliminary data will help to facilitate this goal.

We have reworked the manuscript to state these goals more clearly and to highlight that this is a preliminary analysis of this method:

1. We now state in the abstract conclusion that stratification of ADPKD risk using routine primary care data may be possible rather than that is possible.
2. We have changed the title from “Early identification of people with autosomal dominant polycystic kidney disease: a cross sectional study” to “Identification of people with autosomal dominant polycystic kidney disease using routine data: a cross sectional study”. We feel this reflects the slightly more modest aims of the paper.
3. In the limitations section we now highlight that we are unable to determine which features become apparent before a clinical diagnosis is made and therefore a longitudinal analysis is required to confirm that screening using routine data is possible.
4. The conclusion in the main manuscript has now been updated to the following:

“Whilst the limitations of the available data prevent us from conclusively demonstrating that early ADPKD case finding is possible using routine primary care data, this approach appears promising. A longitudinal analysis using linked primary and secondary care data could more conclusively demonstrate whether or not this method could be used in ADPKD screening. If this approach proves successful it could be used to form the basis for or an adjunct to an ADPKD screening program.”
Reviewer 2

Minor Essential Revisions:

1. Abstract. Line 1: The word "disease" should be moved between "kidney" and (ADPKD).

   This typographical error has now been corrected.

2. Abstract. Line 2: The word "new" should be deleted.

   This typographical error has now been corrected.

3. Abstract. Results. Line 3: The word "area under" is repeated. It should be deleted.

   This typographical error has now been corrected.

4. Background. Line 2: Instead of "....and results....", it would be better to write "... and may result....."

   This sentence has been altered as suggested by the reviewer.

5. Background. Reference 6 at the end of the first paragraph is not the appropriate reference.

   The median survival ages quoted in this sentence are excerpted from reference 6. We have therefore left this reference unchanged at present. If we are still felt to be in error we will gladly amend this.

6. Which formula is used for the calculation of eGFR?

   We used the MDRD equation as this is still the universally used eGFR formula used in UK primary care. We have now made this clear in the methods section of the manuscript.

7. How many ADPKD patients are included in the study? In the abstract it is written as 225. In the text, it is 255.

   This is a further typographical error in the abstract. The number of 255 appearing in the main body of the manuscript is correct. We have amended this and have carefully reviewed the manuscript for further typographical errors.

8. The references should be written in a standard format.

   We have updated the EndNote referencing style template we were using with the latest version from the publishers website and adjusted the references accordingly.
Editorial Request:

1. Please confirm whether you have obtained the relevant permission to access patient database and use patient data. Please ensure that a statement is added within the Methods section of the manuscript.

   We obtained permission and ethical approval for the use of the data for this analysis. We have now added this information to the methods section of the manuscript (under the subsection ‘ethical considerations’).

2. Please remove the revision with track changes document from the additional files to comply with the guidelines.

   We have made this amendment and submit a final version of the manuscript.

3. Formatting - You now have an opportunity to check that your revised manuscript conforms to the journal style (http://www.biomedcentral.com/info/ifora/medicine_journals ). It is important that your files are correctly formatted.

   We have reviewed these guidelines and have made sure our updated manuscript complies with the formatting requirements.