## PEER REVIEW HISTORY



## ARTICLE DETAILS

| | |
|---|---|
| **TITLE (PROVISIONAL)** | Prevalence and burden of chronic kidney disease among the general population and high risk groups in Africa: a systematic review" |
| **AUTHORS** | Abd ElHafeez, Samar; Bolignano, Davide; D'Arrigo, Graziella; Dounousi, Evangelia; Tripepi, Giovanni; Zoccali, Carmine |

## VERSION 1 – REVIEW

| | |
|---|---|
| **REVIEWER** | Dr Dorothea Nitsch<br>LSHTM, UK<br>competing interests:<br>I am currently involved in a study to derive a validated eGFR for Subsaharan Africa and Steering committee member on the DEGREE study (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5210224/) |
| **REVIEW RETURNED** | 12-Jan-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | The study question is clearly very important, and it is very helpful to see what has been done so far on CKD in Africa. The data despite all the inherent limitations (discussed below) are useful for anybody wanting to raise awareness and for having an argument about CKD indeed being a major public health issue in Africa.<br><br>My main concerns are twofold<br><br>1.<br>There is currently no validated eGFR formula for Africa, and there is ongoing discussion of whether the African-American coefficient in the CKDEPI or MDRD equations should be used, or not. The use of this coefficient in the estimation of eGFR from serum creatinine shifts the mean of a eGFR distribution by 20% and therefore can substantially distort all prevalence estimates. Similar issues are known for calibration of creatinine. There is a related issue that comparisons between different populations based on eGFR can be confounded by the age/sex distribution of the sample, as well as by differences in muscle mass which can vary widely depending on nutritional status. Moreover the study quality varies greatly with respect to sampling processes and represenativeness of the studied population.<br><br>Suggestions:<br>Could the authors clarify in the tables whether the African-American correction coefficients were used or not? The authors do mention it in the methods but it would be useful in the tables to make these more informative to the reader. |

| | |
|---|---|
| | Could the authors attempt some form of age/sex standardisation for the geographical comparisons for the high quality general population studies that they have identified? (Figure 2)<br><br>Could the authors discuss biases in estimating eGFR and how this would have affected prevalence estimates in more depth than currently done? I have in the past simulated what a 5-10 ml/min systematic measurement error in the eGFR estimation for a prevalence survey, and prevalences can in the worst case halve or double especially if the mean eGFR is between 70-80ml/min (this is not published, but fairly easy to check using either simulated normal distribution data or some CKD survey data which the authors may have access to). Similar issues apply to proteinuria dipstick studies - could the authors report whether dipsticks were read using an electronic reader (would this influence the study quality rating)?<br><br>For the HIV studies - could the authors discuss to which extent malnutrition at start of ART may affect the prevalence estimates and if yes, in which direction?<br><br>The listed issues around the problem of doing comparative CKD prevalence studies should be mentioned in the abstract and discussed more fully in the main body of the paper. I agree that many biases outlined above mean that what was reported here may well be an underestimate of the true problem, but I would like to see a fuller discussion of why this might be.<br><br>2. Due to the heterogeneity used in the different studies the authors didn't want to pool the results - this is appropriate. Hence I don't understand how the authors can then present an unknown form of summary statistics in graphical format in figures 2&3? I didn't understand how the authors ended up with Figure 2 (how did they pool studies in for a respective geographical area?), and Figure 3. Figure 3 is a pretty pie chart as to the underlying attributed reasons for CKD by local clinicians - how were these estimates combined and how was confounding by age/sex and selection (e.g. access to care/diagnostics) into the CKD studies dealt with? Please clarify the underlying calculations in the statistical section. Please discuss under which assumptions the statistics displayed in these two figures would be meaningful. |

| | |
|---|---|
| **REVIEWER** | Vivek Jha<br>George Institute for Global Health, India |
| **REVIEW RETURNED** | 17-Jan-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | Overall, this is a good effort that for the first time provides comprehensive estimates for the entire continent. Understandably, there are caveats in such analysis because of problems with the quality of source data. It would be good if the the authors could address the following points:<br><br>Why did they limit their search to the two databases, PubMed and Ovid-Medline. What is the estimate of the studies that might have been missed by this strategy, since many local journals might not be covered in these databases. |

| | |
|---|---|
| | What was the basis of assigning "high" quality to a study. Did the authors consider the assay methodology used in these studies and were satisfied that they passed the quality threshold, and were abnormal values confirmed?<br><br>According to their analysis, the pooled CKD prevalence was 9.4% but the prevalence of eGFR<60 and proteinuria was 4.1%, suggesting that about 5.3% were in the >60 eGFR range, a point where classification as CKD is problematic, especially with a single data point.<br><br>Another point that needs explanation is the lower prevalence of CKD in HIV-infected subjects (4.4%), which is at odds with multiple reports that HIV was one of the commonest causes of CKLD in Sub-Saharan Africa.<br><br>The discussion needs to highlight the methodological issues that might introduce uncertainties around the data, and the conclusions need to be softened to reflect this. |

| | |
|---|---|
| **REVIEWER** | Joseph Lunyera, MBChB, MSc.<br>Duke University School of Medicine, Durham, NC 27701, United States |
| **REVIEW RETURNED** | 19-Feb-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | Dr. Zoccali and colleagues systematically reviewed the literature to describe the burden of chronic kidney disease (CKD) in Africa. By building on and extending prior work done by Stanifer and colleagues, Dr. Zoccali's team present a much-needed update on the burden of CKD in Africa. As noted by the authors, this updated systematic review spans the entire continent – thus extending Stanifer et al's work which was limited to sub-Saharan Africa.<br><br>While I endorse this effort, I have major concerns regarding the date on which the search was last run: Dr. Zoccali's team last run their search on September 25th, 2014 - which is more than 2 years from the manuscript submission date. Thus, the authors need to update their search to include more recent publications that are relevant to this topic (as a rule, the search date should not be >12 months older than the date of manuscript submission). Other than the outdated search date, the manuscript is well written, and the authors have addressed the key limitations<br>– particularly as it pertains to pooling the data using meta-analytic techniques.<br><br>If the authors agree to update the search, they should also consider the following:<br>• Use a database (e.g., Web of Science and/or Scopus) to search the reference list of included studies (Dr. Zoccali and colleagues appear to have hand-searched the reference lists of the included studies – which is no longer recommended as the standard of practice when conducting systematic reviews).<br>• Report the inter-rater agreement for study inclusion and quality assessment using cohen's kappa coefficient. |

| | |
|---|---|
| **REVIEWER** | Sirirat Anutrakulchai<br>Division of Nephrology, Department of Internal Medicine, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand 40002 |
| **REVIEW RETURNED** | 28-Feb-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | 1. Some references are not completely described, for example, no mention of volume and page numbers in the reference # 33, 120 and 156.<br>2. Figure 3 is represented the etiologies of CKD but the sum of percentage is less than 100 %, what is the rest ?<br>3. Regarding to limitations, the enrolled papers were studied during the wide-ranged period form 1995 to 2014. Therefore, the variation of CKD prevalence and etiologies might be partly affected by chronological change. Additionally, methods of GFR, serum creatinine and proteinuria assessments were periodically changed when the better suggested methods appeared which make difficulty in comparisons between these papers. |

## VERSION 1 – AUTHOR RESPONSE

**Reviewer 1**
Reviewer Name: Dr Dorothea Nitsch
Institution and Country: LSHTM, UK
Please state any competing interests: I am currently involved in a study to derive a validated eGFR for Subsaharan Africa and Steering committee member on the DEGREE study(https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5210224/)

Please leave your comments for the authors below
Thank you for giving me the opportunity to review this exciting systematic review.
The study question is clearly very important, and it is very helpful to see what has been done so far on CKD in Africa. The data despite all the inherent limitations (discussed below) are useful for anybody wanting to raise awareness and for having an argument about CKD indeed being a major public health issue in Africa.
My main concerns are twofold:

There is currently no validated eGFR formula for Africa, and there is ongoing discussion of whether the African-American coefficient in the CKDEPI or MDRD equations should be used, or not. The use of this coefficient in the estimation of eGFR from serum creatinine shifts the mean of a eGFR distribution by 20% and therefore can substantially distort all prevalence estimates. Similar issues are known for calibration of creatinine. There is a related issue that comparisons between different populations based on eGFR can be confounded by the age/sex distribution of the sample, as well as by differences in muscle mass which can vary widely depending on nutritional status. Moreover the study quality varies greatly with respect to sampling processes and representativeness of the studied population.

Suggestions
Comment :
Could the authors clarify in the tables whether the African-American correction coefficients were used or not? The authors do mention it in the methods but it would be useful in the tables to make these more informative to the reader.

Response: We agree with the referee for his suggestion. We already included the correction factor in the table for the studies used it in estimating GFR. We referred to it in the tables, reference numbers 12 (P:19, table 2) , 13(P:20, table 2) and 14 (P:36, table 3). Also, in the method section we reported that when correction factor was added , we included the prevalence based on it (P:5, L :23).
As regards the heterogeneity in the studies in respect to sampling process and representativeness, we tried to overcome this problem by calculating the pooled prevalence based on the high/medium quality studies. This may ensure similarity of the included studies based on the methodology and sampling technique.

Comment:
Could the authors attempt some form of age/sex standardisation for the geographical comparisons for the high quality general population studies that they have identified? (Figure 2)

Response: Unfortunately, based on the available data we collected from different studies, we can't do the standardization as this requires having information about age and sex strata from each study and this was not available, see Tripepi G. et al. Nephron Clin Pract 2010;116:c322–c325)

Comment:
Could the authors discuss biases in estimating eGFR and how this would have affected prevalence estimates in more depth than currently done? I have in the past simulated what a 5-10 ml/min systematic measurement error in the eGFR estimation for a prevalence survey, and prevalences can in the worst case halve or double especially if the mean eGFR is between 70-80ml/min (this is not published, but fairly easy to check using either simulated normal distribution data or some CKD survey data which the authors may have access to). Similar issues apply to proteinuria dipstick studies - could the authors report whether dipsticks were read using an electronic reader (would this influence the study quality rating)?

Response: We added it to the discussion (P: 11, L:23-25, P:12, L:1-13)

Comment:
For the HIV studies - could the authors discuss to which extent malnutrition at start of ART may affect the prevalence estimates and if yes, in which direction?

Response: we have now discussed the problem (P:14, L:10-15)

Comment: The listed issues around the problem of doing comparative CKD prevalence studies should be mentioned in the The listed issues around the problem of doing comparative CKD prevalence studies should be mentioned in the abstract and discussed more fully in the main body of the paper. I agree that many biases outlined above mean that what was reported here may well be an underestimate of the true problem, but I would like to see a fuller discussion of why this might be.

Response:: thank you for your comment. We added it to the abstract (conclusion section) and to the strengths and limitations. In the discussion we explained in details the problems that we faced in analyzing data from different studies , specially when the data are of poor quality (P:11, L:22-25; P:12, L: 1-18)

Comment:Due to the heterogeneity used in the different studies the authors didn't want to pool the results - this is appropriate. Hence I don't understand how the authors can then present an unknown form of summary statistics in graphical format in figures 2&3? I didn't understand how the authors ended up with Figure 2 (how did they pool studies in for a respective geographical area?), and Figure 3. Figure 3 is a pretty pie chart as to the underlying attributed reasons for CKD by local clinicians - how were these estimates combined and how was confounding by age/sex and selection (e.g. access to care/diagnostics) into the CKD studies dealt with? Please clarify the underlying calculations in the statistical section. Please discuss under which assumptions the statistics displayed in these two figures would be meaningful.

Response: We added an explanation about the way of calculating the pooled prevalence in the method section (P: 6, L:11-16). We agree with the referee that the data from these 2 figures should be interpreted with limitation due to the heterogeneity in the definition of CKD, information bias due to differences in underlying attributed reasons of CKD by local clinicians, and confounding due to age and gender differences among countries and geographical areas. We extensively referred to these limitations in the discussion
For Fig 3, we agree with you that pie chart may not be the appropriate chart and for that we changed it to bar chart.
We have now specified this in the statistical analysis section (P: 6, L:13-16)

**Reviewer: 2**
Reviewer Name: Vivek Jha
Institution and Country: George Institute for Global Health, India
Please state any competing interests: None
Please leave your comments for the authors below

Comment: In this systematic review, ElHafeez et al have done an admirable effort of collecting the published data on the prevalence of CKD in Africa.
Overall, this is a good effort that for the first time provides comprehensive estimates for the entire continent. Understandably, there are caveats in such analysis because of problems with the quality of source data. It would be good if the authors could address the following points:
Why did they limit their search to the two databases, PubMed and Ovid-Medline. What is the estimate of the studies that might have been missed by this strategy, since many local journals might not be covered in these databases.

Response: We agree with you that it would be better if we included more databases. Unfortunately, we don’t have access to the African databases. To mitigate this limitation we carefully screened all references listed in the original papers included in our systematic review to identify those we might have missed. We referred to this as one of the limitations of our review (P:15, L:6-7)

Comment: What was the basis of assigning "high" quality to a study. Did the authors consider the assay methodology used in these studies and were satisfied that they passed the quality threshold, and were abnormal values confirmed?
R: we evaluated the quality of the different studies based on the tools described in method section (P:6, L: 1-6)

According to their analysis, the pooled CKD prevalence was 9.4% but the prevalence of eGFR<60 and proteinuria was 4.1%, suggesting that about 5.3% were in the >60 eGFR range, a point where classification as CKD is problematic, especially with a single data point.

Response: the reported prevalence of 9.4% which was updated to 10.1% after including the recent studies was reported from the general population only, while 4.1% ( now 6.2%, after including the new studies) of the whole groups had CKD defined as eGFR< 60 ml/min/1.73 m2, regardless the study sample criteria either the general population or the high risk groups
Another point that needs explanation is the lower prevalence of CKD in HIV-infected subjects (4.4%), which is at odds with multiple reports that HIV was one of the commonest causes of CKLD in Sub-Saharan Africa.
We agree. Only 18 studies out of the total 42 studies among HIV patients were medium quality and the remaining 24 were all low quality. To minimize the error in our estimate we calculated the prevalence only on the basis of the 18 medium quality studies. Furthermore, the CKD prevalence was calculated along with the KDOQI definition, which demands information on the eGFR and albuminuria/proteinuria. We believe that these two factors (calculation based on medium quality studies and application of KDOQI criteria) may justify the low estimate in our study. In other publications reporting a higher prevalence of CKD in HIV patients, the lack of definition of CKD along KDOQI criteria and/or the inclusion of low quality studies might explain the apparently higher prevalence in the same studies..

Comment: The discussion needs to highlight the methodological issues that might introduce uncertainties around the data, and the conclusions need to be softened to reflect this.

Response: we modified the discussion to highlight the methodological flaws and conclusion was changed accordingly

**Reviewer 3**
Reviewer Name: Joseph Lunyera, MBChB, MSc.
Institution and Country: Duke University School of Medicine, Durham, NC 27701, United States
Please state any competing interests: None declared
Please leave your comments for the authors below
Dr. Zoccali and colleagues systematically reviewed the literature to describe the burden of chronic kidney disease (CKD) in Africa. By building on and extending prior work done by Stanifer and colleagues, Dr. Zoccali's team present a much-needed update on the burden of CKD in Africa. As noted by the authors, this updated systematic review spans the entire continent – thus extending Stanifer et al's work which was limited to sub-Saharan Africa.

Comment: While I endorse this effort, I have major concerns regarding the date on which the search was last run: Dr. Zoccali's team last run their search on September 25th, 2014 - which is more than 2 years from the manuscript submission date. Thus, the authors need to update their search to include more recent publications that are relevant to this topic (as a rule, the search date should not be >12 months older than the date of manuscript submission). Other than the outdated search date, the manuscript is well written, and the authors have addressed the key limitations – particularly as it pertains to pooling the data using meta-analytic techniques.

Response: We agree with you and the search strategy was updated till April 7th, 2017.

Comment: If the authors agree to update the search, they should also consider the following:
• Use a database (e.g., Web of Science and/or Scopus) to search the reference list of included studies (Dr. Zoccali and colleagues appear to have hand-searched the reference lists of the included studies – which is no longer recommended as the standard of practice when conducting systematic reviews).

Response: thank you for your suggestion. We agree with the referee that using web of science or scopus could reduce the cost and time of the researchers. However the manual search of references remains the gold standard according to Chapman et al article" Semi-automating the manual literature search for systematic reviews increases efficiency" Health Info Libr J. 2010 Mar;27(1):22-7

Commment:
• Report the inter-rater agreement for study inclusion and quality assessment using cohen's kappa coefficient.

Response: We added it in the method section (P6 :, L: 12-13) and the findings in the result (P: 7, L:9-10)

**Reviewer 4**
Reviewer Name: Sirirat Anutrakulchai

Institution and Country: Division of Nephrology, Department of Internal Medicine, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand 40002
Please state any competing interests: None declared
Please leave your comments for the authors below

1. Some references are not completely described, for example, no mention of volume and page numbers in the reference # 33, 120 and 156.

Response: We revised and corrected it

2. Figure 3 is represented the etiologies of CKD but the sum of percentage is less than 100 %, what is the rest ?

Response: we agree with the referee. The Fig 3 represents the reported etiologies of CKD from the studies done among CKD and didn't cover all the causes of CKD in Africa. For that, we changed the figure to bar chart

3. Regarding to limitations, the enrolled papers were studied during the wide-ranged period form 1995 to 2014.
Therefore, the variation of CKD prevalence and etiologies might be partly affected by chronological change. Additionally, methods of GFR, serum creatinine and proteinuria assessments were periodically changed when the better suggested methods appeared which make difficulty in comparisons between these papers.

Response: we agree with the referee that there is a continuous change in the CKD definition. We tried to restrict the prevalence reported from our studies to the high/medium quality studies to limit the effect of bias due to the methodological flaws. In addition, there are only 8 studies out of 152 included in our review which have been done before 2002 ( KDOQI)

## VERSION 2 – REVIEW

| | |
|---|---|
| **REVIEWER** | Dorothea Nitsch<br>LSHTM, UK<br>I am currently involved in a study to derive a validated eGFR for Subsaharan Africa and Steering committee member on the DEGREE study (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5210224/) |
| **REVIEW RETURNED** | 04-May-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | As mentioned before, I don't understand how the authors can pool the data if they point out that these are too heterogenous for formal meta-analysis? It may be worth pointing out in the text how within a given study the prevalence changes with the change in estimation formula, with sometimes even doubling of the observed CKD prevalence. I would strongly suggest removing the map (Figure 2) for this reason.<br>I wonder whether the authors wish to present pooled results for causes of CKD stratifying by biopsy status. I suspect that the population put forward for biopsy is quite different to the CKD population overall. |

| | |
|---|---|
| **REVIEWER** | Vivek Jha<br>GIGH, India |
| **REVIEW RETURNED** | 28-Apr-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors have addressed most of the points. A minor point that still needs to be mentioned is the acknowledgement of uncertainty regarding CKD diagnosis with a single isolated value of eGFR without proteinuria. |

| | |
|---|---|
| **REVIEWER** | Joseph Lunyera, MBChB, MSc.<br>Duke University School of Medicine, Durham, NC, United States |
| **REVIEW RETURNED** | 16-May-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | I am generally satisfied with the authors' updated search strategy. However, I wonder why this updated search strategy does not yield references from East Africa - which raises concerns about regional publication bias. For instance, the pooled prevalence of CKD by region (Figure 2) does not reflect the burden of CKD in the East African region because only one study from Madagascar was considered (the study by Ramilitian and colleagues). I am aware of several publications describing the prevalence of CKD in Uganda and Tanzania, but none of which was retrieved in this updated search. Please comment on this concern.<br><br>Despite my concern, I remain persuaded that this systematic review presents a much-needed update on the burden of CKD in Africa. |

| | |
|---|---|
| **REVIEWER** | Sirirat Anutrakulchai<br>Division of Nephrology, Faculty of Medicine, Khon Kaen University, Thailand 40002 |
| **REVIEW RETURNED** | 13-May-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors have already excellently revised the manuscript. I have some questions about the causes of CKD represented by Fig 3, please clarify;<br>1. In Table S2, several papers did not show all etiologies of CKD. Therefore, I assume that unreported causes might be "other causes" (known causes but not reported) or "undetermined" (unknown causes). In my opinion, if the authors clearly define the percentage of "other causes" (separate another bar) and "unknown or undetermined cause" and sum these percentages with the reported causes for all 100 percent, readers will more understand the Fig 3.<br>2. In Table S2, some papers reported the percentage of undetermined causes which varied from 3.3-53.5 %. However, in Fig 3 showed the overall 2.5 % of undetermined cause which is quite low. What is a reason? If the reason is 2.5 % taken from number of reported undetermined persons divided by total CKD population reported in all 42 studies, it may be wrong because several papers did not report the number of undetermined persons. |

## VERSION 2 – AUTHOR RESPONSE

**Reviewer 1**

Reviewer Name: Dorothea Nitsch

Institution and Country: LSHTM, UK

Please state any competing interests: I am currently involved in a study to derive a validated eGFR for Subsaharan Africa and Steering committee member on the DEGREE study (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5210224/)

Please leave your comments for the authors below

Thank you for considering the comments

Comment: As mentioned before, I don't understand how the authors can pool the data if they point out that these are too heterogeneous for formal meta-analysis? It may be worth pointing out in the text how within a given study the prevalence changes with the change in estimation formula, with sometimes even doubling of the observed CKD prevalence. I would strongly suggest removing the map (Figure 2) for this reason.

Response: We referred to the bias introduced by the different equations used for estimating GFR in the discussion (P: 12, L: 5-10). With this proviso, we provided tentative estimates in the various areas in Africa. In the legend to Figure 2 we reiterated that the estimates we made are bound to be imprecise and inaccurate. We hope that the referee may agree that, however imprecise and inaccurate, these estimates have some descriptive value.

Comment: I wonder whether the authors wish to present pooled results for causes of CKD stratifying by biopsy status. I suspect that the population put forward for biopsy is quite different to the CKD population overall.

R: This is a useful suggestion. However, the needed information could not be extracted as some papers reported using the biopsy in a sub-sample without differentiating the causes of CKD based on biopsy status. We hope that the referee may understand our plight.

**Reviewer 2**
Reviewer Name: Vivek Jha
Institution and Country: GIGH, India
Please state any competing interests: None declared

Please leave your comments for the authors below

Comment:
The authors have addressed most of the points. A minor point that still needs to be mentioned is the acknowledgement of uncertainty regarding CKD diagnosis with a single isolated value of eGFR without proteinuria.

Response: We referred to this in the discussion (P:12, L: 14-17) and (P:15, L:15-17)

**Reviewer: 3**
Reviewer Name: Joseph Lunyera, MBChB, MSc.
Institution and Country: Duke University School of Medicine, Durham, NC, United States
Please state any competing interests: None declared

Please leave your comments for the authors below

Comment: I am generally satisfied with the authors' updated search strategy. However, I wonder why this updated search strategy does not yield references from East Africa - which raises concerns about regional publication bias. For instance, the pooled prevalence of CKD by region (Figure 2) does not reflect the burden of CKD in the East African region because only one study from Madagascar was considered (the study by Ramilitian and colleagues). I am aware of several publications describing the prevalence of CKD in Uganda and Tanzania, but none of which was retrieved in this updated search. Please comment on this concern.

Response: Actually, we included 32 studies from the Eastern macro-area: i.e. 19 studies based on the old search before updating the search strategy to include all studies up to April 2017. So, we added 13 new studies after updating the search strategy.

Despite my concern, I remain persuaded that this systematic review presents a much-needed update on the burden of CKD in Africa.

**Reviewer 4**

Reviewer Name: Sirirat Anutrakulchai
Institution and Country: Division of Nephrology, Faculty of Medicine, Khon Kaen University, Thailand 40002
Please state any competing interests: None declared

Please leave your comments for the authors below

The authors have already excellently revised the manuscript. I have some questions about the causes of CKD represented by Fig 3, please clarify;

1. In Table S2, several papers did not show all etiologies of CKD. Therefore, I assume that unreported causes might be "other causes" (known causes but not reported) or "undetermined" (unknown causes). In my opinion, if the authors clearly define the percentage of "other causes" (separate another bar) and "unknown or undetermined cause" and sum these percentages with the reported causes for all 100 percent, readers will more understand the Fig 3.

Response: we agree with the reviewer about his comment but the problem is some papers were focusing mainly on one cause, for example, diabetes mellitus or polycystic kidney disease, and in these papers they reported the percentage of the cause of interest only. In other papers undetermined or other causes were combined. For this reason, we adopted the bar chart to graphically describe etiologies (because the total percentage of the whole causes is not 100)

2. In Table S2, some papers reported the percentage of undetermined causes which varied from 3.3-53.5 %. However, in Fig 3 showed the overall 2.5 % of undetermined cause which is quite low. What is a reason? If the reason is 2.5 % taken from number of reported undetermined persons divided by total CKD population reported in all 42 studies, it may be wrong because several papers did not report the number of undetermined persons.

Response: We recalculated it based on your suggestion and verified the percentages and now it is modified to be 20%

**VERSION 3 – REVIEW**

| | |
|---|---|
| **REVIEWER** | Dorothea Nitsch<br>LSHTM, UK<br>see before. |
| **REVIEW RETURNED** | 12-Jun-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors state "No meta -analysis was conducted in this review due to the huge discrepancy in the definition used to identify CKD, the methods of creatinine measurement, urine protein assessment, and in the quality of the reporting.", yet they say in the abstract: " In the community-level studies, based on available medium and high quality studies, the pooled prevalence of CKD in Africa was 10.1% (95% CI: 9.8%-10.5%). West/Central-West had the highest prevalence (16.5%), followed by Central (16%), Southern (12.2%), Eastern (11.0%), and North (4%) Africa. The prevalence in sub-Saharan Africa was 14.02%.<br>The pooled prevalence of<br>16 CKD in the high risk groups was 5.6% (95% CI: 5.4-5.8%) in HIV (based on available medium and high quality studies), 24.7% (95% CI: 23.6-25.7%) in diabetes (based on all available studies which are of low quality except four of medium quality) and 34.5% (95 % CI: |

| | |
|---|---|
| | 34.04%-36%) in hypertensive patients (based on all available studies which are of low quality except two of medium 20 quality)"<br><br>I still worry that pooling prevalence data when the quality of data is too low to carry out a meta-analysis is not the right approach. Note that there is some marked variation in the ages of people included in the survey data, which will confound the results and make any summary prevalence meaningless. To which extent are regional differences driven by differences in the ages of studied populations?The authors could have reported a range of reported prevalence data for each region, ideally taking into account age differences. |

| | |
|---|---|
| **REVIEWER** | Vivek Jha<br>George Institute for Global Health, India |
| **REVIEW RETURNED** | 24-Jun-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors have addressed my comment. |

| | |
|---|---|
| **REVIEWER** | Joseph Lunyera, MBChB, MSc.<br>Division of General Internal Medicine<br>Duke University School of Medicine<br>Durham, NC, United States |
| **REVIEW RETURNED** | 16-Jun-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | Thanks for clarifying. I have no further concerns. |

| | |
|---|---|
| **REVIEWER** | Sirirat Anutrakulchai<br>Faculty of Medicine, Khon Kaen University<br>Thailand |
| **REVIEW RETURNED** | 08-Jun-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors already presented clearly. |

## VERSION 3 – AUTHOR RESPONSE

**Reviewer 1**

Comment:The authors state "No meta -analysis was conducted in this review due to the huge discrepancy in the definition used to identify CKD, the methods of creatinine measurement, urine protein assessment, and in the quality of the reporting.", yet they say in the abstract: " In the community-level studies, based on available medium and high quality studies, the pooled prevalence of CKD in Africa was 10.1% (95% CI: 9.8%-10.5%). West/Central-West had the highest prevalence (16.5%), followed by Central (16%), Southern (12.2%), Eastern (11.0%), and North (4%) Africa. The prevalence in sub-Saharan Africa was 14.02%. The pooled prevalence of 16 CKD in the high risk groups was 5.6% (95% CI: 5.4-5.8%) in HIV (based on available medium and high quality studies), 24.7% (95% CI: 23.6-25.7%) in diabetes (based on all available studies which are of low quality except four of medium quality) and 34.5% (95 % CI: 34.04%-36%) in hypertensive patients (based on all available studies which are of low quality except two of medium 20 quality)"

I still worry that pooling prevalence data when the quality of data is too low to carry out a meta-analysis is not the right approach. Note that there is some marked variation in the ages of people included in the survey data, which will confound the results and make any summary prevalence meaningless. To which extent are regional differences driven by differences in the ages of studied populations? The authors could have reported a range of reported prevalence data for each region, ideally taking into account age differences.

Response: Thank you for your suggestion. We agree with you about the quality of the data. For that, we restricted the pooled analysis to the high-medium quality studies, when available. As regards the age consideration, this required us to have information about the age strata from each study and this was not available (Tripepi G. et al. Nephron Clin Pract 2010;116:c322–c325). The only available data were the mean or median of the age and this can’t be used to stratify studies based on age categories. These pooled estimate provide a tentative idea about the burden of CKD according to the study population group and the quality of the studies and we mentioned that it should be taken with caution (comment on Figure 2)

**VERSION 4 – REVIEW**

| | |
|---|---|
| **REVIEWER** | Dorothea Nitsch<br>LSHTM UK<br>as before in past reviews |
| **REVIEW RETURNED** | 14-Jul-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | Unfortunately, I am still not happy with the suggested changes for two reasons - firstly CKD prevalence is hugely dependent on the age of the included studies (there is also an issue with gender)<br>-secondly, the validity of the pooled CKD prevalence may vary depending on the definition used, leading to over-or underestimates of true prevalence. The formula used to define CKD prevalence may have differing validity in different regions of Africa, also depending on gender dependent measurement error of CKD status.<br><br>Hence, pooling these studies without taking account of gender and age distributions is to me not meaningful.<br><br>I appreciate that the authors have tried to restrict to studies with similar outcome definitions. However, the authors pool data across a mix of CKD-types, those defined by reduced eGFR, those defined by proteinuria and/or CKD defined by both. In the UK the prevalence of CKD stage 3-5 (eGFR defined) is approximately half of the underlying CKD prevalence when proteinuria is also taken into account. For example, if the authors pool a study that used proteinuria only with a study that used both eGFR and proteinuria then the summary pooled CKD prevalence will greatly depend on the relative proportion of the size of the two studies (i.e. whether the bigger study measured both urine and blood, or not).<br><br>Therefore I think that the regional pooled prevalence figures are misleading as they depend on the age, gender composition, and type of outcome definition used for CKD within each study as well as the respective study size.<br><br>Please restrict to medium/high quality cross-sectional studies without age-restriction (representative of the respective community, not just dependent on occupation, we're looking for a random sample/representative survey - a survey that has many more women than men may have a significant response bias and may be lower quality) that have measured the same entity (e.g. reduced eGFR, or proteinuria, or both) - so 3 figures in total rather than just 1. Then the figure could say (crude, not age-/gender adjusted) pooled prevalence estimate of CKD (specify definition used).<br><br>Bigger emphasis should be given to the gender distributions in the included studies in the quality assessment and the discussion. |

**VERSION 4 – AUTHOR RESPONSE**

**Reviewer 1**
Reviewer Name: Dorothea Nitsch. Institution and Country: LSHTM UK
Please state any competing interests: as before in past reviews
Please leave your comments for the authors below. Thank you for attempting to address my concerns.

Comment: Unfortunately, I am still not happy with the suggested changes for two reasons - firstly CKD prevalence is hugely dependent on the age of the included studies (there is also an issue with gender)

Response: Unfortunately, the majority of the included studies did not report the prevalence of CKD stratified by either age or gender. In order to assess the potential effect of age and gender among studies we performed a series of meta-regressions plotting CKD prevalence as a function of mean age and proportion of males in each study. These analyses showed no statistically significant effect of age and gender on CKD prevalence across studies (see appendix)". However, we agree with the referee that the lack of information on age and gender prevalence is a major limitation that cannot be solved with meta-regression analysis because the limited number of data points might well prevent the appreciation of any underlying effect of age and gender on CKD prevalence. In the main text we have further expanded on this limitation. We now report the full range of the prevalence in various study groups and (in parenthesis) the tentative estimate of the pooled prevalence. In the results (pg 10, L 9-10) and in the discussion (pg 14, L 1-13) we reiterate that the estimate of the pooled prevalence has a major limitation in the fact that most available studies do not provide information on age and gender stratified CKD prevalence . We hope that the referee may agree on this solution. If the referee believes that this solution remains unacceptable, the sole possibility we have is to drop the estimate of the pooled prevalence altogether and just report the range of the prevalence in various countries (individual countries prevalence was reported in Tables 2-6)

Comment: secondly, the validity of the pooled CKD prevalence may vary depending on the definition used, leading to over-or underestimates of true prevalence. The formula used to define CKD prevalence may have differing validity in different regions of Africa, also depending on gender dependent measurement error of CKD status.

Response: thank you for your suggestion. Now we estimated the prevalence in each group (general population, HIV, diabetic, hypertensive) categorized by the definition used among the high/medium quality studies (P: 10-12)

Comment:
Hence, pooling these studies without taking account of gender and age distributions is to me not meaningful.

Response:
I appreciate that the authors have tried to restrict to studies with similar outcome definitions. However, the authors pool data across a mix of CKD-types, those defined by reduced eGFR, those defined by proteinuria and/or CKD defined by both. In the UK the prevalence of CKD stage 3-5 (eGFR defined) is approximately half of the underlying CKD prevalence when proteinuria is also taken into account. For example, if the authors pool a study that used proteinuria only with a study that used both eGFR and proteinuria then the summary pooled CKD prevalence will greatly depend on the relative proportion of the size of the two studies (i.e. whether the bigger study measured both urine and blood, or not).

Comment:
Therefore I think that the regional pooled prevalence figures are misleading as they depend on the age, gender composition, and type of outcome definition used for CKD within each study as well as the respective study size.
Please restrict to medium/high quality cross-sectional studies without age-restriction (representative of the respective community, not just dependent on occupation, we're looking for a random sample/representative survey - a survey that has many more women than men may have a significant response bias and may be lower quality) that have measured the same entity (e.g. reduced eGFR, or proteinuria, or both) - so 3 figures in total rather than just 1. Then the figure could say (crude, not age-/gender adjusted) pooled prevalence estimate of CKD (specify definition used).

Response: We have now calculated the CKD prevalence in each study group categorized by the definition. Moreover, we produced the 4 maps suggested by dr. Nitsch for CKD prevalence among general population. In some areas with some CKD definitions there is only one study or no study at all and perhaps this is the reason why estimates by CKD definition produce results of difficult interpretation which may appear incoherent. We send these maps for dr. Nitsch scrutiny but we would be inclined at not including the same maps in the paper. The prevalence of CKD in individual countries is reported in the Tables. The CKD prevalence ranges (and the tentative pooled estimates) in African macro-areas is reported in the main text. We feel that these maps do not add to the clarity of the manuscript.

Comment:
Bigger emphasis should be given to the gender distributions in the included studies in the quality assessment and the discussion.

Response:We have now further expanded on the limitation of scarce information on age and gender stratified information on CKD prevalence. We dedicate (pg 14.L 1-13) a new paragraph discussing these limitations. At page 9 L 15-19 we comment that we applied validated tools to assess the quality of each study based on its type taking in consideration the sampling techniques and whether the sample was representative of the target population or not. Alongside with study quality, the prevalence of males and females is reported in Tables 2-6 whenever information on gender is given

**VERSION 5 – REVIEW**

| | |
|---|---|
| **REVIEWER** | Dorothea Nitsch<br>LSHTM, UK<br>as before |
| **REVIEW RETURNED** | 31-Aug-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | I would like to congratulate the authors who have stuck admirably with the revision process. By analysing data in defined subgroups there is more clarity about the definitions used and the quality of the data. Thank you for adding supplemental figure 1 which is reassuring. I have no further comments. |