Reviewer's report

Title: Estimation of benchmark dose of urinary cadmium as the reference level for renal dysfunction: A large sample study in five cadmium polluted areas in China

Version: 1
Date: 2 April 2015

Reviewer: Brent D. Kerger

Reviewer's report:

Major Compulsory Revisions:

1. The authors have provided a reasonably detailed description of the methods and results, but some further clarifications are needed. First, the authors must clarify the context for study subject selection or self-selection in regards to the original 12000 subjects that are noted to have been "randomly selected." The potential for inadvertent selection bias in relation to participation rate (how many were queried in order to get the 12000 actually studied?) and other factors as to how the study population was recruited or otherwise identified. Also, the rationale and references should be provided for selection of the age subgroup (35-89) identified for the current analysis, and potential influences of including very elderly individuals (e.g., > 70 yrs) should be considered in the data analysis and implications discussed.

2. Second, although a Chinese government source is described for the 1000 ug/g Cr B2MG criterion used for the analysis, a more detailed basic rationale is needed to explain to the readers what the strengths and weaknesses of this criterion may be in relation to published literature on its implications for clinically important kidney disease.

3. Third, the rationale for dividing the study subjected into the selected subgroups should be explained, and authors should comment on the potential influence of these seemingly arbitrary groupings as opposed to examining the trends in alternative but statistically appropriate ways.

4. Fourth, the authors need to more thoroughly describe and discuss their use of the BMD modeling and their criteria applied to determine what was a "good fit" given the model and parameter choices made. Suggest the authors stick with a single BMR definition rather than using both 5 and 10%, and support their decision based on other published literature or regulatory approaches.

5. The authors should provide descriptive data in Table 1 that identifies the frequency of persons in each subgroup that showed B2MG > 1000 and also frequency of persons with UCd above some credible limit that one might consider clearly elevated in this study population based on statistical analysis and/or based on other published studies. The frequency of individuals meeting both criteria (elevated B2MG and UCd) would also be important.

6. One important element that is missing from the statistical analysis by geographic region is a quantitative correlation of the frequency of 'affected
individuals’ and the lack of correlation between GMs for UCd to B2MG in accordance with those regions the authors believe to be the most Cd contaminated and hence having the highest risk of kidney disease. BMDs in high risk regions should be compared to the presumed low risk regions and implications of the statistical analysis should be discussed in light of BMD modeling of the entire study population.

7. The authors should discuss the possible reasons why the lowest exposed region (Hubei) consistently showed the lowest BMD for both males and females, and explore the relationship between regional indicators of Cd exposure and both BMD estimates as they do or do not correlate to the frequency of elevated B2MG, UCd , or both. The degree of variability identified in the BMD and BMDL estimates should be discussed by region/gender and in the combined analysis. As noted above, the authors should identify a single BMD criterion and not present both 5 and 10% values in the final analysis without also presenting the parallel information in underlying data analyses.

Minor Essential Revisions.

1. There are a number of English grammar problems (where I have indicated the suggested revision in quotes below) and other clarifications that should be fixed as follows.

   Line 25: "However, there is still a lack of information. . ."
   Line 27: "in a Chinese..."
   Line 43: "population and geographic location affect..."
   Line 52: "and sufficiently high long-term exposure to Cd"
   Line 53: "skin contact and inhalation can possibly cause health effects"
   Line 55: "(mainly cancer from inhalation exposures)"
   Line 64: "represents an estimate of the dose"
   Line 80: New paragraph at "Itai-Itai disease"
   Line 84: "effects in the general"
   Line 137: Spell out radio-immunoassay (RIA)
   Line 149: "Environmental Protection Agency"
   Line 184-189: describe ranges with dash rather than ~ between them and use 'frequency' to describe these percentages rather than the term 'ratio'. It is more important to highlight the strong and statistically significant dose-response relationships rather than the frequency range data.
   Line 190: "dysfunction, indicated by"
   Line 191: omit "obviously"
   Line 192: "A dose-response relationship is apparent between..."
   Line 197: "The results for males are"
   Line 198: more context needed for explaining relevance of p > 0.05 in the BMD model, since this is opposite of standard statistical testing.
Line 202: replace "obviously" with "appreciably"
Line 203: Should begin paragraph with "Results for females are presented in Table 4..."
Line 217: "BMD value is significantly affected by geographic region and gender."
Line 223: cite WHO reference for 5 ug/g reference level.
Line 230: "ideal biomarker for"
Line 231: "selection of geographic region significantly affected"
Line 235: "samples"
Line 240: "BMD analysis methods."
Line 242: "Guizhou compared to other provinces"
Line 249: "in the other two"
Line 250: omit "obviously"
Line 252: missing words - apparently citing male and female % but only list females
Line 255: Authors seem to interchangeably use "sample population" for geographic region population- be specific and be consistent as "sample population" is not really descriptive.
Line 258: "subjects may be more"
Line 265-6: Sentence requires clarification: This was done to minimize the effect of study population in single area on BMD evaluation."
Line 268: "populations."
Line 269: "The sample size was also larger than that of Japanese populations"
Line 273: omit "also"
Line 277-79: use term "frequency" instead of "ratio"
Line 288: omit "than"
Line 289: clarify what is meant by "greater quality of subjects"
Line 297: correct JEFCA to JECFA
Line 314: change "sample population" to "geographic location"
Table 2: cite actual p values at bottom rows rather than 0.000 for most.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**
I declare that I have no competing interests.