Reviewer’s report

Title: Metabolic Profiling Detects Systemic Effects of Environmental and Lifestyle Exposure to Cadmium in a Human Population.

Version: 1 Date: 11 March 2012

Reviewer: Carolyn Slupsky

Reviewer’s report:

The article “Metabolic profiling detects systemic effects of environmental and lifestyle exposure to cadmium in a human population” by Ellis et al. investigated changes in metabolism of 178 individuals living near a disused zinc smelter where elevated levels of environmental Cd were present. Using urinary NMR metabonomics combined with measurement of 8-oxo-2'-deoxyguanosine (a marker of oxidative stress) by MS, they found differences in six metabolites (citrate, 3-hydroxyisovalerate, 4-deoxyerythronic acid, DMG, creatinine, creatine). Additionally, they determined that 8-oxo-2'-deoxyguanosine was elevated in individuals with high Cd exposure, and citrate could also be correlated with smoking status. Even with removal of smokers from their dataset, Cd exposure could be detected.

This article is a potentially important article in that it shows that environmental exposure to toxicants may be detected much earlier than previously reported, and perhaps more importantly, may be causing subclinical detrimental health effects.

While I like this article, I am very disappointed that the authors did not take the care to ensure that the conversion to pdf was acceptable. In my version, throughout the manuscript, there were squares in place of symbols that have made review of this article very difficult and time-consuming. I also really had a difficult time with Figure 2. I am not sure if it is the missing symbols, or that there are words missing from the text or what, but in my version I could not understand the figure based on the figure legend.

I think that this article should eventually be accepted, but I hope that the authors will be able to address my comments. Overall my comments are generally minor, although they need to be addressed prior to publication. Certainly the authors will have all of the data and should have no problem addressing them. My hope is that by following these suggestions, the ability of the reader to understand the article will greatly be increased which may hopefully increase the article’s impact.

Other comments:

1. Figure 3A. I am confused as to why the authors chose to use ANOVA for analysis of citrate between the three groups (never smoked, previously smoked, current smoker). Why did they not use the non-parametric Kruskal-Wallis test? They do not mention anywhere in the text (that I could find) that they determined
citrate was normally distributed, or that they transformed it to make it normally distributed. ANOVA is a parametric test, and requires data that is normal. The authors need to clarify whether citrate was normally distributed for each group.

2. Figure 3B. The authors use Welch’s t-test to determine 8-oxo-2'-deoxyguanosine concentration. However, they state in the methods that they logged the concentration of this compound. Did they take the logged concentration and divide by creatine? Also, why divide by creatine, and not creatinine? I think that this needs to be clarified. Also, it needs to be clarified whether the non-parametric test needs to be used (Welch’s t-test) or a parametric test (Student’s t-test). Perhaps more attention to a description of the figure and a brief description of the statistics in the figure legends would help to clarify this.

3. Figure 3B. The authors state n=40. Is that per group, or total? How many per group?

4. Table 1. Authors report R2 and Q2 values. There are several Q2 values that are higher than R2. Which values of R2 are the authors reporting? It does not make sense to me that Q2 should ever be higher than R2. Also, the authors need to clarify how they determined the p-value in this table. I would personally like to see the validation plots as a figure rather than expressed in this table.

5. Also, I would like to see the actual PLS plots showing the separation between the groups, as well as the validation (as mentioned above) so that I can assess model validity more easily.

6. Authors state that they removed 51 samples from the 178 originally profiled due to ethanol and analgesic consumption, yet in Table 2, they seems to be using these samples? The authors need to clarify why they can now use these samples in analysis, or why for one analysis they excluded them, and then included them again (in the first paragraph of the results section, I do not believe that it is sufficient for the authors to state that it was excluded from the initial pattern recognition analyses due to confounding effects of the diuretic or pharmacological effect of the compounds, but then later they include it…. Either they should be removed from analysis entirely, or should remain in the dataset, just removing the peaks corresponding to those compounds). My major concern here is whether including the additional 51 individuals may change some of the other results. In other words, how were these 51 distributed (how many were smokers, non-smokers, high Cd, low Cd, etc…). If they were evenly distributed across the groups, then I may not have a problem if you include their data, but that needs to be better addressed in the text.

7. Upon reading the paper, I have to wonder a bit about those people exposed to second hand smoke. Would that play a part in urinary citrate concentrations?

8. The authors state in several places that their results correlate to systemic changes of Cd exposure. My question is whether it really is systemic? The authors mention on page 5 that Cd exposure causes nephropathy, and that the
kidney accumulates Cd. Could the observed changes simply reflect the changes happening in the kidney? If it is systemic, I would like to see more explanation.

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests