Article details: 2016-0142	
Title	Comprehensive survey of household radon gas levels and risk factors in Southern Alberta
Authors	Fintan K.T. Stanley BSc, Siavash Zarezadeh, Colin D. Dumais BSc, Karin Dumais BSc, Renata MacQueen, Fiona Clement PhD, Aaron A. Goodarzi PhD
Reviewer 1	Dr. Sarah B Henderson BASc PhD
Institution	Environmental Health Services, BC Centre for Disease Control, Vancouver, BC
General comments (author response in bold)	This manuscript describes a radon sampling campaign in Alberta, and related data analyses on the large number of samples collected. It contains a lot of useful information, but the objectives of the manuscript are confusing, the content is disorganized, and it simply is not ready for publication at this time. Major comments: 1) There are no clear Introduction, Methods, Results, and Interpretation sections. While these headings are present in the document, there is so much cross-contamination between the sections that the manuscript is almost impossible to read. For example, there is a results section discussion the impacts of radon remediation in 90 homes, but these 90 homes were not mentioned anywhere in the methods. The results section also gives detailed information on possible reasons that radon concentrations are higher in newer homes, but this information belongs in the interpretation. The methods section on statistical analyses is seven lines long, but much of the necessary information has ended up in the results. I encourage the authors to go carefully through the manuscript and to separate their work into cleaner sections more clearly describing: (1) why they undertook the study; (2) what they did, both with the sampling campaign and they analyses, including high radon homes that were remediated and resampled; (3) what they found; and (4) what their findings may mean for Alberta and in general. **We have significantly revised the organization of our entire manuscript to separate out each section as per the reviewer's request, emphasising the main findings more clearly.** 2) There is a second in the results on the relative risks in Alberta compared with Canadian averages from the cross-country survey conducted by Health Canada. In these comparisons the authors are comparing apples to oranges. The Health Canada survey was a statistically robust random sample, while the survey the authors conducted was a convenience sample. It is fair to discuss the the two in context of each other, but it is not fair to conclude from convenience methods that concentrations are higher Alberta than areas characterized by a random sample. Convenience samples will typically turn up higher concentrations than random samples, because they attract participants who are a bit more worried about radon to begin with. There will be more information on this in papers by Branion-Calles, who used data from multiple surveys to model radon in BC. **We have moved this discussion & substantially altered to meet the request of the editors and reviewers, fully acknowledging that our survey utilized a convenience sample.** 3) A complex modelling exercise with an eventual R2 value of ~1% isn't really worth reporting. Regardless, the descriptive information about building factors associated with the radon concentrations is useful, but ANOVA results and t-tests are more informative than model coefficients (and generally more robust to lack of normality in the data). It is well-known that radon is very hard to model -- the descriptive information is far more useful, particularly for the home age characteristics. This and the remediation results are the most important key messages from the study, and I encourage the authors to highlight both by removing the uninformative modelling results. **We have refocussed our text on home age and remediation data but, since we had been asked for this by the editorial board earlier in the review process, we have left the GLM modelling in the results, we have removed the predicative regression model as we agree that the result was not sufficiently interesting given radon is indeed very hard to model.** 4) The authors state that they log-transformed the data to base 10 for the modelling -- why not base e? Were the normality assumptions of both transformations tested? Base e is the better option for radon in my experience. On this note, all of the radon concentration axes should be log-transformed rather than chopped into pieces as has been done in figures. **Thanks for this suggestion, we have now performed the transformation with base e, as suggested. We had in fact tried and tested both before submitting, and found whether it is base 10 or base e, both transformations meet the assumptions of the model, and the end results are near indistinguishable. With regards to the second point, we feel that displaying the data in figure 1 on a log-scale reduces the visual interpretability of the raw data, and it is preferable to split the axes to show the spread of the data. We feel strongly that for visually conveying the data to as wide an audience as possible that this is the best option. All other figures (where relevant) have been altered to log scales, fitting with the reviewers request.**

	Other comments: 1) Some of the language throughout the manuscript is...er...unscientific. Expressions like "somewhat" and "statistically significant manner" and "rise substantially" are not consistent with accurate scientific writing. Suggest authors read revised draft carefully for such language. **We have revised our manuscript to remove such language wherever possible.** 2) There are many cases stating "Figure X shows that...". The better way is always to describe what Figure X shows, and then refer to it in brackets (Figure X). I tell my students that the text of a paper should stand alone, such that the entire thing makes sense without the Tables and Figures. Likewise, each Table and Figure should stand alone such that each makes sense without the text of the paper. This manuscript needs quite a bit of work to achieve that ideal. **We have revised our manuscript to achieve this ideal.** 3) FSA = Forward Sortation Area (not Forward Sorting Code) is the standard Canadian acronym. **Thanks and done.** 4) Figure 1 is not very useful, even if you are quite familiar with the study area. Suggest removing. Figure 2 is great! **We feel it is useful to show the 'raw data', especially to the stakeholders and participants of our study. As this is the only figure that we have that indicates all data points collected, it is our preference to leave this in the manuscript, assuming the objections are not overwhelmingly negative to this point.** 5) Figure 3 should be on log scale, and box plots could be notched to show significant differences, and also width could be used to indicate number of observations in each box. **We have revised the plots and changed the axis to log scale.** 6) Not clear how (b) and (c) are different in Figure 4. Choose one or the other? Or make a box plot of (c)? Again, (a) would be more informative if the log-transformed values were plotted and the two distributions were shown overtop of each other. **We agree. We have revised and merged Figure 4B+C together. We have also compared the home age distribution in our study to the distribution of home ages across the area.** This will be a nice contribution to the Canadian radon literature when it gets tidied up -- the authors have some really fantastic data to work with! Thanks for letting me review. **Thanks!**