Author's response to reviews

Title: On the validity of area-based income measures to proxy household income

Authors:

Gillian E Hanley (ghanley@chspr.ubc.ca)  
Steve Morgan (morgan@chspr.ubc.ca)

Version: 2 Date: 5 March 2008

Author's response to reviews: see over
February 27, 2008

Dear Editors,

Thank you for having our manuscript, “On the validity of area-based income measures to proxy household income” reviewed for publication in BMC HSR. We are grateful for the reviewer’s constructive comments and suggestions, several of which we were able to address through revisions to the paper. We feel the revised manuscript is significantly improved as a result of this process. Below we outline how we responded and altered the manuscript according to each reviewer comment/suggestion.

We hope that the editors will understand that our analysis was constrained, in part, by privacy concerns of the data stewards for the province of British Columbia and by the privacy rules of Statistics Canada. As our initial analysis plan was to validate area-based income as a categorical variable (our initial thought process was that categorical income variables are most commonly available to health researchers), we did not ask for permissions to run the regressions using the continuous variables. We agree with the reviewers comments that an analysis using area-based and individual-level income as a continuous variable would be very useful and offer additional insights to what we were able to offer here. We hope to obtain permission to use the variables in this way; unfortunately this may take months or even years. Thus, although we recognize the importance of the reviewer’s suggestions to do so, we were not able to run the continuous-variable analyses in this project.

Sincerely,

Gillian Hanley and Steve Morgan

Response to Reviewers:

Reviewer 1-Huge Gravelle

1. It would be interesting to compare it with the results from a regression of mean area drug costs on mean area income.
   We agree that this area-to-area approach is a potentially interesting way of validating measures. Unfortunately, we are unable to do this. Permission to use the research databases from which we draw is given on a project-by-project basis. Owing to privacy concerns of data stewards for the province of British Columbia, our permissions give us access to the estimates of neighborhood incomes but require that we aggregate neighborhoods into anonymized income strata and prohibit us from attempting to identify individual neighborhoods using this information. As such, we cannot run the analysis on an area-to-area basis.

2. It would be informative to have a diagram showing the results from two simple non-parametric regressions of (i) individual drug costs on actual income and (ii) individual drug cost on proxy area income.
Owing to privacy concerns, our data access permissions did not permit us to treat income as a continuous variable for this analysis. We agree that the analysis is potentially instructive and will seek permission to do this in future analyses.

3. The regression reported in Table 5 also included variables controlling for the presence of seniors and household size. The coefficients on these variables should also be included in the results table to show the effect of the choice of income measure. It is of interest to know if the use of proxy income measures also affects the estimated effects of household size or age.
   This is an important suggestion and we have added these coefficients to Table 5. It is interesting to note that the coefficients for presence of seniors in the household and household size do not differ. Both are of the expected sign and are of the same magnitude whether household income or neighborhood income deciles are used.

4. Say whether reported R2 are adjusted,
   The R2 reported in Table 5 is the adjusted R2. Thank-you for catching this oversight and we have made the clarification.

Reviewer 2- Donald Klepser

1. Third sentence of the third paragraph in the Results section
   The authors thank the reviewer for catching this inaccuracy and have corrected it. The text is now consistent with the tables.

2. Authors refer to both a BC PharmaCare and a BC PharmaNet. It is not clear if these are the same program
   We have clarified the difference between the BC PharmaCare and the BC PharmaNet data system. We believe the distinction between the two is now clear.

3. The authors present all income data as deciles. It is likely of interest to readers to know the actual ranges of those deciles.
   Income ranges have been added to Table 4. For parsimony, ranges are not added to all other tables.

Reviewer 3- Bruce Stuart

A more useful sensitivity test using the prescription drug cost data would be to estimate side-by-side simple regressions with the alternate income measures entered as continuous variables. This would tell you whether area-measured income tend to underestimate or overestimate the true effect of income on drug spending.

Once again, we agree that such an analysis may add more information. With the publication of the present analysis, we are hopeful that more attention will be drawn to the potential importance of this. As mentioned in responses above, the present study was constrained by strict privacy concerns that prohibited us from potentially identifying the neighborhoods in which individuals live. We were therefore required to analyze data as a categorical rather than continuous variable.
Other tests of this sort would also be useful if the authors have other variables. For example, if they had basic demographics and household size, they could test whether the bias associated with area-income measures is attenuated in multivariate models.

We had originally included covariates for the presence of a senior in the household as well as household size in our regressions. However we had not included coefficients for these variables in Table 5 and thus it was initially confusing to the reader. We have now included coefficients for these variables in Table 5 and we have also run the regression without these covariates. Initially, we had not included results for a regression run without covariates, and thus we could not detect whether the bias was attenuated in multivariate models. We agree with the reviewer that this important and thus we have now included regression results in Table 5 for regressions of income deciles on total drug costs run with and without covariates. We find that the bias in the variables is relatively robust to the presence of covariates.