Author's response to reviews

Title: Rating Neighborhoods for Older Adult Health: Results from the African American Health Study

Authors:

Elena M Andresen (andresen@phhp.ufl.edu)
Theodore K Malmstrom (malmsttk@slu.edu)
Fredric D Wolinsky (fredric-wolinsky@yiowa.edu)
Mario Schootman (mschootman@im.wustl.edu)
J Philip Miller (jphiliomiller@wustl.edu)
Douglas K Miller (dokmille@iupui.edu)

Version: 2 Date: 16 June 2007

Author's response to reviews: see over
Editors
BMC: Public Health

June 16th, 2007

We are pleased to submit our revised manuscript “Rating Neighborhoods for Older Adult Health: Results from the African American Health Study.”

The paper has been substantially revised. We conducted analyses to answer several questions reviewers had posed, and reorganized and added material as they recommended. Because of differing reviewer opinions about material to add, and our intent to keep the manuscript at a reasonable length, we made one major change and dropped the “outcomes” analyses (Table 4). Suggestions and differences among the reviewers lead us to think that a thorough treatment of outcomes is a very different paper. Such a paper must include a much broader set of analyses and methods than is possible here, and one that would be better answered by the results of follow-up outcomes (intended during our next five years of this cohort study). In addition, the outcomes section focused attention on the cohort subjects, not the neighborhood ratings, and the latter are the subject of this manuscript. Therefore, this manuscript now limits the scope to the methods of developing and testing the rating scale. Specific responses for reviewers are listed below.

Sincerely,

Elena M. Andresen, PhD.
Professor & Chair, Department of Epidemiology & Biostatistics
College of Public Health and Health Professions
University of Florida Health Sciences Center
PO Box 100231
Gainesville, FL 32610
352-273-5359 office; 352-273-5365 fax; andresen@phhp.ufl.edu

Rating Neighborhoods for Older Adult Health: Results from the African American Health Study
Response to reviewers

Reviewer 1

Major Compulsory Revisions

1. *The introduction is a bit disjointed, referring obliquely to social theories of place and the general interest in whether (and how) neighborhoods affect health.* We revised the introduction to focus on measurement issues in observer ratings, rather than trying to provide a thorough rationale for neighborhood effects – which is much more comprehensively done by those we cite.
2. **Narrow the health outcome to one.** We dropped these analyses for the current manuscript, but note that multiple reviewers raised questions and made suggestions that we plan to pursue in future investigations once additional data are available.

3. **I suggest explaining to the reader what is learned at each phase. This can be carried through to the Discussion.** For example, little is made of the comparison of interviewer-rated and respondent-rated neighborhood scores. We now add focus on the interviewer-resident concordance (including a clarification of Table 2), and extended the explanation and discussion of the phases.

4. **Did interviewer effects vary by these two areas?** We think this is an important question, and one we had not considered. We first constructed a model with all 17 interaction terms in the model, and interviewers 10 and 13 emerged with interactive effects. We then constructed interaction models where we tested each of the interviewer interactions separately, and then constructed a model with only the two significant interviewer interactions in the model in table 3 the same two interviewers as in the omnibus test emerged with significant interaction effects.
   
   1) The beta for interviewer 10 was 2.289; the beta for interaction [interviewer by city] was 1.359. Interpretation: Interviewer 10 provided worse neighborhood ratings overall compared to referent and the interaction indicates that his/her ratings were EVEN worse (higher score) for the city catchment area.
   
   2) The beta for interviewer 13 was -0.219; the beta for interaction [interviewer by city] was 1.213. Interpretation: Interviewer provided better neighborhood ratings overall compared to referent, and the interaction indicates that his/her ratings were worse ONLY in the city catchment area.

We were cautious about interpreting this finding because the number of interaction terms made it likely that some “significant findings” would occur by chance. Further, there was no overall direction to interviewer effects (city vs. suburbs). The $R^2$ of the main effect model shown in Table 3 is 0.320, and was increased only to 0.332 with the addition of the two interaction terms, suggesting the interactions did not provide an improved explanatory effect on the model. We added an abbreviated comment on these analyses to the results section.

5. **Table 4 models arrangement.** We dropped the outcomes analyses, but the advice for table construction is a good one we plan to employ in future publications.

Minor

1. **On page 14, the first sentence under “Phase 3” is not clear. I’m not sure what this meant related to your findings.** The organization of the material in “phases” was unclear. We dropped the outcomes analysis, and reorganized and extended the discussion of each phase.

2. **It is not clear what is presented in Table 2.** The table now has more descriptive labels to clarify the analysis and findings – this confusion was noted by multiple reviewers.

**Reviewer 2**

**Introduction**
We revised the introduction, and as requested, discussed the issues arising from the first neighborhood rating scale

Methods

The age range of the study participants and perhaps sex distribution should be stated. Added: There was a narrow range of ages (50-64) by design.

The authors should mention how many participants were enrolled in the inner city and how many in the suburban area. Added.

Were the participants clustered by block? There was minor clustering of the cohort subjects, and this is now described in more detail. Clustering was discussed in a recent publication on the cohort, which is now added as a reference (Schootman et al., 2006). In most studies of neighborhood effects, multiple study participants are nested within their neighborhood, requiring the use of multilevel statistical techniques. In the AAH study sample, most subjects resided on a block on which they were the only subject (65.9%). Only 3.6 percent of block faces contained five or more participants. We were not able to use multilevel statistical techniques between and within block faces because there is not enough clustering of participants within block faces to support a robust multi-level analytic approach. In our recent study of lower body limitations (Schootman et al., 2006), we randomly selected one subject per block face from the block faces with more than 1 subject and repeated the analysis, with no appreciable change in our estimates.

Instrument: section could be shortened. We received conflicting recommendations from reviewers in regard to the selection of scales for testing in this section. We have abbreviated, but hopefully clarified the issues related to the three scales and our pilot, “head-to-head” field test. Please include a reference for the original scale. We have included the only published description of the original scale we modified. We have been unable to locate, after vigorous searching, more comprehensive discussion of development and testing of the original scale used by Sampson and Raudenbush; we have also been unable to find any published reports on validity and reliability of that scale.

The authors should consider including the original rating system and the modified versions as an appendix. We provide the modified seven item scale as an appendix, but include a brief description of all items were considered in the Chicago version in Table 1.

Outcome measures. These have been dropped – we agree with multiple reviewers that this analysis requires substantially more detail. Since the background, methods, and results of this work would divide the paper into two very different aims (both measurement and outcomes), we believe that the outcome analysis should be separated into a publication that provides the depth and thoroughness that an adequate response to the reviewer’s questions would require. Furthermore, the longitudinal outcomes from the neighborhood measure are the most relevant, and these data will be part of the next five year study of the cohort.

Analysis. Reviewer comments on analyses to be considered during revision, and because the outcomes analyses have been dropped (and other phases reorganized), we believe we have resolved the confusion.
Results
The presentation of instruments and results was reorganized and is clearer.

Phase 3 does not present a comparison of the global rating by subjects to the interviewers rating scale instead the table presents the participants rating scale means by the categories of the global rating. We believe that poor table notes and titles made the direct comparison of rater scale scores and global ratings by resident subject unclear in Table 2, and this has now been improved.

The direction of the rating scale system should be stated in the text either on the methods or the results section. This has been added to each section and also as a footnote to each table.

In phase 4, third sentence, the authors stated that there was no association of interviewers experience and scale score (beta coefficient 0.326) perhaps the presentation of the p-value will be more efficient here as the coefficient does not state whether the association was significantly different or not. The p-value for this association is 0.135 and has been added to the text.

The finding of an interviewer effect (Table 3) is troubling and deserve some further examination. Perhaps the authors should examine the age and health status of the participants by interviewers, most of the interviewers with a significant effect have a negative coefficient indicating a better rating so it is possible that younger and/or healthier participants were interviewed by the interviewers with significant effects. The health status of subjects did vary by where they lived (e.g., see the reference Miller et al., 2005), which is a major hypothesis of our cohort study design. However there was no pattern of assigning interviewers by subject type or catchment area, and the age range of subjects (50-64) makes age an unlikely covariate to the interviewer effects. We also tested a model in which we included only interviewers with a large number in both catchment areas (at least ten), and this model did not suggest any smaller percentage of interviewers varied from the reference interviewer.

The description of the results in Table 4. We dropped Table 4. A table with the participants’ characteristics should be included. We added demographic descriptions of the cohort in the description of the cohort. The subjects of the cohort are not the focus of this measurement paper, but the rating scale will be linked to subject outcomes in future waves and the subjects’ characteristics will be an important part of those analyses.

Discussion
We dropped the outcomes analysis and table 4, and also the discussion related to Table 4 analyses and findings. The research using the original Chicago Scale does not include any tests of reliability or validity for our direct comparisons and discussion.

Reviewer: 3

1. Page 5, first paragraph. You mention age variation in neighborhood effects on health. We have trimmed the background and theory sections of the introduction to focus on the main
purpose of this paper, development and testing of a rating scale. Therefore, we did not pursue an augmented response to this issue in the revision.

2. **The background section and mechanisms:** We greatly reduced the background, but appreciate and added the suggested reference.

3. **Age range of the respondents in this survey?** We added demographic descriptions of the cohort – the cohort included 41.8% men at baseline, with a mean age of 56.8 years (range 50-64).

4. **Why did you choose just to focus on African Americans? Add a note commenting on this approach.** We added an explanation of the overall aims of the cohort, and the focus on one race.

5. **Page 5, the last paragraph. This paragraph should appear under the Method section.** Toward the end of this Background section, the reader should get a clear message regarding the key purposes of this study. I suggest you use this space to summarize the main aims of the study. We trimmed the background section of the paper to provide more specific rationale for the aim of this paper, i.e., development and testing of the rating scale. Because we had previously used and published results on a rating scale we decided was inadequate (Andresen et al 2006 *Journal of Aging & Health*), this forms part of the rationale. We moved the longer discussion of pretesting to the methods section, as suggested.

6. **What are the years of the 2nd, third, and fourth waves?** We added this: 2002, 2003, and 2004.

7. **Why did you use “fear for falling” as an outcome variable?** We dropped the outcomes analysis, as discussed above.

8. **Factor loadings below 0.5 are pretty low; number of items with loadings this low.** The factor loads for 7-item scale were: traffic (.470), street condition (.653), noise (.528), beer (.650), cigarettes (.701), garbage (.796), & residential unit condition (.755). We added these results in the text. **Will the alpha values go up if those items with low factor loadings are excluded from the scale?** We agree that the alpha value for the scale is sufficiently high (0.75 for the single block face; 0.73 for both block faces); since only one item dips below a factor loading of 0.50, we have not added details about changes with deletion of any item(s).

9. **Testing interaction effect of neighborhood variables with urban/suburban location.** We did not expect an interaction with neighborhood variables and the two catchment areas – rather, we expected very strong concordance of the neighborhood ratings with the neighborhood catchment areas. This is the focus of the discriminant and concurrent validity analyses. However, based on a suggestion of another reviewer, we tested if the interviewer effect interacted with neighborhood, and found two of 17 interactions significant (Reviewer 1, item 4).

10. **Modeling strategy: you did not use the conventional method like random effects models to analyze the data, nor used robust standard errors to adjust inter-correlations with each block; from what I read you only used one-level logistic regression models in the data analysis. Why?** There was minor clustering of the cohort subjects, and as discussed above with Reviewer 2, we have previously tested clustering, and there is not enough clustering to use multilevel statistical
techniques between and within block faces. In our recent study of lower body limitations (Schootman et al., 2006), we randomly selected one subject per block face from the block faces with more than 1 subject and repeated the analysis, with no appreciable change in our estimates.

11. Did you use weight to get city-wide representative results? We have applied sampling weights to our publications on the health results of subjects, but the focus of this paper is on the rating scale, e.g., factor structure and discriminant validity, and therefore analyses in the current study do not use a weighted strategy.

12. Census-based neighborhood SES measures can be linked to your data. It would be interesting to test how these different sources of measurement of neighborhood contexts are correlated and which one(s) have more discriminate power predicting health. We agree that neighborhood SES measures from Census are useful, and this forms part of the analysis of a paper now in press for the American Journal of Epidemiology on the health outcomes (diabetes) of the subjects themselves. However, the focus of the present manuscript is not on the resident subjects, but on the rating scale itself.

13. Is the neighborhood scale that appears in the tables based on the interviewers’ ratings or the respondents’ ratings or both? Table 2 presents correlations between the interviewers’ ratings and the resident respondents’ ratings. We appreciate that multiple reviewers found this table confusing and have clarified it in response.

14. The finding that neighborhood ratings are not significantly associated with health outcomes after controlling for individual income should be commented in more depth. The analysis of outcomes would require substantive additions to the paper, and we chose to drop it to concentrate on the clearer aspects of measurement that were intended as the focus. We agree that the health outcomes analysis is of interest, and consider this to be an important component of the longitudinal analyses we will conduct in the additional waves of follow-up of the cohort that will begin next year.