Author's response to reviews

Title: Transition to the New Data Collection Standards in the Department of Veterans Affairs

Authors:

Min-Woong Sohn (msohn@northwestern.edu)
Huiyuan Zhang (h-zhang2@northwestern.edu)
Noreen Arnold (nlarnold@att.net)
Kevin Stroupe (k-stroupe@northwestern.edu)
Brent C Taylor (Brent.Taylor2@va.gov)
Timothy J Wilt (wiltx001@umn.edu)
Denise M Hynes (Denise.Hynes@va.gov)

Version: 2 Date: 21 June 2006

Author's response to reviews: see over
June 19, 2006

Dear Editor:

I am pleased to submit a revised manuscript titled “Transition to the new race/ethnicity data collection standards in the Department of Veterans Affairs.” All the comments except for discretionary ones were addressed and the manuscript was revised accordingly. At the end of this letter, I am attaching a file that shows item-by-item responses to the reviewers’ comments.

Please feel free to contact me should you need any additional information from me.

Min-Woong Sohn, Ph.D.

Research Assitant Professor, Northwestern University  
Research Health Scientist, Hines VA Hospital
Reviewer 1

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. According to my calculations, only about 20% of the VA population of interest (all VA healthcare users in 2004) were included in this study (52% of the individuals had a single self reported race value in 2004 could be linked to the older race data x 39% who had valid self-reported race)= ~20%. If this is correct, and only ~ 20% of the population were represented, discussion of the representativeness of this sample and generalizability of the findings should be included in the paper. Otherwise, if I misunderstand, it would be helpful to spell out more clearly the proportion of the parent population who were included in this study sample. The authors briefly touch on the related issue of the representativeness of the sample in the limitations paragraph of the paper, but I would encourage them to expand this discussion further.

This calculation is correct. For representativeness, however, the percentage of all individuals with observer-recorded race values in 1999 – 2002 who could be linked to the self-reported race values in 2004 would be more meaningful, because it was the old race values whose accuracy was evaluated in this paper. The percentage was 28%. To address this, changes in the manuscript were made on pp. 10-11 and 19.

2. Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct) I would encourage the authors to comment on the finding that agreement rates varied by geographic region (and likely, by extension, on the racial/ethnic composition of geographic regions). That is, the rates were higher in the Midwest where there might be more whites, than in the West, where there might be more Native Americans or Hispanics.

This issue was addressed on pp. 17 and 19. We mentioned that the regional variations in the completeness of self-reported race may reflect the variations in the regional distribution of Hispanics, Asians, and other non-African-American minorities and that researchers using self-reported data should be careful about the potential selectivity in the self-reported race/ethnicity data.

3. I would also encourage the authors to clarify whether the fact that the parent study was concerned with urologic diseases is relevant to this paper. I was confused at first that the parent population was urology patients until I read the Study Design section several times to understand that this did not seem to be the case.

We addressed this issue by dropping the mention of the parent study from the text. See p. 6, first paragraph.
1. This paper addresses an important problem for researchers using the race and ethnicity data in the VA databases, the comparability of self-repost data collected post-2003 and observational data collected pre-2003. This demonstration that both data sources appear to be comparable for whites and blacks but not for other minority groups is very helpful to researchers in this area. However, only 39% of veterans had self-report data in 2004, and only 52% of veterans with a single self-report value in 2004 could be linked to data from 1997-2002. In particular, the 25% of self-reported Hispanics with no race data were excluded from the comparison over time. The implications of this acknowledged limitation deserve more attention.

This comment partly overlaps the comment on the limited generalizability by the first reviewer (#1). The implication of Hispanics not reporting race values was that the self-reported data may have self-selection problem that observer-recorded may not have. See pp. 17 and 18-19 for changes made in the text.

2. The second identified objective, how multi-racial data can best be "bridged" to a single value, is addressed descriptively with no real answer provided other than that results differ depending on how you do it.

We change the objective to match what’s described in the text better. See p. 5. Now the sentence reads, “The objective of this study is to examine the effect of this transition on the research use of race/ethnicity data for multi-year trends in the VA, specifically focusing on how comparable race data collected under two different methods are and what the effect of bridging may be on different race categories when one try to map multiracial race values to the old single races for the same individuals.”

3. Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Reference 28 is redundant (also listed as reference 26).

Reference 28 was dropped.

4. How multiple discrepant records were handled should be clarified.

The issue was clarified on p. 10. We mentioned that multiple discrepant records were not used in any comparisons in the study.