Author's response to reviews

Title: Temporal and Geographic variation in the validity and internal consistency of the Nursing Home Resident Assessment Minimum Data Set 2.0

Authors:

Vincent Mor (Vincent_Mor@brown.edu)
Mark A Unruh (Mark_Unruh@brown.edu)
Orna Intrator (Orna_Intrator@brown.edu)
Shubing Cai (Shubing_Cai@brown.edu)

Version: 2 Date: 27 December 2010

Author's response to reviews:

To the Editor:

This manuscript "Temporal and Geographic Variation in the Validity and internal consistency of the Nursing Home Resident Assessment Instrument Minimum Data Set 2.0" represents a summary evaluation of the performance of the nursing home MDS in the US that is widely used for clinical, policy and program evaluation research. Using data from all MDS assessments and all nursing home admissions over the period 1999 thru 2007, we assess the validity of the diagnostic data on the MDS by cross walking it with matched Medicare hospital claims. We also assess the construct validity of several of the scales used to characterize the nursing home populations' functioning and risk in terms of how they predict one year mortality. Finally, by cross-walking the MDS admission and discharge records with Medicare claims and enrollment, including death, data, we assess the completeness and accuracy of this information and its utility as an indicator of hospitalization and mortality. We discuss the findings of these analyses in light of the switch to a new version of the MDS which was introduced in October, 2010 in the US.

Responses to Reviewers :

Temporal and Geographic variation in the validity and internal consistency of the nursing home resident assessment Minimum Data Set 2.0

Note to the Editor:

My authors and I have responded to the suggestions of the reviewers. In so doing, we have added appendices presenting detailed results over multiple years and altered Table 1 to include it in the results rather than in the methods section. Additionally, in reviewing our paper we also realized that we’d not actually provided as much information that the reviewers and readers might want to see relevant to the “completeness” of the MDS record stream relative to information in the Medicare claims (e.g. mortality and SNF admissions). This information has
been added to the results immediately following data on the match rate and sample description and is incorporated into the discussion, particularly regarding the problems of using the MDS discharge record to estimate hospitalizations.

Finally, in accordance with the instructions to the authors, we have added sections on the authors’ competing interests and contributions to the current manuscripts as well as acknowledgements to the funders and to programming staff.

Reviewer: Jones

We appreciate the praise regarding the paper. Preparing this paper has been a long time coming.

Discretionary Revisions:

1. While we appreciate the importance of the issue that the reviewer raises regarding the relative utility of an instrument like the MDS for clinical vs. regulatory or reimbursement purposes, expanding the discussion to include this issue would take the paper well beyond its current, rather technical scope. Furthermore, the issue of how “reliable” clinical data needs to be to guide decision making is a matter related to test-retest reliability and not the internal consistency of multi-item scales. In any case, the standard for determining the adequacy of the diagnostic accuracy of a test is highly dependent upon the kind of decisions that need to be made (e.g. PSA tests predicting prostate cancer vs. an MMSE to help in the clinical diagnosis of dementia).

2. Since both reviewers requested additional test parameters in addition to the positive predictive value, we’ve included both the PPV and the sensitivity of the MDS diagnoses relative to the Medicare claim gold standard in Table 3. We are adding an appendix that includes sensitivity, specificity, PPV and NPV for the reader to consider all the aspects of the information overlap.

Minor Points:

1. The reviewer correctly notes that this sentence is insufficently clear, raising a wide array of additional issues that are beyond the scope of this paper. Consequently, we’ve chosen to drop the sentence altogether.

2. We used a CPS cut point of less than or equal to a value of 3; we used the median for the 28 point ADL scale. We’ve made this explicit in the methods and in the table.

3. We have included the alpha levels for the entire population in the text and added the sample size to Table 5 which presents alpha levels for the scales by the four strata. In the methods we’ve provided an additional rationale for presenting the alpha levels for very different types of residents. In the text we include the alpha for the 28 point ADL scale.

Reviewer: Phillips
Minor Essential Revisions:

1. The reviewer is correct that several studies of the “validity” of the MDS relative to research assistant observations found low correspondence in patient level analyses but facility level averages were consistent with the MDS based measures. Since this seemed to be a “nuance” rather than a major point, we chose not to include it in the background section in the original version, but agree that it does raise a point relevant to general discussions of MDS validity. We’ve added a few sentences and the relevant references.

2. We thank the reviewer for pointing out this lack of clarity. On consideration, this sentence didn’t really add a lot so we’ve dropped it.

3. The issue of language regarding “internal consistency” and alpha reliability presents a challenge in the current paper since we use the term internal consistency to refer to the cross-walk between data elements in the MDS that should be logically strongly connected AND are presenting the results of average inter-item correlations, known as “alpha reliability” according to J.C. Nunnally in his text on Psychometric Theory. Empirically testing the “internal consistency” of multi-item scales is a function of the correlation matrix rather than a comparison of the test-retest reliability of different judges rating the same phenomenon. In essence, in the alpha reliability framework, the items are repeated measures just like the judges are when there is no “gold standard”. However, in recognition of the reviewer’s concern, we’ve used the term “alpha reliability”. The reviewer’s other point referred to the analyses we conducted in which we identified pairs of items in the MDS that should logically have a high degree of correspondence. Several of the pairs include having an ADL score that is “perfect” indicating no impairment which is contrasted with an item like paraplegia; no one should have paraplegia indicated AND be fully independent in ADL. While this may appear to be a trivial example, meeting these tests represents a minimum data quality threshold. For some of these item pair cross-walks there is no expectation that the relationships should be perfect, rather it should be high because deviations from high correspondence should be relatively unusual. In this instance, it is instructive to compare the rates of internal consistency across facilities as a possible indicator of data quality. However, we have dropped several of the comparisons on the grounds that the natural cross-walk between the items is “too far” afield. We have expanded our discussion of the rationale for these paired comparisons in the Methods section. Additionally, we expand our description of the alpha reliability test applied to assess the stability of multi-item scales.

4. We have edited the tense throughout the document to be in the past tense since we clearly already did these analyses.

5. We are using the 28 point ADL long form scale developed by Morris and Fries; we’ve specified as much in the methods section.

6. The issue of the accuracy of a diagnostic test is not exactly the same as comparing two measurement approaches such as we are doing comparing the diagnosis of the MDS to the “gold standard” of the Medicare hospital discharge diagnosis. However, if we consider the MDS an “alternate” test being compared
to the gold standard, Epidemiology texts define the Positive Predictive Value of a diagnostic test as the probability that an individual with a “positive” test (having an MDS diagnosis of CHF, for example) actually has the disease (has a Medicare diagnosis). Since we are assessing the ability of the MDS diagnosis to replicate the diagnosis on the Medicare claim, using PPV seems appropriate. Nonetheless, it is also reasonable to include information on the “sensitivity” of the MDS since that is the probability that someone with a Medicare claim diagnosis will have a positive test (the corresponding MDS diagnosis). Consistent with the other reviewer’s request, we’ve added the specificity to tables 3 and 4 as well. Consequently, we have revised the table and the methods to include sensitivity as well as the PPV.

7. We have added the Sample Size to the Figure titles.

8. Good point, we’ve added the CHESS averages as well as the ADL averages and CPS distribution to Table 2.

9. We’ve added the sample size to the table titles.

10. As noted above, we have expanded the description of these analyses in the Methods section including providing a rationale for each of the item pairs being compared. Thus, the condition of aphasia in the MDS can be cross-walked with making one self understood. Regardless of whether the resident can understand (passive aphasia), if they are not able to make themselves understood (active aphasia), there should be a reasonably high degree of correspondence, although there will be more individuals who are unable to make themselves understood than who have aphasia.

11. We thank the reviewer for this suggestion and we’ve added the population average alpha levels in addition to those for the sub-populations. The rationale for stratifying the analyses by ADL and CPS is that we were so concerned about the possibility that there would be different correlation structures across such very different types of residents; we forgot to include the results for the overall population. These have been added. Furthermore, we include statements concerning the population long form ADL alpha reliability in the text, which as the reviewer suggests is excellent.

12. The reviewer makes a good point and we have modified the section of the discussion regarding use of the MDS for research purposes. Furthermore, we note that even the well being, mood and behavior scales have reasonable alpha levels and the fact that they behave the same way with very different types of residents, is particularly encouraging.

Discretionary Revisions:

13. We’ve altered the methods and results sections to include Table 1 in the results section and have dropped reference to the population under 65 since, as the reviewers rightly point out, our study focuses largely on the 65+ Medicare population.

14. The reviewer correctly notes that the MDS allows for the stipulation of ICD codes in addition to the “check box” option. Cursory examination of the ICD fields in the MDS Repository data suggest that few records have any fields completed
and relatively few have many diagnoses, suggesting that the facility staff are using the check boxes on the MDS and not the fields for the ICD data. Since many facilities have separate paper or computerized medical record, we assume that the true ICD coded diagnoses are located there.

15. We have made the correction in the text.
16. We concur and have corrected the sentence appropriately.

If you have any further questions, please don't hesitate to contact me.

Vince Mor