Title: Functional Status Decline as a Measure of Adverse Events in Home Health Care: an observational study

Authors:

Tanya P Scharpf (tanyascharpf@hotmail.com)
Natalie Colabianchi (natalie.colabianchi@case.edu)
Elizabeth A Madigan (elizabeth.madigan@case.edu)
Duncan Neuhauser (duncan.neuhauser@case.edu)
Timothy Peng (timothy.peng@vnsny.org)
Penny H Feldman (PFeldman@vnsny.org)
John FP Bridges (john.bridges@urz.uni-heidelberg.de)

Version: 2 Date: 4 September 2006

Author's response to reviews: see over
September 4, 2006

Dear the BioMed Central Editorial Team,

We thank both reviewers for their thought provoking comments on the paper ‘Functional Status Decline as a Measure of Adverse Events in Home Health Care: an observational study.’ Our main effort has been to reply to the compulsory and essential comments made by the reviewers and to make the paper more readable for a general audience. We have bolded text throughout the manuscript for which significant changes were made. To follow is our reply to the reviewers’ comments.

Reply to Reviewer: Katherine Berg

** Compulsory/Major * Minor essential

**1. Present background information on the quality initiative used by CMS for home care and a framework for understanding how adverse event measures fit within the initiative

Response:
Thank you for pointing this out. Shaughnessy et al. developed a comprehensive framework using a two-stage process for assessing the quality of home health care which separates the effects of natural disease progression from the effects of substandard care. The process and the measures they developed were used by the Centers of Medicare and Medicaid in the USA as the model on which to base home health care quality improvement [1,2] The aforementioned sentences have been added directly to the manuscript itself.

**2. If there is an expectation that adverse events can be predicted, a multi-nominal model for competing risks of hospitalization and death and substantial functional decline is more appropriate.

Response:
We agree with the reviewer that this approach would have some value, however, after working with this data we find that some of the events are so rare that the maximum likelihood model, especially the complicated multi-nominal model, will not converge. Also, with the rarer events there is less variation across episodes. Previous quality research in institutionalized settings (e.g., hospitals and nursing homes) has focused on mortality as the preferred outcome measure (Aiken et al.2002; Krumholz et al.1999; Iezzoni et al.1996; Normand et al.1995). However, mortality as an adverse event in home health care is extremely rare compared to studies of 30-day post-operative mortality comparing hospitals or surgeon quality. In the current study of 49 437 episodes of care, only two unexpected deaths occurred. Therefore, functional status may be a more common and appropriate measure of home health care patient outcomes.

*3 There should be greater clarification in the use of predictive validity and test-retest reliability. The terms are generally used for measures whereas the authors use them to denote the validity and stability of multivariate models.

Response:
Thank you for pointing this out. The measures are being evaluated by assessing the c-statistics and Hosmer-Lemeshow chi square values produced by the models; this is noted as such throughout the manuscript.
Reviewer: Gideon Caplan

**1.** Validity, in my book, is the degree to which a test measures what it is supposed to measure. That is, a comparison between a test and a gold standard. I may have completely misunderstood, but I cannot see where that is done.

Response:
Thank you for this comment. We evaluated the ability of the indices to predict functional decline based on evaluating c-statistic scores and Hosmer-Lemeshow chi squares from the indices models. We agree that this is not quite the same as validity and have corrected our language throughout the manuscript to read ‘the indices ability to predict functional decline’.

**2.** The authors have done an evaluation of test-retest reliability, but that is generally taken to mean that the test is applied twice to the same sample. I cannot see where it is clearly stated that the two sets of patient episodes are necessarily for the same patients.

Response:
Thank you for identifying this conceptual ambiguity. This is corrected throughout the manuscript by instead using the term ‘consistency over time’.

**3.** I would use the whole ADL index, rather than only one, two or three ADLs. But that is the authors choice.

Response:
The whole ADL index is available for use, but we need to identify an outcome that can be simply monitored in practice. Essentially we are testing whether the method chosen by the Centers for Medicare and Medicaid in the USA are optimal. The indices require declines in a minimum of three, two, or one ADL. However, episodes which exhibited substantial declines in all fives ADLs would certainly fall into the index representing declines in ‘three or more ADLs’. It is not desirable to require substantial declines in all the ADLs since this is extremely rare; patients exhibiting such poor physical functioning in all ADLs would be unlikely to remain in home care.

*4. The writing is clear enough. On p.6, last paragraph, I would like to see substantial defined. I assumed this meant two points, but from the discussion, it may possibly have been three points.

Response:
Thank you for this comment. This has been corrected directly in the manuscript by redefining that substantial refers to a minimum two-unit decline.

*5. I believe there may be a significant typo at the top of page 8; in line two “two” should be “three”.

Response:
Thank you for pointing this out, it was a typo and has been corrected in the revised manuscript.

Thank you for your time and consideration,
Tanya P. Scharpf

e-mail: tanyascharpf@hotmail.com
phone: (216)382-7204
cell: (216)973-8308