Reviewer's report

Title: Unmet need and psychological distress predict emergency department visits in community-dwelling elderly women: a prospective cohort study

Version: 1 Date: 18 December 2010

Reviewer: Namkee G Choi

Reviewer's report:

Overall, the manuscript is easy to read and well documented. However, further clarifications in the following areas are necessary:

1. (Major compulsory) The authors employed a logistic regression analysis with two primary independent and many control variables; however, there is no theoretical framework and hypotheses. A theory-based approach to conceptualizing their research questions and clearly specified hypotheses is desirable.

2. (Major compulsory) Age inclusion criterion was specified as 75 years and older in the abstract and in the third line in the Source of Data section. However, the subsequent descriptions of survey/interview data collection are all about 65 years and older. Please clarify these discrepancies.

3. (Major compulsory) In the measurement section, it was unclear how the “no need” was scored. (See: “Once each participant’s level of need (no need, met need, or unmet need) was determined for each PADL and IADL, we aggregated the results into counts of met and unmet PADL and IADL need. We categorized IADL unmet need and the IADL net need as 0, 1, and 2 or more.”) I do not see any description of the scoring for “no need.” This is troublesome also because it is not at all clear how the “no need” was included in the logistic regression analysis. The analysis section needs to be clearer in its description of the variable attributes. Moreover, the last paragraph on adjusting for disability is very confusing to read. For example, in the second to the last sentence, “0 PADL met need” was entered in the equation as 0, while 1 IADL met need was entered as 1, so if a subject’s PADL need was not met, her disability score is low? Why can you just state that the disability score is the number of ADL/IADL impairments? I do not understand why this sentence is needed at all, as I do not see the disability score used anywhere in the analysis. The disability-related variables entered in the logistic regression model are the numbers of met and unmet needs.

4. (Minor essential) Place of residence: The authors may need to say a bit more about the control variables. For example, what does senior housing refer to? Do they differ from what we call assisted living facilities in the US? Or are they like low-income, public housing for older adults?

5. (Discretionary) Self-rated health: What was the original scale before it was dichotomized as good or poor?
6. (Minor essential) ED visits: it would help if you can provide count data on the number of ED visits, even if you collapsed the variable into a dichotomous one. If there is a substantial range, it may help to conduct a zero-inflated negative binomial analysis (it is a count outcome) to see if the number of ED visits may be a significant factor.

7. (Discretionary) Discussion and interpretation: The authors speculated that the significance of met PADL as a predictor of ED visit may be because family member and homecare workers are more likely to provide transportation for the older adults to ED. This may be so, but I wonder if further data on how these older adults were transported to ED (e.g., private car or ambulance) may be needed before speculating on this.

**Level of interest:** An article of limited interest

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.