Reviewer's report

Title: Falls in advanced old age: recalled falls and prospective follow-up of over-90-year-olds in the Cambridge City over-75s Cohort study

Version: 1 Date: 21 August 2007

Reviewer: David Ganz

Reviewer's report:

General

This is an interesting manuscript that addresses a growing cohort of individuals -- the "oldest old," about whom little is known. I think this manuscript is an important contribution to the literature and is generally well written but requires significant revision. My detailed comments follow below.

-----------------------------------------------------------------------

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

1) I am having some difficulty understanding the study's methods. There seems to be an attempt to compare retrospective to prospective data collection, as has been the case in previous studies. However, the authors seem to be comparing patients' baseline recall of falls in the year prior to baseline, to prospective collection of data on falls in the year subsequent to baseline. Because the cohort is aging over time and life events may occur in one year but not the other, to me it does not make sense to compare the two separate years. In previous work, the same year has been looked at both prospectively and retrosepectively, by having patients recall their falls at the END of the 1 year of prospective follow-up, or at intervals during the follow-up (reference 4 contains details). Figure 1 seems to be trying to get at this but the Methods section does not explain how the data in Figure 1 were arrived at. In addition, the superimposed graph of prevalence vs. incidence and the category "no falls recalled" make this Figure quite unclear.

In sum, if patients in the study were not asked at the end of the year of prospective data collection to recall falls in the past year, I think the whole prospective vs. retrospective comparison should be dropped, since it applies to two different years. The study still provides important epidemiologic data of interest to researchers, and can be reported descriptively.

2) Page 7. Poisson (or negative binomial) regression is likely to be inappropriate for this data if falls are clustered (i.e., one fall tends to occur right after a previous fall rather than at random). A key assumption for Poisson data is that there is a random process generating rare events that are independent of one another. That assumption seems to be failed here.

3) Page 7. Please elaborate on the variables available for insertion into
regression models, and why these variables were chosen. A significant concern in this manuscript are the multiple tests run on the data without any adjustment for multiple comparisons and without pre-specified hypotheses. These make the subsequent results difficult to interpret. This must somehow be addressed in the Methods section.

4) Page 7. In situations where the outcome of interest is common (such as number of fallers in this cohort), the odds ratio is known to overestimate the relative risk. In cohort studies the relative risk can be calculated directly using crude data and should be. For multivariable logistic regression (not being used here but could be), methods exist to adjust the odds ratio to get a corrected relative risk (see for example Zhang and Yu, Journal of the American Medical Association 1998; vol 280; pg. 1690). More sophisticated statistical methods also exist to generate a relative risk from logistic regression, such as using predicted probabilities. Given the availability of these methods and the known tendency of the odds ratio to overestimate the relative risk when the outcome of interest is common, relative risks should be reported rather than odds ratios in all cases in this manuscript. Whichever methods you choose, please report them in the Methods section.

5) Page 12. The sentence states, "Of these over-90-year-old men and women 60% fell at least once in the year after interview, closely matching the proportion who remembered falling in the year before interview (58%)." As discussed earlier, comparing two separate years is problematic. There are multiple reasons why this is so, including study attrition causing only healthier patients to have data points at both times (probably less of a problem in this sample), the noise introduced into the data by health histories that change from year to year, and the relatively small sample size (although this is an issue with many falls studies). You could provide the same data to the reader (descriptively useful) without trying to make inference about it, especially as the results are mixed for single fallers, repeat fallers, and total number of falls.

6) Discussion section. Please shorten this section significantly and focus on the limitations you think are most important. The Methods section, however, should be enlarged to provide more detail.

----------------------------------------------------------------------------------------------------------------------

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

1) In the abstract, conclusion section, the authors write that "Recall methods may underestimate numbers of repeated falls but not the proportions of fallers." However, in the results section, they state that the "proportion reported to have fallen more than once was lower using retrospective recall than prospective reports (34% vs. 45%), as were fall rates..." The two sentences seem to be at odds. Please clarify (or refocus the abstract entirely to be consistent with prior comments).
2) Page 6, sixth line from bottom. Sentence reading "If resident in a care home..." needs to be fixed.

3) Page 14, "prevalence does not increase in proportion to length of recall period..." A period is needed at the end of this sentence.

4) Table 1. Please present crude relative risks instead of unadjusted odds ratios. It does not make sense to present incidence rates for recalled falls because these data are recalled rather than prospectively collected, so I would remove the incidence information.

5) Table 2. Please present crude relative risks rather than unadjusted odds ratios. Please remove negative binomial regression estimates of incidence rate ratios (labeled as relative risk in the "incidence" column) for reasons discussed above (inapplicability of negative binomial regression to clustered events).

6) Figure 1. The prevalence data appear to apply to the previous year, yet the incidence data apply to the current year. Thus the two sets of data should not be superimposed. If the authors choose to keep either the prevalence or the incidence data (incidence data are more likely to be useful), they should provide additional detail as to how the data were derived (e.g., how does "period of recall" fit with incidence data?).

7) Figure 2 is quite useful and a nice contribution.

8) Figure 3. Please label the y-axis of the two graphs so that we know what we are looking at. I think this should be "Percent free of falls in subsequent year" or something to that effect.

--------------------------

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article whose findings are important to those with closely related research interests

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.