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: 2 Date: 19 December 2007

Reviewer: M. Clare Robertson

Reviewer's report:

Thank you for the opportunity to review this revised paper. The authors have addressed my comments and those of the other reviewers very carefully and thoroughly. The paper is much improved as a result and reads well with some interesting new information provided about the sample.

Major Compulsory Revisions

The author must respond to these before a decision on publication can be reached. For example, additional necessary experiments or controls, statistical mistakes, errors in interpretation.

I strongly support the use of negative binomial regression models for modelling the number of falls in the follow up year, and the number of falls reported in the previous year for that matter (separate models). Negative binomial regression models are commonly used in the falls literature and the authors do provide a published justification for their use in this field (reference 11). These models do adjust for overdispersion and my understanding is therefore that they account at least to some extent for the clustering of falls within individuals. Figure 2 illustrates nicely that the distribution of falls in the follow up year appears to meet a negative binomial distribution.

I have only one concern. Please can the description of the statistical methods used be explained a little more clearly still? It is not quite clear when negative binomial regression or logistic regression has been used especially when the results of multivariable models are reported. Perhaps subheadings in the Statistical Methods section would help (eg 'Prevalence' of falling; 'Incidence' of falls; Multivariable models).

I am familiar with the cluster(varname) statement being used in Stata when adjusting negative binomial regression models for clustering in cluster randomised controlled trials, but I am unclear about its use in the models the authors report. Perhaps the authors could state the independent and dependent variables used in the models and the variable used in the cluster statement, if indeed the cluster statement is necessary. The reason I am puzzled is that, if there is one record (row of data) per individual as I would expect, clustering on the individual would have no effect on the size of the confidence intervals of the relative rate estimates (incidence rate ratios) produced by the models.
Also, as the results from a negative binomial regression model are expressed as incidence rate ratios, this term should be used where these models are reported in the Results section and in Tables 1 and 2. The abbreviation RR should perhaps be avoided for these models, as it may be interpreted incorrectly as a relative risk ratio (rather than a relative rate ratio).

- Minor Essential Revisions

The author can be trusted to make these. For example, missing labels on figures, the wrong use of a term, spelling mistakes.

1. Measurements / Falls during follow-up section: I suggest using “If the participants were resident in a care home, they...” to avoid the use of both singular “was” and plural “they” in this sentence.

2. Results section, bottom of page 9: I suggest “how long after baseline before any fall event occurred.” or similar rather than using “fall arose.”

- Discretionary Revisions

These are recommendations for improvement which the author can choose to ignore. For example clarifications, data that would be useful but not essential.

What next?: Accept after minor essential revisions

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests.