Reviewer's report

Title: Timing of surgery for hip fracture and in-hospital mortality. A retrospective population-based cohort study in the Spanish National Health System.

Version: 1 Date: 22 August 2011

Reviewer: Reijo Sund

Reviewer's report:

The authors study the potential association between the timing of surgery for hip fracture and in-hospital mortality with large Spanish data. The study is quite interesting, although there have been a lot of studies on the issue in recent years. There are, however, some issues that should be considered before publication:

1. Literature review on previous studies should be revised. Studies measuring time from admission to operation should be separated from those measuring time from fracture to operation. Published meta-analyses on the issue could be cited for more comprehensive list of older papers. Also paper by Sund and Liski (Qual Saf Health Care 2005;14:371–377, doi: 10.1136/qshc.2004.012831) should be cited as it gives a methodological explanation for controversial results on the issue.

2. In-hospital mortality is not very good outcome measure if the effect of operative delay is studied. Here it is the only option, but the authors should evaluate the usefulness of this measure even more carefully: Is it comparable across hospitals? (i.e. are the hip fracture patients operated and rehabilitated in the same hospital and then discharged home? Is the length of stay nearly constant or are the persons with most problems/complications staying longer in hospital without discharge so that their in-hospital mortality will be overemphasized?) What about discharged patients who die shortly after discharge? In addition, surprisingly large proportion of patients who die very soon after fracture (and operation) may be such that they would have died no matter how good treatment they had received.

3. Please comment on the quality of coding of secondary diagnoses as it may have huge impact on the severity measures used.

4. Why backward-forward logistic regression was used instead of a manual choice of the most important variables? Is it reasonable to keep both fracture and operation type variables separately in the model as they are not independent of each other (operation type will be selected after knowing the exact fracture type from x-rays).

5. Results section should be revised so that reference group is always explicitly mentioned (x% more or less likely _than what_?).
6. Please discuss why weekend surgery is associated with higher mortality of hip fracture patients in Spanish data as that seems not to be the case in all countries (see. e.g. Bell & Redelmeier NEJM 2001;245(9)663-8 or Sund and Liski 2005).

7. It is reported that 25% of patients were operated early. Was that constant proportion across hospitals? It is known that, at least in Finland, there if huge variation between the proportions of early operated patients in different hospitals (Sund and Liski 2005). A simple performance assessment interpretation of hospital specific delay rates is that the percentage of late surgery patients can be reduced to a potentially achievable level (that would less than 10% in Finland). This level can also be interpreted as the upper limit for the proportion of acceptable delayed patients. Correspondingly, the expected proportion of unacceptable delayed patients is the proportion of late surgery patients exceeding this upper limit, and thus the overall hospital level mortality of hip fracture patients should increase with an increasing proportion of late surgery patients given that the longer operative delay would have an adverse effect on mortality.

In this sense, the authors should also report a scatter plot between proportions of delayed patients and mortality. There should not be any trend if timing of surgery has no effect on in-hospital mortality. Please see http://www.bmj.com/content/332/7547/947/reply.

8. There were data from several hospitals in this study. It is likely (see bullet 7) that the association between proportion of early operated patients and in-hospital mortality varies between hospitals. This means that a (multi-level) model that allows varying effect (i.e. random-intercept model is not enough) should be used. Use such a model or discuss the related bias (which actually seems to be erroneously ignored in most studies)!

9. Add some explanations and/or more detailed column headings to the tables. Now it is difficult to know what percentages mean in different places.

10. Total n is 56500 in Tables 1 and 3 but 56482 in Tables 2 and 5. Why?

11. Please discuss the impact of medical and non-medical (lack of personnel/operating rooms) reasons of delay to the potential association.

12. Optional: Confirm your main finding with an instrumental variable approach (use day of week as an instrument) similarly as in Sund and Liski 2005.

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

**Quality of written English:** Needs some language corrections before being published

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

I declare that I have no competing interests.