Author’s response to reviews

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

Authors:

Julián Librero (julian.librero@uv.es)
Salvador Peiró (peiro_bor@gva.es)
Edith Leutscher (ed.leutscherv@coma.es)
Juan Merlo (Juan.Merlo@med.lu.se)
Enrique Bernal-Delgado (ebernal.iacs@aragon.es)
Manuel Ridao (ridao_man@gva.es)
Natalia Martínez (nmartinez.iacs@aragon.es)
Gabriel Sanfélix-Gimeno (sanfelix_gab@gva.es)

Version: 2 Date: 19 October 2011

Author’s response to reviews: see over
Dear Editor,

Please find enclosed the revised version of our manuscript “Timing of surgery for hip fracture and in-hospital mortality. A retrospective population-based cohort study in the Spanish National Health System.” (MS: 1546736131528395), submitted for your consideration for publication in BMC Health Services Research.

In response to the concerns and suggestions expressed by the referees and the Editorial Office, we have made the following changes:

**Reviewer #1 (Andrea Giusti).** The article is a well-designed and well-written retrospective analysis about time to surgery and mortality after hip fracture. The sample is adequate as well as the Methods and the statistical Analyses. The Results are interesting and may really represent the basis for future research in this field, even if they confirm our uncertain about the effect on mortality of early surgery after hip fracture. I have only some minor comments:

[Reviewer #1 #01]. Page 6: I suggest to explain some of the exclusion criteria. In particular, the Authors should explain the reason why they excluded patients with a length of stay longer than 30 days and those admitted to hospital with less than 30 hip fractures.

Most of the exclusion criteria (such as age, multiple trauma or pathologic fractures) were intended to ensure a more homogeneous premorbid health status. Elective admissions were excluded because hip fractures in these hospitalizations suggest inpatient fractures and the time from admission to surgery could be not related to the surgery for hip fracture.

It is expected that (at least some) delayed patients will be sicker patients needing to be stabilised medically before surgical treatment. The propensity score and the multivariate models try to control the severity differences between both types of patients, but in our opinion excessive delays could correspond to very different types of patients whose risk differences cannot be adequately adjusted with the available covariates. Therefore, we excluded these patients to reduce the heterogeneity between early and delayed surgery groups. The 30 days term was chosen (not without a certain arbitrariness) because we consider delays beyond 30 days for non-medical reasons to be very unlikely and that these exclusion criteria would not involve (quite the opposite) a bias in the analysis.

Finally, we excluded hospitals with less than 30 hip fracture discharges in whole the period analyzed (74 cases in 4 years from 19 hospitals) because we suspect these hospitals do not have trauma and orthopaedic wards and the cases correspond to patients transferred from/to other hospitals.
Following the reviewers suggestion we have added some sentences explaining the exclusion criteria in the M&M section.

[Reviewer #1 #02]. Tables 2 and 5: As I understand from the Tables, Authors included only variables that were significant in the multivariate analyses. I suggest to include in these Tables also the variables included in the multivariate models that resulted not significant.

Non-significant variables in the multivariate analysis are directly excluded from the statistical programs when stepwise methods are used. Because they are not significantly different to the basal value their OR is, by definition, equal to 1 and their inclusion in tables is very unusual. In the previous version of the manuscript these variables are included as footnotes. We could force the entry of these variables into the model but this would alter the OR values of other variables without a clear logic.

No amendments have been made regarding this comment.

[Reviewer #1 #03]. Page 12: The Authors suggest that unobserved factors related to the care itself are responsible for better results and explain only few of these potential unobserved factors. Most of these are related to general protocols worldwide implemented such as thrombosis prophylaxis. I believe that there are other factors such as the implementation of orthogeriatric services, as well as our better knowledge of the "Hip Fracture Syndrome" in the elderly. If the Authors are in agreement they should comment on that.

We agree with the reviewer that several factors (including organizational changes) have contributed to improving clinical outcomes in hip fracture in addition to those cited in the previous version of our manuscript.

Following the reviewers suggestion we have added some of them, including the implementation of geriatric co-management and better knowledge of the hip fracture syndrome.

[Reviewer #1 #04]. Limitations: Another limitation arises by the fact that it is well established that ICD-9 discharge diagnosis for Hip Fracture may be miscoded up to the 3%. Please include this limitation with relevant references (Giusti A et al., Bone 2011; Maggi S OI).

We agree with the reviewers [see also response to reviewer #2 #03] that accuracy and completeness of ICD9CM codes (both primary and secondary diagnoses) is always an important concern in outcome research using administrative databases.[1] It is known

---

that the quality of diagnostic data varies according to countries and regions, hospitals and clinical conditions. Although Spanish Regional Departments of Health have carried out some internal audits on the Minimum Basic Data Set in their respective regions, there are very few published studies on data quality in Spain and, to our knowledge, none of them revised the hip fracture coding.

➔ We have included a paragraph about this shortcoming in the Discussion section with the corresponding references. Anyway, notes that for the main objective of our study regarding problems of data completeness or inaccuracy would suppose an important bias only if there were differentials between early or delayed surgery groups.

Reviewer #2 (Reijo Sund): 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:

[Reviewer #2 #01]. 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

Our study did not aim to perform a systematic review of literature on hip fracture surgery (a goal that could be a project in itself). In fact, our non-systematic review was based on the 17 references cited in a previous review published in 2006 and updating later studies till 2009, finding a total of the 30 references used in the previous version of our manuscript. We acknowledge that this non-systematic review contains important absences, including several recent meta-analyses and the really interesting paper from the reviewer (very useful in the discussion of some results and not identified in the previous paper of Majumdar, nor in our search because it starts in 2006).

➔ We have revised our literature search to include published meta-analyses and other references, including the paper cited by the referee.

The reviewer also suggests separating “studies measuring time from admission to operation ... from those measuring time from fracture to operation”. We do not understand the purpose of this separation in the context of our study. With the exception of in-hospital fractures (in patients admitted from other reasons), hip fractures come to hospital very quickly (typically within a few hours) because pain and loss of function leads to the seeking immediately help. When the majority of studies use ranges from 1 to 2 days, these few hours do not seem particularly relevant. Vidal et al studied this


question specifically and concluded “The gap from admission to surgery may be used as a surrogate of the actual delay from fracture to surgery when studying in-hospital HF mortality”. [4]

» Finally, we have opted to maintain the previous separation of references in accordance with the results of the corresponding study (early surgery beneficial, harmful or not significant).

[Reviewer #2 #02]. 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.

We agree with the reviewer that in-hospital mortality is an endpoint with many limitations that gives a biased picture of mortality (see Jencks et al[5] for a comprehensive explanation of this bias) and we discussed (certainly in a very brief way) this issue in the previous manuscript. The use of inpatient death as an outcome measure involves several problems:

1. More aggressive discharge policies can reduce “mortality” rates if patients are discharged “alive” to die at home. These discharge policies could be, as the referee suggests, different between hospitals and also between years and between patient conditions. Therefore in-hospital mortality introduces heterogeneity and could bias our results if patients with poor outcomes (associated with one or another of the surgical alternatives evaluated) were discharged early. We have no information about patients who die shortly after discharge although in Spain hospitals use a code of “in extremis discharge” for patients discharged alive in terminal conditions (in our study these patients were considered deaths). Otherwise, rehabilitation in Spain is initiated in the hospital setting, but patients are usually discharged to continue it in the outpatient setting and we do not think this issue is important in our context.

2. Longer length-of-stay (LOS) increases the exposition period (the probability of identifying the outcome) for some outcomes such as mortality or complications. Because delayed surgery is associated with longer LOS a greater identification of deaths in this alternative is probable. Also, because –as the referee suggests- persons with more problems/complications stay in hospital longer, in-hospital mortality will be overemphasized in these groups. If poor patient condition was associated with the surgical intervention (or with the decision to delay the intervention), this bias could contribute to overemphasizing mortality in the group with delayed surgery.


3. There is a complex interaction between LOS, complications and mortality because patients with complications have longer LOS, but in turn longer LOS increases the risk of complications. Also the recording of secondary diagnosis could be associated with LOS affecting the covariate adjustments. Therefore, controlling the biases introduced by the use of in-hospital mortality as the endpoint (for example, using LOS as a regressor) is not easy to correct. 

Since in-hospital mortality was the only outcome measure available in our study we cannot control most of these problems (although some of them should have been adjusted by the propensity score).

➡️ **We have opted to broaden the discussion about this limitation in the corresponding section.**

Otherwise, patients who died very soon after fracture and before the intervention were excluded from our study (in fact, all patients without surgery were excluded). Patients who died very soon after the operation were assigned to the corresponding group of early or delayed surgery. It is possible that they would have died no matter how good the treatment they had received, but also other causes of death (for example, surgery or anaesthesia complications) are possible for these patients, including causes related with the decisions about surgery timing. Therefore we think that our analysis (similar to the intention-to-treat analysis) is adequate.

➡️ **No changes have been made regarding this last comment.**

**[Reviewer #2 #03]. Please comment on the quality of coding of secondary diagnoses as it may have huge impact on the severity measures used.**

As stated in a previous response (see response [Reviewer #1 #04]), there are very few published studies on MBDS data quality in Spain and, to our knowledge, none of them have reviewed patients with hip fractures. Years ago, our group studied the quality of secondary diagnosis (in one Spanish region and in conditions other than hip fractures) finding several problems of accuracy, completeness and information biases.[6] For example, in severe conditions chronic comorbidities diagnoses tend to act towards protecting from (not increasing the risk of) death. We speculate that chronic conditions were over-recorded when there were no other more severe conditions (acting as markers that there were no other serious conditions).

For the main objective of our study, problems of data completeness or inaccuracy would suppose a bias only if there was a differential between early or delayed surgery groups and/or with the main study endpoint. While we do not expect biases related to timing to surgery itself, we consider feasible the presence of differential information biases

---

according to certain patient characteristics (age, severity), their management (complications, adverse effects) or the main outcome (alive, death).

⇒ We have opted to broaden the discussion about this limitation in the corresponding section.

[Reviewer #2 #04]. 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).

As noted in the previous version of the manuscript “we used a backward-forward multivariate logistic regression model to analyze the relationship between selected explanatory factors (age, sex, type of fracture, type of surgery, Charlson Comorbidities Index scores, RMI scores and discharge year) and …”. Therefore, we combine a “manual choice of the most important variables” with the backward-forward method to exclude non-significant covariates.

⇒ No changes were made regarding this comment.

Although the type of fracture is related to the type of operation (usually, intracapsular fractures are treated by hemi-arthroplasties or complete arthroplasties, while trochanteric fractures are treated with reduction and external fixation) the type of fracture is not the sole determinant of the type of intervention. For example, in non-displaced intracapsular fractures or in young patients can be treated by fixation with preservation of the femoral head. Therefore, we opted to introduce both variables into the logistic models. In these models, trochanteric fractures and the use of arthroplasties were negatively associated with early surgery but not with in-hospital mortality.

⇒ No changes were made regarding this comment.

[Reviewer #2 #05]. Results section should be revised so that reference group is always explicitly mentioned (x% more or less likely _than what_?).

⇒ The writing of the section has been revised in accordance with the reviewer’s comment.

[Reviewer #2 #06]. 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).

We have revised the analysis and we have opted to change the analysis with the weekend surgery variable for the analysis with the weekend admission variable. Higher mortality in weekend surgery is probably explained by the presence of severe complications requiring emergency surgery (including emergency surgery in patients admitted over previous days). We think that to explore the relationship between weekend admission and mortality is more relevant to our study than the previous
analysis. The new results (non-adjusted) show no differences in mortality between weekday and weekend admissions (4.29 vs. 4.23, p=0.755; see Table 3). Also, in the stratified analysis mortality for weekend admission does not show differences between early and delayed surgery (4.25% vs. 4.23%; p=0.949). For weekdays the early surgery group show a lower mortality (3.73 vs. 4.49, p<0.001). Anyway, an adequate analysis of weekend “risks” should consider the severity of patients conditions and the control of similar periods (as holidays), issues not addressed in our study. Actually we are trying to obtain the 2002-2010 series to analyze the weekend impact on mortality for several conditions including hip fracture.

We have corrected the tables and text, substituting data on weekend surgery with data on weekend admission.

[Reviewer #2 #07]. 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.

Following the reviewer’s suggestion we have prepared a scatter plot between proportions of delayed patients and mortality (see figure below). Trends were estimated using a lineal OLS regression (green line), and a smoother estimation by means of a hospital weighted-polynomial regression (red line). As expected by the reviewer, the overall hospital-level mortality of hip fracture patients increases (discreetly) with an increasing proportion of delayed surgery patients, but the proportion of delayed surgery patients was non-linearly associated with a higher mortality rate (with a sharp decline in mortality after exceeding 80-85% of patients with delayed surgery). These results are very suggestive that the effect of delay in operating surgery on mortality is mainly due to unavoidable delay in more severe patients who are unfit for surgery.

We have included this analysis in the web annexe with the corresponding references in the main text of the manuscript.
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)!

Following the reviewer’s suggestion we have constructed a multi-level analysis model to study variation between hospitals in the risk of death after hip fracture surgery (multilevel logistic random intercept model) and to analyze the effect (constant vs. heterogeneous) of early surgery on mortality across hospitals (multilevel logistic random coefficient model). The results of this analysis show that: 1) Around 2% of variation in the risk of death could be attributable to the hospital-level; 2) There is no evidence that the delay in surgery effect differs across hospitals. These results do not alter the study’s conclusions.

➔ We have opted for including the multilevel analysis in the web annexe with the corresponding references in the main text of the manuscript.

[Reviewer #2 #09]. Add some explanations and/or more detailed column headings to the tables. Now it is difficult to know what percentages mean in different places.

➔ We have changed column headings in accordance with the reviewer’s comment.
As noted in the footnotes of descriptive tables (1, 3 and 4), there were 18 cases with missing values in the variable sex. These 18 cases are excluded by the logistic analysis (Tables 2 and 5).

➔ We have included the corresponding footnote in Tables 2 and 5.

[Reviewer #2 #11]. Please discuss the impact of medical and non-medical (lack of personnel/operating rooms) reasons of delay to the potential association.

➔ We have included a paragraph in the discussion section commenting on reasons for delay and potential associations with mortality.

[Reviewer #2 #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.

Following the reviewer’s suggestion we used the weekdays of admission as an instrumental variable for time to surgery. First, we checked that day-of-the-week admission was related with time-to-surgery but not with in-hospital mortality. Second, we estimated the impact of delay in a baseline probit model without the day-of-admission. Finally, we replicated this previous model using day-of-the-week admission as an instrumental variable. The results of these analyses show that correction of the possible endogeneity of the surgery delay variable by means of the day-of-the-week instrumental variable does not alter the study’s conclusions.

➔ We include this analysis in the web annexe with the corresponding references in the main text of the manuscript.

[Other editorial requirements #01]. Tables: Please ensure that the order in which your tables are cited is the same as the order in which they are provided. Every table must be cited in the text, using Arabic numerals. Please do not use ranges when listing tables. Tables must not be subdivided, or contain tables within tables. Please note that we are unable to display vertical lines or text within tables, no display merged cells: please re-layout your table without these elements. Tables should be formatted using the Table tool in your word processor. Please ensure the table title is above the table and the legend is below the table. For more information, see the instructions for authors on the journal website.

➔ Tables have been re-structured in accordance with the editorial requirements.

[Other editorial requirements #02]. Figure cropping: It is important for the final layout of the manuscript that the figures are cropped as closely as possible to minimise white space around the image. For more information, see the instructions for authors: http://www.biomedcentral.com/info/ifora/figures.
Figures have been revised in accordance with the editorial requirements.

[Other editorial requirements #03]. Please also ensure that your revised manuscript conforms to the journal style (http://www.biomedcentral.com/info/ifora/medicine_journals). It is important that your files are correctly formatted.

We have revised the entire manuscript in accordance with the journal’s style.

We very much appreciate the comments made by the referees, which we believe have been extremely useful in improving our manuscript.

Thank you for your interest in our paper. Please do not hesitate to contact us for any further clarification.

Yours sincerely,

Salvador Peiró, MD PhD
On behalf of all the authors