Author's response to reviews

Title: World Health Organization fracture risk assessment tool in the assessment of fractures after falls in hospital

Authors:

Shin-ichi Toyabe (toyabe@med.niigata-u.ac.jp)

Version: 2 Date: 29 September 2009

Author's response to reviews: see over
Editor,

BMC Health Services Research

September 29, 2009

Dear Editor,

Thank you very much for giving us an opportunity to revise our manuscript entitled “World Health Organization fracture risk assessment tool in the assessment of fractures after falls in hospital” (manuscript ID: MS: 4102538282896851), which was previously submitted to BMC Health Services Research. We have revised the manuscript in accordance with the comments by the referees and have also provided point-by-point replies to questions from the referees in this letter. All of the corrections in the manuscript have been underlined.

Sincerely yours,

Shin-ichi Toyabe, MD.
Crisis Management Office,
Niigata University Medical and Dental Hospital,
Asahimachi-Dori 1-754, Chuo-Ku, Niigata City 951-8520, Japan
TEL: +81-25-227-0375
FAX: +81-25-227-0375
E-mail: toyabe@med.niigata-u.ac.jp
To reviewer 1

**Major Compulsory Revisions #1:** This was a retrospective study of over 20,000 patients aged 40-90 years admitted to a single hospital over 3 years. The participant characteristics are not described in adequate detail. One assumes that most are elderly, but the study includes patients as young as 40 years. More information should be provided including whether acute or elective admissions were included.

**Our reply:** We added more information on the study subjects to the ‘Methods’ section of our manuscript. The new information includes admission type (acute or elective), gender and age of the patients.

**Major Compulsory Revision #2:** How would the two risk tools perform in a much more selected population that is at higher risk e.g. over the age of 65 or 70 years? This is a suggestion made by the author in the discussion and it should be possible to do it within the current dataset.

**Our reply:** We analyzed risk factors for fallers and for fractures after falls in the subjects who were more than 65 years old, and the results are shown in Table 5. The results were almost same as those in the case of subjects who were aged from 40 to 90 years.

**Major Compulsory Revision #3:** The paper concludes that STRATIFY is reasonably good at predicting falls and FRAX is reasonably good at predicting fractures in the fallers, but the data aren’t analysed to see how the combination would work in practice. For example, it appears that FRAX probabilities have only been calculated in the fallers.
As most of the fractures occurred on the first fall, the need for FRAX is negated. How would a combination of STRATIFY and FRAX work in the whole study population at admission (or better still in a preselected older segment of the population)?

**Our reply:** We appreciate your expert advice. According to your advice, both STRATIFY score and FRAX score were entered into a multiple logistic model and multivariate Cox proportional hazards model. The results are shown in Table 4C and 4D. Both STRATIFY score and FRAX score were calculated at admission not only for fallers but for all patients, and the scores were calculated again when fall events are occurred. The assessment to calculate these two scores was performed for all patients at admission, and the assessment was done again when the fall events occurred. This information was added to the ‘Methods’ section.

**Minor Essential Revisions #1:** The FRAX output is consistently referred to as a score but are more properly called 10-year fracture probabilities. It is not clear whether the numbers refer to major fracture or hip fracture probabilities alone.

**Our reply:** We used ten-year probability of major osteoporotic fracture according to body mass index. We added the information to the “Methods” section.
To reviewer 2

Revision #1: Introduction. The authors correctly point out that falls in hospitals are problematic, that the use of falls risk screening tools is common, but that the consequences of falls are rarely considered in this. I agree with the authors to this point however would encourage them to consider strengthening their argument for the importance of this by highlighting that approximately 2/3rd of falls do not result in physical injury and the costs attributable to falls are heavily skewed towards those that result in fracture. (Minor essential revision)

Our reply: Thank you for your helpful advice. According to your advice, we revised the ‘Introduction’ section.

Revision #2: Method. The methods section is very brief. A sub-heading for study design is needed. Greater description of the hospital, the types of wards it has and the community it serves is needed. A subheading and specific description of the reliability and validity of the FRAX and STRATIFY is needed. Justification should also be provided as to why the STRATIFY was selected. Two systematic reviews have examined this instrument (Haines et al 2007, Oliver et al 2008) and have found the STRATIFY to have poor predictive accuracy. Indeed, Oliver who created the STRATIFY tool now actively discourages its use. This should be recognised in the present manuscript. (Major essential revision)

Our reply: We enriched the description in the ‘Methods’ section and used subheadings for the study design and risk assessment tools for falls and fractures. We added statements on poor predictive power for falls to the ‘Methods’ and ‘Discussion’ sections.
Revision #3: Analysis. If FRAX and STRATIFY were completed only on admission, why was there need to model time-dependent covariates in Cox models? No mention is made of the ROC analyses that are later presented. Surely there was missing data in this study. How was this handled? (Minor essential revision)

Our reply: The risk assessment for falls and fractures was performed at admission of the patients and was done again when the fall event occurred. Values of several risk factors might change at the second or later assessment. For example, the factor ‘history of falls’ must be positive at the second or later assessment. We added this information to the ‘Methods’ section. An explanation of ROC analysis and an explanation of treatment of missing data were also added to the ‘Methods’ section.

Revision #4: Results. I am not sure why the predictive accuracy of the STRATIFY is being presented in the results as this does not illuminate what is novel about this particular study. If the authors wish to retain this aspect of the paper, this should be justified with a research aim and appropriate review of relevant literature (such as the systematic reviews mentioned above). They also need to consider the accuracy of this instrument in terms of its ability to predict the event rate of falls as the limitations of using standard sensitivity and specificity calculations in this field have previously been articulated [Journal of Clinical Epidemiology 2006;59(11):1155-1161.] (Minor essential revision)

Our reply: Our aims are (i) to analyze risk factors for fractures after falls and (ii) to examine the performance of the STRATIFY tool in a Japanese setting. We added this information to the ‘Introduction’ section. We also added information on the limitation of
using the standard approach for calculating sensitivity and specificity to the
‘Discussion’ section.

Revision #5: Results. My preference would be to replace the term monovariate with univariate, and multivariate logistic with multiple logistic. (Discretionary revision)
Our reply: We replaced the term monovariate with univariate and replaced the term multiple logistic analysis with multiple logistic analyses.

Revision #6: Results. A FRAX cut point of 10 was employed in some analyses, though it was not specified whether this cut point was specified a priori. The dangers of “searching” screening instruments with multiple possible cut points for the cut point that optimises the predictive accuracy of the instrument have previously been presented in this field in the systematic review by Haines in 2007 mentioned previously which was cited by the authors. My concern is that the results presented may be over-optimistic. Normally a cut point is established using a development dataset, and is then tested on a separate validation dataset. (Minor essential revision)
Our reply: We appreciate your expert advice. According to your advice, we established cut-off values for STRATIFY and FRAX scores using a development dataset and tested the cut-off values using a validation dataset based on the 10-fold cross validation method. We added this information to the ‘Methods’ and ‘Results’ sections.

Revision #7: Results. The formatting of tables needs addressing eg. use of capitals, specify regression coefficient instead of B. Also, use of terms such as exp(B) in the text, although as a researcher I understand what this means, are not helpful for clinicians who
are inexperienced with statistics. (Discretionary revision)

**Our reply:** We revised the formatting of tables and corrected statistical terms in the text according to your advice.

**Revision #8:** Results. I would also like to see more raw data presented rather than just regression tables. (Discretionary revision)

**Our reply:** We added raw data to Table 1 and Table 3.

**Revision #9:** Discussion. Some of the grammar needs review eg. low PPV is considerable implications…(Minor essential)

**Our reply:** Our manuscript was proofread again by a native English speaker.

**Revision #10:** Discussion. A problem with this study is that there were only 18 fractures in this study, thus there are considerable problems with statistical power for many of the risk factors examined. My feeling is that this is the reason why cohort studies focussing on this topic are scant. The point raised in the second paragraph is essentially that hospitals should employ a fracture prevention program instead of just a falls prevention program. I would tend to agree though I feel the researchers need to be more specific about how they would envisage the role of instruments such as FRAX and STRATIFY in such an approach. The authors also need to consider the implications of such an approach on nursing workload and the likely acceptance of staff to complete yet another screening tool. The authors did not consider in this study whether nursing staff clinical judgement could be as accurate as the two instruments considered for the purpose of falls prediction and fracture risk. Surely these instruments would need to be superior to
simple clinical judgement prior to a recommendation that they should be introduced into clinical practice (what is the point of adding an instrument if it does not provide better information than what is easily available through clinical judgement). (Minor essential)

**Our reply:** The problem of small sample size in terms of fracture after falls was discussed as a limitation of our study in the ‘Discussion’ section. Thank you for your comment on the role of risk assessment tools, which increases nursing workload but might have inferior performance compared with simple clinical judgement. We discussed this point in the ‘Discussion’ section.

**Revision #11: Discussion.** The comment that reduction in LOS might be an effective strategy in Japan appears somewhat unusual. First, even if patients were discharged sooner, what is to stop them from falling and fracturing once they have left the hospital (ie. censoring just means we do not see the fracture, yet it still occurs)? Second, reducing LOS is easier said than done and many elaborate health service programs have failed in this end. The number of fractures in the study (and subsequent width of confidence intervals) needs to be listed and discussed as a limitation. (Major essential)

**Our reply:** We appreciate your valuable comment. We deleted the description concerning reduction of LOS. Actually, LOS turned out to be not significant in multivariate analyses when we corrected errors of coding and values of some variables and performed the analyses again.