**Author's response to reviews**

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

**Authors:**

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

**Version:** 3 **Date:** 5 February 2010

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

BMC Health Services Research

February 5, 2010

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

Minor essential revision: A number of spelling errors remain and the legend for Table 3 is incorrect (refers to falls rather than fracture after falls).

Our reply: Thank you for your comments. Our manuscript was checked again by a native English speaker, and we corrected the error in the legend for Table 3.
To reviewer 2

**Major essential revision #1:** The 10 fold cross-validation method is listed in the analysis approach with reference given to a text on genetic micro-array analysis. No description is provided for the reader of what this approach actually entails, and it is difficult to see from the results section how employment of this technique has changed the results. The authors must fully describe this technique and how they have applied this to their dataset. In my initial review I raised the issue of development and validation datasets in relation to scanning for cut-off points leading to overly-optimistic results. As the STRATIFY has pre-established cut-off points that were applied this concern is not relevant to this instrument. Rather it is for the FRAX instrument where I am concerned that scanning for cut-off points has occurred as it has not previously been employed in this setting for this purpose. It appears from the methodology that the cut-off for this has been selected based on reviewing the ROC curve. Hence a validation dataset that is independent of the dataset that was used to develop this cut-off point is required as results generated from the dataset from which this cut-off point was selected are likely to be overoptimistic.

**Our reply:** We have described in detail the 10-fold cross-validation method and how we applied this method to our dataset in the Methods section.

**Major essential revision #2:** The authors now state that the STRATIFY has poor accuracy, but I question the standard by which this appraisal has been made. The AUC for the STRATIFY in predicting fallers was higher than that of the FRAX in predicting fractures, so I am afraid that you cannot conclude on one hand that the STRATIFY was
poor in predicting fallers while at the same time saying the FRAX is suitable for use in clinical practice. In my initial review I highlighted the results in other studies that were poorer than the initial study and the published opinion of the author who created this instrument. I did so because I wanted the authors to justify why they selected this instrument for investigation in the first place. I did not want this to influence the authors’ interpretation of their own results and their conclusions. Previous authors have argued that instruments with imperfect results are of no use though this is not always the case [Journals of Gerontology: Medical Sciences. 63(5):543]. Recent economic modelling has demonstrated that instruments with predictive accuracy at similar levels to that observed for the STRATIFY in the present study are still capable of making falls prevention programs operate more efficiently [Medical Care. 2009;47(4):448-56]. In light of this information, the authors need to reconsider how they have appraised both the STRATIFY and the FRAX and they need to be consistent in their interpretation of results.

**Our reply:** We appreciate your expert advice. Certainly, performance of the STRATIFY tool for predicting inpatient falls in our study (sensitivity = 0.67, specificity = 0.75, Youden index = 0.42) was almost the same as the results of previous research. We read arguments for and against the usefulness of the STRATIFY tool [Journal of Gerontology 63(5):543] and reconsider the usefulness of the STRATIFY tool in a Japanese hospital setting. As you kindly advised us, the accuracy of judgement by using the STRATIFY tool in a Japanese hospital setting was equally high relative to the accuracy of clinical judgement of physiotherapists [Med Care 47(4):448]. We corrected the conclusion regarding the usefulness of the STRATIFY tool.
Minor revision: Discussion section – the authors suggest a program of providing hip protectors yet there is evidence that many hospital patients may be non-compliant with this intervention [Age & Ageing. 2006 35(5):520-523]. This should at least be acknowledged as a limitation of this approach.

Our reply: We added information on the problem of compliance with hip protectors to the Discussion section.

Written English: Several of the revised sections require grammatical review. For example “Most of fallers did not suffer from injuries, but only 2.8% of the fallers suffered peripheral fractures in our study” should instead read “A majority of patients who fell did not injure themselves and only 2.8% of patients who fell incurred a peripheral fracture”. I would suggest again a thorough proof reading of this document.

Our reply: Thank you for your suggestion. According to your suggestion, our manuscript was proofread again by a native English speaker.