Author's response to reviews

Title: Development and Validation of a Questionnaire for Analyzing Real-life Falls in Long-term Care Captured on Video

Authors:

Yijian Yang (yijiany@sfu.ca)
Rebecca Schonnop (rcs10@sfu.ca)
Fabio Feldman (Fabio.Feldman@fraserhealth.ca)
Stephen N Robinovitch (steve@sfu.ca)

Version: 3 Date: 2 April 2013

Author's response to reviews: see over
Response to Reviewers:

BMC Geriatrics MS: 1184294381796400: “Development and Validation of a Questionnaire for Analyzing Real-life Falls in Long-term Care Captured on Video”

Comments from the Editors (Prof. Ian Cameron)

1. Structure: Please check the instructions for authors on the journal website to ensure that your manuscript follows the correct structure for this journal and article type.

Response: We have revised our manuscript structure to ensure compatibility with the instructions for authors on the journal website, and have added sections on “Keywords,” “List of abbreviations,” and “Competing interests.”

2. Competing interests: Manuscripts should include a Competing interests section.

Response: We have added “Competing interests” on page 14 as follows:

The authors declared that they have no competing interests.

3. Figure titles: All figures must have a figure title listed after the references in the manuscript file. The figure file should not include the title or number (e.g. Figure 1... etc.). The figures are numbered automatically in the order in which they are uploaded.

Response: We have now added a figure title (for our single figure) in the manuscript file.

4. Figure cropping: It is important for the final layout of the manuscript that the figures are cropped as closely as possible to minimize white space around the image.

Response: In the version uploaded, we have carefully cropped the figure to minimize the white space.

Comments from Reviewers:

Reviewer's report

Title: Development and Validation of a Questionnaire for Analyzing Real-life Falls in Long-term Care Captured on Video
Version: 2 Date: 23 January 2013
Reviewer: Clemens C. Becker
Reviewer's report:

The submitted article is relevant and adds information to a recent publication in the Lancet. In this article the authors report a new approach of video footage analysis in long-term care. The
submitted article describes the video rater process in more detail and presents data on reliability, retest and inter-rater reliability in particular. The paper at this stage has several shortcomings and should be revised.

1. General remark and wording: the authors make very general comments on the causality of falls and interventions that might be designed and tested based upon their findings and material. It should be clear from the abstract that this material is limited to certain situations and sites in long-term care and by no means represents the full spectrum of fall situation of older persons. The review will mark some of these statements. It is clear that for most of the falls there was no or little clinical information on the observed individuals.

Response: We agree that our abstract should clarify that our results are limited to certain situations and sites in long-term care. We now provide these details in our revision (page 2, Abstract) as follows.

In the Background section of the Abstract:

Falls are the number one cause of injuries in older adults, and are particularly common in LTC. Lack of objective evidence on the mechanisms of falls in this setting is a major barrier to prevention. Video capture of real-life falls can help to address this barrier, if valid tools are available for data analysis.

In the Method section of the Abstract:

Over three years, we video-captured 221 falls experienced by 130 individuals in common areas (e.g., dining rooms, hallways, and lounges) of two LTC facilities.

In the Conclusion section of the Abstract:

Our results provide strong evidence of the reliability of the FVAQ for classifying biomechanical, behavioural, situational, and environmental aspects of falls captured on video in common areas in LTC. Application of this tool should reveal new and important strategies for the prevention and treatment of falls and fall-related injuries in this setting.

2. The term cause of a fall (causality) is misleading as the intrinsic component can only be guesstimated at best.

Response: We agree, and have now replaced the term “cause of fall” with “biomechanical cause of imbalance” throughout. We also clarify, in first presenting the FVAQ (page 6), that video analysis cannot reveal the intrinsic (physiological) contributions to the cause of the fall, as follows:

While falls result from interactions between physiological (intrinsic), environmental, and situational factors, video analysis itself cannot reveal physiological causes of falls (or the intentions of the faller). Instead, the FVAQ provides meaningful categorization of
biomechanical features that may be important to consider, along with clinical data, in improving our understanding of the cause and prevention of falls.

We also add the following statement in the last paragraph of the Introduction (Page 4, paragraph 1) to better establish the value of combining the FVAQ with clinical data to guide prevention strategies:

Video technology provides a means for capturing footage of real-life falls in high-risk environments such as LTC [14-16], and providing information on the biomechanical and situational aspects of falls in these settings. This information can complement clinical data (on disease diagnoses, medications, and functional status) in revealing the mechanisms of falls, and in designing and selecting prevention efforts at a population or individual level.

3. The authors have the privilege to have access to this material. It is also clear that most countries will not allow similar studies or routines in long-term care because of the current legislation and other restrictions. This means that the findings and the hypothesis derived from the material will be hard or impossible to test under different circumstances. This does not reduce the value of the material but it highlights that the information and approach should be phrased with modesty.

Response: We agree these are important considerations, which we now address in the limitations section of the Discussion (page 15, paragraph 1) as follows:

We recognize that currently, there is limited partnering between researchers and care providers in LTC for video capture of falls. We hope that our model for data collection and analysis facilitates growth in the applications of this tool to LTC and other high risk settings, such as hospitals or senior centres [15, 40].

4. Introduction: this section is well written and balanced. There is only one question. The authors cite the sideway hip fracture paper (ref. 26). That seems to be an assumption by the original authors. Why could this not be a backward fall with turning during the fall to the side?

Response: We agree that, like other studies, Greenspan’s study (reference 28) did not separately examine the initial fall direction and the landing configuration, or recognize the probability of rotation (change in fall direction) during descent. We clarify our approach on page 8, paragraph 1 as follows:

We considered initial fall direction separately from body configuration at landing, to account for body rotation during descent.

Methods:
5. The presented paper excludes the resting and recovery phase of the fall, why?

Response: We agree that post-fall behaviour is important, including the ability to rise after falling (Noury et al., 2008). However, as observed in our videos, for the vast majority of falls in
common areas in LTC, residents are assisted by care staff to rise after falling. As such, a study of post-fall behaviour in this setting enters the domain of patient-care provider interactions, beyond the scope of our current study. We now discuss these issues in the Methods (page 6 - 7) as follows:

We did not consider post-fall behaviour, such as the ability to rise after falling (Noury et al., 2008), since preliminary viewing of videos indicated that, for the vast majority of falls in common areas in LTC, residents are assisted by care staff to rise after falling. As such, a study of post-fall behaviour in this setting enters the domain of patient-care provider interactions, beyond the scope of our current study.

We also discuss the opportunity for future expansion of the questionnaire to examine post-fall behaviour, in the limitations section of the Discussion (page 15, paragraph 1):

Further “analysis packages” may build on the core template provided by the FVAQ, to probe issues such as pre-fall or post-fall behaviour, additional aspects of balance recovery or fall protective responses, or questions of known or suspected relevance to specific clinical subgroups or environments. Additional iterations should be based on a consensus process between researchers and stakeholders to agree on the right questions and response categories, and establish acceptable approaches for data collection and linking to health information.

6. The paper (and the Lancet paper) do not report the consequences of the falls this limits the findings.

Response: Our questionnaire was based on fall characteristics that can be observed from the videos. We did include one question related to fall consequences—“Perceived site of greatest injury risk/impact severity,” which showed only moderate reliability (agreement). We now clarify in the Methods (on page 7, paragraph 1) that we did not focus on the consequences of falls.

7. Why did the authors analyze 15 falls? What was the rationale to choose 15 and not more (power analysis etc.)? The author should list the selected falls according to their classification. With such a low number certain fall types might have been missed or only 1-2 examples might have been analyzed. Please, list situations and selection bias.

Response: We now clarify in the Methods (on page 9, paragraph 2) that our sample size was based on published guidelines for observer agreement studies (Freedman et al., 1993), as follows:

Our sample size was based on published guidelines for observer agreement studies [35]. We estimated a priori that (for a given question) the average percentage of agreement between the two teams would be 85 percent (or 15 percent disagreement). In order to detect a desired 90% confidence interval of between 0 and 30 percent disagreement, we calculated a minimal required sample of 15 observations.
We agree our sample size made it likely that we would have zero or low representation within specific categories of responses for some questions. We have clarified this issue by adding a table (Table 3) that lists (for each group) the number of responses for each category from nine key questions.

We also address the issue of selection bias on page 9, paragraph 1, as follows:

15 fall videos were selected randomly (using a random number generator to minimize bias) from our database.

Results:

8. The term cause of the fall and activity at the time of the fall are misleading. Although this is quite often used a fall is the interaction between the person, the environment and the exposure of a person in a given environment. The term activity is also misleading. A walking activity is most of the time driving by a goal such as walking to a dining hall. Different walking or transfer activities can have very different circumstances. This contextual information are likely to be correctly recorded in a post fall interview (e.g. walking to, or during a risk factor assessment prior to the fall).

Response: We agree falls result from complex interactions between physiological, environment and situational factors, and that activity at the time of falling can be considered at various levels (action, intention, movement style or quality). We now address these issues in the Methods (on page 6), as follows:

While falls result from interactions between physiological (intrinsic), environmental, and situational factors, video analysis itself cannot reveal physiological causes of falls (or the intentions of the faller). Instead, the FVAQ provides meaningful categorization of biomechanical features that may be important to consider, along with clinical data, in improving our understanding of the cause and prevention of falls.

As discussed in our response to comment 2 above, we now replaced the term “cause of fall” with “biomechanical cause of imbalance” throughout the manuscript.

We also clarify our approach to classifying the activity at the time of falling (on page 7, paragraph 2) as follows:

We classified activity at the time of falling into general categories, without consideration of the intent of the action (e.g., “walking”, as opposed to “walking to the dining room”).

9. Please specify if the findings on fall direction, stepping responses and impact to specific body sites where also possible with the cameras with a low frame rate (e.g. 5/min)? It is not clear how the frame rate and the camera position influenced the rater process.

Response: We agree on the value of providing additional details on how video quality may have affected reliability. We now clarify (Page 4, paragraph 2) that all cameras recorded at 15 frames per second (an earlier version of our data collection system included cameras at lower
frame rates, but these have since been upgraded to 15 frames per second). We also now describe the number of video frames between fall initiation and impact in the Methods (on page 9, paragraph 1), as follows:

The time interval between fall initiation (loss of balance) to fall impact ranged from about 700 ms (for a rapid trip) to 3000 ms (for a fall related to incorrect weight shifting); and the corresponding number of video frames ranged from 10 -45.

We also comment on the issue of video quality in the Discussion, on page 13, paragraph 1 as follows:

Teams rated their probability in selecting the correct answer between 84 - 100% (depending on the question), reflecting their strong confidence, and the adequacy of our video collection techniques, in identifying key features of the fall (barring significant occlusion of body parts from the camera view, which did not occur).

And on page 13, paragraph 2 as follows:

In completing the FVAQ, the team often faced challenges related to camera resolution, distance between the faller and the camera. In each case, only a single camera recorded the fall. Clearly, improvements in the number and resolution of cameras should improve the reliability of most questions in the FVAQ. However, of the six poorly scoring questions, only one - type of footwear – was clearly related to video quality (e.g., distance between the faller and the camera).

10. For the less positive agreement it is again likely that some of the information can be extracted from reports such as footwear. For the aspects that cannot be captured from the reports such as contribution of clutter, reach-to-grasp responses and perceived site of injury risk please again comment on sampling rates and camera positions.

Response: We now provide a more well-rounded discussion on the potential causes (including video quality and camera position) for poor reliability on specific questions of the VFAQ (including clutter, reach-to-grasp responses and perceived site of injury risk) on page 13, paragraph 2, as follows:

In completing the FVAQ, the team often faced challenges related to camera resolution, distance between the faller and the camera. In each case, only a single camera recorded the fall. Clearly, improvements in the number and resolution of cameras should improve the reliability of most questions in the FVAQ. However, of the six poorly scoring questions, only one - type of footwear – was clearly related to video quality (e.g., distance between the faller and the camera). More complex challenges arose for other items, which might be addressed through refinements to the questions and/or instruction manual for improved clarity. For example, we observed poor reliability for contribution of clutter in causing the fall. This may more relate to the ambiguity in our definition of clutter, or the challenge of attributing casual links between falls and environmental features [25], aside from cases of obvious trips over obstacles (which made up only 13% of our sample).
We observed moderate reliability for site of greatest perceived injury risk/impact severity. This may relate to difficulties in judging the injury potential of impacts to multiple body sites (on average, impact was reported to occur to 4 body sites). Reach-to-grasp responses showed good agreement but only a moderate Kappa value, perhaps due to its low frequency of occurrence creating a high probability for chance agreement [37].

11. The authors mention that definitions and examples for each category were provided in a comprehensive instruction manual. Is this accessible? The training process can only be poorly be judged if this not the case.

Response: We agree, and have now provided a short version of our FVAQ instruction manual as a supplementary document with this manuscript.

12. The phenotypic causes: “trip/stumble”, “slip”, “incorrect transfer/shift of body weight,” “collapse/loss of consciousness” and “loss of support with external object” and hit and bump should be discussed in a consensus process. It is likely that this is not a full list or derived from the camera placement perspective. A fall on a staircase, for instance, would not fit into this category.

Response: We agree with the reviewer on these issues, and refer to our above responses (to comments 2 and 8) concerning our categorization of the cause of falls.

We also agree on the value of highlighting the need for a consensus process on these issues, which we describe in the Discussion (on page 15, paragraph 1), as follows:

We hope that our model for data collection and analysis facilitates growth in the applications of this tool to LTC and other high risk settings, such as hospitals or senior centres [15, 41]. Further “analysis packages” may build on the core template provided by the FVAQ, to probe issues such as pre-fall or post-fall behaviour, additional aspects of balance recovery or fall protective responses, or questions of known or suspected relevance to specific clinical subgroups or environments. Additional iterations should be based on a consensus process between researchers and stakeholders to agree on the right questions and response categories, and establish acceptable approaches for data collection and linking to health information.

Concerning the issue of falls on stairs, we now explain in the Methods (on page 4, paragraph 2):

In both facilities, no stairs were located in the areas accessible to residents.

We also state in the Discussion (on page 14, paragraph 3) that:

[We designed the FVAQ to focus on the situational and environmental context of falls in common areas of two LTC facilities (e.g., hallways, dining rooms, and living rooms). Accordingly, it may not capture the range of mechanisms of falls in bedrooms, bathrooms, and stairways, or among healthier older adults living in the community.
13. The activity list is not complete. Lying, for instance is missing.

Response: We now clarify in Table 3, and in the Methods (on page 7, paragraph 2) that our activity list did include lying (grouped with seated):

The most common reported activities leading to falls are walking, and transferring to or from a seated or lying position [9, 22, 24, 26, 27]. The FVAQ included these along with “standing” (Table 3)

14. The list of protective responses such as stepping, grasping, and feet-in-place swaying or upper limb “windmilling” is interesting. Again this should be discussed in consensus process.

Response: We agree that the set of balance recovery responses we explored was limited, and now highlight in the Discussion (on page 15, paragraph 1) the value of future work in this area, as follows:

Further “analysis packages” may build on the core template provided by the FVAQ, to probe issues such as pre-fall or post-fall behaviour, additional aspects of balance recovery or fall protective responses, or questions of known or suspected relevance to specific clinical subgroups or environments. Additional iterations should be based on a consensus process between researchers and stakeholders to agree on the right questions and response categories, and establish acceptable approaches for data collection and linking to health information.

15. Occurrence of contact to key body sites (head, pelvis, torso, hand/wrist, elbow/forearm, knee, and shoulder. It is likely that raters observed multiple impacts. How consistent were the findings of multiple vs. single impact?

Response: We agree this is worth reporting, and now address the issue of impacts to multiple body sites in the Results (on page 12, paragraph 2) as follows:

Most falls were reported to involve impacts to multiple body sites (head, torso, pelvis, knee, hand, elbow, and shoulder). In inter-rater testing, the mean number of impact sites was 4.0 (SD = 1.9) for one team, and 4.2 (SD = 1.8) for the other, with positive correlation between teams in the number of impacting sites ($R^2 = 0.84$; $p<0.001$).

Discussion:
16. The wording and the structure of discussion section is superior to the abstract and the wording of some other sections. Most limitations are mentioned. This could lead the rephrasing of the abstract.

Response: We thank the reviewer for these suggestions, and have rewritten key aspects of the Abstract to mirror the text of the Discussion.
Reviewer's report

Title: Development and Validation of a Questionnaire for Analyzing Real-life Falls in Long-term Care Captured on Video

Version: 2 Date: 4 December 2012
Reviewer: Geoff Murray

Reviewer's report:

Major Compulsory Revisions:

1. The main measures of this study are the inter-rater and intra-rater reliability. Co-author SNR trained the observers using an instruction manual. Clearly the training and instruction manual are significant factors impacting on inter-rater and intra-rater reliability. For example the clutter had poor reliability. Is the poor reliability because the instruction manual did not clearly define clutter or is it for some other reason? I do think the authors should have included some discussion on factors that impacted on inter-rater and intra-rater reliability as these are the main numeric outcomes of the study.

Response: We agree on the value of providing additional details on the factors that may have impacted inter-rater and intra-rater reliability of items in our questionnaire. We now include our instruction manual as a supplementary document to the manuscript. Furthermore, we pay increased attention in the Discussion to the value of refining the wording of our questions and definitions, and improvements in video quality (on page 13, paragraph 2):

In completing the FVAQ, the team often faced challenges related to camera resolution, distance between the faller and the camera. In each case, only a single camera recorded the fall. Clearly, improvements in the number and resolution of cameras should improve the reliability of most questions in the FVAQ. However, of the six poorly scoring questions, only one - type of footwear – was clearly related to video quality (e.g., distance between the faller and the camera). More complex challenges arose for other items, which might be addressed through refinements to the questions and/or instruction manual for improved clarity. For example, we observed poor reliability for contribution of clutter in causing the fall. This may more relate to the ambiguity in our definition of clutter, or the challenge of attributing casual links between falls and environmental features [25], aside from cases of obvious trips over obstacles (which made up only 13% of our sample). We observed moderate reliability for site of greatest perceived injury risk/impact severity. This may relate to difficulties in judging the injury potential of impacts to multiple body sites (on average, impact was reported to occur to 4 body sites). Reach-to-grasp responses showed good agreement but only a moderate Kappa value, perhaps due to its low frequency of occurrence creating a high probability for chance agreement [37].

2. The questionnaire included an estimate of what the observer thought was the probability of his/her answer being correct, but as far as I can see this data was not included in the results. Fairly obviously if the observers rated their probability of being correct as low, then the
likelihood of good inter-rater and intra-rater reliability would be less. I do think the authors should include this data in Table 2 and their discussion.

Response: We agree with this suggestion, and have now included mean probabilities in selecting the correct answer to Table 2. We also report correlations between agreement and probability. We add the following text to the Inter-Rater Reliability section of the Results (on page 10, paragraph 3):

The mean probability reported by teams in selecting the correct answer ranged from 84% - 100% for one team, and from 90% - 100% for the other team. There was significant correlation between agreement in responses and probability in the answer being correct ($R^2 = 0.37; p=0.001$).

We also add the following to the Intra-Rater Reliability section of the Results (on page 11, paragraph 2):

The mean probability reported by teams in selecting the correct answers ranged from 90% - 100% for the baseline analysis, and from 85% - 100% for the repeat analysis. Again, there was significant correlation between agreement in responses and probability in the answer being correct ($R^2 = 0.31; p=0.005$).

Finally, we add the following to the Discussion (on page 13, paragraph 1):

Teams rated their probability in selecting the correct answer between 84 - 100% (depending on the question), reflecting their strong confidence, and the adequacy of our video collection techniques, in identifying key features of the fall (barring significant occlusion of body parts from the camera view, which did not occur). A significant correlation existed between agreement and probability, although probability explained only 37% and 31% of the variance in inter-rater and intra-rater agreement, respectively.

Minor Essential Revisions: Nil

Discretionary Revisions:

1. I think the authors should state the level of care received by participants: i.e. low level, high level or mixed. They should also provide base line data such as the average age of participants, M:F ratio.

Response: We did not have access to the health records for all participants in this study, as not all of them have provided consent. We have provided information regarding the age, gender, and other characteristics in our revision (page 5, paragraph 2) as follows:

Resident characteristics. Residents of New Vista had an average age of 81 years (SD=13), and 67% were female. Residents at Delta View had an average age of 82 years (SD=10), and 61% were female. Among the 15 participants included in this study, the mean age was 82 years (SD=12), and 47% (n=7) were women. As described previously
[16], among residents captured falling who provided us with consent to access their health records, 34% had Alzheimer's disease, 13% had diabetes, 31% had hypertension, 19% had stroke, and 6% had Parkinson's disease. These prevalence data were similar to those observed among fallers not captured on video, and to the overall profile of residents at the two LTC facilities.

2. If possible I think the authors could state the percentage of falls that were captured as a proportion of the total falls in the facilities during the period of study (A question in my mind is whether this model of analysis of falls is generalizable to other settings such as bathrooms)

Response: We agree and now provide this information (on page 5, paragraph 1) as follows:

In 2010 at Delta View, 45% of falls occurred in common areas, of which 65% were captured on video. In 2010 at New Vista, 34% of falls occurred in common areas, of which we captured 28% on video.

We also address the applicability of our analysis to other settings in the Discussion (on page 14, paragraph 3) as follows:

[W]e designed the FVAQ to focus on the situational and environmental context of falls in common areas of two LTC facilities (e.g., hallways, dining rooms, and living rooms). Accordingly, it may not capture the range of mechanisms of falls in bedrooms, bathrooms, and stairways, or among healthier older adults living in the community.

3. I would like to know more information about the cameras used - were they infra red cameras for example for resolution of images recorded at night (In this regard I was wondering how lighting was actually assessed). The cameras recorded at 4 to 15 frames a second. Can the authors state the average number of frames that were analysed in each fall, and perhaps the range? How was the inter-rater and intra-rater reliability affected by the number of frames?

Response: We now clarify (Page 4, paragraph 2) that all cameras recorded at 15 frames per second (an earlier version of our data collection system included cameras at lower frame rates, but these have since been upgraded to 15 frames per second). We also now describe the number of video frames between fall initiation and impact in the Methods (on page 9, paragraph 1), as follows:

The time interval between fall initiation (loss of balance) to fall impact ranged from about 700 ms (for a rapid trip) to 3000 ms (for a fall related to incorrect weight shifting); and the corresponding number of video frames ranged from 10 -45.

We also comment on the issue of video quality in the Discussion, on page 13, paragraph 1 as follows:

Teams rated their probability in selecting the correct answer between 84 - 100% (depending on the question), reflecting their strong confidence, and the adequacy of our
video collection techniques, in identifying key features of the fall (barring significant occlusion of body parts from the camera view, which did not occur).

And on page 13, paragraph 2 as follows:

In completing the FVAQ, the team often faced challenges related to camera resolution, distance between the faller and the camera. In each case, only a single camera recorded the fall. Clearly, improvements in the number and resolution of cameras should improve the reliability of most questions in the FVAQ. However, of the six poorly scoring questions, only one - type of footwear – was clearly related to video quality (e.g., distance between the faller and the camera).

4. I would also like to know a little more about circumstances of filming e.g. what was the average distance from the camera to where the fall occurred, and range of distances and perhaps how filming circumstances related to inter-rater and intra-rater reliability.

Response: We also now clarify there was a reasonably even distribution of falls occurring near, mid-distance, and far from the camera (Page 9, paragraph 1). The potential effect of camera distance on reliability is addressed in our response to comment 3 above.

Four (27%) of the falls occurred relatively close to the camera, 5 (33%) occurred at a far distance, and 6 (40%) occurred at a moderate distance.

5. The authors examined 3 stages of falls - initiation, descent and impact. Presumably they had opportunity to also examine observable factors before initiation of the fall e.g was the resident's gait or stance unsteady, where they transferring without an aid that may have been present. The authors also presumably had opportunity to observe factors after the impact stage e.g how long before help arrived, did the residents expose themselves to risk by attempting to rise without assistance. The authors could consider this in their discussion. I noted that to some degree the authors acknowledged this limitation in the second last paragraph.

Response: We agree of the value of these suggestions, and have added the following text to the Discussion (on page 15, paragraph 1):

Further “analysis packages” may build on the core template provided by the FVAQ, to probe issues such as pre-fall or post-fall behaviour, additional aspects of balance recovery or fall protective responses, or questions of known or suspected relevance to specific clinical subgroups or environments. Additional iterations should be based on a consensus process between researchers and stakeholders to agree on the right questions and response categories, and establish acceptable approaches for data collection and linking to health information.

We also add text to the Methods (at the bottom of page 6), explaining our rationale for restricting our current focus to fall initiation to impact, as follows:
We did not consider post-fall behaviour, such as the ability to rise after falling (Noury et al., 2008), since preliminary viewing of videos indicated that, for the vast majority of falls in common areas in LTC, residents are assisted by care staff to rise after falling. As such, a study of post-fall behaviour in this setting enters the domain of patient-care provider interactions, beyond the scope of our current study.
Reviewer's report

Title: Development and Validation of a Questionnaire for Analyzing Real-life Falls in Long-term Care Captured on Video

Version: 2 Date: 15 January 2013

Reviewer: Jian Sheng Chen

Reviewer's report:

1. This validation study of a fall video analysis questionnaire (FVAQ) concludes that the FVAQ is a useful tool for revealing fall circumstances and has good reliability (reproducibility). However, validity of the FVAQ is also important but has not been assessed in the study. For example, perceived site of the greatest risk for injury based on a video reading might not be correct. This could lead to wrong conclusion in a fall-related injury study. A good fall assessment tool should have both good reliability and good validity. Also, the reliability in this study may have been inflated by calculating the Kappa coefficient based on video readings from two three-member teams rather than two individual observers. Moreover, the results might be influenced by quality of the videos.

Response: We agree on the importance of discussing how our study was limited to examining internal reliability, and not external validity. We now address this issue at the beginning of the limitations section of the Discussion (on page 14, paragraph 3) as follows:

We focused on assessing the internal reliability (reproducibility of results) of the FVAQ. Additional studies are required to examine external validity, for example by relating FVAQ responses to data from fall incident reports, observed injuries, risk for future falls, and the nature of future falls.

We also clarify our rationale for our team (versus individual observer) approach to analysis on page 6, as follows:

We designed the FVAQ to be completed by a team of evaluators, to reduce the biases inherent in individual evaluators and allow interdisciplinary perspectives [18].

Furthermore, we now comment on the issue of video quality in the Discussion, on page 13, paragraph 1 as follows:

Teams rated their probability in selecting the correct answer between 84 - 100% (depending on the question), reflecting their strong confidence, and the adequacy of our video collection techniques, in identifying key features of the fall (barring significant occlusion of body parts from the camera view, which did not occur).

And on page 13, paragraph 2 as follows:
In completing the FVAQ, the team often faced challenges related to camera resolution, distance between the faller and the camera. In each case, only a single camera recorded the fall. Clearly, improvements in the number and resolution of cameras should improve the reliability of most questions in the FVAQ. However, of the six poorly scoring questions, only one - type of footwear – was clearly related to video quality (e.g., distance between the faller and the camera).

Major Compulsory Revisions
2. It will be good if this study can be carried out based on readings from individuals on good quality videos with comprehensive fall incident reports. So, the results of reliability and validity (against fall incident report) are directly related to the FVAQ and can be used to compare with other fall assessment tools. Results from readings of two three-member teams may be used to demonstrate ways of improving the reliability.

Response: We agree on the need to comment on these issues. As described in our response to comment 1 above, we now describe in the Discussion (on page 14, paragraph 3) that our study was limited to examining internal reliability, and not external validity:

We focused on assessing the internal reliability (reproducibility of results) of the FVAQ. Additional studies are required to examine external validity, for example by relating FVAQ responses to data from fall incident reports, observed injuries, risk for future falls, and the nature of future falls.

We also discuss the need for a consensus process to refine and expand the application of the FVAQ in fall prevention (on page 15, paragraph 1) as follows:

Further “analysis packages” may build on the core template provided by the FVAQ, to probe issues such as pre-fall or post-fall behaviour, additional aspects of balance recovery or fall protective responses, or questions of known or suspected relevance to specific clinical subgroups or environments. Additional iterations should be based on a consensus process between researchers and stakeholders to agree on the right questions and response categories, and establish acceptable approaches for data collection and linking to health information.

Minor Essential Revisions
3. The authors should consider assessing impacts of video quality by comparing readings from poor to good quality videos.

Response: We agree on the value of providing additional details on how video quality may have affected reliability. We now clarify (Page 4, paragraph 2) that all cameras recorded at 15 frames per second (an earlier version of our data collection system included cameras at lower frame rates, but these have since been upgraded to 15 frames per second). We also now describe the number of video frames between fall initiation and impact in the Methods (on page 9, paragraph 1), as follows:
The time interval between fall initiation (loss of balance) to fall impact ranged from about 700 ms (for a rapid trip) to 3000 ms (for a fall related to incorrect weight shifting); and the corresponding number of video frames ranged from 10 -45.

We also comment on the issue of video quality in the Discussion, on page 13, paragraph 1 as follows:

Teams rated their probability in selecting the correct answer between 84 - 100% (depending on the question), reflecting their strong confidence, and the adequacy of our video collection techniques, in identifying key features of the fall (barring significant occlusion of body parts from the camera view, which did not occur).

And on page 13, paragraph 2 as follows:

In completing the FVAQ, the team often faced challenges related to camera resolution, distance between the faller and the camera. In each case, only a single camera recorded the fall. Clearly, improvements in the number and resolution of cameras should improve the reliability of most questions in the FVAQ. However, of the six poorly scoring questions, only one - type of footwear – was clearly related to video quality (e.g., distance between the faller and the camera).

Discretionary Revisions
4. A good example of the FVAQ’s application in the discussion should add weights to the paper.

Response: We agree with this suggestion, and have now expanded our discussion of the application of our FVAQ (in page 14, paragraph 2), as follows:

The 24 questions on the FVAQ probe previously hidden aspects of falls and contribute new information to guide fall prevention efforts. For example, information on the biomechanical causes of imbalance and activities leading to falls (both of which exhibited strong reliability) helps to guide improved fall risk assessment and balance training protocols, along with efforts to reduce environmental hazards and create safer movement environments [16]. Information on fall severity (impacting body parts) can provide insight on injury mechanisms and help guide the design of protective padding (e.g., hip protectors [38]) and compliant “safety” flooring [39]. Attempts to prevent or lessen the injury potential of the fall (through balance recover by stepping, or arresting the fall with the upper limbs) are important neurological markers, that may also help in guiding exercise-based fall injury prevention programs.
Reviewer's report

Title: Development and Validation of a Questionnaire for Analyzing Real-life Falls in Long-term Care Captured on Video

Version: 2 Date: 19 January 2013

Reviewer: Doungkamol Sindhusake

Reviewer's report:
No revisions are required.