Reviewer’s report

Title: Health-related quality of life after vertebral or hip fracture: a seven-year follow-up study

Version: 1 Date: 10 March 2009

Reviewer: Lisa Lix

Reviewer’s report:

This manuscript provides a wealth of information about an observational study to investigate the effect of fracture on health-related quality of life (HRQOL). The objectives of the paper are well-defined and the manuscript is generally well written. My main concerns with the manuscript lie in the choices of statistical methods and the discussion of study strengths and limitations.

Major Compulsory Revisions

1. As Figure 1 demonstrates, there were 95 study participants with a vertebral or hip fracture at baseline, 91 study participants at the two-year follow-up, and 67 at the seven-year follow-up, which corresponds to a loss to follow-up of 29.5% of the original study cohort. The authors need to carefully consider the mechanism by which drop-out occurred and provide comparisons of differences between study completers and drop-outs. They should also consider the mechanisms by which missing data may occur, as described by Little and Rubin (2002): missing completely at random, missing at random, and missing not at random. If the missing data pattern in the current study is missing not at random, then the analyses of HRQOL change will be biased.

2. The authors used repeated-measures ANOVA to test for change over time, but only for those individuals who participated in all three waves of data collection. There are a number of limitations associated with this approach to the analysis. First, this analysis makes stringent assumptions about the covariance structure of the repeated measurements, which are unlikely to be satisfied in many applications. The model cannot accommodate time-varying covariates, which may be associated with HRQOL and should be controlled for in the analysis of change over time. A random-effects regression models is a flexible method for the analysis of data characterized by missing values, increasing variability across measurement occasions, and covariates that change with time. Furthermore, the analysis should not be limited just to those individuals who completed all three waves of data collection; a random-effects model can accommodate missingness.

3. The use of Pearson correlation coefficients (see Table 4) assumes a normal distribution of responses and a linear association between variables. How did the authors evaluate these assumptions? Isn’t comorbidity measured using a dichotomous scale? If so, then a Pearson correlation, which assumes an interval scale, is not an appropriate choice.
4. The authors conducted multiple linear regression analyses to identify predictors of each of eight HRQOL domains. The investigated predictors were balance, hand grip, spinal deformity index, age, bone mineral density, physical activity, new comorbidity, and fall frequency. These analyses are a form of data mining, in which the authors are searching through the data to identify statistically significant predictors. Unless the sample size is large, this approach to modeling may result in overfitting and a failure of the model to validate on a new set of study participants. I recommend that the authors make substantial changes to this section by: (a) focusing on a smaller set of explanatory variables that are justified based on prior theory/research, and (b) validating their analysis results. Given that sample size is small, a cross-validation analysis cannot be conducted. Therefore, the authors should investigate the use of an empirical re-sampling technique, such as the bootstrap, to conduct the validation.

5. The authors used a one-way ANOVA to test for differences in the adjusted means for the vertebral and hip fracture groups. The utility of this analysis for inferring differences in the populations from which these samples were drawn is limited because the authors have not controlled for covariates such as co-morbidity and recent new fracture, which may differ between the two groups and may be confounded with the dependent variable.

6. The Bonferroni method, which was adopted to control the familywise Type I error rate to $\#$, is not the optimal approach when the tests that comprise the family are correlated. It will have lower power than alternate procedures that account for this correlation. Further information can be found in Blair et al. (1996).

7. The analysis of longitudinal HRQL is challenged by the potential for study participants to experience response shift over time. Response shift is defined as “a change in the meaning of one’s self-evaluation of a target construct as a result of: (a) change in the respondent’s internal standards of measurement (i.e., scale recalibration), (b) change in the respondent’s valuation of component domains constituting the target construct, or (c) redefinition of the target construct (i.e., reconceptualization)”. (Schwarz & Sprangers, 1999). The authors should discuss the implications of response shift for their interpretation of study results.

Minor Essential Revisions

8. The first section of the Results, which describes patient characteristics, should be shortened substantially. It repeats much of the information already reported in Table 1. For example, in the sentence that begins “At baseline, 24/42 (57%) in the vertebral group…” the frequencies do not need to be reported in addition to the percentages.

9. The interpretation of the p-values in Tables 2 and 3 is not clear to me. Does the p-value correspond to a test of the omnibus, or overall hypothesis? If so, this should be stated in the table note.

Discretionary Revisions
10. In the enumeration of study strengths and limitations, the authors should discuss their rationale for not including an osteoporosis-specific measure of HRQOL, such as the Osteoporosis Quality of Life Measure (Badia et al., 1997) in the data collection for the vertebral and hip fracture groups.

References

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

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests.