Reviewer's report

Title: Dynamic relationship between multiple START assessments and violent incidents over time: a prospective cohort study

Version: 3  Date: 18 July 2014

Reviewer: Harry Kennedy

Reviewer's report:

MANUSCRIPT # 3447157311361252

Research article
Dynamic relationship between multiple START assessments and violent incidents over time: a prospective cohort study
Richard Whittington, Johan Håkon Bjørngaard, Andrew Brown, Rajan Nathan, Stephen Noblellt and Beverley Quinn
BMC Psychiatry (Submitted: 2014-05-23)

1. Is the question posed by the authors well defined?
   Yes. This is an interesting and important general point concerning all forms of risk assessment in medicine and psychiatry.

2. Are the methods appropriate and well described?
   The methods are appropriate but could be described in greater detail and with greater clarity – as outlined below. There is also a problem with under-powering of the sample which is in itself interesting from a methods point of view, since the calculation of power for studies such as this is not well established.

3. Are the data sound?
   The data are sound, though the study is under-powered.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
   Yes.

5. Are the discussion and conclusions well balanced and adequately supported by the data?
   I have set out some points that should be acknowledged in the discussion and in the abstract. In particular there are some interpretative points that should be balanced as suggested below.

6. Are limitations of the work clearly stated?
   I have set out some additional issues that should be acknowledged as limitations.
7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?

Yes.

8. Do the title and abstract accurately convey what has been found?

The abstract needs to be more balanced in its interpretation of the results.

9. Is the writing acceptable?

Yes.

MAJOR COMPULSORY REVISIONS

Page 2

53 Conclusion: risk factors which are considered to have the capacity to fluctuate dynamically

54 did not do so substantially in this group of mentally disordered offenders.

Comment: This may be overly pessimistic. The authors found that the START strengths and vulnerabilities scores predicted violence. They have also demonstrated that protective factors were better predictors than vulnerability factors at the individual level. The authors found that in this group of forensic in-patients, there was little change in these risk factors over time.

DISCRETIONARY REVISIONS

Page 9

211v It should be noted that the sample size in Figure 1 decreases as the
START number

212 increases since relatively few patients had a large number of assessments. The sample size

213 at the 1st. and 10th. assessments for Vulnerabilities and Strengths respectively were 45/48

214 and 18/20. Scores beyond the 20th. assessment are based on two and then only one patient.

Comment: Was the number included in the regression analyses confined to those who were present throughout the period analysed? For example, the analysis should be confined to the first ten analyses, and only the 45 for whom data on vulnerabilities at each time point are available should be included, and likewise only the 18 for whom data on strengths are available. If all the data available at all time points is to be used, then a survival analysis such as Cox regression would be more appropriate, using censored data.

215 Those who had been discharged by the study end-date (January 2011) had a marked

216 reduction in their Vulnerabilities score (mean= -3.9; 95%CI -7.9, +0.2)

Comment: this is not a statistically significant difference since the 95% CI
overlaps zero. It would be more accurate to say “…had a marked trend toward reduction.”

226 A standard single-point ROC analysis of the predictive validity of the first START assessment and subsequent involvement in violence was conducted. Four patients were excluded from this analysis as their first incident occurred prior to their first START assessment. Twenty two (48%) of the remaining 46 patients were violent at least once subsequent to their first START assessment. The average (mean) period between the 230 assessment and the violence was 231 days (sd 199 days, range 2-677 days). The AUC for Vulnerabilities was 0.69 (CI 0.52-0.85, p=0.037) and that for Strengths (reversed for prediction of non-violence) was 0.75 (CI 0.59-0.89, p=0.005).

Comment: have the authors considered as a secondary analysis, calculating the ROC for the START immediately prior to each of the 22 first violent incidents, using the baseline START for the remaining 24 who were not violent? This might be expected to produce a significantly higher ROC if the START is dynamically responsive in a useful way. ROC scores can be compared for statistical difference using a programme published by the Pinel institute. At the most basic, was there any evidence of an increase in START vulnerability scores or decrease in START strengths at the last assessment prior to the first violent incident? This is a relatively simple analysis but would be a useful lead in to the next section.

240 It can be seen that 4 patients here are relatively high risk at their first assessment (Vulnerability score #25) and one is relatively low risk (score ~10). Patients 3 and 4 change rapidly and by their last assessment are scoring below 10 whilst Patients 1, 2 and 4 remain relatively stable in their risk level, even in some cases over 10 assessments.

Comment: this appears to show that the START is responsive to change, when there is change to detect. The problem with the study appears to be that too few of the patients actually had any change. This is a ‘power’ problem though it may also arise from the design of the study. For example most in-patient violence occurs very early in an admission, and very seldom towards the end of an admission, even in forensic units. The observations are weighted towards ‘late phase’ assessments and behaviours.

251 Table 2 presents the associations between changes in START scores and
the risk of a subsequent violent incident. In the total sample a ten point increase in START Vulnerabilities score was associated with increased risk (odds ratio, OR) of 3.1 (95% CI, 1.47-7.46). Within patients the degree of increased risk was lower with an OR of 1.77, but this estimate is made with low precision due to reduced power (95% CI, 0.56-5.54). The START Strengths score was not associated with the risk of a subsequent violent incident in the total sample (OR=0.91, 95% CI, 0.34-2.47). In the within patient analysis, there was an increased risk of a violent incident (OR=2.26, 95% CI, 0.38-13.42), but the confidence intervals were again wide in this fixed effect model due to the small sub-sample.

Comment: please use the methods section to explain in greater detail what observations are being used here, and what the comparators are. Comment: please state clearly that none of these odds ratios are statistically significant.

Also, in this setting, the nurses making the assessments would participate in discharge discussions as part of the multidisciplinary team but would not make the final discharge decision.

Comment: this is an important methodological point and should have been stated in the methods section – describe who made the assessments, whether the same person made serial assessments, and similarly how decisions to discharge were made, and by whom. For example, patients detained in England and Wales under the Mental Health Act can be discharged by the treating consultant psychiatrist, but those detained subject to restrictions on discharge (imposed by the courts for more serious or persistent offenders) can only be discharged by a Mental Health Review Tribunal or with the permission of the Minister for Justice.

The key finding here is that risk factors which are considered to have the capacity to fluctuate dynamically did not do so substantially in this group of mentally disordered offenders.

Comment: this is not the only explanation. Lines 240-243 above actually show that the START was sensitive to change in the few patients who had dynamic fluctuations in risk. Most patients were stable. It would be more accurate to say...
“did so only in a few patients, and did not do so in most of the patients included in this group…….”

285…..The ICC values for both domains were approaching 1, indicating that very high
286 levels of variance in scores could be attributed to stability within patients

Comment: it would be more accurate to say that very high proportions of the variance in scores could be attributed to stability within patients. The problem is that there was too little variance.

344 Figure 1: mean START score at each assessment

Comment: since there are reducing numbers included at each time point, and since the more stable, recovering patients are lost at each successive time point due to discharge, it is not surprising that the vulnerability mean score rises steeply towards the latter part of this time series. A plot of serial measures over time for a constant cohort should be given – e.g. those for whom all ten of the first ten assessments were available.

349 Figure 2: individual variation in risk factor stability

Comment: this is very informative.

359 Table 2. Odds ratio with 95% confidence interval for a violent episode after START
360 assessment (<31 days after assessment) in the whole sample and within patients (patient
361 fixed-effect models).

Comment: the column headed “within patient assessment” needs to be explained much more clearly, both within the methods section and in the labelling of the table.

REFERENCES

Comment: these are not in the correct format for the journal. Please see instructions for authors.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests.