Reviewer's report

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

**Version:** 3  
**Date:** 17 July 2014  
**Reviewer:** Jay P Singh

**Reviewer’s report:**

***Major Compulsory Revision**  
**Minor Essential Revision**  
*Discretionary Revision*

**ABSTRACT**

- **Please define the acronym “START”**.
- **Using “n” to signify both the number of assessments as well as the number of violent incidents is hard to follow.**
- **Please standardize whether after colons the first word is capitalized or not (e.g., “Dynamic” versus “routine”).**
- **Numbers 10 and over should be expressed as numerals (e.g., “ten” to “10”)**
- **Please provide p-values for odds ratios.**
- **Please standardize whether “strengths” and “vulnerabilities” are capitalized or not.**
- **Please add a second sentence to the “Conclusion” explaining the main implication(s) of this finding**

**INTRODUCTION**

- ***The suggestion that “risk status” does not answer the “when?” question conflicts with static risk assessment tools such as the VRAG, Static-99, and other well-known actuarial schemes that provide risk estimates for specific time periods.**
- **The phrase “substantial empirical literature” is up to conjecture. In a field where literally hundreds of articles are published a year, having five articles published on a tool is not many.**

**METHOD**

- **Numbers 9 and below should be expressed as numerals (e.g., “1” to “one”).**
- ***What was the rationale for 14-items having been the threshold of the number of START items that needed to be completed?**
- ***The amount of missing data and prorating required suggests that the findings of the study, while clinically relevant, may not be definitive for the START.**
• ***Please discuss the START in terms of its unique features (key and critical factors, T.H.R.E.A.T assessment, prediction of multiple risks, summary risk rating).

• **Please ensure that all acronyms are defined upon first use (e.g., ROC).

• ***As lengths of follow-up seem to have varied, single-point ROC curve analyses were not appropriate. An alternative method that accounts for such variability (e.g., Cox regression or survival ROC curve analysis) may be beneficial.

• ***START Vulnerability and Strength scores, according to the tool’s authors, are to be used to guide the development of a summary risk rating rather than to predict outcomes in and of themselves. However, I do not see any mention of summary risk ratings (i.e., Low, Moderate, High). Hence, the clinical applicability of the analyses must be brought into question, especially the ROC curve analyses.

• **Please include standard reporting information such as the directionality of the statistical tests and whether such tests were one- or two-tailed (as well as the alpha level).

RESULTS

• **Please include at least one decimal place for percentages.

• **Please include dispersion parameters for medians (e.g., inter-quartile ranges).

• **Why are there periods after “1st.”, “10th.”, and “20th”?

• **Please provide proper reporting of statistical tests rather than just reporting the statistical significance level (e.g., “=0.018”).

• *Please hyphenate “Twenty two”

• *Please use either “mean” or “average”

• **Please standardize whether standard deviation will be reported as “SD” or “sd”

• **Please standardize whether “95%” will be reported along with “CI”

• ***Why were 10-point increases/decreases used in the logistic regression analyses? And as time at risk varied, Cox regression rather than logistic regression would have been a more appropriate method.

DISCUSSION AND CONCLUSION

• ***I am not able to comprehensively comment on this section, as the statistical analyses used to support the arguments are currently not reliable.

• **Single sentence paragraphs should be avoided.

• *Please make use of subheadings to ease reading.

• *The odds ratios and AUCs found should arguably be compared to those of other tools whose predictive validity has been meta-analytically investigated.

REFERENCES
• **Please include DOI numbers, where available.

FIGURES
• ***As the number of individuals on whom the START was administered changes dramatically from the 1st to the 25th assessment, this figure’s graphic display means little. Statistical adjustment for the number of individuals included needs to be incorporated if this figure is to remain and be interpreted.

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

Quality of written English: Needs some language corrections before being published

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

Declaration of competing interests:
I am the founder and CEO of the Global Institute of Forensic Research, which specializes in the development of risk assessment executive research digests, continuing education in forensic mental health and corrections, and software solutions.