PubPeer.com Post – October 23, 2015

Major Problems with Swartout, Koss et al. (2015),

"Trajectory Analysis of the Campus Serial Rapist Assumption,"

***JAMA Pediatrics,* published online July 13, 2015.**

Post by Jim Hopper, PhD

Investigation and analysis by me (Jim Hopper, PhD) and another researcher hired as an independent consultant (Allison Tracy, PhD) have revealed that this paper has major problems that call into question its scientific validity:

1. **Unreported problems that invalidate the latent class trajectory analyses,** e.g., their model is under-identified and under-powered, and has unacceptably low rates of classification certainty (as low as 51% for the "derivation" dataset and 73% for the "validation" dataset).
2. **False statement about their analyses.** They claim in the paper that they did not use the (minimal) senior year data in their latent trajectory analysis, when in fact they did, as can be seen from their own Mplus code and by comparing its results with those reported in the paper. This inclusion was not an oversight: the output from that model matches the fit statistics given in the article (but not perfectly), and the plots produced by the syntax match the graphs in the article's figure. And the impact of including the additional data is not trivial: omitting the senior year data resulted in statistically significant evidence that the method used to handle missing data was inappropriate and may have biased the results. In addition, when the senior year data actually were omitted, not only was the model (again) under-identified and under-powered, but the resulting latent class trajectories no longer included a "decreasing" group.
3. **Problems with the integrity and validity of their data,** including (a) incorrect values resulting from recoding values for "missing" and "no response" into "never [raped]," (b) about 110 erroneous missing values for the sophomore year rape variable and about 70 erroneous missing values for the junior year rape variable, and (c) other mismatches between the data produced by Swartout's own syntax and the analysis dataset he used and provided. These problems are much more extensive, and much more significant, than what is mentioned in the correction published on October 19, 2015.

Also, the authors operationalized "serial rape" in questionable ways that (even if their model and data were valid) call into questions their conclusions, including:

1. **Exclusion of data on attempted rape.** The difference between completed and attempted rape is often just a matter of luck. Attempted rape is still a crime and experiences of attempted rape can be quite traumatic to victims.

2. **Exclusion of data on the frequency of rapes** (and attempted rapes) self-reported by student participants at each time period. For example, does it really make sense to define someone who reports having committed more than 2 rapes, or even more than 5 rapes (within an 8- or 12-month period), as not a serial rapist?
3. **Use of latent class trajectory analysis** to address what they refer to as the "serial rapist assumption," which is actually two simple propositions: the majority of rapists are serial rapists, and the vast majority of rapes are committed by serial rapists.

We have prepared comprehensive documentation to support and flesh out the claims above:

1. An Initial Critical Response to Swartout et al.'s (2015) paper in *JAMA Pediatrics,* "A Trajectory Analysis of the Campus Serial Rapist Assumption," by Jim Hopper, David Lisak, & Allison Tracy. This document includes:
    - An introduction, by Jim Hopper and David Lisak, which provides important contextual information, including about how the paper has been promoted by the authors and their colleagues, and about this critique.
    - A presentation by Jim Hopper, "What Does the 'Derivation Dataset' Used in Swartout et al. 2015 Tell Us About Serial Rape? Simple Frequency Analyses." This shows the massive contradiction between what Swartout and colleagues claim about their data and analyses, on the one hand, and what their data actually reveal, on the other.
    - A document by Allison Tracy, "Methodological Critique: Executive Summary," which explains the major problems with Swartout and colleagues' latent class analyses, and with the validity and integrity of their "derivation dataset." Its appendix is a statement of peer review of Dr. Tracy's work, provided by another researcher (Hanno Petras, PhD) who has extensive experience with latent class growth models.
    - Note: I have bundled these into a single document because the introduction and its contextual information are critical, because the three sub-documents are best understood together, and to make it easy for people to share the entire document via email and other channels.
2. Technical Report of Methodological Critique of Swartout et al., 2015, by Allison Tracy. This comprehensive (75-page) report, including extensive appendices containing statistical software coding syntax and analysis outputs, documents in great detail all of the analyses conducted by Dr. Tracy and all of her findings.
3. Mplus code used to conduct the analyses reported by Swartout and colleagues in their paper (use mouse to save zip file), provided by Kevin Swartout to Jim Hopper and Allison Tracy.
4. Analysis dataset used by Swartout (use mouse to save zip file), provided by Swartout to Hopper and Tracy.

5. This link to the publicly available, full version of the dataset used to generate the dichotomous variables of the "derivation dataset" employed in the analyses for the paper. (Notes: The dataset is "DS2: Male Data." The Codebook includes mislabeled X_____ variables [starting on page 512 of the PDF], such that acts the male research participants committed AGAINST females are incorrectly described as acts committed against them BY other males. Also, as can be seen on page 512 and in the final line of recode commands at the top of page 513, "no response" [0] and "missing" [9] values were recoded as "never" [1] having committed the act.)
6. SPSS coding syntax used for Dr. Hopper's straightforward and simple frequency analyses on the "derivation" dataset (use mouse to save sps file).
7. Letter to the editor of JAMA Pediatrics, by Jim Hopper, David Lisak and Allison Tracy. This is the letter submitted to the editor. At this point we cannot know if the letter will be published, and if it is, whether *JAMA Pediatrics* will edit it first.