Reviewer's report

Title: Simultaneous evaluation of abstinence and relapse using a Markov chain model in smokers enrolled in a two-year randomized trial

Version: 1 Date: 13 February 2012

Reviewer: Kevin Delucchi

Reviewer's report:

This manuscript presents an interesting analysis of data from a smoking cessation trial. The use of Markov modeling is fairly novel and its use here is the main strength of the paper. The presentation is generally clear and technically correct. The data are sufficient in size and quality to support the findings and the discussion is reasonable and does not go beyond the results. There are several weak points to the manuscript, however.

It is not true that most smoking studies apply cross-sectional analysis to longitudinal data anymore. Where it is done, it is usually an appropriate application in that the research question involves a comparison of abstinence rates at a key assessment point. A second problem with the premise of this work is that in many studies the goal is to achieve abstinence. While relapse is common, it is often not the primary focus and once one knows the abstinence rates, obviously one knows the relapse rates. The authors should make it clear that this approach is not the same as (but may be complimentary to) a longitudinal mixed-effects model.

The paragraphs discussing missing data are correct but that information is widely known now and I’m not certain it is worth the space. Also, it is unclear why so much print is given this topic when the authors deleted subjects. While the percent of data deleted is not large, it runs counter to good statistical practice.

The analysis used “the backward elimination procedure.” What exactly is this? If it means removing variables based on an algorithm, such stepwise procedures are a poor choice. Frank Harrell’s text on regression spells out the issues which include the F test statistics do not have the claimed distribution, the standard errors of the parameter estimates are too small and p-values are too low, due to multiple comparisons.

The main problem I have with the analysis is that it is fairly constrained. Why did the authors not consider more complex sets of Markov model? An example of such an approach can be found in Delucchi and Weisner’s paper modeling the presence of problem drinking (Delucchi & Weisner, 2007). The authors also fail to cite other work using similar modeling in smoking (Hughes, Keely, Fagerstrom, & Callas, 2005; Martin, Velicer, & Fava, 1996; Velicer, Martin, & Collins, 1996). Finally, it is not clearly explained to the average reader what information this form of modeling provides over and above the more standard approach.

drinking: A 7-year, 5-wave study. Alcoholism-Clinical And Experimental Research, 31(6), 61A-61A.


**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