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: 28 February 2012

Reviewer: Peter R Killeen

Reviewer's report:

This research methodology is important and timely. The authors' first paragraph is right-on. As is their second. Critics will object to the memoryless quality of the Markov model (MM); but it is a more innocuous simplification than many others that are made by others. The focus on missingness, and how to handle it, is, i think, overdone; it is important how we impute status to drop-outs, of course. In the Ellerbeck study that provided the data for this analysis, significant differences vanish if it is assumed that those who are non-responders have reverted to smoking (a more reasonable imputation, imho). Only in the 'secondary analyses', with 'no imputation', were significant differences found. Thus, the nominal success or failure of a major study hinges on how one treats drop-outs.

Minor essential revisions.

page 4, do a better job of clarifying the kinds of missingness. (Do not rely on citations 3 and 4 to do that job for you).

Make contact with the work of CK Enders (e.g., 2006) on modern techniques for missing data imputation.

Make contact with the hoary Kaplan-Meier "missing at random" bias.

Eq 2, and the general treatment, is elegant and thorough.

Perhaps too much so; I would have preferred a model-comparison approach, with time-invariant transition probabilities (over all groups) as a "null hypothesis", and then evaluation of the information gained by recognizing the different groups. But i understand that is not the authors' tradition.

I recognize that the authors are married to the KanQuit data set; but i think they do their modeling an injustice, by restricting it to these data. Because of the liberality of the intervention, the intervention effects are weak. (Even in the best of cases, of course, smoking interventions are relatively weak). Nonetheless, nothing new is learned from all of the analyses. The authors should do a better job of resonating all of the points that they do make with prior discoveries--that is make better contact with the traditional literature. They should do themselves a favor and apply this modeling to data-sets with stronger effects (e.g., perhaps the Piper et al studies of varenicline, etc).

They should also make better contact with applications of MM to cessation data. For instance, they talk about the cost of interventions, but there are many good articles using MM to analyse the monetary benefits from successful cessation
programs, harm reduction, etc. (e.g., Hurley et al; Orme et al; etc etc etc). They certainly must make contact with the BENESCO model.

I have applied my MM using the average published parameters for individuals in cessation programs, and placebo controls, to the data from the minimal PM group and maximal HDM group (bottom rows of my Table 1). It does a surprisingly good job of predicting these data (with the exception of the 24 month PM group). I attach pictures. The authors, excellent statisticians, can take this farther than i, if they care to. (Of course, given the radical difference between validated cotinine levels and self reports (the former being half the latter) in the source data, much of this attention to detail may be mere sophistry).

Level of interest: An article of importance in its field

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 declare that i have no competing interests.

I have published an alternative model; but i think it complements rather than competes with the present article. That interest is transparent in my review.