Reviewer's report

Title: Latent variables and structural equation models to model longitudinal relationships: an illustration in nutritional epidemiology

Version: 1 Date: 22 January 2010

Reviewer: Emil Kupek

Reviewer's report:

The manuscript illustrates the use of structural equations modelling (SEM) to determine direct and indirect effects of baseline adiposity on adiposity change 2 years after. Adiposity is defined as latent variable based on 4 observed anthropometric variables which measure different aspects (locations) of body fat and are highly correlated. The mediating variable is a test score reflecting “cognitive restraint” with respect to eating behaviour. The test was applied both for baseline and follow-up evaluations.

The theme is relevant for nutritional epidemiology but needs some clarifications and additional statistical analysis to justify its conclusions. Overall, this is an interesting methodological issue for many researchers who deal with indirectly measured outcomes without a gold standard and want to assess a mediating effect in a longitudinal perspective.

Major Compulsory Revisions

1. Explain the advantages of adiposity as a latent variable as opposed to four observed variables IN STATISTICAL TERMS. This is a central idea of the paper and deserves a better elaboration than just saying “One way to take advantage of all these measurements is to combine them into an adiposity latent variable within a structural model. This approach avoids the drawbacks of either arbitrarily choosing a single adiposity measurement or performing separate analyses on each fat-mass indicator, which leads to multiple testing problems and interpretation difficulties when analyses are not sufficiently consistent.”

First, why would separate analyses be such a “drawback”? It is perfectly justified to analyse how each of 4 different adiposity measures change in time and how cognitive restraint influences these changes, remembering that the measures reflect (by and large) different locations of fat in the body. A table presenting the results with this standard regression approach compared to the latent variable analysis would provide some grounds to compare the two.

Second, it is not clear to me what was meant by “multiple testing problems”. The analysis presented is not a multi-group ANOVA comparison where such problems may arise (but there are ways of dealing with them). As long as your observed outcomes are not exchangeable, I see no multiple testing problem in the way it is usually defined.

Third, what was meant by the phrase “when analyses are not sufficiently
consistent” and what type of “interpretation difficulties” would it imply?

2. Maximum likelihood estimation assumes multivariate normality; departure from it may have large impact on standard errors and all the inferences based on these parameters. Did you test this assumption?

3. There are two basic components in your SEM: the factor analysis for latent adiposity (measurement model) and the relationship between the variables in the model (structural model). It is of great importance that the measurement model fits well before you can move to the inferences about the structural relationship. For example, what were the communalities, KMO and percentages of the variance explained by the latent factor for the measurement models? How did the transformation such as log affect these measures?

4. Did you try a measurement model for adiposity after 2 years instead of change scores? As far as I can see, the latter should be more variable than the former due to the multiplicity of factors influencing the change. Were there problems with the model converging when original follow-up measurements were used?

5. To justify the equality constrains on p.9, you can test formally whether their imposition reduced the deviance beyond a pre-specified level; can you show the results?

6. On p.9, RMSEA for males (0.16) is not very convincing. Can you show which part of the model (measurement or structural) did not fit well? Why don’t you show 95% confidence interval for RMSEA?

7. Did you try a sub-group analysis by baseline adiposity level, such as BMI<=20 versus BMI>20 or perhaps using BMI of 25 as a cut-off point? It seems reasonable to assume that cognitive restraint may influence adiposity differently for people who are overweight or obese compared to those who are not.

Minor Essential Revisions

1. In conclusion, both in abstract and in discussion, the statement that “The latent variable modeling approach enabled presentation of synthetic results rather than four parallel analyses...” should state exactly what was meant by “four parallel analyses”. I suppose it meant four separate multivariate regressions for the change in each anthropometric measure of adiposity but am not sure how the interclass correlations would be treated (e.g. multi-level analysis), whether stratification by sex would be maintained, etc.

2. On p.9, the phrase “closely linked” can mean many things; please specify (e.g. “highly correlated”).

3. On p.9, substitute “was” for “were” in the sentence “The model fit for the baseline measurements and their changes was only slightly modified...”.

4. On p.10, 2nd paragraph, “and” seems more appropriate than “or” in “As expected, adiposity or CRS changes...”

5. On p. 13, use “the” instead of “common” in “The indirect effect through CRS change is at least partly due to common regression to the mean...”

6. What was the effect size for the hypothesis of interest? What does it mean in practical terms?

7. In table 3, you can add t-values and flag statistically significant ones.

Discretionary Revisions

1. Title may sound better if the phrase “to model” is substituted by “for”.

2. On p.3, the phrase “in which the common problem is how to handle questionnaires” is too vague and probably meant “how to deal with psychometric properties of the questionnaires”.

3. On p.7, the phrase “the effect of baseline adiposity on CRS change was adjusted for baseline CRS and thus freed from the factors confounding the cross-sectional effect” seems too strong because you can never be sure of controlling all the confounders as the phrase implies. Even adjusting for the baseline values works only for the linear relationship between the baseline and follow-up. However, many test scores have a threshold indicating a clinically relevant change above that point, which makes the relationship a non-linear one. For example, it seems plausible that cognitive restraint exerts a clinically relevant effect on adiposity only above certain level of the test score. Was this relationship tested?

4. On the bottom of p.7, the word “deterministic” may well be dropped out. The interpretation of the direct effect which follows may be better suited for discussion.

5. On p.8, 2nd paragraph, you may drop “in practice”.

Level of interest: An article of importance in its field

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

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests.