Reviewer’s report

Title: Socioeconomic status, urbanization and risk behaviours in Mexican youth: an analysis of 3 cross-sectional surveys

Version: 3 Date: 3 February 2011

Reviewer: Isaac Rhew

Reviewer’s report:

This study examined whether population size of locality was associated with likelihood of drug use and sexual behaviors among youth living in Mexico. The authors also examined whether these associations differed according to household socioeconomic status. This study may have impressive findings with potentially significant public health implications. The authors found that youth residing in larger localities compared to youth from rural localities had a lower likelihood for smoking and alcohol consumption. Among youth from low SES households, urban residence was associated with a higher likelihood of sexual initiation and lower likelihood of condom use. This study had adequate sample size, and appeared to consist of a diverse sampling of communities. However, I found it difficult to follow the study’s methodologic approaches as described by the investigators particularly involving their sampling procedures for localities and households. The sampling approaches could have important implications for study interpretation and generalizability, and thus need to be clarified. There also were a number of minor issues that should be addressed.

Major revisions:

1. The sampling approach appears very complex. First, it would be helpful to provide a bit more information about the Opportunitades program and eligibility for it. From my reading of the text, eligibility for the program played an important role in sampling localities and households. For example, authors discuss that localities were selected from high poverty concentration areas for “small urban” and “rural” localities while localities from “urban zones” were chosen from “non-intervention” areas.

2. The authors should expand on the potential implications of their sampling procedures in the Discussion.

3. The complex survey commands should be described further. The authors should be clear about how sample weighting was employed and the rationale for doing so. For example, it would be helpful to briefly describe what types of localities and households were over-sampled and what the distribution of these localities and households is in the general population.

4. Although authors mention estimating ICCs, it is not explicitly clear whether they used multilevel models in their logistic regression analyses to account for clustering. They should clarify this.

5. Authors should present possible explanations for their findings in the
Discussion.

Minor revisions:

1. There are numerous grammatical errors. The manuscript should be edited carefully.
2. 3rd paragraph of the Background is confusing.
3. At end of Background, specify the direction of association that is hypothesized (e.g. greater locality size is associated with higher likelihood of alcohol, tobacco, and illicit drug use).
4. In Background, expand on the advantages of defining urbanization using locality size rather than standard urban-rural dichotomous definitions.
5. It would help readers if authors define “locality.”
6. The authors should state whether the localities in the sample were all geographically distinct, or if some were contiguous. If there were contiguous localities, it would be helpful to know the number (%).
7. Additional information on what indicators comprised the poverty score would be useful.
8. Authors should describe how they decided to define the four urbanization categories.
9. It would be helpful to have more information on how drug use outcomes were measured. In particular, how is “current” use defined? Was this how participants were asked? If so, individuals would answer this item in different ways and could lead to more noise in the data.
10. Authors should state why is an interaction between marital status and gender is included in the model.
11. From my reading, it appears that rural localities were only sampled in 2003. Thus, inclusion of year in the model would make those covariates collinear and the model would not be fit. Were other rural localities sampled in 2001? If so, this further underscores the need to clarify the sampling procedures.
12. In the Methods, the authors should explicitly state how the locality size categories were modeled in logistic regression models (e.g. indicator variables vs. grouped linear term).
13. Additional discussion of how the sampling might affect findings is warranted.
14. In Discussion, authors state that locality size association with health behavior is mediated by SES. I believe they mean moderated. If not, authors should provide support for mediation.

Discretionary revisions:

1. In the Background, authors present detailed cigarette and illicit drug prevalence estimates, but not alcohol. This could be useful since this was a major outcome of interest.
2. The authors could limit some of the early paragraphs in the Background, and
focus more on the rationale for conducting the current study.

3. Organizationally, perhaps it would be helpful to separate out the locality sampling and then describe household sampling for better flow and clarity in the text.

4. Individual demographic characteristics and associations with health behaviors were not specified as aims of the study; however, there is substantial text devoted to discussion of these. If these findings cannot be described in relation to locality size, then I would recommend limiting their discussion of these findings because the focus on locality size becomes lost.

Level of interest: An article of importance in its field

Quality of written English: Not suitable for publication unless extensively edited

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests.