Author’s response to reviews

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

Authors:

Juan Pablo Gutiérrez (jpolutier@insp.mx)
Erika E Atienzo (eeatienzo@insp.mx)

Version: 4 Date: 27 June 2011

Author’s response to reviews:

June 27, 2011

Melissa Norton, MD
Editor-in-Chief
Editorial Office,
BMC Public Health

Dear editorial board:

We are submitting the revised version of the manuscript number 2712974524208514: "Socioeconomic status, urbanicity and risk behaviours in Mexican youth: an analysis of three cross-sectional surveys". The reviewers’ constructive and insightful comments are greatly appreciated. Our point-by-point responses to the comments of the four reviewers are below.

We hope the revised manuscript will have approval for publication in BMC Public Health. We look forward to hearing from you soon,

Erika E. Atienzo, MS (Corresponding author)
e-mail: eeatienzo@insp.mx
Mailing address: Instituto Nacional de Salud Pública. Dirección de Salud Reproductiva.
Av. Universidad No. 655 Col. Santa María Ahuacatitlán. CP. 62100.
Cuernavaca Morelos, México.
Phone: +52 (777) 329 30 00 ext. 3113

Reviewer 1: Danielle Ompad
Reviewer’s report:
Summary:
The authors explore the relation between locality size, SES and risk behaviors among Mexican youth. Locality size was associated with tobacco and alcohol consumption for all SES groups. For condom use, locality size was associated with odds of condom use among the most impoverished youth in urban localities. This is a very interesting paper.

Comments:

Major Compulsory Revisions

1. There is some sense on p5 that rural is “bad” and urban is “good.” But as your results, and those of others, have shown the relation is more complex.
We have made some changes in the text and we believe that the sense of “bad” and “good” has been diminished.

2. You make the point the urbanization, as the process of increasing the urban share of the total populations, is an important focus as a potential exposure. However, you measure are measuring urbanicity – that is, the extent to which a place is currently urban. Despite the issue with terminology, the analysis is valuable.
We agree with the reviewer, and thus we have replaced the term “Urbanization” by “Urbanicity” where possible.

3. More detail on the household inclusion criteria is needed.
We have added more information regarding sampling procedures on page 8 and 9.

4. More detail on the components of the poverty index is needed. It is one of your main variables.
We have added more information and details regarding the poverty index on page 11.

5. Your definitions of locality are based on population size. Are these the standard definitions for the census? If so, please provide a reference. If not, why are you not also considering area (i.e., population density) rather than just population size.
In Mexico, the official definition for locality size is based on population size using a cut-off value of 2,500 inhabitants defined by the National Institute of Statistics, Geography & Informatics (INEGI), which is the Census Bureau in Mexico. We have included a new reference in the text regarding this in the background section (page 6) and in Data management section on page 11.

6. Did you need to use sampling weights for the analyses?
No, we did not use sampling weights.

7. The third paragraph on p10 is not helpful. Draw parallels between studies in the discussion if helpful.
In the revised version we have removed the sentence: “The present work aims to
analyze the difference in risk associated to locality size; however it is worth saying that the data reported herein is consistent with that of prior publications”.

8. Figure 2 replicates table 2. Please remove.

We agree with the reviewer in that Figure 2 replicates data fully described in Table 2. We did this as we believe that a figure that graphically illustrates the association between the principal variables explored may be an interesting contribution for some readers, especially if we consider that there are no restrictions regarding the extension of the article in this Journal. However, taking into consideration the reviewer comment, we have decided to add Figure 2 as a supplementary material, thus only the most interested readers will still have access to it.

9. You make the point that married individuals are less likely to use condoms with comment on p13. I do not think you are implying that this would be risky behavior, but you should make that explicit.

We have eliminated this sentence from the discussion section.

10. I am not sure what you mean on p14 when you say, “However, it should be noted that adolescents’ localities are clearly not influenced by behaviors in general.” I think you are saying that adolescents’ behaviors do not shape their communities. However, previous research suggests otherwise. Please clarify what you mean here.

We mean that locality size, or the number of inhabitants in a locality, is not influenced by behaviours; rather, the direction of such relation goes in the opposite direction. However, in order to avoid confusion to the readers we have deleted such sentence.

11. The findings are interesting, but the manuscript would be stronger if you could comment more on the potential mechanisms for the associations.

We have added new text to the manuscript to discuss this (see Discussion section, page 16).

Minor Essential Revisions

12. Some help is needed with English grammar.

Attending reviewers’ comments, we have revised and edited the text.

13. It is unclear in the background (p6) whether there is no universal definition of urban in Mexico or worldwide. There should be a reference either way.

We edited the sentence in text and now it is read “No universal definition exists in terms of what an urban area is, and definitions vary among countries and over time.” We added a reference [REF # 23] for this sentence (page 6).

14. In the background (p6), you provide a definition for rural localities in Mexico, but not for others.

The National Institute of Statistics, Geography & Informatics (INEGI) of Mexico
uses an official cut-off value of 2,500 inhabitants for defining rural or urban localities. We have edited the text on page 6 and now in reads “In Mexico, the National Institute of Statistics, Geography and Informatics (INEGI) classifies localities with less than 2,500 inhabitants as rural, and the rest are classified as urban. Among urban localities, there is a further categorisation, with small-urban localities defined as up to 50,000 and metropolitan areas as having more than one million inhabitants [24].

15. Table 1 is overwhelming because there are so many numbers. Consider presenting the demographics and risk behaviors in separate tables.

We have edited Table 1 and removed confidence intervals, intra-cluster coefficients and p-values for the comparison between groups. We also replaced the original “Total” column that was separated by Male-Female with the total values in the sample. We decided to keep the original table and attach it as a supplementary material (additional file 1) for the interested readers.

Reviewer 2: Eduardo Lazcano-Ponce

Reviewer's report:

It present a manuscript through three nested surveys in the social program "opportunities" that fight poverty in Mexico, it describes the socio-economic status, urbanization and risk behavior among adolescents. The information may be useful to characterize adolescents living in marginalized areas of Mexico, however, the document has several methodological issues that need to clarify and / or modify.

Specific comments:

The introduction is extremely long and the description of methods and results are very sketchy.

We have somewhat reduced the extension of the background and improved the description of the Methodology section, particularly in aspects related to sampling procedures.

Methods:

3 were applied to population-based surveys should clarify, they are comparable in terms of:

1. Were the same questionnaires applied?

The questionnaires applied were not identical as they were particularly designed to each population (rural, semi-urban and urban); however they are very comparable. Taking into consideration this comment, we edited the next sentence on page 10 (first paragraph): “For all three surveys, data were collected at the household level using comparable questionnaires. In addition to the questionnaire for household informants, young household members were interviewed in relation to their education, labour characteristics, and health-related behaviours. For this analysis, we only used variables measured in the same way for the three surveys.”
2. Was the same type of sample population performed?
No. The sampling procedures varied according to the surveyed population (rural, semi-urban or urban). We have improved the information regarding sampling procedures for each survey. Such information can be found on page 8 of the manuscript.

3. Is the analysis adjusted for the weights that were used in the sample?
No, we did not use sampling weights.

4. Are the age groups similar?
Yes. The three surveys collected information from the same age group.

5. The authors did not state what was the primary objective of such surveys.
We have included on page 8 (first paragraph) of the manuscript the following sentence: “The surveys were part of an evaluation of a national social development program called Oportunidades (formerly Progresa), a conditional cash transfer program in Mexico. Oportunidades began in 1997 in rural areas of Mexico and was extended to small urban areas in 2001 and to large urban areas in 2002. Several rounds of surveys have been implemented since then to gather data regarding health, education, socioeconomic aspects and living conditions to evaluate the program.”

The main problem identified in the manuscript is the definition of major risk variables:
# It is not defined time exposure to alcohol and tobacco consumption, and creates problems in comparing the data with recent surveys of smoking and youth, the national health survey and the survey of addictions and other population-based Mexico.

We have included on page 10 the operationalisation of risk variables assessed: “The dependent variables are risk behaviours. Adolescents were asked if they had smoked at any time, if they currently smoked, and if they currently drank alcohol even on an occasional basis. Regarding sexual behaviour, they were asked for their age at first intercourse, which was then used to classify adolescents who had had sexual relations and those who had not. Among sexually active adolescents, condom use was assessed in relation to the last sexual intercourse.”

We expect that this information will allow the comparison of our results with other recent studies in Mexico.

# The main problem is that there is no definition of the meaning of the urbanization variable as it is interpretable.

We agree with the reviewer, and thus have made some changes in the manuscript in order to address this point. For instance, we have replaced the term “Urbanisation” for “Urbanicity” (where possible), because locality size is more properly an indicator or urbanicity, not urbanisation. With this change, the
results regarding influence of locality size could be interpreted as the influence of urbanicity on health (risk behaviours).

The discussion is poor and out of context with recent references in Mexico in relation to exposure variables studied. I believe that the information is useful but should be vastly improved the manuscript of reference.

We revised and edited the Discussion section. We believe that the new version of the manuscript address the comment made by the reviewer. For instance, we added new references regarding recent studies in Mexico (see references 40-44, 49 in the reference list at the end of the revised manuscript).

Reviewer 3: Isaac Rhew
Reviewer's report:
Major revisions:

The sampling approach appears very complex. First, it would be helpful to provide a bit more information about the Oportunidades 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.

We have added more information and details regarding sampling procedures in the three surveys as well as some information regarding Oportunidades program. This information can be found at page 8 and 9.

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

We have included the next paragraph as part of the limitations of our study in the discussion section at the end of page 17 and beginning of page 18: “Lastly, the sampling procedures for the surveys may affect the validity of our results. As the sample was composed of poor adolescents, our findings do not necessarily reflect the behaviours of adolescents in general. Furthermore, the samples were selected to represent poor families that could be beneficiaries of the program, which means that this study may not be representative of all adolescents from poor households. For instance, one criterion for eligibility to the program was that the locality have access to health and education. However, it is possible that adolescents from poor localities were excluded from the sample due to a lack of access to health and education, thus having different risks because they live in a more isolated context.”

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.

We did not use sampling weights in our analysis. The survey commands used
were SVY commands in Stata that allows accounting for complex survey design in the analyses. We have added this information on the statistical analysis section on page 12 (First paragraph): “The analyses were undertaken using the SVY commands in Stata, which are complex survey commands that allow for the adjustment of standard errors in all clusters of individuals. In the surveys described herein, the clusters were the primary sampling units (localities in the semi-urban and rural zones, and blocks in the urban zones).”

The surveys did not oversample any specific type of households.

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.

The final logistic regression analyses did account for cluster sampling. We have clarified this on page 12 (second paragraph): “Subsequently, a multivariate logistic regression model that accounted for clustering was adjusted for each dependent variable including age as an independent variable, educational level and an interaction between marital status and gender”

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

We added information regarding the mechanism of our findings on page 16 (Discussion section).

Minor revisions:
1. There are numerous grammatical errors. The manuscript should be edited carefully.

We have revised and edited the grammar in the manuscript.

2. 3rd paragraph of the Background is confusing.

We have shortened the paragraph and now it reads: “Young people constitute a critical group in social transformations, so their welfare will have an impact on development. Although they tend to be conceived as healthy, the consequences of their behaviours are known to surface in the long run.” (page 3 last paragraph).

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).

Because of the limited evidence regarding the associations described in our study, our hypothesis was merely correlational and thus no direction was stated. Our hypothesis was specified at the end of the Background section on page 7.

4. In Background, expand on the advantages of defining urbanization using locality size rather than standard urban-rural dichotomous definitions.

We have included at the end of page 6 the next paragraph: “However, a dichotomy of rural versus urban seems limited when identifying features of the environment that may be related to adolescents’ risk behaviours because
individuals from small towns and large cities would be placed in the same category, but behavioural differences related to the setting would be expected to exist. Therefore, a more accurate indicator showing the gradations of urbanicity may be valuable.”

5. It would help readers if authors define “locality.”
We added the definition for “locality” at the end of page 7, in the Methods section: “(with geographical areas defined as the lowest administrative division in Mexico)”.

6. The authors should state whether the localities in the sample were all geographically distinct, or if some were proximate. If there were contiguous localities, it would be helpful to know the number (%).
All the localities in our sample are distinct geographically limited areas. We find the observation made by the reviewer very interesting; nevertheless we can’t appropriately address this point given that we don’t have this information. Distance between localities was not an inclusion criterion for these surveys. Thus, such analysis would imply an extraordinary effort since exploring the number of contiguous localities in our sample goes beyond the purpose of this study, and because we have as long as 704 localities.

7. Additional information on what indicators comprised the poverty score would be useful.
We have included more information regarding the procedure to construct the poverty index on page 11.

8. Authors should describe how they decided to define the four urbanization categories.
We decided to use categories employed in the last Mexican Census. We have included the next sentence in page 11: “We used this classification because it is the official domain used in the last census for classifying locality size [37].”

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.
We have included on page 10 the operationalisation of risk variables assessed: “The dependent variables are risk behaviours. Adolescents were asked if they had smoked at any time, if they currently smoked, and if they currently drank alcohol even on an occasional basis. Regarding sexual behaviour, they were asked for their age at first intercourse, which was then used to classify adolescents who had had sexual relations and those who had not. Among sexually active adolescents, condom use was assessed in relation to the last sexual intercourse.”

10. Authors should state why is an interaction between marital status and gender is included in the model.
On the statistical analysis section (second paragraph on page 12) we added the next: “Subsequently, a multivariate logistic regression model that accounted for clustering was adjusted for each dependent variable including age as an independent variable, educational level and an interaction between marital status and gender (as the potential of being married is expected to be different for males and females).”

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 be not be fit. Were other rural localities sampled in 2001? If so, this further underscores the need to clarify the sampling procedures.

Rural and urban localities were sampled in 2003, whereas semi-urban were sampled in 2001. Our 4 categories of locality size are based on external data, meaning that we are not using the original categories employed by the Program in defining the localities to be sampled (Rural, Semi-urban or Urban). Once we have created the categorical variable for locality size, we found that (as expected) our rural localities and those sampled by the program were the same (#2,500 inhabitants), however, the semi-urban and urban localities as defined by the program (sampled in 2003 and 2001 respectively) were distributed into our small semi-urban (>2,500 and #15,000 inhabitants), large semi-urban (>15,000 and #100,000 inhabitants) and urban category (>100,000 inhabitants). The inclusion of year in the model does not make this variable and locality size collinear, given that using our classification of locality size the category of 2003 arrange individuals from the 4 categories of locality size and the category of 2001 includes individuals from semi-urban and large semi-urban localities. We added more details regarding sampling procedures on page 8 and 9 from the manuscript.

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).

We have included the next sentence on the Statistical Analysis section on page 12: “Locality size and SES categories, and the interaction between each, were modelled as indicator variables, excluding the reference category.”

13. Additional discussion of how the sampling might affect findings is warranted.

We have included the next paragraph as part of the limitations of our study in the discussion section on page 17 and 18: “Lastly, the sampling procedures for the surveys may affect the validity of our results. As the sample was composed of poor adolescents, our findings do not necessarily reflect the behaviours of adolescents in general. Furthermore, the samples were selected to represent poor families that could be beneficiaries of the program, which means that this study may not be representative of all adolescents from poor households. For instance, one criterion for eligibility to the program was that the locality have access to health and education. However, it is possible that adolescents from poor localities were excluded from the sample due to a lack of access to health and education, thus having different risks because they live in a more isolated
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.

We agree with the reviewer, and thus we have replaced the terms “Modulated” and “Mediated” for “Moderated”.

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.

We included some information regarding alcohol consumption on the Background (second paragraph, page 4): “Approximately 11% of males and 7% of females reported being heavy drinkers (5 or more & 4 or more drinks per occasion, respectively”.

2. The authors could limit some of the early paragraphs in the Background, and focus more on the rationale for conducting the current study.

We believe that the changes we have made along with other modifications in the structure of the background section will help to bring the focus on the rationale of our 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.

We have added more details regarding sampling procedures, and this will help the reader to better understand the logic of the sampling in the surveys used (page 8 and 9).

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.

We have reorganized the structure of the discussion section and limited some phrases regarding additional results.

Reviewer 4: Joanna Stewart
Reviewers report:
Major
1 Table 1 should simply be a presentation of the data. It is important to report the characteristics of the sample however the tests of significance should be removed from this table. It is not appropriate to test all the individual comparisons of subcategories of variables separately. Also the different locality sizes were sampled in different surveys with different methods of selection based on
deprivation. All comparisons of locality size differences need to control for deprivation as well as other confounders. The hypotheses of interest are tested more correctly in the multivariable analysis that is reported later. At present the table is very difficult to read. The confidence intervals could be removed as well. This would aid clarity considerably. They are meaningless for the combined samples and of questionable use within sample given the different sampling methods for different samples and the restrictions on who was sampled.

We have edited Table 1 and removed confidence intervals, intra-cluster coefficients and p-values for the comparison between groups. We also replaced the original “Total” column that was separated by Male-Female with the total values in the sample. However, we decided to keep the original table and attach it as an additional file for the most interested readers.

2. Table 2. It is unclear why the results of the two interactions of interest are reported differently. The overall effect of the interaction is reported for marital status/gender but does not seem to be for SES/locality size. However the odds ratios of the outcomes for locality size are reported within SES strata implying that an overall interaction of SES and locality size was found. Given this was the case it is appropriate to report the odds ratios within SES strata as has been done. However for gender/marital status it is unclear what the odds ratios reported are. Have they been reported split in both directions? If this is the case one should be removed. In line with the SES/locality size results, they should be split on marital status or gender with an odds ratio for, say males compared to females presented for single and for married.

We agree that the results of the two interactions are reported differently. Our main purpose in this study is to assess the influence of locality size on risk behaviours, considering the moderating effect of SES. Because of that, the results of the interaction of locality size and SES are presented only in one direction as the reviewer properly observe. However for the interaction of gender and marital status, we decided to present both directions as we consider that some readers would find interesting to see, on one hand, the effects of gender on risk behaviours and, on the other hand, the effect of marital status on risk. By presenting only one direction, let say, the influence of gender on behaviours across single or married adolescents, we can’t automatically infer the influence of marital status on risk behaviours across male or female, and vice versa. We find extremely convenient to present both results, although we totally understand this concern and thus we are willing to make the suggested change if, after reading our motivation, the reviewer still considers that such change needs to be done to improve our work.

1. Requires editing for English, in particular the abstract and background section. We have revised and edited the grammar in the manuscript.

2. On the assumption that the analysis of condom use in table 2 was restricted to those that were sexually active this should be noted.

We have included a note on table 2 to indicate that analysis for condom use was restricted to those sexually active adolescents.