Author's response to reviews

Title: Demographic, psychosocial, and genetic risk associated with smokeless tobacco use among Mexican heritage youth

Authors:

Anna V Wilkinson (anna.v.wilkinson@uth.tmc.edu)
Laura M Koehly (koehlyl@mail.nih.gov)
Elizabeth A Vandewater (elizabeth.a.vandewater@uth.tmc.edu)
Robert K Yu (rkyu@mdanderson.org)
Susan P Fisher-Hoch (Susan.P.Fisher-Hoch@uth.tmc.edu)
Alexander V Prokhorov (aprokhor@mdanderson.org)
Harold W Kohl (Harold.W.Kohl@uth.tmc.edu)
Margaret R Spitz (spitz@bcm.edu)
Sanjay Sanjay Shete (sshete@mdanderson.org)

Version: 3
Date: 4 May 2015

Author's response to reviews: see over
Dear Editors

We thank the reviewers for their constructive comments and careful review. Below we detail our responses to their questions. We provide the lines numbers where we have revised the text in response to the reviewers’ comments and have included the revised text in red in the manuscript. Thank you for the opportunity to revise our manuscript.

Reviewer 1:

1- I could not find anywhere in the text of the original number of SNPs investigated prior to analysis that resulted in SNPs with significant results. Correction for multiple comparisons was performed?

We apologize for this omission. We have included the original number of SNPs investigated. Please see “Following this approach, a total of 672 SNPs on 58 candidate genes were identified and examined in the current study” on lines 151-52.

We did not complete corrections for multiple comparisons due to the exploratory nature of the analyses. Exploratory analyses were performed to find unexpected, yet useful information that can be rigorously tested in future studies. In such studies, it is acceptable to provide unadjusted p-values. We have noted this in the discussion and provide a reference (Schochet, 2008).

2- The article lacks a detailed description with respect to the state of the art in genetic studies on STU, considering both GWAS as well as candidate genes approaches?

To the best of our knowledge, based on extensive literature searches, there are no publications that examine the relationship between the behavioral trait of smokeless tobacco use and genetic variance, be they based on a GWAS or candidate gene approach. Studies that have been published focus on genetic risk for tobacco-related cancer and the disease process, not the acquisition of the behavior.

3- Writing sometimes suggests a determinism that is not suitable for such a complex. This is even more questionable when authors propose that unreplicated genetic data could be used to “identify youth at increased risk for STU”.

We appreciate and share your concern as we do not want to overstate the potential of these results. Rather our intention is to suggest that there are benefits to using widest range of data available (i.e. from cells to society) when assessing individual risk for smokeless tobacco use. We have revised the conclusion in the abstract to say: These data suggest that the use of genetic risk, along with psychosocial, demographic, and behavioral risk factors may increase our ability to identify youth at increased risk for STU, which in turn may improve our ability to effectively target prevention messages to Mexican heritage youth.

4- The preparation of the “genetic risk score” follows a circular logic that obviously results in a positive finding considering that it is based in the same SNPs that were selected exactly because they presented significant findings in the same sample, without any validation procedure.

We agree that the “genetic risk score” follows a circular logic, and as the reviewer noted the results will be positive due to the circular logic. We have revised table 4 to present results based on individual SNPS rather than the risk score.
5- Adjustment of the effect of STU by cigarette use is inappropriate because the pathophysiology of both types of substance abuse is at least partly shared.

Thank you for your comment. We agree with you and have removed the analysis and reference to the analysis from the manuscript.

Reviewer 2:

1. Would be helpful to have some statistics in the results section of the abstract. Also, counts should be accompanied with percentages.

   Thank you for the suggestion. We have added statistics to the results section of the abstract (and have included both counts and percentages).

2. Be careful about the use of the word “risk.” Increased likelihood (or odds) should be used.

   We have revised the language throughout.

3. Line 70 should have e.g. (for example) instead of i.e.

   Updated, please see line 65.

4. Line 74, Confidence intervals are not needed.

   We have removed the confidence intervals from the national and state level prevalence estimates.

5. Line 107, consent and assent should be mentioned. Also compensation for time should be mentioned.

   We have clarified that informed consent and assent were obtained and that participants were provided with a gift certificate to compensate for their time. Please see lines 97 and 102.

6. Major problem with the analyses. Line 109, need to discuss attrition analyses. Was the final sample similar to the original sample? What type of missingness? How do you plan to handle the missingness? What is the advantage of using multiple waves of data examining the same outcome and predictors? What variables are changing each year (other than adding individuals to the ever STU group)? Couldn't the analyses focus on Wave 3 when you have the highest prevalence of STU. The major concern is the use of three waves of data collection with the same analyses.

   We debated whether to present analyses from all three waves of data, or one wave. We initially choose to present the three waves of data as such an approach facilitates a comparison of the relative influence of each covariate at each wave of data collection. However, such an approach is difficult in terms of detailing attrition as you note.

   In response to your comment we have restricted the analyses to the Wave 2 data, gathered in 2008-09. We examine Wave 2 rather than Wave 1, because Wave 2 permits an analysis of social disinhibition, a construct that was not assessed at Wave 1. We examine Wave 2, rather than Wave 3, because data on one variable (N=100 on subjective social status). Therefore to maximize the sample size and use the most available data on most variables, we focus on 2008-09 only.
Missing outcome and/or covariate data are common occurrences in cohort studies. The standard analytic approach implemented in most popular statistical analysis software (e.g. SAS, STATA) is to remove participants with missing data. Such an analysis is called complete case analysis. There are several imputation methods each with, sometimes untestable, assumptions. Therefore, we chose to follow the standard analytic approach. The data are said to be missing completely at random, if failure to observe the data occur completely at random and does not depend on any other variables (e.g. outcome variables of interest). The sample characteristics of those missing covariates and those with complete data were not different (which also can be seen by the consistency of our results when cross sectional analyses of each wave of data were conducted). Therefore, we do not believe the attrition is associated with outcome of interest.

7. Line 113 use dichotomous or binary instead of two level variable.

Updated, please see line 111.

8. Line 118, please include the number and percentage of individuals who were inconsistent regarding their STU.

We have noted in the limitations that roughly 20% of the youth reported inconsistent STU. However consistent with tobacco research among adolescents, the first response of yes is taken at face value, and therefore all reports of yes were analyzed as a yes. Please see lines 293-94.

9. Line 161, 3 different cross sectional analyses for the same sample? Does Bonferroni correction apply since the sample is not independent?

In response to point 6, the data from only one wave (2008-09) are presented. Also, see our response to Reviewer 1 about multiple corrections.

10. Line 226 two periods, correct. Also, would be helpful to include the statistic that goes with the statement. Include odds ratio when describing results e.g. 235

Corrected, thank you!

11. Line 237 has 9% instead of 95% CI

Updated – thank you!

12. Line 247 change risk to appropriate word

Updated – thank you!

13. Is this studied properly powered given the percentage of STU?

The study power is limited. We have noted this in the limitations section. Please see lines 291-3: A third limitation pertains to the relatively small number of individuals who reported lifetime STU implying limited statistical power for this preliminary study.

14. The discussion section cannot be fully evaluated until the methodological issues are addressed.

We thank you for the opportunity to revise our manuscript and look forward to hearing from you. Sincerely, Anna