Author’s response to reviews

Title: Correlates of cognitive susceptibility to smoking among Mexican origin youth residing in Houston, Texas

Authors:

Anna V Wilkinson (awilkins@mdanderson.org)
Andrew J Waters (andrew.waters@usuhs.mil)
Vandita Vasudevan (vvasudev@mdanderson.org)
Melissa L Bondy (mbondy@mdanderson.org)
Alexander V Prokhorov (aprokhor@mdanderson.org)
Margaret R Spitz (mspitz@mdanderson.org)

Version: 2 Date: 25 July 2008

Author’s response to reviews: see over
July 25, 2008

Dear Dr Manning:

Thank you very much for the opportunity to revise our manuscript entitled “Correlates of susceptibility to smoking among Mexican origin youth residing in Houston, Texas.” We found the reviewers comments to be very helpful and we have been able to respond appropriately to all the concerns raised. Below we outline the specific changes.

Reviewer: Thomas Valente

“First, … puzzled by the repeated use of the term cognitive susceptibility and would recommend dropping the modifier “cognitive.”” The measure we used to assess susceptibility was designed by Pierce et al., and was referred to as “cognitive susceptibility to smoking” by the authors. However, we understand and agree with the reviewer’s point. When we reference the measure for the first time in the text and in table 1, we refer to it as “cognitive susceptibility to smoking,” but subsequently discuss smoking susceptibility and will drop the modifier throughout the text and from the title.

“Second… discussion of the logistic regression results on page 11 refers to … Can you clarify?” We apologize for the lack of clarity. The reviewer’s interpretation of table 5 is correct and we have modified the text in the results and discussion sections to be consistent.

“Third, why describe outcome expectations as a cognitive variable? …. I would rephrase.” Again, we agree with the reviewer’s assessment and have rephrased the sentence,

“Overall, the strongest risk factor we identified was holding positive expectations about smoking, although both family- and school-based characteristics were important, too.“

We have corrected the typo.

Reviewer: Eduardo Lazcano-Ponce

“In the introduction, the authors explicit precise values from previous investigation reports. In this respect, it will be more important to interpret them according with the size of sample and power of comparability of the results.” The first survey that we reference in the introduction was sponsored by the Centers for Disease Control and Prevention and the second by the Texas Department of Health. Both surveys are explicitly designed to monitor the health and health behaviors of youth at the national and state level respectively. Therefore both surveys, by design, have sufficient sample size and adequate power to permit comparisons by ethnicity. We have revised the introduction to reflect these important details and apologize for the lack of clarity in our original submission.
Moreover, results from two large population-based surveys of middle and high school students lend support to the predictive validity of the susceptibility construct [8, 9].

The authors indicate the accomplishment of the study … consequently, rate of answer is not described.” We apologize for not including this information. We have expanded the discussion of the sample size and response rates.

“A total of 3,000 households were identified as eligible from the cohort database. Of these 1,425 were contacted. Just over 90% of all parents with age-eligible children who were contacted agreed to enroll their child in the study.”

The methods do not specify the inclusion criteria used to establish the Hispanic origin…” We apologize for not including this detail in this manuscript. We have expanded the description of the cohort study from where the study participants are drawn to include:

“Participants are self-identified Mexican Americans and Mexicans of any age and sex who reside in predominantly…”

Variables included in the multivariable model are unclear.” We apologize for the lack of clarity and have revised all variable names that were inconsistent across tables to be consistent. We hope this change will enable the reader to start with table 1, in which we describe the measures, and easily read the data presented in tables 2 through 5. In other words, the variables listed in table 4, which presents the results from the multivariable model, are described in detail in Table 1. In both tables they have the same variable label, so we hope this will allow the reader to easily determine which variables maintained significance in the multivariable models and were therefore retained.

Although there are many variables that are not associated, it is desirable to know their quantified effect in the models.” We completed an unconditional backward elimination stepwise logistic regression analysis to develop the model in Table 4. In backward stepwise regression, the analysis begins with a full or saturated model and variables are eliminated from the model in an iterative process. The fit of the model is tested after the elimination of each variable to ensure that the model still adequately fits the data. When no more variables can be eliminated from the model, the analysis has been completed. Therefore, the risk factor variables that are maintained in the model are those that demonstrate a significant association with the outcome variable, susceptibility. As a result, it is not possible to describe the quantitative effects of the variables that were found to not be associated in the multivariable models.

In table 5, the authors mention a model of accumulated risk. Nevertheless a variable model was not developed.” We apologize for the lack of clarity. The accumulated risk model was developed using the methodology outlined below:

“Finally, we created a cumulative risk score for each participant by summing the risk factors identified from the logistic regression (see Table 4). We dichotomized most risk factors assessed as continuous variables (positive outcome expectations, peer norms, temptations to try cigarettes, and
subjective social status) at the median. However, we dichotomized detentions as none or at least one. We then assigned each risk factor the value of 1 and created a risk index by summing risks [14]. Cumulative risk scores ranged between 0 and 8; no participants obtained the highest possible score of 9. We then completed an unconditional logistic regression.”

It is important for the authors to indicate how colinearity was avoided.” And “The discussion is pertinent… There is a potentially multiple correlation exists with the aforementioned variables… Which of these were used for inclusion in the variable models?” Prior to completing the multivariable analyses we examined the correlation coefficients among all variables that demonstrated a significant (p<0.05) bivariate association with susceptibility. As a result of this analysis, we found two variables (both measures of peer smoking behavior) were highly correlated (r=0.68; p< 0.01). Therefore we created a single indicator variable from these two measures. We then reexamined the correlations coefficients and found that none were above r=0.25. We have rewritten this section of the methods.

“To identify the variables that independently predicted susceptibility, we performed an unconditional stepwise logistic regression analysis. Prior to completing this multivariable analysis, we examined the correlations among the risk factors that demonstrated a significant bivariate association (p < 0.05) with susceptibility. Because the two peer influence variables (“three best friends smoke,” “some friends smoke;” see Table 1) were correlated (r = 0.64; p < 0.01), we created a dichotomous summary measure (“any friends smoke”) to use in the multivariate models. Participants who responded “none” on the "How many of your friends smoke?" question and “0” on the "How many of your closest friends smoke?" question were coded as 0; all other participants were coded as 1. However, the correlations among all other risk factors were low. Therefore all risk factors that demonstrated a significant bivariate association (p < 0.05) with susceptibility were entered into the model. All risk factors measured as scaled variables (e.g. attitudes toward smoking, family and peer norms, parental acculturation, and subjective social status) were entered as continuous variables, while the categorical variables were entered as dummy variables.”

In addition, after completing the model presented in Table 4, we examined the Variance Inflation Factor (VIF) for each variable included in the model. The VIF is an index which measures how much the variance of a coefficient (square of the standard deviation) is increased because of collinearity. The highest VIF obtained was 1.19 indicating little collinearity among the risk factors we examined. To the description of the statistical analyses we added:

“To determine the extent of multicollinearity among the risk factor variables, we examined the variance inflation factor (VIF) for each variable retained in the model.”

To the description of the results of the multivariate analysis we added:

“The highest VIF obtained was 1.19 indicating that collinearity was not present among the nine risk factors retained in the model.”
We feel that the manuscript is improved as a result of incorporating the reviewers’ suggestions and hope that you agree. We thank you for the opportunity to revise and resubmit our manuscript.

Sincerely,

Anna Wilkinson, PhD
Assistant Professor
Department of Epidemiology
University of Texas MD Anderson Cancer Center