

Tobacco Use in American Youth: An Analysis of the 2014 American National Youth Tobacco Survey

Fiona McLean

Abstract

Tobacco use in American youth is a major health concern. To better understand consumption patterns of tobacco in this population, it was investigated if regular use of chewing tobacco, snuff or dip is more common among Americans of European ancestry than for Hispanic-Americans and African-Americans and if regular use of hookah or water pipes differs between the sexes. Data from the 2014 American National Youth Tobacco Survey was used to help answer these questions. It was discovered that the use of chewing tobacco is significantly higher in Caucasian Americans, compared to African Americans and Hispanic Americans, even once one accounts for the fact that white Americans more likely to live in rural areas and chewing tobacco is a rural phenomenon. It was also discovered that there is no statistically significant difference in the use of hookah or water pipes between the sexes, provided age, ethnicity, and other demographic characteristics are similar. It is hoped that these conclusions will help to better inform tobacco related decisions involving the youth of America.

Introduction

The use of tobacco is the leading cause of preventable disease and death in the United states. Tobacco use and addiction mainly begin in adolescence, making the use of tobacco a major health concern for the youth (Arrazola et al. 2015). In order to understand tobacco use among the youth, two main research questions were proposed and investigated. It was investigated whether regular use of chewing tobacco, snuff or dip is more common among Americans of European ancestry than for Hispanic-Americans and African-Americans in the United States, for young adults aged 9-19. The difference in the likelihood of having used a hookah or water pipe on at least one occasion between men and women, provided their age, ethnicity, and other demographic characteristics are similar was also considered. The 2014 American National Youth Tobacco Survey, a survey conducted by the CDC, which provides representative data regarding American youth's beliefs and behaviors towards tobacco, was used to investigate the above questions ("National Youth Tobacco Survey (Nyts)" 2019).

Methods

The data set used in the analysis is the 2014 American National Youth Tobacco Survey, a survey conducted by the CDC on young adults in the United States. The survey contains responses from 22,007 young adults on 162 questions regarding their use, perception of, and experience with tobacco.

As an exploratory tool, tables of ages and grades were created, as well as tables comparing race, sex, and urban or rural residence compared to tobacco use. Data for 9 year olds was removed due to suspicious results concerning the grade level of many 9 year olds respondents. The variable age was centered around the median, 14 years old. It is assumed that the responses are independent of one another.

In order to determine if (1) regular use of chewing tobacco, snuff or dip is more common among Americans of European ancestry than for Hispanic-Americans and African-Americans and if (2) the likelihood of having used a hookah or water pipe differs between sexes, logistic models were built. As a teenager either regularly uses these tobacco products (either chewing tobacco or a hookah) or does not, a logistic regression model is appropriate to answer both questions.

For research question (1), if regular use of chewing tobacco, snuff or dip is more common among Americans of European ancestry vs. Hispanic and African Americans, European ancestry was set as the baseline for comparison.

The model was constructed as below:

$Y_i \sim \text{Bernoulli}(\mu_i)$, where:

- Y_i is the indicator for a respondent using chewing tobacco, snuff or dip regularly
- μ_i is the probability of a respondent using chewing tobacco, snuff or dip regularly

The logistic regression model is:

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = \beta_0 + \beta_1 I_{Sex} + \beta_2 I_{RaceBlack} + \beta_3 I_{RaceHispanic} + \beta_4 I_{RaceAsian} + \beta_5 I_{RaceNative} + \beta_6 I_{RacePacific} + \beta_7 I_{Rural} + \beta_8 (Age) + \beta_9 (Age \times I_{Sex})$$

Race, Rural, Sex, Age and the interaction of Age with Sex was included to allow analysis of the differences of chewing tobacco between races, while controlling for these extraneous variables. It was particularly important to control for the urban vs. rural location of a respondent, as chewing tobacco is regarded to be a rural phenomenon. The interaction term between Age and Sex was included, as it is statistically significant and the interaction plot below shows that there is likely an interaction between these terms. Other interaction terms were not included, as they were not as significant in the model and including them did not improve model fit substantially. Therefore this model was chosen as it fits the data well, controls for concerning factors, and is relatively simple.

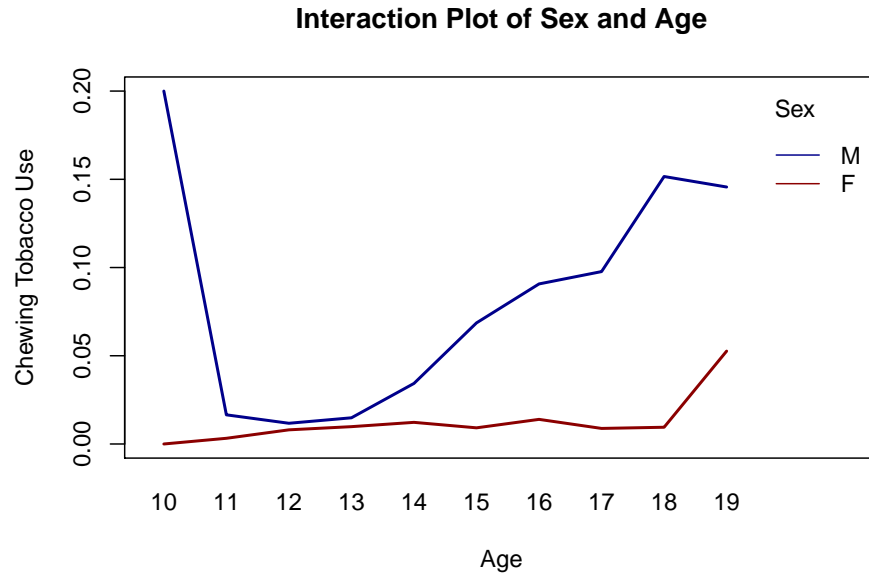


Figure 1: The interaction plot suggests that there may be an interaction between sex and age, and that including an interaction term in the model is wise.

For research question (2) the differences in the likelihood of having used a hookah or water pipe between sexes was investigated. Male sex was set as the baseline for comparison.

The model was constructed as below:

$Y_i \sim \text{Bernoulli}(\mu_i)$, where:

- Y_i is the indicator for a respondent using a hookah or water pipe regularly
- μ_i is the probability of a respondent using a hookah or water pipe regularly

The logistic regression model is:

$$\log \left(\frac{\mu_i}{1-\mu_i} \right) = \beta_0 + \beta_1 I_{Sex} + \beta_2 I_{RaceBlack} + \beta_3 I_{RaceHispanic} + \beta_4 I_{RaceAsian} + \beta_5 I_{RaceNative} + \beta_6 I_{RacePacific} + \beta_7 I_{Rural} + \beta_8 (Age)$$

Age, Rural, Sex, and Race were included to allow analysis of the differences of hookah or water pipe use between sexes, while controlling for these extraneous variables. Interaction terms were not included in this model, as inclusion of interaction terms did not increase model fit enough to account for the added complexity.

Results

Model 1

The percentage of chewing tobacco, snuff, or dip users was compared among the races in question. There is reason to believe that Caucasian Americans may have a higher rate of tobacco use as seen in Table: 1 by the higher percentage of chewing tobacco use for Caucasians. However, it is important to consider that Caucasian Americans live in rural areas, and chewing tobacco is considered a rural phenomenon. It is also important to ensure that the difference in tobacco use is statistically significant, and not just a result of the sample.

Race	False	True	Total	% Chew Tobacco
white	9130	504	9634	5.23%
black	3196	39	3235	1.21%
hispanic	5561	120	5681	2.11%

Table 1: The percentage of teenagers using chewing tobacco for each race.

The logistic regression model (1) described above suggests **statistically significant differences (at the 5% confidence level) in the likelihood of regular use of chewing tobacco, snuff, or dip use when comparing young Americans of European ancestry to those of Hispanic-Americans and African-Americans of the same demographic characteristics** in the United States.

	Estimate	Std. Error	z value	Pr(> z)	Odds	2.5 %	97.5 %
(Intercept)	-3.494	0.095	-36.660	0.000	0.030	-3.684	-3.310
Raceblack	-1.498	0.172	-8.688	0.000	0.224	-1.852	-1.174
Racehispanic	-0.734	0.108	-6.821	0.000	0.480	-0.949	-0.527
Raceasian	-1.525	0.343	-4.447	0.000	0.218	-2.273	-0.913
Racenative	0.127	0.287	0.443	0.658	1.135	-0.480	0.654
Racepacific	0.890	0.398	2.233	0.026	2.434	0.035	1.616
RuralUrbanRural	0.937	0.090	10.449	0.000	2.552	0.763	1.114
AgeC	0.398	0.024	16.525	0.000	1.488	0.351	0.445
SexF	-1.352	0.127	-10.672	0.000	0.259	-1.606	-1.109
AgeC:SexF	-0.306	0.055	-5.554	0.000	0.737	-0.414	-0.198

Table 2: The table provides the output for the logistic regression model. As the p-value is small and 0 is not in the confidence interval for the races in question, we can conclude that there are statistically significant differences in the likelihood of the use of chewing tobacco in African Americans and Hispanics when compared to white American teenagers.

The odds for each race of chewing tobacco were calculated. The odds of white teenagers chewing tobacco is higher than the odds of chewing tobacco compared to other races.

	Urban Est.	Urban Lower	Urban Upper	Rural Est.	Rural Lower	Rural Upper
White Odds	0.030	0.025	0.037	0.078	0.054	0.111
Black Odds	0.007	0.004	0.011	0.017	0.008	0.034
Hispanic Odds	0.015	0.010	0.022	0.037	0.021	0.066

Table 3: The table provides the odds of using chewing tobacco in each demographic group of interest. It is clear that the odds for White teenagers is higher than those of Black and Hispanic in both urban and rural settings.

The odds of chewing tobacco, snuff, or dip for Caucasian youth are 4.47 times the odds of African-American youth of a similar demographic. The odds of chewing tobacco, snuff, or dip for Caucasian youth are 2.08 times the odds of Hispanic youth of a similar demographic.

Table: 4 shows probability estimates of chewing tobacco, snuff, or dip for each racial group, depending on urban or rural location. Clearly Caucasian's have the highest probability of chewing tobacco, snuff, or dip compared to Hispanic and African American youth. As expected, the probability of chewing tobacco, snuff, or dip is higher in a rural setting for all races compared to an urban setting.

	Urban	Rural
White	2.90%	7.20%
Black	0.70%	1.70%
Hispanic	1.40%	3.60%

Table 4: The table provides the probability of tobacco use for each race and urban/rural setting.

Model 2

The percentage of hookah use was compared among the sexes. As seen in Table: 5 the percentage of females using hookah or pipes is less than 1% smaller than the percentage of males using hookah or pipes. A model is used to determine if the difference in use of hookah and pipes is significant between the two sexes, once other characteristics such as race, age, and rural location are considered.

Race	False	True	Total	% Hookah
M	8613	1409	10022	14.06%
F	8443	1368	9811	13.94%

Table 5: The percentage of teenagers using hookah or pipes for each sex is shown.

The logistic regression model (2) described above suggests **that there does not exist statistically significant differences (at the 5% confidence level) in the likelihood of regular use of hookah or pipes when comparing young male Americans to young female Americans of the same demographic characteristics in the United States.**

	Estimate	Std. Error	z value	Pr(> z)	Odds	2.5 %	97.5 %
(Intercept)	-2.146	0.047	-45.216	0.000	0.117	-2.240	-2.054
SexF	0.041	0.043	0.942	0.346	1.041	-0.044	0.125
RuralUrbanRural	-0.387	0.044	-8.692	0.000	0.679	-0.474	-0.300
Raceblack	-0.645	0.071	-9.071	0.000	0.525	-0.785	-0.507
Racehispanic	0.347	0.049	7.148	0.000	1.415	0.252	0.443
Raceasian	-0.647	0.119	-5.451	0.000	0.523	-0.886	-0.420
Racenative	0.163	0.190	0.857	0.392	1.177	-0.226	0.522
Racepacific	1.011	0.271	3.737	0.000	2.748	0.463	1.528
AgeC	0.420	0.012	36.201	0.000	1.522	0.397	0.443

Table 6: The table provides the output for the logistic regression model. As the p-value for sexF is much greater than our set significance level of .05 and 0 is in the confidence interval, we can conclude that there are not statistically significant differences in the likelihood of the use of hookah and pipes between the two genders.

Table 7 refers the odds of hookah and pipe use for each sex. The odds of using hookah or pipes for men are 0.96 times the odds for a woman of similar demographic.

	Estimate	Lower Bound	Upper Bound
Female Odds	0.122	0.102	0.145
Male Odds	0.117	0.106	0.128

Table 7: The table provides the odds of hookah use for each gender. It is clear that the odds do not differ greatly between the two genders, and that there is significant overlap in the confidence intervals.

Discussion

There exists statistically significant differences in the likelihood of regular use of chewing tobacco, snuff, or dip use when comparing young Americans of European ancestry to those of Hispanic-Americans and African-Americans of the same demographic characteristics. When comparing race and rural or urban location, it was discovered that white rural teenagers are the most likely to chew tobacco, while African American urban teenagers are the least likely. In all races, urban use of chewing tobacco, snuff, or dip was lower than rural use. In order to combat the use of chewing tobacco in the youth, it is recommended that reduction efforts are focused on Americans of European ancestry, as they are more likely to chew tobacco regardless of location.

There does not exist statistically significant differences in the likelihood of regular use of hookah or pipes when comparing young male Americans to young female Americans of the same demographic characteristics. While race age, and urban vs. rural location were found to be significant predictors of hookah or pipe use, sex was not found to be significant. Therefore, it is recommended that efforts to decrease hookah and pipe use should be equally dispersed across both sexes, as women are as likely as men to use hookah or pipes.

References

Arrazola, René A, Tushar Singh, Catherine G Corey, Corinne G Husten, Linda J Neff, Benjamin J Apelberg, Rebecca E Bunnell, et al. 2015. "Tobacco Use Among Middle and High School Students—United States, 2011–2014." *MMWR. Morbidity and Mortality Weekly Report* 64 (14). Centers for Disease Control; Prevention: 381.

"National Youth Tobacco Survey (Nyts)." 2019. https://www.cdc.gov/tobacco/data_statistics/surveys/nyts/index.htm.