Investigation of Impact on Smoking due to Vaporizing in High Schools

Justin Williams, MSPH
Florida A&M University
Institute of Public Health

Introduction:

In the year 1948, a British physiologist named Richard Doll published the first major studies that proved smoking could cause serious health damage, and published his findings that showed a close link between smoking and lung cancer in the British Medical Journal in 1950 (Doll, 1950). This was a groundbreaking event, because up until the 20th century, this was the first time tobacco use was scrutinized and health implications had been called into question (Gilman and Xun, 2004). The United States Surgeon General's Report on Smoking and Health 'doubled-down' on Richard Doll's reporting in 1964 by conducted its own study, and determined the dangerous correlation in the relationship between smoking and cancer, in a 20 year study (NIH, 1964). With that being known, the tobacco companies have attempted to re-brand tobacco use over the years in the form of "vaporization." "Vaping" is the consumption of nicotine in the liquid form, heating the substance to the point steam is produced and inhaled. Due to this repackaging, the assumption implies that "vaping" is not as harmful as smoking cigarettes and therefore must be safe to consume. This trend has already started to infiltrate the high schools (in some regions middle schools) with the misconception that one form of nicotine is not being swapped out for another.

Physiopathology:

The reason why nicotine is so addictive is because a surge of endorphins in the reward circuits of the brain causes a brief euphoria when nicotine is administered (NIDA: Tobacco, Nicotine and E-Cigarettes, 2018). Nicotine is absorbed into the bloodstream through the lining of the mouth and the lungs and travels to the brain in a matter of seconds (U.S. Department of Health and Human Services, 2014). Like other drugs of abuse, nicotine increases levels of the neurotransmitter dopamine in these reward circuits, which reinforces the behavior of taking the drug, regardless of how it is consumed (NIDA: Tobacco, Nicotine and E-Cigarettes, 2018).

Methods:

In order to further determine the addictive characteristics of tobacco, an analysis of a self-created dataset will be conducted. The data utilized to assess the effects of smoking on high school level participants came from the 2015 Youth Risk Behavior Surveillance System (YRBSS, 2015). A three-stage cluster same design was used to produce a representative sample of 9th through 12th grade students (YRBSS, 2015). The survey was inclusive of all regular public, Catholic, and other private school students, in grades 9 through 12, in the 50 states of America and District of Columbia.

The national high school YRBS survey questions and dichotomous variables utilized to gain prospective of smoking consumption among high school students (grades 9 through 12) were:

Q2. What is your sex?

A. Female

B. Male

/*Q2. = Qsnd = gender, A. = 1 = Female, B. = 2 = Male in SAS code results*/

Q4. Are you Hispanic or Latino?

A. Yes
B. No

/*Q4 = Qfth = ethnicity, A. = 1 = Yes, B. = 2 = No in SAS code results*/

Q5. What is your race?

A. American Indian or Alaska Native
B. Asian
C. Black or African American
D. Native Hawaiian or Other Pacific Islander
E. White

/*Q5 = newrace = race , A. = 1 = Am. Indian, B. = 2 = Asian, C. = 3 = Black, D. =4 = Hawaiian, E. = 5 = White in SAS code results*/

Q33. During the past 30 days, on how many days did you smoke cigarettes?

A. 0 days

B. 1 or 2 days

C. 3 to 5 days

D. 6 to 9 days

E. 10 to 19 days

F. 20 to 29 days

G. All 30 days

/*Q33. = newQ = smoking, A. = 1 = 0 days, B. = 2 = 1.5 days, C. = 3 = 4 days, D. = 4 = 7.5 days, E. = 5 = 14.5 days, F. = 6 = 24.5 days, G. = 7 = 30 days in SAS code results*/

Q39. Have you ever used an electronic vapor product?

A. Yes
B. No

/*Q39. = Qthn = vaping, A. = 1 = Yes, B. = 2 = No in SAS code results*/

These questions and how they were answered were used for specific bio-statistical testing that provide greater insight on patterns that would explain why/why not teens choose to smoke cigarettes.

Results:

The information provided by YBSS questions were converted into SAS to provide statistical analysis testing and graphs. The Chi-square test is used to determine whether the variation in data due to chance or due to one of the variables that we are testing.

/*Chi-square analysis*/
/*Q5 = newrace = race , A. = 1 = Am. Indian, B. = 2 = Asian, C. = 3 = Black, D. =4 = Hawaiian, E. = 5 = White in SAS code results*/
/*Q33. = Smkng = smoking, A. = 1 = 0 days, B. = 2 = 1.5 days, C. = 3 = 4 days, D. = 4 = 7.5 days, E. = 5 = 14.5 days, F. = 6 = 24.5 days, G. = 7 = 30 days in SAS code results*/

The FREQ Procedure

Table of Smkng by newrace

Frequency Percent Row Pct Col Pct	newrace					
Smkng	1	2	3	4	5	Total
1	399 3.44 3.89 88.08	658 5.68 6.41 93.07	1688 14.57 16.45 92.44	200 1.73 1.95 88.89	7319 63.19 71.31 87.42	10264 88.61
2	16 0.14 3.54 3.53	16 0.14 3.54 2.26	46 0.40 10.18 2.52	10 0.09 2.21 4.44	364 3.14 80.53 4.35	452 3.90
3	6 0.05 3.30 1.32	2 0.02 1.10 0.28	18 0.16 9.89 0.99	2 0.02 1.10 0.89	154 1.33 84.62 1.84	182 1.57
4	6 0.05 4.20 1.32	4 0.03 2.80 0.57	10 0.09 6.99 0.55	7 0.06 4.90 3.11	116 1.00 81.12 1.39	143 1.23
5	7 0.06 5.26 1.55	6 0.05 4.51 0.85	12 0.10 9.02 0.66	0 0.00 0.00 0.00	108 0.93 81.20 1.29	133 1.15
6	5 0.04 4.46 1.10	4 0.03 3.57 0.57	11 0.09 9.82 0.60	1 0.01 0.89 0.44	91 0.79 81.25 1.09	112 0.97
7	14 0.12 4.71 3.09	17 0.15 5.72 2.40	41 0.35 13.80 2.25	5 0.04 1.68 2.22	220 1.90 74.07 2.63	297 2.56
Total	453 3.91	707 6.10	1826 15.76	225 1.94	8372 72.28	11583 100.00

Frequency Missing = 4041

Statistics for Table of Smkng by newrace

Statistic	DF	Value	Prob
Chi-Square	24	74.2646	<.0001
Likelihood Ratio Chi-Square	24	85.1230	<.0001
Mantel-Haenszel Chi-Square	1	11.8368	0.0006
Phi Coefficient		0.0801	
Contingency Coefficient		0.0798	
Cramer's V		0.0400	

Effective Sample Size = 11583
Frequency Missing = 4041

/*Chi-square analysis*/
/*Q4 = Ethn = ethnicity, A. = 1 = Yes, B. = 2 = No in SAS code results*/
/*Q33. = Smkng= smoking, A. = 1 = 0 days, B. = 2 = 1.5 days, C. = 3 = 4 days, D. = 4 = 7.5 days, E. = 5 = 14.5 days, F. = 6 = 24.5 days, G. = 7 = 30 days in SAS code results*/

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of Smkng by Ethn		
	Ethn		
Smkng	1	2	Total
1	4434 30.02 33.72 90.77	8717 59.01 66.28 88.17	13151 89.03
2	186 1.26 33.04 3.81	377 2.55 66.96 3.81	563 3.81
3	76 0.51 31.02 1.56	169 1.14 68.98 1.71	245 1.66
4	49 0.33 28.65 1.00	122 0.83 71.35 1.23	171 1.16
5	42 0.28 26.09 0.86	119 0.81 73.91 1.20	161 1.09
6	19 0.13 14.39 0.39	113 0.76 85.61 1.14	132 0.89
7	79 0.53 22.64 1.62	270 1.83 77.36 2.73	349 2.36
Total	4885 33.07	9887 66.93	14772 100.00
Frequency Missing = 852			

Statistics for Table of Smkng by Ethn

Statistic	DF	Value	Prob
Chi-Square	6	45.9657	<.0001
Likelihood Ratio Chi-Square	6	50.5295	<.0001
Mantel-Haenszel Chi-Square	1	40.0557	<.0001
Phi Coefficient		0.0558	
Contingency Coefficient		0.0557	
Cramer's V		0.0558	

Effective Sample Size = 14772
Frequency Missing = 852

/*Chi-square analysis*/
/*Q2. = Gend = gender, A. = 1 = Female, B. = 2 = Male in SAS code results*/
/*Q33. = Smkng = smoking, A. = 1 = 0 days, B. = 2 = 1.5 days, C. = 3 = 4 days, D. = 4 = 7.5 days, E. = 5 = 14.5 days, F. = 6 = 24.5 days, G. = 7 = 30 days in SAS code results*/

The FREQ Procedure

Frequency
Percent
Row Pct
Col Pct

Table of Smkng by Gend

Smkng	Gend 1	Gend 2	Total
1	6779 45.53 51.12 90.24	6481 43.53 48.88 87.87	13260 89.07
2	258 1.73 45.50 3.43	309 2.08 54.50 4.19	567 3.81
3	111 0.75 45.12 1.48	135 0.91 54.88 1.83	246 1.65
4	80 0.54 46.24 1.06	93 0.62 53.76 1.26	173 1.16
5	70 0.47 43.48 0.93	91 0.61 56.52 1.23	161 1.08
6	72 0.48 54.96 0.96	59 0.40 45.04 0.80	131 0.88
7	142 0.95 40.57 1.89	208 1.40 59.43 2.82	350 2.35
Total	7512 50.46	7376 49.54	14888 100.00

Frequency Missing = 736

Statistics for Table of Smkng by Gend

Statistic	DF	Value	Prob
Chi-Square	6	29.8378	<.0001
Likelihood Ratio Chi-Square	6	29.9315	<.0001
Mantel-Haenszel Chi-Square	1	18.0026	<.0001
Phi Coefficient		0.0448	
Contingency Coefficient		0.0447	
Cramer's V		0.0448	

Effective Sample Size = 14888
Frequency Missing = 736

/*Chi-square analysis*/

/*Q39. = Vap = vaping, A. = 1 = Yes, B. = 2 = No in SAS code results*/

/*Q33. =Smkng = smoking, A. = 1 = 0 days, B. = 2 = 1.5 days, C. = 3 = 4 days, D. = 4 = 7.5 days, E. = 5 = 14.5 days, F. = 6 = 24.5 days, G. = 7 = 30 days in SAS code results*/

Frequency Percent Row Pct Col Pct	Table of Smkng by Vap		
	Vap		
Smkng	1	2	Total
1	5301 36.34 40.74 78.86	7712 52.87 59.26 98.04	13013 89.20
2	496 3.40 88.57 7.38	64 0.44 11.43 0.81	560 3.84
3	209 1.43 86.72 3.11	32 0.22 13.28 0.41	241 1.65
4	148 1.01 89.70 2.20	17 0.12 10.30 0.22	165 1.13
5	142 0.97 91.61 2.11	13 0.09 8.39 0.17	155 1.06
6	119 0.82 95.97 1.77	5 0.03 4.03 0.06	124 0.85
7	307 2.10 93.03 4.57	23 0.16 6.97 0.29	330 2.26
Total	6722 46.08	7866 53.92	14588 100.00
Frequency Missing = 1036			

Statistics for Table of Smkng by Vap

Statistic	DF	Value	Prob
Chi-Square	6	1389.3711	<.0001
Likelihood Ratio Chi-Square	6	1548.5385	<.0001
Mantel-Haenszel Chi-Square	1	1002.6928	<.0001
Phi Coefficient		0.3086	
Contingency Coefficient		0.2949	
Cramer's V		0.3086	

Effective Sample Size = 14588
Frequency Missing = 1036

The Chi-square value (<.0001) determines that there is some significance in our determination that race may play a significant role in the probability of the chances that a teenager will participate in smoking cigarettes. Unfortunately, the results for this test may be a bit skewed due to the fact that 90% of the teenagers choose not to smoke. The same estimation of skewness can apply to ethnicity as well, even though the Chi-square p value indicates a significant value (.0074). Again, the Chi-square p value (<.0001) indicates that "vaping" may play a significant role in whether a teenager would choose to smoke cigarettes. Although, due to the skewed numbers of the 91% of teenagers that chose not to don't smoke in the first place, it is hard to say if the data in this case can be trusted.

The regression model could be utilized to inform us of the impact predictive variables such as race, ethnicity, gender, and history of vaporizing would have on the dependent variable, which would be a measure of smoking consumption.

/*Multiple regression analysis, newrace = race, Gend = gender, Ethn = ethnicity, Vap = vaporizing*/

The REG Procedure
Model: MODEL1
Dependent Variable: Smkng

Number of Observations Read	15624
Number of Observations Used	11092
Number of Observations with Missing Values	4532

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	1222.90042	305.72510	238.67	<.0001
Error	11087	14202	1.28095		
Corrected Total	11091	15425			

Root MSE	1.13179	R-Square	0.0793
Dependent Mean	1.34683	Adj R-Sq	0.0789
Coeff Var	84.03378		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	1.87923	0.07612	24.69	<.0001
newrace	1	0.02983	0.00925	3.23	0.0013
Ethn	1	0.16491	0.02649	6.22	<.0001
Gend	1	0.03618	0.02152	1.68	0.0927
Vap	1	-0.65409	0.02165	-30.21	<.0001

As we can see with the results of running a multiple regression analysis test there can be a great deal of information that can be extracted the survey questions. First of all, the R^2 test represents the percentage of the dependent variable that is attributed by the independent or predicted variables. Therefore, with an R2 estimate of 0.0662, the determination of whether a teenager would be experiment with smoking at 7% will be attributed to race,

ethnicity, gender and vaping. All independent variables expect gender (p value=.9502) play a significant role in the determination of whether a teen will smoke or not. Also, vaporizing indicates that it would have a negative effect on the determination of whether a teen will smoke. Meaning, the more a teenager "vapes" the less likely they are to experiment with smoking cigarettes.

Conclusion:

Given the information we were able to come up with, it appears that smoking does not seem to be gender specific, seeing that our significance test for gender (p value = 0.9502) in our regression test was greater than 0.05. Vaporizing appears to reduce the likelihood of whether a teenager will take up smoking. However, due to the fact that "vaping" is nothing more that inhaling liquefied nicotine it is debatable that using a vaporizer is any less hazardous than smoking cigarettes. Thus, with the data that was collected of over 14,000 teens, the fact that 91% of them don't smoke at all proved to good and bad. Bad, for the fact that without a more even distribution of smoking habits it is difficult to say if the information that was calculated is significantly accurate. Good, because only 9% of our teen population are experimenting with smoking cigarettes. This information is encouraging. With enough research and asking the right questions that number has the potential to increase, once we can determine what keeps the other 91% from smoking and expose that information to the other 9%.

References:

1.) Sopori, M. (2002). Effects of Cigarette Smoke on the Immune System. Nature Review Immunology, 2, 372-377. doi: 10.1038/nri803

2.) Gottsegen, J. (1940) *Tobacco: A Study of Its Consumption in the United States*. Pitman Publishing Corperation, New York, 22, 4, 808-809.

3.) Gilman, S. and Xun, Z. (2004). Smoke: A Global History of Smoking. Chicago, IL: University of Chicago Press.

4.) Doll, R. amd Hill, A. (1950). Smoking and Carcinoma of the Lung. Preliminary Report. British Medical Journal. **2**(4682): 739–48. doi:10.1136/bmj.2.4682.739

5.) U.S. Department of Health and Human Services. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General, 2014. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2014.

6.) Center of Disease Control (CDC). (2015). Youth Risk Behavior Surveillance Systerm (YRBSS) [data survey].

7.) National Institute of Drug Abuse (NIDA). Tobacco, Nicotine, and E-Cigarettes. Last edited January 2018, bhttps://www.drugabuse.gov/publications/tobacco-nicotine-e-cigarettes/nicotine-addictive

8.) Wikipedia. History of Tobacco. Last edited on 29 April 2018, https://en.wikipedia.org/wiki/History_of_tobacco

SAS Code:

```
dm 'log' clear;
libname jw 'f:\yrbs';

data jw.smoking;
 set jw.yrbs2015;

data jw.smokingj;
set jw.smoking;
if q33='A' then dcs=0;
if q33='B' then dcs=1.5;
if q33='C' then dcs=4;
if q33='D' then dcs=7.5;
if q33='E' then dcs=14.5;
if q33='F' then dcs=24.5;
if q33='G' then dcs=30;
if q1='A' then dcs=12;
if q1='B' then dcs=13;
if q1='C' then dcs=14;
if q1='D' then dcs=15;
if q1='E' then dcs=16;
if q1='F' then dcs=17;
if q1='G' then dcs=18;
if q2='A' then dcs=1;
if q2='B' then dcs=2;
if q4='A' then dcs=1;
if q4='B' then dcs=2;
if q5='A' then newrace=1;
if q5='B' then newrace=2;
if q5='C' then newrace=3;
if q5='D' then newrace=4;
if q5='E' then newrace=5;
if q39='A' then dcs=1;
if q39='B' then dcs=2;
if q32='A' then dcs=0;
if q32='B' then dcs=8;
if q32='C' then dcs=9.5;
if q32='D' then dcs=11.5;
if q32='E' then dcs=13.5;
if q32='F' then dcs=15.5;
if q32='G' then dcs=17;
if q33 = '1' then newQ = 1;
if q33 = '2' then newQ = 2;
if q33 = '3' then newQ = 3;
if q33 = '4' then newQ = 4;
if q33 = '5' then newQ = 5;
if q33 = '6' then newQ = 6;
if q33 = '7' then newQ = .;
if q33 = '9' then newQ = .;
if q1 = '12' then Qfst = 1;
if q1 = '13' then Qfst = 2;
if q1 = '14' then Qfst = 3;
if q1 = '15' then Qfst = 4;
if q1 = '16' then Qfst = 5;
if q1 = '17' then Qfst = 6;
```

```
if q1 = '18' then Qfst = 7;
if q2 = '1' then Qsnd = 1;
if q2 = '2' then Qsnd = 2;
if q4 = '1' then Qfth = 1;
if q4 = '2' then Qfth = 2;
if q39 = '1' then Qthn = 1;
if q39 = '2' then Qthn = 2;
if q32 = '0' then Qtht = 1;
if q32 = '8' then Qtht = 2;
if q32 = '9.5' then Qtht = 3;
if q32 = '11.5' then Qtht = 4;
if q32 = '13.5' then Qtht = 5;
if q32 = '15.5' then Qtht = 6;
if q32 = '17' then Qtht = 7;
run;

proc freq data=jw.smokingj;
table Qfst Qsnd Qfth Qtht newQ Qthn newrace;
 run;

proc print data=jw.smokingj;
var q1 q2 q4 q5 q32 q33 q39;
 run;


proc contents data=jw.smokingj;
run;

proc print data=jw.smokingj;


proc reg data = jw.smokingj;
model newQ = newrace Qsnd Qfth Qthn;
run;

proc freq data = jw.smokingj;
table newQ * newrace /chisq;
run;

proc freq data = jw.smokingj;
table newQ * Qsnd /chisq;
run;

proc freq data = jw.smokingj;
table newQ * Qthn /chisq;
run;

proc freq data = jw.smokingj;
table newQ * Qfth /chisq;
run;
```