



Article

# Machine Learning-Based Nicotine Addiction Prediction Models for Youth E-Cigarette and Waterpipe (Hookah) Users

Jeeyae Choi <sup>1,\*</sup>, Hee-Tae Jung <sup>2</sup>, Anastasiya Ferrell <sup>1</sup>, Seoyoon Woo <sup>1</sup> and Linda Haddad <sup>1</sup>

<sup>1</sup> School of Nursing, University of North Carolina, Wilmington, NC 28403, USA; ferrella@uncw.edu (A.F.); woos@uncw.edu (S.W.); haddadl@uncw.edu (L.H.)

<sup>2</sup> College of Information and Computer Science, University of Massachusetts, Amherst, MA 01002, USA; heetae@umass.edu

\* Correspondence: choij@uncw.edu; Tel.: +1-910-962-2487

**Abstract:** Despite the harmful effect on health, e-cigarette and hookah smoking in youth in the U.S. has increased. Developing tailored e-cigarette and hookah cessation programs for youth is imperative. The aim of this study was to identify predictor variables such as social, mental, and environmental determinants that cause nicotine addiction in youth e-cigarette or hookah users and build nicotine addiction prediction models using machine learning algorithms. A total of 6511 participants were identified as ever having used e-cigarettes or hookah from the National Youth Tobacco Survey (2019) datasets. Prediction models were built by Random Forest with ReliefF and Least Absolute Shrinkage and Selection Operator (LASSO). ReliefF identified important predictor variables, and the Davies–Bouldin clustering evaluation index selected the optimal number of predictors for Random Forest. A total of 193 predictor variables were included in the final analysis. Performance of prediction models was measured by Root Mean Square Error (RMSE) and Confusion Matrix. The results suggested high performance of prediction. Identified predictor variables were aligned with previous research. The noble predictors found, such as ‘witnessed e-cigarette use in their household’ and ‘perception of their tobacco use’, could be used in public awareness or targeted e-cigarette and hookah youth education and for policymakers.

**Keywords:** smoking; machine learning; nicotine addiction; youth e-cigarette use; youth waterpipe use



**Citation:** Choi, J.; Jung, H.-T.; Ferrell, A.; Woo, S.; Haddad, L. Machine Learning-Based Nicotine Addiction Prediction Models for Youth E-Cigarette and Waterpipe (Hookah) Users. *J. Clin. Med.* **2021**, *10*, 972. <https://doi.org/10.3390/jcm10050972>

Academic Editor:  
Venerando Rapisarda

Received: 22 January 2021  
Accepted: 22 February 2021  
Published: 2 March 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Electronic cigarette (e-cigarette) use in youth in the U.S. has increased rapidly in the past ten years. The U.S. Surgeon General and the FDA Commissioner have declared it as an epidemic [1,2]. A survey conducted between 2014 and 2018 showed the rate of e-cigarette use in young adults (18–24 years) to be 7.6%; this is interpreted as approximately one in five adolescents currently using e-cigarettes [3,4]. These numbers are truly alarming and tend to be higher in the state-level surveys. Nicotine in e-cigarettes is highly addictive [5]. Its addictive nature is reported to be as much as that of cocaine and heroin [6]. Most e-cigarettes and other vaping products contain some level of nicotine [7]. For example, individual pods of Juul, one of the most popular vaping products with youth, contain nicotine concentrations equivalent to a pack of cigarettes [8]. The earlier people start nicotine use, the harder it is to quit later in life. This is especially true for e-cigarette users under the age of 25 as their brains and nervous systems are not fully developed and are more vulnerable to nicotine addiction [9]. Recent studies have shown that in addition to highly toxic nicotine, vaping products also contain carcinogens, heavy metals, and other harmful substances [10].

While e-cigarettes and similar electronic nicotine delivery systems are promoted as less harmful than smoking tobacco products, they are far from harmless. Over 2600 lung-injury cases have been identified as E-Cigarette or Vaping Product Use-Associated Lung Injuries (also known as EVALI) in the U.S. (August 2019–January 2020) [11]. Of those

cases, >15% were 17 years old or younger [12]. Vaping products and hookah smoking cause nicotine addiction in non-smokers and various health issues (e.g., may lead to cancer, changes in cardiac functions, respiratory problems) [13]. Ingestion or direct contact with these products is toxic and can be fatal [14,15].

E-cigarettes' threat to health comes from their common ingredients (e.g., oxidizing agents, carcinogens, heavy metals) [16]. For starters, vegetable glycerol and propylene glycol, which are frequent additives in e-liquids, can create additional harmful chemicals—aldehydes—when heated [16]. Exposure to these chemicals can increase the gut microbiome's vulnerability to leaking, leading to loss of essential molecules and microbes. Depletion of this microbiome can promote chronic inflammation, which in turn puts the body at risk of cancers and buildup of plaque around artery walls [17]. Eventually, this buildup can lead to angina, stroke, and myocardial infarction [18–20]. Although there is not enough evidence to demonstrate e-cigarettes' direct and long-term damage to the cardiovascular system, e-cigarette use has been associated with increased heart rate, blood pressure, thrombosis, inflammation, oxidative stress, and vascular dysfunction [21,22]. Additionally, acrolein and carbonyls in e-liquids are high enough to lead to cardiovascular toxicity. Even if none of these solvents were present in e-liquids, nicotine by itself is known to increase cardiac arrhythmias [23]. E-cigarettes' effect on pulmonary health reveals users' further vulnerability from this product.

Flavorings in e-cigarettes are toxic and can permanently damage human lungs, causing disbalance in the levels of the lungs' inflammation proteins and the killing of lung cells altogether [24,25]. Heavy metals (e.g., nickel, lead, cadmium) in e-cigarette liquids are additional damaging irritants to the lungs and the airway overall. Although these metals are present in trace amounts, they can change human DNA and cause cell mutation, which in turn can lead to cancer [21]. Constituents of e-cigarette liquids such as pulegone, acrolein, formaldehyde, tobacco-specific nitrosamines, and oxidants can also raise the risk of cancer [21]. E-cigarettes have been on the U.S. market for a little over a decade [26]. This limits the amount of longitudinal data available to show the long-term effects of this product. However, the known effects of e-cigarette's ingredients on their own should be sufficient to anticipate potential risks.

Recent studies show increased vulnerability to COVID-19 in adult and youth e-cigarette vapers and smokers [27–29]. Although there is no published evidence in the connection of hookah use and susceptibility to COVID-19, the social nature of this activity and the composition of a hookah device were noted by the World Health Organization (WHO) as potential risks of transmission [30]. While the evidence is still emerging on the health effects and safety of vaping, there is limited knowledge for e-cigarette and hookah cessation programs. Youth are often restricted from adult tobacco cessation programs or holistic interventions. In addition, prediction models based on identified predictor variables can be used for early intervention and prevention programs effectively. The aim of this study was to identify predictor variables such as social, mental, and environmental determinants that cause nicotine addiction in youth e-cigarette or hookah users and build nicotine addiction prediction models using machine learning algorithms.

## 2. Materials and Methods

We applied machine learning, which is one of the data mining techniques to process and analyze data. We used the publicly available 2019 National Youth Tobacco Survey (NYTS) dataset collected electronically and selected e-cigarette or hookah users who had tried the products at least once in their lifetime. The mode imputation for missing data and two machine learning algorithms, Random Forest with ReliefF and Least Absolute Shrinkage and Selection Operator (LASSO), were applied to identify nicotine addiction predictor variables and build prediction models. MATLAB 2020b (Natick, MA, USA) was used for training models and 10-fold cross-validation for evaluating their performance.

### 2.1. The National Youth Tobacco Survey (NYTS)

The NYTS was developed to collect necessary datasets for supporting the design, implementation, and evaluation of tobacco prevention and control programs nationally. This survey's datasets are more comprehensive than those in the Youth Risk Behaviors Surveillance Systems, especially on tobacco-related indicators for middle- and high-school students. They include: (1) tobacco use (e.g., cigarettes, cigars, smokeless tobacco, electronic cigarettes, hookahs, roll-your-own cigarettes, pipes, snus, dissolvable tobacco, bidis, and heated tobacco products), (2) exposure to secondhand tobacco smoke and e-cigarette aerosol, (3) smoking cessation, (4) minors' ability to purchase or obtain tobacco products, (5) knowledge and attitudes about tobacco, and (6) familiarity with pro-tobacco and anti-tobacco media messages [31]. The NYTS data were collected between 15 February 2019 and 24 May 2019. A total of 19,018 students completed the tablet-based or web-based questionnaires, yielding an 85.8% student response rate. The 2019 NYTS was reviewed and approved by the Office of Management and Budget, the contracted data collectors' institutional review board (IRB), and the Center for Disease Control and Prevention (CDC)'s IRB.

### 2.2. Data Preprocessing and Defining Labels

We removed irrelevant variables, such as student login and primary sampling unit numbers, and variables asking for the text form of answers (e.g., a brand name of cigarettes used). A total of 193 variables were included in the final analysis after one variable was used as a label—severity of nicotine addiction.

The missing data rate of 27.0% was calculated based on missing data ( $n = 339,770$ ) and total expected data ( $N = 1,256,623$ ). We used mode imputation for missing data. The mode imputation is an easy-to-use and straightforward way to replace missing values by filling in the mode [32].

We used the time to the first cigarette of the day (TTFC) tool to interpret the severity of subjects' nicotine dependence. TTFC is a strong predictor of quitting behavior and can independently predict the maintenance of quitting attempts by assessing the severity of nicotine addiction [33,34]. According to studies indicating "time to the first cigarette of the day", the following nicotine dependence variables were labeled as (1) heavy nicotine addiction when a response is within 5–30 min, (2) moderate nicotine addiction when it is longer than 30 min to 1 h, (3) light nicotine addiction when it is longer than 1 h but less than 24 h, and (4) non-nicotine addiction when a response is either no use of cigarette or using it rarely [33,35–38].

### 2.3. Machine Learning Algorithms and Feature Selection

#### 2.3.1. Least Absolute Shrinkage and Selection Operator (LASSO)

The LASSO is a regression analysis method. It selects variables and performs regularization to increase the prediction accuracy [39]. There are many advantages of using LASSO, especially as it excludes irrelevant variables to improve interpretation of the prediction models [40]. We applied LASSO to build the prediction models since it has been used for various types of data including from surveys [41,42].

#### 2.3.2. Random Forest with ReliefF Variable Selection

Random Forest is an ensemble algorithm that has a computational efficiency over larger datasets. This algorithm randomly selects a subset of variables and constructs many decision trees. Strengths of Random Forest are low bias, high variance, and low correlation between constructed trees [43–45]. Sixty-four decision trees were set in Random Forest, which was recommended in a study as the optimal number of trees [45].

The ReliefF method is an enhancement of the Relief algorithm, which is one of the best-known variable selection methods. Relief makes highly accurate and effective variable estimation. However, it shows low performance on incomplete data since its best estimation is attained through weighting. ReliefF is enhanced to overcome these limitations [46]. We used ReliefF to select variables and the Davies–Bouldin clustering evaluation index [47] to

decide the optimal number of variables. Then, the Random Forest algorithm was applied to build the prediction models with the optimal numbers of variables.

#### 2.4. Evaluation of Performance of Prediction Models

Root Mean Square Error (RMSE) is a standard way of measuring the errors of prediction models, and the Confusion Matrix is a specific table describing the performance of a classification model. We computed the RMSE to evaluate the performance of the prediction models and visualize the performance of developed prediction models in the Confusion Matrix.

### 3. Results

A total of 6511 participants (mean age =  $15.36 \pm 1.85$  years old) were identified as ever being e-cigarette or hookah users. Most of them were non-Hispanic ( $n = 4540$ , 69.73%) and white ( $n = 4504$ , 69.18%). They lived with someone who smokes cigarettes ( $n = 2131$ , 32.73%), e-cigarettes ( $n = 1560$ , 23.96%), or hookahs ( $n = 256$ , 3.92%). Two-thirds of participants indicated English as their first language ( $n = 4244$ , 65.18%) and did not have any disability ( $n = 4384$ , 74.24%). Table 1 shows the characteristics of participants in this study.

**Table 1.** Characteristics of study participants (N = 6511).

Variables/Categories		N (%)
Age	9	15 (0.23)
	10	1 (0.02)
	11	88 (1.35)
	12	379 (5.82)
	13	698 (10.72)
	14	932 (14.31)
	15	1078 (16.56)
	16	1190 (18.28)
	17	1323 (20.32)
	18	746 (11.46)
	19	56 (0.86)
	NA	5 (0.08)
Grade	6th	330 (5.07)
	7th	610 (9.37)
	8th	844 (12.96)
	9th	1058 (16.25)
	10th	1151 (17.68)
	11th	1242 (19.08)
	12th	1252 (19.23)
	NA	24 (0.37)
Gender	Male	3446 (52.93)
	Female	3042 (46.72)
	NA	23 (0.35)
Ethnicity	Non-Hispanic	4540 (69.73)
	Mexican	1052 (16.16)
	Puerto Rican	190 (2.92)
	Other Hispanic	90 (1.38)
Race	American Indian	554 (8.51)
	Asian	396 (6.08)
	Black	1194 (18.34)
	Hawaiian	261 (4.01)
	White	4504 (69.18)

**Table 1.** *Cont.*

Variables/Categories		N (%)
<b>Lives with someone who uses</b>	Cigarettes	2131 (32.73)
	Cigars	580 (8.91)
	Chewing tobacco	697 (10.70)
	E-cigarettes	1560 (23.96)
	Hookahs	256 (3.92)
	Pipes	168 (2.58)
	Snus	117 (1.80)
	Dissolvable tobacco	105 (1.61)
	Bidis	98 (1.51)
	Heated tobacco	131 (2.01)
No tobacco products	2964 (45.62)	
<b>English is their first language</b>	Yes	4244 (65.18)
	No	2106 (32.35)
	NA	161 (2.47)
<b>Has disability</b>	Yes	1482 (22.76)
	No	4384 (74.24)
	NA	195 (2.99)

NA = Not answered.

### 3.1. Evaluation of Performance of the Prediction Model

Table 2 is a summary of performance of prediction, and Figure 1 shows the Confusion Matrices of LASSO and Random Forest. RMSE was the average RMSE from ten trained prediction models. Accuracy was calculated using data in each Confusion Matrix. All correctly predicted numbers in diagonal cells were summed and divided by the total number of participants (N = 6511).

**Table 2.** Performance of prediction.

Algorithm	Root Mean Square Error (RMSE) (SD)	Accuracy
LASSO	0.7509 ( $\pm 0.0287$ )	0.6370
Random Forest	0.7436 ( $\pm 0.0401$ )	0.7342

Table 3 is a summary of accuracy comparison based on Confusion Matrices in classes (different severity of nicotine addiction groups). Since numbers in diagonal cells in Confusion Matrices indicate correct predictions in each class, we tabulated them for accuracy comparison between two machine learning algorithms. For example, we simply divided correct prediction numbers in Class 1 ( $n = 4480$ ) by the total number of participants (N = 6511).

**Table 3.** Comparison of accuracy based on Confusion Matrices (N = 6511).

Class	LASSO (%)	Random Forest (%)
1: Non-nicotine addicted	4480 (68.81)	5104 (78.39)
2: Lightly addicted	269 (4.13)	57 (0.88)
3: Moderately addicted	104 (1.60)	4 (0.06)
4: Heavily addicted	36 (0.55)	328 (5.04)



**Figure 1.** Confusion Matrices from LASSO (A) and Random Forest (B). Class 1 is a non-nicotine addicted group, Class 2 a lightly addicted group, Class 3 a moderately addicted group, and Class 4 a heavily addicted group. Note: Diagonal numbers from top left to bottom right are the numbers the model correctly identified a class.

### 3.2. Predictor Variables Used in the Nicotine Addiction Prediction Models

LASSO and Random Forest identified an average of 85 and 41 predictor variables to build the prediction models, respectively. All 41 predictor variables except one were also identified by LASSO. Table 4 shows the 40 predictor variables identified by both LASSO and Random Forest.

**Table 4.** Predictor variables used both by LASSO and Random Forest.

List of Predictor Variables Identified by both LASSO and Random Forest	
1	How old were you when you first tried cigarette smoking, even one or two puffs?
2	How strongly do you agree with the statement ‘All tobacco products are dangerous’?
3	During the past 30 days, on how many days did you smoke cigarettes?
4	When was the last time you smoked a cigarette, even one or two puffs?
5	During the past 30 days, what brand of cigarettes did you usually smoke?
6	Menthol cigarettes are cigarettes that taste like mint. During the past 30 days, were the cigarettes that you usually smoked menthol?
7	How old were you when you first tried smoking a cigar, cigarillo, or little cigar, even one or two puffs?
8	During the past 30 days, on how many days did you smoke cigars, cigarillos, or little cigars?
9	Have you ever used chewing tobacco, snuff, or dip, such as Copenhagen, Grizzly, Skoal, or Longhorn, even just a small amount?
10	How old were you when you used chewing tobacco, snuff, or dip for the first time?
11	In total, on how many days have you used e-cigarettes in your entire life?
12	During the past 30 days, on how many days did you use e-cigarettes?
13	During the past 30 days, what brand of e-cigarettes did you usually use?
14	Have you ever tried a “heated tobacco product”, even just one time?
15	During the past 30 days, on how many days did you use any tobacco product(s)?
16	During the past 30 days, have you had a strong craving or felt like you really needed to use a tobacco product of any kind?
17	Are you seriously thinking about quitting the use of all tobacco products?
18	During the past 12 months, how many times have you stopped using all tobacco products for one day or longer because you were trying to quit all tobacco products for good?
19	Are you seriously thinking about quitting cigarettes?
20	During the past 12 months, how many times have you stopped smoking cigarettes for one day or longer because you were trying to quit smoking cigarettes for good?



Table 4. Cont.

List of Predictor Variables Identified by both LASSO and Random Forest	
21	During the past 30 days, did anyone refuse to sell you any tobacco products because of your age?
22	During the past 30 days, how often did you see a warning label on a cigar, cigarillo, or little cigar package?
23	During the past 30 days, how often did you see a warning label on a package of hookah tobacco?
24	How much do you think people harm themselves when they smoke cigarettes some days but not every day?
25	During the past 7 days, on how many days did someone smoke tobacco products in your home while you were there?
26	During the past 7 days, on how many days did you ride in a vehicle when someone was smoking a tobacco product?
27	Because of a physical, mental, or emotional condition, do you have serious difficulty concentrating, remembering, or making decisions?
28	During the past 30 days, on the days you smoked, about how many cigarettes did you smoke per day?
29	During the past 30 days, on the days that you smoked, about how many cigars, cigarillos, or little cigars did you smoke per day?
30	If one of your best friends were to offer you a cigar, cigarillo, or little cigar, would you smoke it?
31	During the past 30 days, on how many days did you use chewing tobacco, snuff, or dip?
32	Which of the following best describes the type of e-cigarette you have used in the past 30 days? If you have used more than one type, please think about the one you use most often.
33	How old were you when you first tried smoking tobacco in a hookah or waterpipe, even one or two puffs?
34	Do you think that you will try smoking tobacco in a hookah or waterpipe soon?
35	Do you think you will smoke tobacco in a hookah or waterpipe in the next year?
36	If one of your best friends were to offer you a hookah or waterpipe with tobacco, would you try it?
37	How easy do you think it is for people your age to buy tobacco products online?
38	During the past 30 days, how often did you see a warning label on an e-cigarette package?
39	How much do you think people harm themselves when they use chewing tobacco, snuff, dip, snus, or dissolvable tobacco products, some days but not every day?
40	During the past 30 days, on how many days did you smell the vapor from someone who was using an e-cigarette in an indoor or outdoor public place?

#### 4. Discussion

We identified predictor variables and built the nicotine addiction prediction models using machine learning algorithms for youth e-cigarette or hookah users. The models built by LASSO and Random Forest (RF) with ReliefF show high predictive performance. RMSE of both models were almost the same (LASSO: 0.7509 vs. RF: 0.7436). Since there are no standard ranges of good RMSE values, it was defined based on the study sample size ( $N = 6511$ ) and closer to zero. Thus, we interpreted that the RMSE values indicated high performance of prediction. Accuracy of the models by RF (0.7342) was higher than that of by LASSO (0.6370) overall. However, the models from each machine learning algorithm showed different strengths of prediction performance. The Confusion Matrix of LASSO showed a higher performance on predicting Class 2 (a lightly addicted group) and Class 3 (a moderately addicted group) than RF. Meanwhile, the Confusion Matrix of RF showed a higher performance on predicting Class 1 (a non-nicotine addicted group) and Class 4 (a heavily addicted group) than LASSO.

##### 4.1. Predictor Variables Aligned with the Literature

The predictor variables identified by both machine learning algorithms confirmed many socio-environmental determinants, such as peer pressure and brand name of tobacco products [48–51], as important predictor variables. Studies indicate there are several factors closely associated with youth nicotine addiction, including age of starting tobacco use, gender, grade level, ethnicity/race, living with current tobacco users, English not being the first language at home, and having a disability/mental health condition [42,52–58]. The results show that LASSO's 85 predictor variables include age of starting tobacco use, gender, grade level, English not being the first language at home, and having a disability or mental condition. However, RF only includes age of starting tobacco use and having a disability or mental health condition in its 41 predictor variables. Although current studies indicate that living with current tobacco users and English not being the first language at home are predictor variables [56,58], both LASSO and RF-based prediction models do not

use them in the models. These may be justified by half the survey participants not living with smokers and two-thirds of them speaking English as their first language at home.

#### 4.2. Predictor Variables Identified in Detail

Items 13 and 32 in Table 4 reflect the evidence that specific types and brands of e-cigarette products can determine higher nicotine dependence [59]. For example, pod-based e-cigarettes have been found to result in stronger nicotine addiction [60]. Compared to other tobacco products, e-cigarettes have a minimal requirement for a warning label [61]. Warnings of nicotine addiction on e-cigarette packaging can motivate users to quit [62] and it can be the case with government's recommended labels. However, warning labels provided by tobacco companies, on the other hand, are shown to be harder to read and understand [63]. Therefore, item 38 in Table 4 reflects the struggle between health warnings' reach of e-cigarette users and participants' motivation to quit this tobacco product. Alternatively, cues of a substance can trigger a person to use this substance. Response to these triggers can create a behavioral conditioning, which can lead to a reinforced nicotine addiction [64]. Therefore, when a participant frequently smells e-cigarette vapor (item 40), this person is likely to struggle avoiding e-cigarette use and is likely to progress to nicotine dependence. Overall, items listed in Table 4 provide helpful guidance in selection of variables for future research.

#### 4.3. Implication for Practice and Future Study

This study also confirms prior published evidence and will aid future investigators in a more educated use of variables. Tobacco companies are known to divide their specific product marketing to separate populations (e.g., menthol cigarettes, smokeless tobacco) [65,66]. E-cigarette marketing was long suspected of targeting white, middle-class, hipster youth and creating a separate culture of tobacco users [67,68]. Thus, it is not surprising that students who tried e-cigarettes at least once are predominately from white, non-Hispanic households that are either native to the U.S. or are acculturated to the English-speaking society. Although it was expected that e-cigarette use among youth is more prevalent in older adolescents based on prior 2011–2018 trends [3], what is most surprising is the similarity in the rates of e-cigarette experimentation among female and male students. Prior research shows e-cigarette use rates tend to be higher among self-identified male students than those who identify themselves as female [69]. Considering that exposure to e-cigarette marketing is significantly associated with adolescents' ever and current e-cigarette use [70,71], this change in gender differences may be accounted to tobacco companies' increased e-cigarette marketing [72]. Further research is needed to identify any connection between gender differences in e-cigarette marketing expenditures, adolescents' exposure, and their e-cigarette use. Prior studies have demonstrated the role modeling of tobacco use among family on adolescents' smoking and vaping [57,73]. Our results confirm that role modeling is still present among e-cigarette experimenters (e.g., more than half of these respondents (54.38%) reported living with a tobacco user). However, an interesting finding is that less than a quarter of participants witnessed e-cigarette use in their household (23.96%). This trend is worth investigating further. It would be important to distinguish whether participants' experimentation with e-cigarettes occurs due to the role modeling of tobacco use in general or whether it is related to participants' perception of their tobacco use as a safer or better alternative to that of their family members. Although e-cigarettes are promoted as safer tobacco alternatives, research has shown that e-cigarette use among youth is a gateway to smoking and can lead to higher frequency of smoking among youth [74,75].

#### 4.4. Limitations and Strengths

We were able to build machine learning-based nicotine addiction prediction models for middle and high schoolers using publicly available survey data. Although the survey showed a good return rate and the size of data was large, there were still limitations. Data were self-reported; validity of data is contingent upon respondents' honesty and



memory. The degree of under-reporting and over-reporting of information was unknown. This survey was cross-sectional, measuring responses at a single snapshot in time, hence overlooking the development of nicotine addiction and dependency behaviors as they develop over time. If other variable selections and machine learning algorithms were used, there may have been different predictor variables showing different optimal performance. Our results were based on responses from ever e-cigarette and hookah users, collected prior to the federal Tobacco 21 legislation [76] as the data from the 2020 NYTS was not yet released. It is possible that responses collected post-December 2019 and those during the COVID-19 outbreak in the U.S. may further alter this list of predictors. Nevertheless, the major strength of the study results is the predictor variables identified by the nicotine addiction prediction models that could be used to design and develop nicotine addiction early intervention and prevention programs for youth.

## 5. Conclusions

We successfully used two machine learning algorithms (LASSO and RF with ReliefF) to identify nicotine addiction predictor values and built high-performing nicotine addiction prediction models for youth e-cigarette and Hookah users.

Results of this study can guide the focus of future tobacco prevention programs and cessation studies. For example, participants' perceptions about the dangers of e-cigarette use indicate the need to expand adolescents' awareness about these products. Adolescents' preference for menthol-containing products can be an opportunity to include these products in the awareness campaign. Age at tobacco initiation highlights the point at which or immediately before which it is crucial to introduce the tobacco prevention programs. Experimentation with heated tobacco products reveals adolescents' continued fascination with novel tobacco products and new technology, which implies adolescents' potential vulnerability to these products and the possibility of heated tobacco products' displacement of currently popular electronic cigarettes. The addition of these products to prevention programs and their stricter control by the FDA may help prevent the perpetuation of the current e-cigarette epidemic among youth.

Success and number of attempts in cessation highlight the need for the expansion and reconsideration of tobacco cessation help for underage groups. Adolescents' access to tobacco products, their exposure to warning labels, and proximity to others using tobacco products show the importance of stricter tobacco control in these areas. Besides, results may be used in targeted e-cigarette and hookah youth education and by policymakers in kindergarten to 12th grade education.

**Author Contributions:** Conceptualization, J.C., H.-T.J., and A.F.; methodology, J.C., H.-T.J., A.F., and S.W.; software, H.-T.J.; validation, J.C. and H.-T.J.; formal analysis, J.C., H.-T.J., and A.F.; investigation, J.C., H.-T.J., and A.F.; resources, J.C., A.F., and S.W.; data curation, J.C. and H.-T.J.; writing—original draft preparation, J.C.; writing—review and editing, J.C., H.-T.J., A.F., S.W., and L.H.; visualization, J.C. and H.-T.J.; supervision, J.C. and L.H.; project administration, J.C., H.-T.J., and A.F.; funding acquisition, J.C., H.-T.J., and A.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Irwin Belk Professor Faculty Mentoring Funds at the University of North Carolina Wilmington.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study, due to survey data used in this study are publicly available.

**Informed Consent Statement:** Patient consent was waived due to data used in this study are publicly available.

**Data Availability Statement:** Publicly available datasets were analyzed in this study. This data can be found here: [https://www.cdc.gov/tobacco/data\\_statistics/surveys/nyts/index.htm](https://www.cdc.gov/tobacco/data_statistics/surveys/nyts/index.htm) (accessed on 22 January 2021).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Center for Disease Control and Prevention (CDC). Surgeon General's Advisory on E-cigarette Use of Among Youth. Available online: [https://www.cdc.gov/tobacco/basic\\_information/e-cigarettes/surgeon-general-advisory/index.html](https://www.cdc.gov/tobacco/basic_information/e-cigarettes/surgeon-general-advisory/index.html) (accessed on 1 September 2019).
2. Food and Drug Administration (FDA). Statement from FDA Commissioner Scott Gottlieb, M.D., on New Steps to Address Epidemic of Youth E-Cigarette Use. Available online: <https://www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-new-steps-address-epidemic-youth-e-cigarette-use> (accessed on 1 September 2019).
3. Gentzke, A.; Creamer, M.; Cullen, K.; Ambrose, B.K.; Willis, G.; Jamal, A.; King, B. Vital signs: Tobacco product use among middle and high school students—United States, 2011–2018. *Morb. Mortal. Wkly. Rep.* **2019**, *68*, 157. [CrossRef]
4. Williford, J.; Zablotsky, B.; Zelaya, C. QuickStats: Percentage of adults aged 18–24 years who currently smoke cigarettes or who currently use electronic cigarettes, by year—National health interview survey, United States, 2014–2018. *Morb. Mortal. Wkly. Rep.* **2019**, *68*, 39.
5. National Institute on Drug Abuse (NIDA). *Tobacco, Nicotine, and E-Cigarettes Research Report*; National Institute of Health (NIH): Bethesda, MD, USA, 2020.
6. U.S. Department of Health and Human Services. Nicotine addiction: Past and present. In *How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General*; Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion: Atlanta, GA, USA, 2010.
7. Romberg, A.R.; Lo, E.J.M.; Cuccia, A.F.; Willett, J.G.; Xiao, H.; Hair, E.C.; Vallone, D.M.; Marynak, K.; King, B.A. Patterns of nicotine concentrations in electronic cigarettes sold in the United States 2013–2018. *Drug Alcohol Depend.* **2019**, *203*, 1–7. [CrossRef]
8. Fadus, M.C.; Smith, T.T.; Squeglia, L.M. The rise of e-cigarettes, pod mod devices, and JUUL among youth: Factors influencing use, health implications, and downstream effects. *Drug Alcohol Depend.* **2019**, *201*, 85–93. [CrossRef] [PubMed]
9. Center for Disease Control and Prevention (CDC). E-Cigarettes: Talk to Youth about the Risks: Nicotine Can Harm Developing Brains. Available online: <https://www.cdc.gov/tobacco/features/back-to-school/e-cigarettes-talk-to-youth-about-risks/index.html> (accessed on 12 January 2021).
10. Siegel, D.A.; Jatlaoui, T.C.; Koumans, E.H.; Kiernan, E.A.; Layer, M.; Cates, J.E.; Kimball, A. *Update: Interim Guidance for Health Care Providers Evaluating and Caring for Patients with Suspected E-Cigarette, or Vaping, Product Use Associated Lung Injury—United States, October 2019*; Center for Disease Control and Prevention: Atlanta, GA, USA, 2019; pp. 919–927.
11. Centers for Disease Control and Prevention (CDC). Outbreak of Lung Injury Associated with the Use of E-Cigarette, or Vaping, Products. Available online: [https://www.cdc.gov/tobacco/basic\\_information/e-cigarettes/severe-lung-disease.html#latest-information](https://www.cdc.gov/tobacco/basic_information/e-cigarettes/severe-lung-disease.html#latest-information) (accessed on 14 January 2021).
12. Ellington, S.; Salvatore, P.P.; Ko, J.; Danielson, M.; Kim, L.; Cyrus, A.; Rose, D. *Update: Product, Substance-Use, and Demographic Characteristics of Hospitalized Patients in a Nationwide Outbreak of E-Cigarette, or Vaping, Product Use-Associated Lung Injury—United States, August 2019–January 2020*; Center for Disease Control and Prevention: Atlanta, GA, USA, 2020; pp. 44–49.
13. Munzel, T.; Hahad, O.; Kuntic, M.; Keaney, J.F.; Deanfield, J.E.; Daiber, A. Effects of tobacco cigarettes, e-cigarettes, and waterpipe smoking on endothelial function and clinical outcomes. *Eur. Heart J.* **2020**, *41*, 4057–4070. [CrossRef] [PubMed]
14. Committee on the review of the health effects of electronic nicotine delivery systems. *Public Health Consequences of E-Cigarettes*; National Academies Press (US): Washington, DC, USA, 2018.
15. Leventhal, A.M.; Stone, M.D.; Andrabi, N.; Barrington-Trimis, J.; Strong, D.R.; Sussman, S.; Audrain-McGovern, J. Association of e-cigarette vaping and progression to heavier patterns of cigarette smoking. *JAMA* **2016**, *316*, 1918–1920. [CrossRef] [PubMed]
16. Cheng, T. Chemical evaluation of electronic cigarettes. *Tob. Control* **2014**, *23*, ii11–ii17. [CrossRef] [PubMed]
17. Sharma, A.; Lee, J.; Fonseca, A.G.; Moshensky, A.; Kothari, T.; Sayed, I.M.; Ibeawuchi, S.-R.; Pranadinata, R.F.; Ear, J.; Sahoo, D. E-cigarettes compromise the gut barrier and trigger inflammation. *Iscience* **2021**, *24*, 1–15. [CrossRef] [PubMed]
18. Ran, Y.; Wang, Y.; Zhu, M.; Wu, X.; Malhotra, A.; Lei, X.; Zhang, F.; Wang, X.; Xie, S.; Zhou, J. Higher plaque burden of middle cerebral artery is associated with recurrent ischemic stroke: A quantitative magnetic resonance imaging study. *Stroke* **2020**, *51*, 659–662. [CrossRef] [PubMed]
19. Weber, C.; Deseive, S.; Brim, G.; Stocker, T.; Broersen, A.; Kitslaar, P.; Martinoff, S.; Massberg, S.; Hadamitzky, M.; Hausleiter, J. Coronary plaque volume and predictors for fast plaque progression assessed by serial coronary CT angiography—A single-center observational study. *Eur. J. Radiol.* **2020**, *123*, 108805. [CrossRef] [PubMed]
20. Vahed, S.Z.; Barzegari, A.; Zuluaga, M.; Letourneur, D.; Pavon-Djavid, G. Myocardial infarction and gut microbiota: An incidental connection. *Pharmacol. Res.* **2018**, *129*, 308–317. [CrossRef]
21. Seiler-Ramadas, R.; Sandner, I.; Haider, S.; Grabovac, I.; Dorner, T.E. Health effects of electronic cigarette (e-cigarette) use on organ systems and its implications for public health. *Wien. Klin. Wochenschr.* **2020**, *20*, 1–8. [CrossRef]
22. Middlekauff, H.R. Cardiovascular impact of electronic-cigarette use. *Trends Cardiovasc. Med.* **2020**, *30*, 133–140. [CrossRef]
23. Bhatnagar, A. E-cigarettes and cardiovascular disease risk: Evaluation of evidence, policy implications, and recommendations. *Curr. Cardiovasc. Risk Rep.* **2016**, *10*, 1–10. [CrossRef]
24. Ween, M.P.; Moshensky, A.; Thredgold, L.L.; Bastian, N.A.; Hamon, R.; Badieli, A.; Nguyen, P.T.; Herewane, K.; Jersmann, H.; Bojanowski, C.M. E-cigarettes and health risks: More to the flavour than just the name. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **2020**. [CrossRef] [PubMed]

25. Jabba, S.V.; Diaz, A.N.; Erythropel, H.C.; Zimmerman, J.B.; Jordt, S.-E. Chemical adducts of reactive flavor aldehydes formed in e-cigarette liquids are cytotoxic and inhibit mitochondrial function in respiratory epithelial cells. *Nicotine Tob. Res.* **2020**, *22*, S25–S34. [[CrossRef](#)] [[PubMed](#)]
26. Gholap, V.; Halquist, M.S. Historical perspective of proactive and reactive regulations of E-cigarettes to combat nicotine addiction. *J. Addict. Med.* **2020**, *14*, 443–445. [[CrossRef](#)] [[PubMed](#)]
27. Gaiha, S.M.; Cheng, J.; Halpern-Felsher, B. Association between youth smoking, electronic cigarette use, and COVID-19. *J. Adolesc. Health* **2020**, *67*, 519–523. [[CrossRef](#)] [[PubMed](#)]
28. Hopkinson, N.S.; Rossi, N.; El-Sayed Moustafa, J.; Laverty, A.A.; Quint, J.K.; Freidin, M.; Visconti, A.; Murray, B.; Modat, M.; Ourselin, S.; et al. Current smoking and COVID-19 risk: Results from a population symptom app in over 2.4 million people. *Thorax* **2021**. [[CrossRef](#)] [[PubMed](#)]
29. Li, D.; Croft, D.P.; Ossip, D.J.; Xie, Z. The association between statewide vaping prevalence and COVID-19. *Prev. Med. Rep.* **2020**, *20*, 101254. [[CrossRef](#)] [[PubMed](#)]
30. World Health Organization (WHO). Tobacco and Waterpipe Use Increases the Risk of COVID-19. Available online: <http://www.emro.who.int/tfi/know-the-truth/tobacco-and-waterpipe-users-are-at-increased-risk-of-covid-19-infection.html> (accessed on 12 January 2021).
31. Office on Smoking and Health. *2019 National Youth Tobacco Survey: Methodology Report*; National Center for Chronic Disease Prevention and Health Promotion: Atlanta, GA, USA, 2019.
32. Peng, Q. *Comparison between Denoising Autoencoders and Random Forest for Imputation of Mixed Data from Electronic Medical Records*; University of California Los Angeles: Los Angeles, CA, USA, 2018.
33. Borland, R.; Yong, H.H.; O'Connor, R.J.; Hyland, A.; Thompson, M.E. The reliability and predictive validity of the heaviness of smoking index and its two components: Findings from the international tobacco control four country study. *Nicotine Tob. Res.* **2010**, *12* (Suppl. 1), S45–S50. [[CrossRef](#)]
34. Branstetter, S.A.; Muscat, J.E.; Mercincavage, M. Time to first cigarette: A potential clinical screening tool for nicotine dependence. *J. Addict. Med.* **2020**, *14*, 409–414. [[CrossRef](#)]
35. Foulds, J.; Veldheer, S.; Yingst, J.; Hrabovsky, S.; Wilson, S.J.; Nichols, T.T.; Eissenberg, T. Development of a questionnaire for assessing dependence on electronic cigarettes among a large sample of ex-smoking E-cigarette users. *Nicotine Tob. Res.* **2015**, *17*, 186–192. [[CrossRef](#)]
36. Liu, G.; Wasserman, E.; Kong, L.; Foulds, J. A comparison of nicotine dependence among exclusive E-cigarette and cigarette users in the PATH study. *Prev. Med.* **2017**, *104*, 86–91. [[CrossRef](#)]
37. Rhoades, D.A.; Comiford, A.L.; Dvorak, J.D.; Ding, K.; Hopkins, M.; Spicer, P.; Wagener, T.L.; Doescher, M.P. Vaping patterns, nicotine dependence and reasons for vaping among American Indian dual users of cigarettes and electronic cigarettes. *BMC Public Health* **2019**, *19*, 1211. [[CrossRef](#)] [[PubMed](#)]
38. Vogel, E.A.; Prochaska, J.J.; Rubinstein, M.L. Measuring e-cigarette addiction among adolescents. *Tob. Control* **2020**, *29*, 258–262. [[CrossRef](#)]
39. Friedman, J.H.; Hastie, T.; Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **2010**, *33*, 1. [[CrossRef](#)]
40. Paper in Business Analytics Feature Selection Using LASSO. Available online: [https://beta.vu.nl/nl/Images/werkstuk-fonti\\_tcm235-836234.pdf](https://beta.vu.nl/nl/Images/werkstuk-fonti_tcm235-836234.pdf) (accessed on 22 January 2021).
41. Signorino, C.S.; Kirchner, A. Using LASSO to model interactions and nonlinearities in survey data. *Surv. Pract.* **2018**, *11*, 1–10. [[CrossRef](#)]
42. McConville, K.S.; Breidt, F.J.; Lee, T.C.M.; Moisen, G.G. Model-assisted survey regression estimation with the Lasso. *J. Surv. Stat. Methodol.* **2017**, *5*, 131–158. [[CrossRef](#)]
43. Chen, W.; Xie, X.; Wang, J.; Pradhan, B.; Hong, H.; Bui, D.T.; Duan, Z.; Ma, J. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena* **2017**, *151*, 147–160. [[CrossRef](#)]
44. Kesler, S.R.; Rao, A.; Blayney, D.W.; Oakley-Girvan, I.A.; Karuturi, M.; Palesh, O. Predicting long-term cognitive outcome following breast cancer with pre-treatment resting state fmri and random forest machine learning. *Front. Hum. Neurosci.* **2017**, *11*, 555. [[CrossRef](#)]
45. Oshiro, T.; Santoro, P.; Baranauskas, J. How Many Trees in a Random Forest? In *Machine Learning and Data Mining in Pattern Recognition*; Perner, Ed.; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7376.
46. Peker, M.; Balli, S.; Sağbaşı, E.A. Predicting human actions using a hybrid of relief feature selection and kernel-based extreme learning machine. In *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications*; IGI Global: Hershey, PA, USA, 2020; pp. 307–325. [[CrossRef](#)]
47. Davies, D.L.; Bouldin, D.W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, *1*, 224–227. [[CrossRef](#)]
48. Bilgic, N.; Gunay, T. Evaluation of effectiveness of peer education on smoking behavior among high school students. *Saudi. Med. J.* **2018**, *39*, 74–80. [[CrossRef](#)]
49. Perks, S.; Armour, B.; Agaku, I. Cigarette brand preference and pro-tobacco advertising among middle and high school students—United States, 2012–2016. *Morb. Mortal. Wkly. Rep.* **2018**, *67*, 119–124. [[CrossRef](#)]

50. Ramji, R.; Nilsson, M.; Arnetz, B.; Wiklund, Y.; Arnetz, J. Taking a stand: An untapped strategy to reduce waterpipe smoking in adolescents. *Subst. Use Misuse* **2019**, *54*, 514–524. [[CrossRef](#)] [[PubMed](#)]
51. Treur, J.L.; Rozema, A.D.; Mathijssen, J.J.P.; van Oers, H.; Vink, J.M. E-cigarette and waterpipe use in two adolescent cohorts: Cross-sectional and longitudinal associations with conventional cigarette smoking. *Eur. J. Epidemiol.* **2018**, *33*, 323–334. [[CrossRef](#)]
52. Higgins, S.T.; Kurti, A.N.; Redner, R.; White, T.J.; Gaalema, D.E.; Roberts, M.E.; Doogan, N.J.; Tidey, J.W.; Miller, M.E.; Stanton, C.A.; et al. A literature review on prevalence of gender differences and intersections with other vulnerabilities to tobacco use in the United States, 2004–2014. *Prev. Med.* **2015**, *80*, 89–100. [[CrossRef](#)]
53. Yu, E.; Lippert, A.M. Race/ethnicity modifies the association between school prevalence of e-cigarette use and student-level use: Results from the 2014 US National Youth Tobacco Survey. *Health Place* **2017**, *46*, 114–120. [[CrossRef](#)]
54. Sharapova, S.; Reyes-Guzman, C.; Singh, T.; Phillips, E.; Marynak, K.L.; Agaku, I. Age of tobacco use initiation and association with current use and nicotine dependence among US middle and high school students, 2014–2016. *Tob. Control* **2020**, *29*, 49–54. [[CrossRef](#)] [[PubMed](#)]
55. Wang, T.; Neff, L.; Park-Lee, E.; Ren, C.; Cullen, K.; King, B. E-cigarette use among middle and high school students—United States, 2020. *Morb. Mortal. Wkly. Rep.* **2020**, *69*, 1310–1312. [[CrossRef](#)]
56. Alcala, H.E.; Albert, S.L.; Ortega, A.N. E-cigarette use and disparities by race, citizenship status and language among adolescents. *Addict. Behav.* **2016**, *57*, 30–34. [[CrossRef](#)] [[PubMed](#)]
57. Coleman, M.; Donaldson, C.; Crano, W.D.; Pike, J.; Stacy, A. Associations between family and peer e-cigarette use with adolescent tobacco and marijuana usage: A longitudinal path analytic approach. *Nicotine Tob. Res.* **2020**. [[CrossRef](#)]
58. Kwon, E.; Seo, D.C.; Lin, H.C.; Chen, Z. Predictors of youth e-cigarette use susceptibility in a U.S. nationally representative sample. *Addict. Behav.* **2018**, *82*, 79–85. [[CrossRef](#)] [[PubMed](#)]
59. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Population Health and Public Health Practice; Committee on the Review of the Health Effects of Electronic Nicotine Delivery Systems. *Public Health Consequences of E-Cigarettes*; National Academies Press (US): Washington, DC, USA, 2018. Available online: <https://www.ncbi.nlm.nih.gov/books/NBK507191/> (accessed on 12 February 2021).
60. Lee, S.J.; Rees, V.W.; Yossefy, N.; Emmons, K.M.; Tan, A.S.L. Youth and young adult use of pod-based electronic cigarettes from 2015 to 2019: A systematic review. *JAMA Pediatrics* **2020**, *174*, 714–720. [[CrossRef](#)] [[PubMed](#)]
61. Food and Drug Administration (FDA). Retailers: Chart of Required Warning Statements on Tobacco Product Packaging and Advertising. Available online: <https://www.fda.gov/tobacco-products/retail-sales-tobacco-products/retailers-chart-required-warning-statements-tobacco-product-packaging-and-advertising> (accessed on 12 January 2021).
62. Brewer, N.T.; Jeong, M.; Hall, M.G.; Baig, S.A.; Mendel, J.R.; Lazard, A.J.; Noar, S.M.; Kameny, M.R.; Ribisl, K.M. Impact of e-cigarette health warnings on motivation to vape and smoke. *Tob. Control* **2019**. [[CrossRef](#)]
63. Lee, H.Y.; Lin, H.C.; Seo, D.C.; Lohrmann, D.K. The effect of e-cigarette warning labels on college students' perception of e-cigarettes and intention to use e-cigarettes. *Addict. Behav.* **2018**, *76*, 106–112. [[CrossRef](#)]
64. Benowitz, N.L. Nicotine addiction. *N. Engl. J. Med.* **2010**, *362*, 2295–2303. [[CrossRef](#)] [[PubMed](#)]
65. Center for Disease Control and Prevention (CDC). African Americans and Tobacco Use. Available online: <https://www.cdc.gov/tobacco/disparities/african-americans/index.htm> (accessed on 12 January 2021).
66. Chaffee, B.; Couch, E.; Walsh, M. Smokeless tobacco in sport and use among adolescents. In *Center for Tobacco Control Research and Education*; University of California San Francisco: San Francisco, CA, USA, 2015.
67. Hess, C.A.; Antin, T.M.; Annechino, R.; Hunt, G. Perceptions of e-cigarettes among black youth in California. *Int. J. Environ. Res. Public Health* **2017**, *14*, 60. [[CrossRef](#)] [[PubMed](#)]
68. Jackler, R.; Chau, C.; Getachew, B.; Whitcomb, M.; Lee-Heidenreich, J.; Bhatt, A.; Kim-O'Sullivan, S.; Hoffman, Z.; Jackler, L.; Ramamurthi, D. JUUL advertising over its first three years on the market. In *Stanford Research into the Impact of Tobacco Advertising*; Stanford University School of Medicine: Stanford, CA, USA, 2019.
69. Kong, G.; Kuguru, K.E.; Krishnan-Sarin, S. Gender differences in U.S. Adolescent e-cigarette use. *Curr. Addict. Rep.* **2017**, *4*, 422–430. [[CrossRef](#)]
70. Ashford, K.; Rayens, E.; Wiggins, A.T.; Rayens, M.K.; Fallin, A.; Sayre, M.M. Advertising exposure and use of e-cigarettes among female current and former tobacco users of childbearing age. *Public Health Nurs.* **2017**, *34*, 430–436. [[CrossRef](#)]
71. Auf, R.; Trepka, M.J.; Selim, M.; Taleb, Z.B.; De La Rosa, M.; Cano, M.Á. E-cigarette marketing exposure and combustible tobacco use among adolescents in the United States. *Addict. Behav.* **2018**, *78*, 74–79. [[CrossRef](#)]
72. Collins, L.; Glasser, A.M.; Abudayyeh, H.; Pearson, J.L.; Villanti, A.C. E-cigarette marketing and communication: How e-cigarette companies market e-cigarettes and the public engages with e-cigarette information. *Nicotine Tob. Res.* **2019**, *21*, 14–24. [[CrossRef](#)] [[PubMed](#)]
73. Vuolo, M.; Staff, J. Parent and child cigarette use: A longitudinal, multigenerational study. *Pediatrics* **2013**, *132*, e568–e577. [[CrossRef](#)] [[PubMed](#)]
74. Barrington-Trimis, J.L.; Yang, Z.; Schiff, S.; Unger, J.; Cruz, T.B.; Urman, R.; Cho, J.; Samet, J.M.; Leventhal, A.M.; Berhane, K.; et al. E-cigarette product characteristics and subsequent frequency of cigarette smoking. *Pediatrics* **2020**, *145*. [[CrossRef](#)] [[PubMed](#)]

75. Owotomo, O.; Stritzel, H.; McCabe, S.E.; Boyd, C.J.; Maslowsky, J. Smoking intention and progression from e-cigarette use to cigarette smoking. *Pediatrics* **2020**, *146*. [[CrossRef](#)]
76. Food and Drug Administration (FDA). Tobacco 21. Available online: <https://www.fda.gov/tobacco-products/retail-sales-tobacco-products/tobacco-21> (accessed on 12 January 2021).