



Article

Exploring Impact of Marijuana (Cannabis) Abuse on Adults Using Machine Learning

Jeeyae Choi ^{1,*} , Joohyun Chung ² and Jeungok Choi ²

¹ School of Nursing, University of North Carolina, Wilmington, NC 28403, USA

² College of Nursing, University of Massachusetts, Amherst, MA 01002, USA; joohyunchung@umass.edu (J.C.); jeungokc@nursing.umass.edu (J.C.)

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

Abstract: Marijuana is the most common illicit substance globally. The rate of marijuana use is increasing in young adults in the US. The current environment of legalizing marijuana use is further contributing to an increase of users. The purpose of this study was to explore the characteristics of adults who abuse marijuana (20–49 years old) and analyze behavior and social relation variables related to depression and suicide risk using machine-learning algorithms. A total of 698 participants were identified from the 2019 National Survey on Drug Use and Health survey as marijuana dependent in the previous year. Principal Component Analysis and Chi-square were used to select features (variables) and mean imputation method was applied for missing data. Logistic regression, Random Forest, and K-Nearest Neighbor machine-learning algorithms were used to build depression and suicide risk prediction models. The results showed unique characteristics of the group and well-performing prediction models with influential risk variables. Identified risk variables were aligned with previous studies and suggested the development of marijuana abuse prevention programs targeting 20–29 year olds with a regular depression and suicide screening. Further study is suggested for identifying specific barriers to receiving timely treatment for depression and suicide risk.

Keywords: marijuana abuse; machine learning; substance abuse; health impact assessment



Citation: Choi, J.; Chung, J.; Choi, J. Exploring Impact of Marijuana (Cannabis) Abuse on Adults Using Machine Learning. *Int. J. Environ. Res. Public Health* **2021**, *18*, 10357. <https://doi.org/10.3390/ijerph181910357>

Academic Editor: Paul B. Tchounwou

Received: 28 August 2021

Accepted: 27 September 2021

Published: 1 October 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

Marijuana is the most commonly used illegal substance worldwide [1–5]. In recent decades, the rate of marijuana users has increased in young adults and pregnant women in the US [6]. The Substance Abuse Center for Behavioral Health Statistics and Quality reported that more than 11.8 million young adults used marijuana in 2018 [7]. Alarming signs of marijuana use among high schoolers were reported in 2019. Twenty-eight percent of second-year high school students indicated that they had used marijuana in the previous year and 35% of senior students indicated they used it during the year prior to 2019 [8]. The current climate of legalizing and decriminalizing marijuana use would contribute to further increments [3,9,10].

Marijuana use could lead to addiction, although it is often questioned and treated as not a serious addiction problem [11]. However, marijuana is highly addictive with serious adverse effects [3,12–14]. Studies indicate marijuana use is associated with various psychosocial and medical problems in young adults [10,15,16]. Berenson explained the association between marijuana, mental illness, and violence [17]. Zehra and colleagues' study showed repeated marijuana use could cause neurobiological changes in the brain, resulting in addictive behavior and withdrawal symptoms [18]. A study confirmed that cognitive function impairment is caused by marijuana addiction [19]. In addition, marijuana use lowers birth weight and has destructive impacts on fetal growth in pregnant women [20]. The American College of Obstetricians and Gynecologists (ACOG) announced that a nursing mother's marijuana use could affect a baby's brain development because it

can be accumulated in breast milk to high concentrations. ACOG is discouraging nursing mothers from using marijuana [21].

Studies indicated that cannabis (marijuana) use is strongly associated with depression and anxiety disorders [22–24]. A study of medical marijuana showed that the use of cannabis during adolescence is related to depression and suicidality in adult life and regular use of cannabis at a young age increases the risk of depression [25]. Suicidal behavior is affected by biological, psychiatric, and psychosocial determinants and substance use disorders. Suicidal ideation and attempts are associated with illicit substances such as cocaine, marijuana, and methamphetamine [26]. Research has also shown that marijuana aggregates mental health and influences repeated attempts of suicide in high schoolers [27,28].

Cannabidiol (CBD) and tetrahydrocannabinol (THC) are two fundamental chemicals in marijuana. CBD is shown to reduce anxiety, depression, and seizures, while THC is known to produce high sensation and intoxicating effects [8,29]. Medical marijuana (cannabis) is extracted from plants of the genus *cannabis* to treat specific diseases and symptoms using CBD's therapeutic effects [30,31]. For example, it shows improvement in seizures, Dravet syndrome, Lennox-Gastaut syndrome, pain, spasticity, inflammatory bowel disease, nausea, anorexia, depression, and anxiety [29,30,32]. However, researchers and clinicians remain concerned about the risks caused by medical marijuana and insist upon assessment of its impact on adults [18,33]. Bridgeman and Abazia acknowledged the beneficial effects of CBD in medical marijuana yet simultaneously emphasized how their efficacy in alleviating symptoms or disease was not well established [30]. There are increased concerns about the amount of THC that causes intoxication and addiction in medical marijuana due to no regulation [34,35]. Pure CBD has been proven to be safe in a research environment, which may not be possible to replicate in commercial CBD products [32]. Cascardo reported that many clinicians are not comfortable with their supervising role when there is limited clinical research available [14].

Marijuana is the most commonly used substance after alcohol and tobacco in the US [8]. Recent legalization of marijuana uses for medical or recreational purpose decreases the perception of risks of marijuana, expecting an increase of marijuana users. Currently, there is a lack of knowledge on how marijuana is affecting psychosocial and medical conditions of adults. The purpose of this study was to explore and analyze characteristics of adults who abuse marijuana (20–49 years old) on behaviors and social relations impacting depression and suicide risk using machine-learning algorithms.

2. Materials and Methods

2.1. Research Design

An exploratory machine-learning approach was adopted and applied to the National Survey on Drug Use and Health data sets. Specific machine-learning algorithms used in the study were logistic regression, Random Forest, and K- Nearest Neighbor (K-NN).

2.2. Data Source

We used the publicly available 2019 National Survey on Drug Use and Health (NSDUH) data sets collected electronically. The primary purpose of this survey was to measure the prevalence and find correlations between substance (illicit drugs, alcohol, and tobacco) use and mental health issues in the US. This was the 39th in a series and participants were citizens of the U.S., including those living on military bases, who were 12 years of age or older at the time of the survey.

The Center for Behavioral Health Statistics and Quality (CBHSQ) within the Substance Abuse and Mental Health Services Administration (SAMHSA) conducted the 2019 NSDUH. Data were collected from 50 states and the District of Columbia using an audio computer-assisted self-interviewing method. The data are publicly available and had already undergone a confidentiality review. The full analytic file of all participants was treated using a statistical disclosure limitation method called MASSC, which consists of the

following four major steps: (1) Micro Agglomeration, (2) optimal probabilistic Substitution, (3) optimal probabilistic Subsampling, and (4) optimal sampling weight Calibration. Personal Identifiable Information (PII) such as name, phone number, and address on the file was removed. A total of 56,136 participants completed the NSDUH in 2019, which was reviewed and approved by RTI's Institutional Review Board under federal regulation [36,37].

2.3. Data Collection

We selected participants who were 20 to 49 years old and classified as having marijuana dependency based on participants' responses related to marijuana uses and behaviors. We applied the mean imputation for missing data and analyzed them using SAS v9.4 and SAS Enterprise Miner v15.1 (SAS Institute Inc., Cary, NC, USA). The missing data rate was 21.0% based on missing data ($n = 48,957$) and expected data ($n = 233,132$).

To explore detailed characteristics of adult marijuana users' behavior and social relation, we focused on depression and suicide risk, mental health conditions. We identified and measured the effects of important variables using Principal Component Analysis and Chi-square. Depression and suicide risk prediction models were built using three machine-learning algorithms (logistic regression, Random Forest, and K-NN) and their performance was measured. The data were split into two parts for training (70%) and testing (30%) in the ratio of 70:30 for evaluating models' performances. Figure 1 shows the process of building the prediction models.

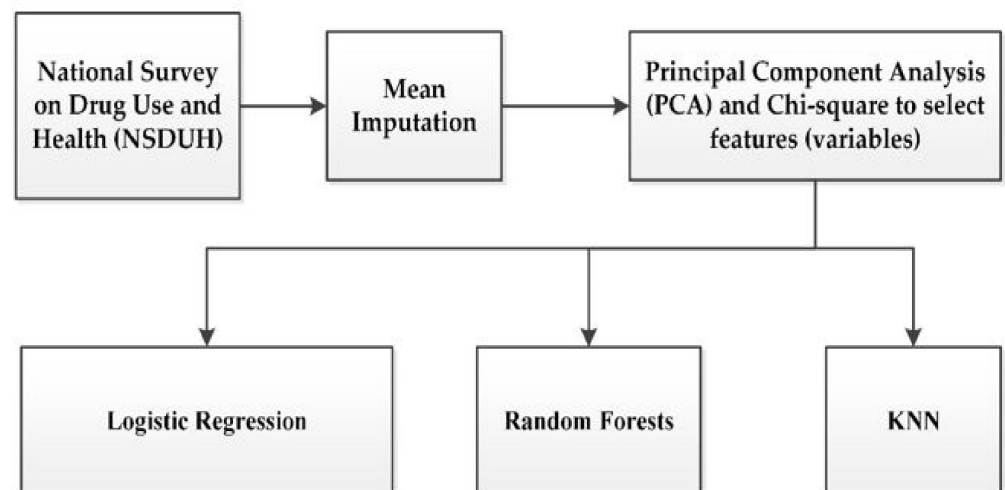


Figure 1. Overall process of building predictive models.

2.4. Data Preprocessing and Defining Labels

2.4.1. Feature (Variable) Selection

We removed irrelevant variables such as respondent identification and recoded variables. In a total of 2741 variables, 851 irrelevant variables were removed, and 1890 variables remained. Principal Component Analysis (PCA) and Chi-square [38] were used to select features (variables). PCA is one of the popular tools for feature selection and reduction of dimensionality in machine learning [39,40]. It removes correlated features, improves algorithm performance, and reduces overfitting [41].

2.4.2. Imputation for Missing Data

We used the mean imputation method for missing data because it is straightforward and preserves the mean of the observed data when the data are missing at random [42,43]. A total of 334 variables were reviewed, and 93 variables were included in the final model.

2.4.3. Labeling of Risk for Depression

The NSDUH assessed major depressive episodes with six questions, which were: (1) How often did you feel nervous in the past 30 days? (2) How often did you feel hopeless in the past 30 days? (3) How often did you feel restless in the past 30 days? (4) How often did you feel sad and that nothing could cheer you up in the past 30 days? (5) How often did you feel that everything was an effort in the past 30 days? and (6) How often did you feel down on yourself, no good or worthless in the past 30 days? When a participant responded as 'all the time' to any of these six variables, it was considered as they had depression in this study.

2.4.4. Labeling of Suicide Risk

Three suicide-related variables were used to label. They were: (1) Thought of suicide at any time in the past 12 months, that is from [the date 12 months prior] up to and including today, did you seriously think about trying to kill yourself? (2) Plan of suicide, during the past 12 months, did you make any plans to kill yourself? and (3) Attempt to suicide, during the past 12 months, did you try to kill yourself? Any participants indicating 'yes' to these questions were considered as risk for suicide in this study.

2.5. Machine Learning Algorithms

2.5.1. Logistic Regression

Logistic regression is the most widely used machine-learning algorithm for binary outcomes [44]. It is based on logistic function, one type of sigmoid function, which converts real-valued continuous inputs into categorical values. This algorithm assumes a linear relationship between the logarithm of the odds of the outcome and the predictors as equivalent [45].

2.5.2. Random Forest (RF)

Random Forest is an ensemble machine-learning algorithm that has a computational efficiency over larger data sets. This algorithm randomly selects a subset of variables and constructs many decision trees. Every individual tree splits its nodes to get a class prediction. Strengths of RF are low bias, high variance, and low correlation between constructed trees [46,47].

2.5.3. K-Nearest Neighbor (KNN)

KNN is a simple machine-learning algorithm that finds the closest data from a query data point depending on k value. Data are classified by the distance to others. It is a non-parametric algorithm and its calculation time is short due to no training period being needed [48,49].

2.6. Measurement of Prediction Model Performances

Sensitivity, Specificity, Accuracy, AUC (area under the curve), Precision, and F1 score were used to measure the performances of prediction models and were calculated based on confusion matrix. Sensitivity is referred to as the true positive rate. From a confusion matrix, the four terms were defined as (1) True positive (TP) = the number of cases correctly identified as the presence of the outcome, (2) False positive (FP) = the number of cases incorrectly identified as the presence of the outcome, (3) True negative (TN) = the number of cases correctly identified as the non-presence of the outcome, (4) False negative (FN) = the number of cases incorrectly identified as the non-presence of the outcome.

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \quad (1)$$

Specificity is the true negative rate (TNR), that is, the proportion of the actual cases that are correctly predicted as negative.

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP})$$

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

Accuracy is the ratio of correctly classified observations to the total number of observations.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

F1 score is a measure of the weighted average of precision and recall.

$$\text{F1 score} = 2 \times (\text{recall} \times \text{precision}) / (\text{recall} + \text{precision})$$

The Area Under the Curve (AUC) is the measure of the ability of a classifier to distinguish between classes and is used as a summary of the receiver operating characteristic (ROC) curve.

3. Results

A total of 698 participants were identified as marijuana dependent in the past year. The majority were between 20–29 years old ($n = 548$, 78.51%), had never been married ($n = 573$, 82.09%), employed ($n = 487$, 82.96%), and covered by any health insurance ($n = 560$, 81.04%). More than half were male ($n = 421$, 60.32%), half were NonHispanic white ($n = 364$, 52.15%), about one-third had some college credit but no degree ($n = 249$, 35.67%), and about one-quarter of them reported that their family income was more than \$75,000 ($n = 175$, 25.07%). Table 1 shows characteristics of participants in this study.

Table 1. Characteristics of participants who abuse marijuana ($n = 698$).

	Characteristics	n (n = 698)	%
Gender	Female	277	39.68
	Male	421	60.32
Age	20–29	548	78.51
	30–34	70	10.03
	35–49	80	11.46
Race	NonHispanic White	364	52.15
	NonHispanic Black or African American	100	14.33
	NonHispanic Native American/Alaska Native	19	2.72
	NonHispanic Native Hawaiian/Other Pacific Islander	3	0.43
	NonHispanic Asian	25	3.58
	NonHispanic more than one race	48	6.88
	Hispanic	139	19.91
	Education	5th–12th grade completed, no diploma	59
High school diploma/GED		204	29.23
Some college credit, but no degree		249	35.67
Associate degree		64	9.17
College graduate or higher		122	17.48
Family income	Less than \$10,000	90	12.89
	\$10,000–\$19,999	94	13.47
	\$20,000–\$29,999	71	10.17
	\$30,000–\$39,999	76	10.89
	\$40,000–\$49,999	88	12.61
	\$50,000–\$74,999	104	14.90
	\$75,000 or more	175	25.07

Table 1. Cont.

Characteristics		n (n = 698)	%
Marital status	Married	83	11.89
	Widowed	2	0.29
	Divorced or Separated	40	5.73
	Never Been Married	573	82.09
Employment	Employed	487	69.77
	Unemployed	100	14.33
	No response	111	15.90
Health Insurance	Covered by any Health Insurance	560	80.23
	Not covered	131	18.77
	No response	7	1.00

3.1. Features (Variables) Identified

The relative importance plots of the input variables were ranked by the Chi-square criteria and are depicted in Figures 2 and 3.

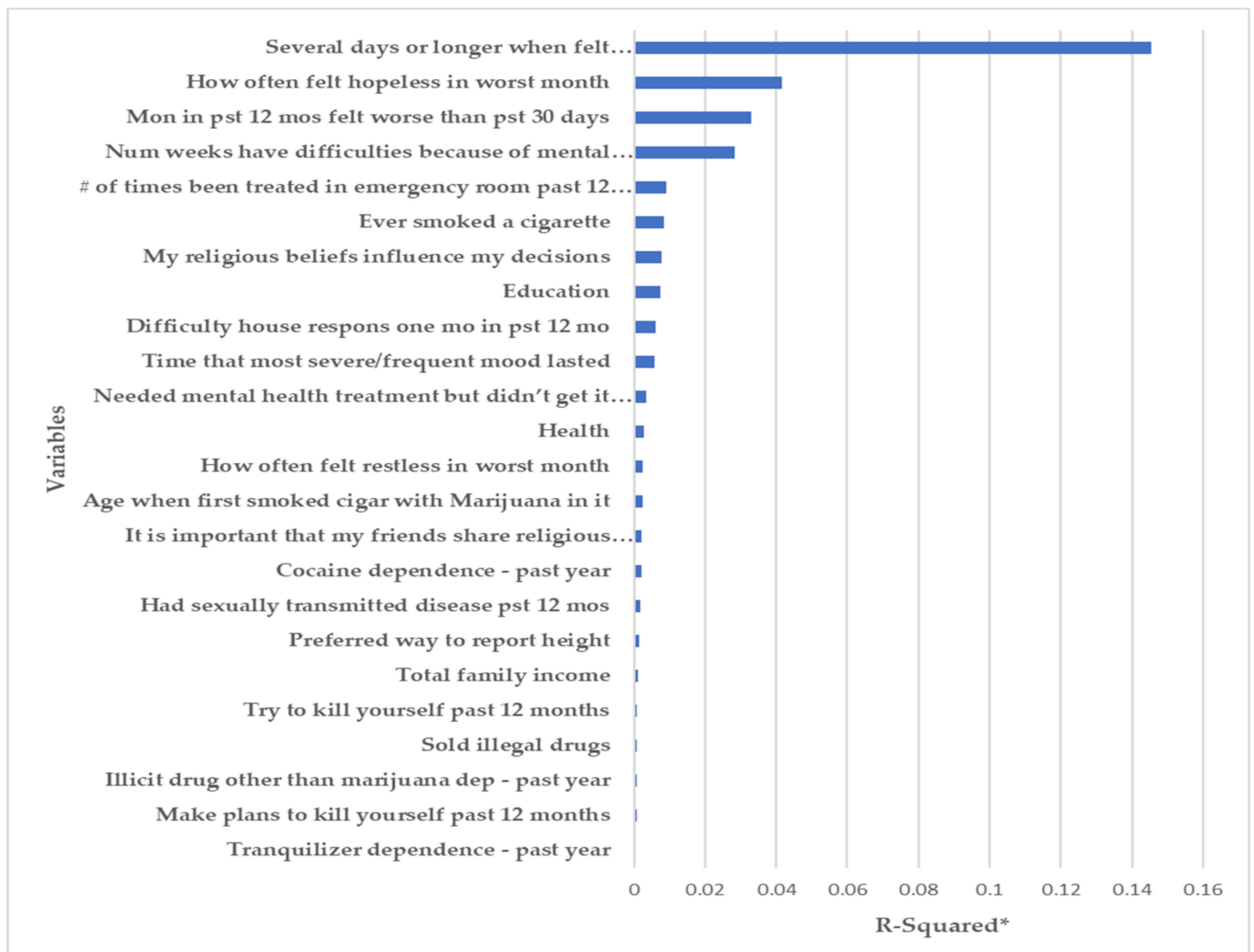


Figure 2. Importance plot for depression (* a larger R-squared value means that a variable explains a larger percentage of the variation in the outcome variable).

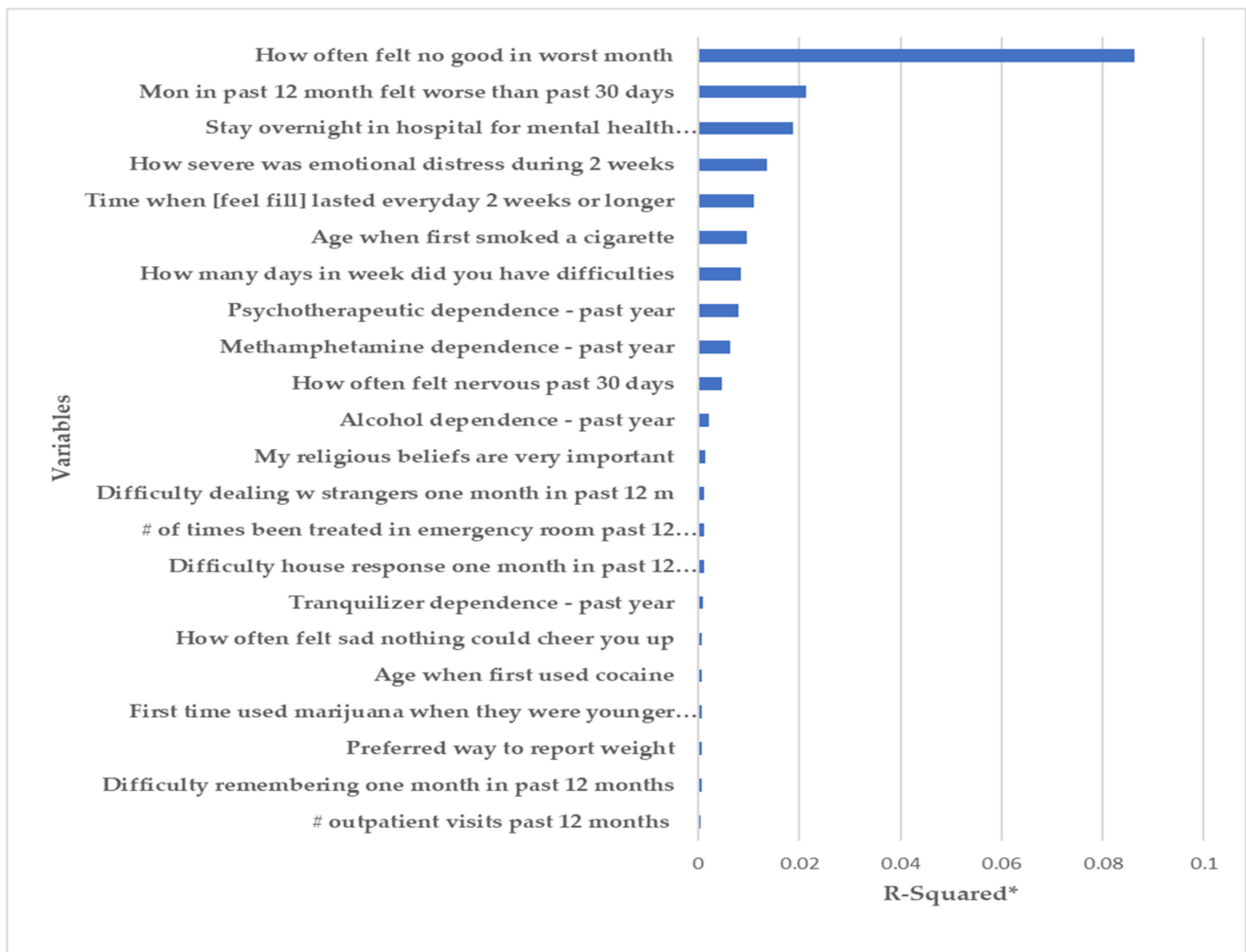


Figure 3. Importance plot for the suicide risk (* a larger R-squared value means that a variable explains a larger percentage of the variation in the outcome variable).

3.1.1. Depression

A total of 271 out of 698 participants (38.8%) indicated 'all the time' to risk for depression questions. The relative importance plot (Figure 2) shows that "Several days or longer when felt sad/empty/depressed" was the highest ranked variable for depression risk prediction models. "How often felt hopeless in the worst month" and "Months in past 12 months felt worse than past 30 days" were ranked as the next highest variables.

3.1.2. Suicide Risk

A total of 180 out of 698 participants (25.8%) reported 'yes' to questions related to suicide risk. The relative importance plot (Figure 3) shows that "How often felt hopeless in worst month" was the highest ranked variable for suicide risk prediction models. "Months in past 12 months felt worse than past 30 days" and "Stay overnight in hospital for mental health treatment past 12 months" were ranked as the next highest variables.

3.2. Measurement of Prediction Model Performances

Sensitivity, Specificity, Accuracy, AUC (area under the curve), Precision, and F1 score were calculated. The comparative results of three algorithms (logistic Regression, RF, KNN) with 89 attributes are summarized in Tables 2 and 3.

4. Discussion

We identified demographic characteristics of adults (20–49 years old) who abuse marijuana. The identified demographic characteristics for most of them were 20–29 years old, never been married, employed, and covered by a health insurance, as aligned with other study findings [50,51]. Overall depression and suicide risk prediction models built by RF and KNN showed good performance (Accuracy = [0.740–0.998], AUC = [0.816–1.0]) but logistic regression showed poor performance. Among the three machine-learning algorithms, models built by RF showed excellent performance (Accuracy = [0.773–0.998], AUC = [0.857–1.0]).

4.1. Impact on Mental Condition

In 2019, the US Census Bureau estimated 328.2 million people in the U.S. and 129.61 million of 20–49 age group (39.5%) [52]. Analyzing the social relations and behaviors of this same group in this study is valuable. The risk variables identified in both prediction models were (1) months in past 12 months felt worse than past 30 days, (2) number of times been treated in the emergency room past 12 months, (3) difficulty in taking care of household responsibilities 1 month in past 12 months, and (4) tranquilizer dependence past year. These indicate that people had been under emotional distress for at least 12 months but had not received proper treatment. These findings are particularly notable since most were employed and covered by health insurance, suggesting finance and healthcare accessibility were not major barriers. The “religious belief” was identified as an influential social relation variable in both prediction models.

Marijuana use such as “smoking cigarette with marijuana in it” and “first time use of marijuana was younger than 20 years old” were identified as variables to influence depression or suicide risk, but not as high as expected. This implies that other factors than marijuana use affected their depression and suicide risk mental state more. Alcohol consumption is known to be associated with depression and suicide risk [27,50,53,54], although its affects as an influential variable to predict depression were beyond the scope of the current study.

Both prediction models showed cocaine and methamphetamine (meth) as a risk variable in depression and suicide, respectively. They are the same type of stimulant but have different effects on the human body. Cocaine is a plant-driven substance and meth is synthesized using various chemicals. Meth increases more dopamine than cocaine, resulting in stronger and longer lasting effects. This study results showed that cocaine was also linked with depression and both cocaine and meth were linked with suicide risk, which are aligned with previous studies [27,55–57].

4.2. Implication for Practice

This study supports earlier published evidence and will aid future investigators in applying a better-informed use of variables. Results revealed that most participants were employed and covered by health insurance; however, they still did not seek or receive proper care to prevent depression or suicide risk. In addition, “religious belief” was identified as a risk variable in both depression and suicide risk prediction models but its impact on this mental condition is not clear. Further study is strongly needed to find reasons and level of impact.

4.3. Limitations

National Survey on Drug Use and Health (NSDUH) data are based on self-reports of drug use and dependencies. Although NSDUH procedures were designed to strengthen participants’ honesty and recall, the degree of underreporting and overreporting of information was unknown. This survey is cross-sectional, measuring responses at a single snapshot in time, hence, overlooking the development of abuse and dependency behaviors as they develop over time. Although the excluded population was only 3%, their inclusion

may generate different results. If other feature selections and prediction model algorithms were used, there may be different risk variables showing different performances.

5. Conclusions

We successfully identified unique characteristics of adults (20–49 years old) who abused marijuana, using publicly available data sets. Well-performing depression and suicide risk prediction models were built using three machine learning algorithms, logistics regression, RF, and KNN. The identified most influential risk variables in models could guide the focus of future marijuana abuse prevention studies. For example, development of marijuana abuse prevention programs targeting the ages of 20–29 group with a regular depression and suicide risk screening. Further study is needed to identify specific barriers of receiving timely treatments for depression and suicide risks when they are well financed and cover a health insurance and level of religious belief impact on depression and suicide risk. Results show that machine learning is a useful tool to explore the impact of marijuana abuse on adults (20–49 years old).

Author Contributions: Conceptualization, J.C. (Jeeyae Choi), J.C. (Joohyun Chung), and J.C. (Jeungok Choi); methodology, J.C. (Jeeyae Choi), J.C. (Joohyun Chung), and J.C. (Jeungok Choi); software, J.C. (Jeeyae Choi) and J.C. (Joohyun Chung); validation, J.C. (Jeeyae Choi), J.C. (Joohyun Chung), and J.C. (Jeungok Choi); formal analysis, J.C. (Jeeyae Choi) and J.C. (Joohyun Chung); investigation, J.C. (Jeeyae Choi), J.C. (Joohyun Chung), and J.C. (Jeungok Choi); resources, J.C. (Jeeyae Choi), J.C. (Joohyun Chung), and J.C. (Jeungok Choi); data curation, J.C. (Jeeyae Choi) and J.C. (Joohyun Chung); writing—original draft preparation, J.C. (Jeeyae Choi) and J.C. (Joohyun Chung); writing—review and editing, J.C. (Jeeyae Choi), J.C. (Joohyun Chung), and J.C. (Jeungok Choi); visualization, J.C. (Jeeyae Choi) and J.C. (Joohyun Chung); supervision, J.C. (Jeeyae Choi); project administration, J.C. (Jeeyae Choi), J.C. (Joohyun Chung), and J.C. (Jeungok Choi); funding acquisition, no funding support. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

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

Data Availability Statement: Publicly available data sets were analyzed in this study. The data can be found here: <https://www.samhsa.gov/data/release/2019-national-survey-drug-use-and-health-nsduh-releases> (accessed on 15 March 2021).

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

References

1. Ethan, X.; Logan, A.; Liam, M.; Leonard, J. Impact of Marijuana (Cannabis) on Health, Safety and Economy. *Int. Digit. Organ. Sci. Res.* **2020**, *5*, 43–52.
2. Lee, J.; Petlakh, K. Progression to drug use from adolescent initiation of marijuana among South Korean inmates: A propensity score matching technique. *Int. J. Comp. Appl. Crim. Justice* **2019**, *44*, 221–230. [[CrossRef](#)]
3. Mametja, M.; Ross, E. Decriminalized, Not Legalized: A Pilot Study of South African University Students' Views on the Use, Impact, Legalization and Decriminalization of Marijuana. *J. Drug Issues* **2020**, *50*, 490–506. [[CrossRef](#)]
4. Reis, J.P.; Auer, R.; Bancks, M.P.; Goff, D.C.; Lewis, C.E.; Pletcher, M.J.; Rana, J.S.; Shikany, J.M.; Sidney, S. Cumulative Lifetime Marijuana Use and Incident Cardiovascular Disease in Middle Age: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Am. J. Public Health* **2017**, *107*, 601–606. [[CrossRef](#)] [[PubMed](#)]
5. Zibokere, E.; Ekiye, E. Theatre as an Agent of Change: Mobilising Against Marijuana Addiction in Tombia Ekpetiama Community in Bayelsa State. *East Afr. J. Interdiscip. Stud.* **2020**, *2*, 1–11. [[CrossRef](#)]
6. National Institute on Drug Abuse. What Is the Scope of Marijuana Use in the United States? Available online: <https://www.drugabuse.gov/publications/research-reports/marijuana/what-scope-marijuana-use-in-united-states> (accessed on 16 June 2021).
7. Substance Abuse Center for Behavioral Health Statistics and Quality. The 2018 National Survey on Drug Use and Health: Detailed Tables. Available online: <https://www.samhsa.gov/data/report/2018-nsduh-detailed-tables> (accessed on 23 June 2021).

8. National Institute on Drug Abuse. Marijuana Research Report 2020. Available online: <https://www.drugabuse.gov/publications/research-reports/marijuana/what-marijuana> (accessed on 22 June 2021).
9. Krauss, M.J.; Rajbhandari, B.; Sowles, S.J.; Spitznagel, E.L.; Cavazos-Rehg, P. A latent class analysis of poly-marijuana use among young adults. *Addict. Behav.* **2017**, *75*, 159–165. [[CrossRef](#)]
10. Pearson, M.R.; Liese, B.S.; Dvorak, R.D. College student marijuana involvement: Perceptions, use, and consequences across 11 college campuses. *Addict. Behav.* **2016**, *66*, 83–89. [[CrossRef](#)]
11. Miller, N.S.; Oberbarnscheidt, T.; Gold, M.S. Marijuana Addictive Disorders and DSM-5 Substance-Related Disorders. *J. Addict. Res. Ther.* **2017**, *8*, 11. [[CrossRef](#)]
12. Popova, L.; McDonald, E.A.; Sidhu, S.; Barry, R.; Maruyama, T.A.R.; Sheon, N.M.; Ling, P.M. Perceived harms and benefits of tobacco, marijuana, and electronic vaporizers among young adults in Colorado: Implications for health education and research. *Addiction* **2017**, *112*, 1821–1829. [[CrossRef](#)]
13. Wong, K.V. Using Marijuana to Cure Marijuana Addiction. *Clin. Exp. Psychol.* **2017**, *3*, 2. [[CrossRef](#)]
14. Cascardo, D. Medical Marijuana: To Prescribe or Not to Prescribe? That Is the Question. *Med. Pract. Manag.* **2020**, *14*, 201–204.
15. Di Forti, M.; Quattrone, D.; Freeman, T.; Tripoli, G.; Gayer-Anderson, C.; Quigley, H.; Rodriguez, V.; Jongsma, H.E.; Ferraro, L.; La Cascia, C.; et al. The contribution of cannabis use to variation in the incidence of psychotic disorder across Europe (EU-GEI): A multicentre case-control study. *Lancet Psychiatry* **2019**, *6*, 427–436. [[CrossRef](#)]
16. National Institute on Drug Abuse. Marijuana DrugFacts 2019. Available online: <https://www.drugabuse.gov/publications/drugfacts/marijuana> (accessed on 16 June 2021).
17. Berenson, A. *Tell Your Children: The Truth About Marijuana, Mental Illness, and Violence*; Free Press: New York, NY, USA, 2019.
18. Zehra, A.; Burns, J.; Liu, C.K.; Manza, P.; Wiers, C.E.; Volkow, N.D.; Wang, G.-J. Cannabis Addiction and the Brain: A Review. *J. Neuroimmune Pharmacol.* **2018**, *13*, 438–452. [[CrossRef](#)]
19. Zizlavsky, S.; Alia, D.; Suwento, R.; Siste, K.; Bardosono, S. P300 auditory wave image and its relation to cognitive function in subjects with marijuana addiction: A cross-sectional study in Cipinang and Pondok Bambu Penitentiary, Jakarta. *J. Phys. Conf. Ser.* **2018**, *1073*, 042039. [[CrossRef](#)]
20. National Institute on Drug Abuse. *Sex and Gender Differences in Substance Use—Substance Use in Women Research Report*; National Institute on Drug Abuse: Baltimore, MD, USA, 2020.
21. American College of Obstetricians and Gynecologists. ACOG Committee Opinion. *Obstet. Gynecol.* **2017**, *130*, e205–e209.
22. Gobbi, G.; Atkin, T.; Zytynski, T.; Wang, S.; Askari, S.; Boruff, J.; Ware, M.; Marmorstein, N.; Cipriani, A.; Dendukuri, N.; et al. Association of Cannabis Use in Adolescence and Risk of Depression, Anxiety, and Suicidality in Young Adulthood. *JAMA Psychiatry* **2019**, *76*, 426–434. [[CrossRef](#)] [[PubMed](#)]
23. Hasin, D.S.; Saha, T.D.; Kerridge, B.T.; Goldstein, R.; Chou, S.P.; Zhang, H.; Jung, J.; Pickering, R.P.; Ruan, W.J.; Smith, S.M.; et al. Prevalence of Marijuana Use Disorders in the United States Between 2001–2002 and 2012–2013. *JAMA Psychiatry* **2015**, *72*, 1235–1242. [[CrossRef](#)]
24. Onaemo, V.N.; Fawehinmi, T.O.; D’Arcy, C. Comorbid Cannabis Use Disorder with Major Depression and Generalized Anxiety Disorder: A Systematic Review with Meta-analysis of Nationally Representative Epidemiological Surveys. *J. Affect. Disord.* **2020**, *281*, 467–475. [[CrossRef](#)] [[PubMed](#)]
25. Gold, M. Medicinal Marijuana, Stress, Anxiety, and Depression: Primum non nocere. *Mo. Med.* **2020**, *117*, 406–411. [[PubMed](#)]
26. Britton, P.C.; Conner, K.R. Suicide Attempts within 12 Months of Treatment for Substance Use Disorders. *Suicide Life Threat. Behav.* **2010**, *40*, 14–21. [[CrossRef](#)]
27. Culbreth, R.; Swahn, M.H.; Osborne, M.; Brandenberger, K.; Kota, K. Substance use and deaths by suicide: A latent class analysis of the National Violent Death Reporting System. *Prev. Med.* **2021**, *150*, 106682. [[CrossRef](#)]
28. Tetteh, J.; Ekem-Ferguson, G.; Swaray, S.M.; Kugbey, N.; Quarshie, E.N.-B.; Yawson, A.E. Marijuana use and repeated attempted suicide among senior high school students in Ghana: Evidence from the WHO Global School-Based Student Health Survey, 2012. *Gen. Psychiatry* **2020**, *33*, e100311. [[CrossRef](#)]
29. Holland, K. CBD vs. THC: What’s the Difference?—2020. Available online: <https://www.healthline.com/health/cbd-vs-thc> (accessed on 22 June 2021).
30. Bridgeman, M.; Abazia, D.T. Medicinal Cannabis: History, Pharmacology, And Implications for the Acute Care Setting. *P T Peer Rev. J. Formul. Manag.* **2017**, *42*, 180–188.
31. Encyclopaedia Britannica. Medical Cannabis. Available online: <https://www.britannica.com/science/medical-cannabis> (accessed on 30 June 2021).
32. Zagorski, N. Be Prepared to Discuss CBD Products With Patients. *Psychiatr. News* **2020**, *55*. [[CrossRef](#)]
33. Mudan, A.; DeRoos, F.; Perrone, J. Medical Marijuana Miscalculation. *N. Engl. J. Med.* **2019**, *381*, 1086–1087. [[CrossRef](#)]
34. Hazekamp, A. The Trouble with CBD Oil. *Med. Cannabis Cannabinoids* **2018**, *1*, 65–72. [[CrossRef](#)]
35. Stuyt, E. The Problem with the Current High Potency THC Marijuana from the Perspective of an Addiction Psychiatrist. *Mo. Med.* **2018**, *115*, 482–486.
36. Center for Behavioral Health Statistics and Quality. *2019 National Survey on Drug Use and Health Public Use File Codebook*; Substance Abuse and Mental Health Services Administration: Rockville, MD, USA, 2020.
37. Research Triangle Institute International. *2019 National Survey on Drug Use and Health: Field Interviewer Manual—Field Interviewer Computer Manual*; Research Triangle Institute International: Rockville, MD, USA, 2018.

38. Alshalabi, H.; Tiun, S.; Omar, N.; Albared, M. Experiments on the Use of Feature Selection and Machine Learning Methods in Automatic Malay Text Categorization. *Procedia Technol.* **2013**, *11*, 748–754. [CrossRef]
39. Mahmoudi, M.R.; Heydari, M.H.; Qasem, S.N.; Mosavi, A.; Band, S.S. Principal component analysis to study the relations between the spread rates of COVID-19 in high risks countries. *Alex. Eng. J.* **2021**, *60*, 457–464. [CrossRef]
40. Song, F.; Guo, Z.; Mei, D. Feature Selection Using Principal Component Analysis. In *2010 International Conference on System Science, Engineering Design and Manufacturing Informatization*; Hubei, Y., Ed.; IEEE Computer Society: Los Angeles, CA, USA, 2010.
41. Advantages and Disadvantages of Principal Component Analysis in Machine Learning. Available online: http://theprofessionalspoint.blogspot.com/2019/03/advantages-and-disadvantages-of_4.html (accessed on 1 July 2021).
42. Nijman, S.W.J.; Hoogland, J.; Groenhof, T.K.J.; Brandjes, M.; Jacobs, J.J.L.; Bots, M.L.; Asselbergs, F.W.; Moons, K.G.M.; Debray, T.P.A. Real-time imputation of missing predictor values in clinical practice. *Eur. Heart J. Digit. Health* **2020**, *2*, 154–164. [CrossRef]
43. Spiess, M.; Kleinke, K.; Reinecke, J. Proper Multiple Imputation of Clustered or Panel Data. In *Advances in Longitudinal Survey Methodology*; Wiley: Hoboken, NJ, USA, 2021.
44. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: New York, NY, USA, 2013.
45. Levy, J.J.; O'Malley, A.J. Don't dismiss logistic regression: The case for sensible extraction of interactions in the era of machine learning. *BMC Med. Res. Methodol.* **2020**, *20*, 171. [CrossRef]
46. 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]
47. 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]
48. Tibaduiza, D.; Torres-Arredondo, M.; Vitola, J.; Anaya, M.; Pozo, F. A Damage Classification Approach for Structural Health Monitoring Using Machine Learning. *Complexity* **2018**, *2018*, 5081283. [CrossRef]
49. Xing, W.; Bei, Y. Medical Health Big Data Classification Based on KNN Classification Algorithm. *IEEE Access* **2020**, *8*, 28808–28819. [CrossRef]
50. Abdel-Rahman, O. Cannabis use among Canadian adults with cancer (2007–2016): Results from a national survey. *Expert Rev. Pharm. Outcomes Res.* **2020**, 1025–1029. [CrossRef]
51. Anim, E. *Investigating the Relationship between Marital Status, Marijuana Use and State Marijuana Laws in the USA*; Northern Arizona University: Flagpole, AZ, USA, 2020.
52. Resident Population of the United States by Sex and Age as of July 1, 2019 (in Million). Available online: <https://www.statista.com/statistics/241488/population-of-the-us-by-sex-and-age/> (accessed on 1 July 2021).
53. Stanton, R.; To, Q.G.; Khalesi, S.; Williams, S.L.; Alley, S.J.; Thwaite, T.L.; Fenning, A.S.; Vandelanotte, C. Depression, Anxiety and Stress during COVID-19: Associations with Changes in Physical Activity, Sleep, Tobacco and Alcohol Use in Australian Adults. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4065. [CrossRef]
54. Amiri, S.; Behnezhad, S. Alcohol use and risk of suicide: A systematic review and Meta-analysis. *J. Addict. Dis.* **2020**, *38*, 200–213. [CrossRef] [PubMed]
55. Desai, R.; Thakkar, S.; Patel, H.P.; Tan, B.E.-X.; Damarlapally, N.; Haque, F.A.; Farheen, N.; DeWitt, N.; Savani, S.; Parisha, F.J.; et al. Higher odds and rising trends in arrhythmia among young cannabis users with comorbid depression. *Eur. J. Intern. Med.* **2020**, *80*, 24–28. [CrossRef] [PubMed]
56. Mochrie, K.D.; Whited, M.C.; Cellucci, T.; Freeman, T.; Corson, A.T. ADHD, depression, and substance abuse risk among beginning college students. *J. Am. Coll. Health* **2018**, *68*, 6–10. [CrossRef] [PubMed]
57. National Institute on Drug Abuse. *Rising Stimulant Deaths Show That We Face More than Just an Opioid Crisis*; NIH: Bethesda, MD, USA, 2020.