Published in final edited form as:

Subst Use Misuse. 2019 ; 54(1): 89–96. doi:10.1080/10826084.2018.1496454.

The “Real” Number of Washington State Adolescents Using Marijuana, and Why: A Misclassification Analysis

Sean M. Murphy, Ph.D^{1,*} and Robert Rosenman, Ph.D²

¹Weill Cornell Medical College, New York, New York 10065, USA

²Washington State University, Pullman, Washington 99164, USA

Abstract

Background: Although causality is difficult to establish, the regular use of marijuana has been associated with many adverse physiological and sociological consequences. There is considerable concern regarding marijuana use among adolescents, as the likelihood of adverse consequences increases significantly for this age group. The most comprehensive data for identifying risk factors for adolescent marijuana use is typically self-report, which may be over- or under-reported for a variety of reasons, including stigmatization, peer-pressure, or fear of being discovered.

Objectives: To identify the prevalence of adolescent marijuana use in Washington State, and the associated risk and protective factors, while controlling for and estimating the extent of misreporting, and its determinants.

Method: Data come from the 2014 Washington State Healthy Youth Survey. We accounted for missingness using chained multivariate imputation equations, resulting in 33,320 complete observations. Our model was estimated using a maximum likelihood multiple regression designed to control for systematic misclassification in binary dependent variables.

Results: Approximately 12% of Washington adolescents claimed to have used marijuana in the past 30 days. Our estimates indicate this figure is likely closer to 18%. Determinants of use included use of other substances, gender, age, and measures of deviant social influences, personality/attitude, school and family bonds, bullying, and acquisition ease. Determinants of misreporting included use of other substances, gender, parental education, and family bonds.

Conclusions: Failing to control for misreporting considerably underestimates the prevalence of marijuana use among adolescents. Our model allows us to better identify at-risk adolescents and inform focused prevention efforts.

Keywords

Marijuana; misclassification; prevalence; adolescent; risk and protective factors

*Corresponding author at: Department of Healthcare Policy and Research, 425 East 61st Street, Suite 301, New York, NY 10065. Phone: 646-962-9710. mmm2010@med.cornell.edu.

Declaration of Interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Introduction

Although causality is difficult to establish, the regular use of marijuana has been associated with many adverse physiological and sociological consequences. These consequences include increased likelihood of developing a marijuana-use, or other substance-use disorder; of having anxiety, depression, or other psychoses among those with a predisposition; of being involved in a motor-vehicle accident, especially for those simultaneously using alcohol; and possibly compromised cardiovascular and cerebrovascular systems (Volkow, Baler, Compton, & Weiss, 2014). For adolescents, the risk of developing a marijuana use disorder is roughly 2 to 4 times higher (Volkow et al., 2014). Hence, prevention efforts targeting adolescents would appear to be a wise investment of public-health funds. One problem with developing such efforts is that the most comprehensive data for identifying risk factors for adolescent marijuana use is typically self-reported, which may be over- or under-reported for a variety of reasons, including stigmatization, peer-pressure, or fear of being discovered. If misreporting occurs, not only will estimates of use prevalence be wrong, potentially suppressing public support and diverting resources away from their most efficient use, but estimates of risk and protective factors that are based on this information will also be biased, rendering intervention programs less effective.

There have been few studies that have looked at the underreporting of marijuana use by adolescents, and those available have focused on at-risk youths. Harrison and Hughes (1997) concluded that self-reported marijuana use is generally reliable, although the accuracy is greater among those who admit using than among those who do not. She reported that over 10% of those who denied using marijuana in the Drug Use Forecasting study (Harrison, 1989) showed it in their urinalysis, while 89 percent of those with a positive urinalysis admitted using marijuana. In another example, Bessa et al. (2010), looking only at pregnant adolescents in their third trimester, found none of the 4% in their sample who had used marijuana reported doing so. Harris et al. (2008) found over 5% of adolescents in a drug treatment program did not report marijuana use. But among a group referred for drug use assessment, only 41 of 73 adolescents for whom a urinalysis indicated recent marijuana use self-reported as doing so (Williams & Nowatzki, 2005). Inconsistencies in self-reported substance use do not appear to be random, and thus can impact outcome and performance measurement (Harris, Griffin, McCaffrey, & Morral, 2008). In fact, misreporting is often tied to demographic characteristics and may vary by consumption-based subgroups of the population (Livingston & Callinan, 2015).

A number of strategies have been identified to reduce, but not eliminate, misreporting/misclassification bias associated with self-reported data on sensitive topics such as substance use (Krumpal, 2013). For example, research has shown that misreporting on surveys can be reduced by increasing privacy (e.g., self-administration), increasing anonymity, and emphasizing the importance of the data being collected. Additional methods for ensuring accurate responses include inserting questions designed to identify those who are being deceitful (e.g. asking about the use of non-existent drugs).

There are at least two alternative approaches to correct prevalence measures for potential misreporting. The Hui-Walter estimation methodology relies on repeated measures to adjust

reported prevalence for misreporting (Biemer & Witt, 1997). Alternatively, verified outside samples of prevalence can be used to adjust reported prevalence in a specific sample (Wright, Gfroerer, & Epstein, 1997). Both these approaches provide estimates of the extent of false reporting, but neither allows the analyst to correct self-reported use; hence, no gain is made in understanding who uses and why, nor who is likely to misreport use.

The objective of this paper is to find the prevalence of adolescent marijuana use in Washington State, and the associated risk and protective factors, after accounting for misreporting/misclassification bias, and to identify factors that contribute to misreporting. We do so by using a multiple-regression-based approach designed to control for, and quantify, misclassification in binary dependent variables, while also allowing for the determination of factors associated with the misclassification (Tennekoon & Rosenman, 2016). Among substance use disorder studies, this technique has been successfully applied to adolescent non-medical opioid use (Murphy, Friesner, & Rosenman, 2015) and tobacco use (Tennekoon & Rosenman, 2015).

Methods

Data

Data come from the 2014 Washington State Healthy Youth Survey. Permission to use the data was granted by the Washington State Department of Social and Health Services, the Washington State Institutional Review Board (IRB) and the [blinded for review] IRB. Healthy Youth Survey questions are obtained from four surveys that have been used extensively throughout the United States for many years, and have demonstrated reliability and validity; the surveys are: Monitoring the Future survey, Youth Risk Behavior Survey, Global Youth Tobacco Survey, and the Communities that Care Survey (Washington State Department of Health, 2014). The Healthy Youth Survey is administered in even years to Washington State public school students in grades 6, 8, 10, and 12 across the state. Public schools are randomly selected to participate by the Washington State Department of Health; schools that were not selected are also welcome to participate. Parents and students are notified about the survey prior to administration and given the opportunity to refuse participation. Students can also choose not to participate immediately before administration. The response rate varied by grade with a high of 79% for grades 6 and 8, a rate of 67% for grade 10, and a rate of 50% for grade 12, for an overall response rate of 68%. Numerous methods are employed to increase the validity and reliability of the survey results. With regard to validity, in addition to using items from the established youth surveys mentioned above, observations are removed if there is evidence of blatant dishonesty (e.g., respondents report misusing a fictitious drug, report being in the wrong grade, provide inconsistent answers). On average, only about 4% of completed surveys are removed. In terms of reliability, students are informed about the importance of the survey results, and the survey is self-administered using standardized procedures in a safe and confidential environment (e.g., administered during one class period, identifying information is not collected). Additionally, the Healthy Youth Survey is periodically assessed for generalizability. Survey results were tested for generalizability in 2002, 2004, 2008, and 2010 and found to be

representative of Washington State public school students (Washington State Department of Health, 2016).

Sample

The initial study population consisted of 35,262 students who were attending one of the public schools randomly selected to participate by the Washington State Department of Health. Of those 35,262 students, 34,122 answered the question about whether they had used marijuana in the previous 30 days, our primary variable of interest. Missing data in our covariates ranged from <1% (gender; and cigarette, alcohol, and illicit drug – other than marijuana and opioids – consumption) to 69% (extent to which friends would agree with marijuana consumption); however, the average amount of missingness was 23%. We conducted Little's Missing Completely at Random (MCAR) test (Little, 1988) and rejected the hypothesis that the data are MCAR, which indicates that using listwise deletion would result in missing variable bias. Therefore, we accounted for missingness in our covariates through multivariate imputation using chained equations (MICE) (Raghunathan, Lepkowski, Van Hoewyk, & Solenberger, 2001; Van Buuren, Boshuizen, & Knook, 1999). MICE is a multiple imputation procedure in which a series of multivariable regression models are run multiple times, one for each variable with missing observations. Each variable with missing observations is modeled as a function of all the other variables in the data, and the flexibility of MICE allows each variable with missing observations to be modeled according to its appropriate distribution. Our final imputed dataset consisted of 33,320 complete observations.

Measures

Our primary outcome is a binary measure of whether the respondent had used marijuana in the past 30 days. Covariates were chosen based on seven domains identified as important in the adolescent development literature (Ellickson & McGuigan, 2000).

- *School bonds.* This domain was represented using an ordinal measure of the respondent's grades, where 1 represents "mostly Fs" and 5 represents "mostly As".
- *Family bonds,* were measured using a binary variable indicating whether the respondent felt that he or she could ask their mom or dad for help with a personal problem.
- *Other problem behaviors,* included binary measures of whether the respondent had smoked cigarettes, drank alcohol, misused prescription opioids, or used *other* illicit substances (including non-prescribed opiates, methamphetamine, inhalants, cocaine, steroids and heroin) in the past 30 days.
- *Exposure to deviant social influences,* included an ordinal variable measuring the extent to which the respondent believed his or her friends would view marijuana use as wrong, where 1 represented "very wrong" and 4 represented "not at all wrong", and a binary variable indicating whether the respondent had been involved with a gang in the past 12 months.

- *Personality and attitude measures* included a binary measure of whether the respondent in the past 12 months had felt “so sad or hopeless almost every day for two weeks or more in a row that [they] stopped doing some usual activities?”, a binary variable indicating whether the respondent had seriously considered committing suicide in the past 12 months, and an ordinal variable indicating the respondent’s level of engagement in after-school activities in an average week, where 1 equaled 0 days and 3 equaled 3 or more days.
- *School and neighborhood characteristics*, included binary measures of whether the school lets parents know when students do something well, whether the respondent had been bullied in the past 30 days, whether the school was in an urban area, and whether the area’s laws and norms were considered “favorable to drug use”. This domain was also represented by an ordinal variable measuring the ease with which the student believed he or she could obtain marijuana, where 1 represented “very hard” and 4 represented “very easy”.
- *Demographic and socioeconomic factors*, included age, gender, race, and binary variables indicating whether the respondent’s mother or father graduated from a 4-year college.

Analysis

We estimated our model using a maximum likelihood multiple regression designed to control for systematic misclassification in binary dependent variables (Tennekoon & Rosenman, 2016). The model is an extension of Hausman et al. (1998); in addition to controlling for misclassification and estimating the probabilities that a value observed as zero should have been a one, and vice-versa, it allows the misclassification probabilities to be a function of multiple factors. The model has been successfully used to estimate misclassification in adolescent self-reported prescription-opioid misuse (Murphy et al., 2015), smoking behavior (Tennekoon & Rosenman, 2013), among other uses. The model has been described in detail elsewhere for the interested reader (Tennekoon & Rosenman, 2016), but adding misclassification probabilities and their potential determinants structurally alters the likelihood function in a known way and allows us to estimate the extent to which a given observation (e.g., one where the dependent variable was observed as a 0) reflects those observed with the alternative outcome (i.e., ones where the dependent variable was observed as a 1). Following the estimation of the model, the predicted number of marijuana users can be calculated without misclassification bias. All analyses were conducted using Stata® 15.1.

The likelihood that an adolescent had used marijuana in the prior 30 days was estimated as a function of the variables listed above. Given that no prior study had evaluated determinants of misclassification of self-reported marijuana use, we modeled the likelihood that an adolescent reported no use, but likely had, as a function of characteristics believed to be indicative of social desirability bias: age, gender, race, parental education, other substance use, grades, and ability to confide in parents. The likelihood that an adolescent reported marijuana use, but likely had not, was modeled as a function of these same factors, as well as the extent to which the respondent believed his or her friends would view marijuana use as wrong.

Results

Descriptive statistics are displayed in Table 1. The mean age of adolescents in our sample was 15, half were female, and half were white/Caucasian. Just over a third had a mother or father with a degree from a 4-year college or obtained an advanced graduate degree, and 91% resided in an urban area. Twelve percent of adolescents self-reported that they had used marijuana in the 30 days prior to the survey, 15% had consumed alcohol, 6% had smoked cigarettes, 3% had used prescription opioids to get high, and 3% had used another illicit substance. Additionally, just under a third of youths reported feelings of major depression in the 12 months prior, and a quarter had been bullied in the prior 30 days.

For contrast, we first provide the results from a multiple probit regression without misclassification adjustment in Table 2. Table 3 contains the results from the multiple regression approach designed to control for systematic misclassification in binary dependent variables. Overall, our corrected estimates indicate that 7% of adolescents ($n \sim 1,935$) who reported that they did not use marijuana in the prior 30 days, actually had, while only 1% of those who reported that they had used marijuana, actually had not done so. Table 4 displays the predicted prevalence values from the probit and misclassification multiple regression models. The predicted prevalence from the probit model is almost identical to that from the raw data at 12.3% (95% CI: 11.1%, 13.5%). The predicted prevalence from the misclassification model is 18.1% (95% CI: 15.5%, 20.8%).

After controlling for misclassification, using the conventional $p\text{-value}=0.05$, significant risk factors for adolescent marijuana use included increased age, being male, cigarette and other substance use (i.e., alcohol, prescription opioids, and other illicit substances), ease of obtaining marijuana, having friends who approve of marijuana use, lower academic grades, gang affiliation, and having major depressive thoughts or suicidal ideation. Cigarette and the other substance use variables had the strongest association with marijuana use. Youths who participated in after-school activities and those who had been bullied were less likely to have used marijuana in the past 30 days. Comparing the results in Table 2 to those in Table 3, controlling for misclassification increased the magnitudes of *all* significant estimated effects, especially those of *other problem behaviors* and gang affiliation. Additionally, after controlling for misclassification, race/ethnicity was no longer a determinant of marijuana use, and the ability to confide in parents became a potential protective factor with a $p\text{-value}$ of 0.06. The results also allow us to identify factors associated with misclassified responses. Youths who felt that they could speak to their mom or dad about personal problems; or consumed other illicit substances, alcohol or cigarettes; were less likely to report that they had not used marijuana when they, in fact, had (equation 2). The mean probability of this type of misreporting was 7%. Cigarette and other-illicit-substance use increased the likelihood of an adolescent misreporting that they had used marijuana in the prior 30 days when they had not, while being female and having a mother with a college degree made doing so less likely (equation 3). With a $p\text{-value}$ of 0.06, Whites/Caucasians were also potentially more likely to report using marijuana when they did not. The mean probability for this type of misreporting was only 1%.

Age and school year merit some additional discussion, as the Washington State Healthy Youth Survey includes students in the 6th, 8th, 10th and 12th grades. Due to severe multicollinearity, age is included in the model, while school year is not. We estimate older students – hence those likely to be in the higher grades – are more likely to use marijuana. However, we did not find, at conventional p-values, that age contributed to either type of misclassification. For marijuana, at least, it seems that younger students and older students likely report use truthfully.

Discussion

Approximately 12% of Washington adolescents claimed to have used marijuana in the past 30 days. Our estimates indicate this figure should be 18%, suggesting that adolescent marijuana use may be a larger public-health issue than previously thought. Moreover, controlling for misclassification resulted in changes in magnitude and significance of risk factors for adolescent marijuana use. Many of the risk and protective factors that we identified are similar to those reported in a recent review of the extant literature (Substance Abuse and Mental Health Services Administration, 2014). For example, being male, poor academic performance, alcohol and tobacco use, friends who use marijuana, affiliating with deviant peers, feelings of major depression, and ease of obtaining marijuana are all risk factors common to our study and the literature. We also identified a few possible inconsistencies between our study and those reviewed in the HSA report. For example, parental education was found to be a protective factor in the papers it reviewed, but was insignificant in our study. The HSA report noted two studies that found competitive athletes to be more likely to engage in marijuana use, whereas we found those who regularly participated in after-school activities (not necessarily just sports) were less likely to use marijuana, although our estimated effect was small. Similarly, at least one prior study found community laws favorable to marijuana use to be a risk factor, while it was not statistically significant in our model. Additional risk factors that we identified, but were not specifically mentioned in the review of the literature, were suicidal ideation, recent misuse of prescription opioids, and recent use of other illicit substances.

One value of correcting for misreporting as we did is to inform on causal factors for misreporting. Seven percent of adolescents who stated that they had not used marijuana in the recent past are predicted to have used. Those who reported alcohol, cigarette, or *other* illicit substance use in the prior 30 days, and individuals who felt that they could speak with their parents about important issues, were significantly less likely to have misreported in this manner. Similarly, in our previous paper on adolescent non-medical prescription-opioid use we found that individuals who reported marijuana or *other* illicit substance use were less likely to fail to report misuse of prescription opioids (Murphy et al., 2015).

Our estimates also indicate that individuals who reported recent cigarette or *other* illicit substance use were more likely to report that they had used marijuana in the previous 30 days, when they most likely had not, while females and those whose mom had a 4-year college degree were less likely to misreport. In our aforementioned study on adolescent prescription opioid misuse, individuals who reported recent use of *other* illicit substances were also more likely to have reported recent prescription opioid misuse, when the evidence

indicates they had not. However, this type of misreporting did not appear to be much of an issue for adolescent marijuana use, as the mean probability of its occurrence was only 1%.

A second value of correcting for misreporting is to obtain better information about risk and protective factors. Table 5 compares the estimates for using marijuana from the probit model to the Misclassification Model. There is only one change in the sign of a risk or protective factor. Our measure of Family Bonds, the ability to confide in parents, changes from a positive coefficient with a p-value of 0.55 to a negative coefficient with a p-value of 0.06. However, as we noted earlier, students able to confide in their parents are less likely to have used marijuana, but report they did not.

A bigger impact along these lines is in the strength of risk and protective factors. Grades is the only protective factor for whom the magnitude is underestimated by probit with a p-value of the difference less than 0.05. More telling, the magnitudes of several risk factors, including all related to other problem behaviors and exposure to deviant social influences, are underestimated by probit, again using a p-value of 0.05. In addition, age is a significantly stronger risk factor than probit estimates.

Strengths and limitations

The comprehensive nature of our data is a strength of this study. Additionally, the state-selected sample, upon which we estimated our model, has been shown to be representative of Washington State adolescents attending public schools.

There are two important limitations associated with our study. First, due to our data being self-report, many of our covariates of interest were missing values. We dealt with the missingness using multivariate imputation chained equations (MICE) (Ragunathan et al., 2001; Van Buuren et al., 1999). As can be seen in Table 1, the raw and imputed datasets are almost identical with regard to the descriptive statistics. Second, although we believe our estimates are generalizable to Washington State public schools students, we are unable to generalize outside of that population.

Conclusion

Failing to control for misclassification considerably underestimates the prevalence of marijuana use among adolescents. Moreover, misreported use biases the importance of various risk and protective factors such as age, family bonds, and other problem behaviors. Our analysis indicates that relatively few students who do not use marijuana report they do, but over one-third of users report they do not. Marijuana users who can confide in their parents are less likely to misreport their use, as are (perhaps somewhat surprisingly), students who smoke or consume other illicit substances. Correcting for potential misclassification in survey responses allows policymakers to better identify at-risk adolescents, and informs focused prevention efforts.

Acknowledgments

Financial support: Washington State University Office of Research; NIDA Center for Health Economics of Treatment Interventions for Substance Use Disorder, HCV and HIV (CHERISH, P30DA040500).

References

- Bessa MA, Mitsuhiro SS, Chalem E, Barros MM, Guinsburg R, & Laranjeira R (2010). Underreporting of use of cocaine and marijuana during the third trimester of gestation among pregnant adolescents. *Addictive Behaviors*, 35(3), 266–269. [PubMed: 19896774]
- Biemer PP, & Witt M (1997). Repeated measures estimation of measurement bias for self-reported drug use with applications to the National Household Survey on Drug Abuse. *NIDA Res Monogr*, 167, 439–476. [PubMed: 9243573]
- Ellickson PL, & McGuigan KA (2000). Early predictors of adolescent violence. *American Journal of Public Health*, 90(4), 566–572. [PubMed: 10754971]
- Harris KM, Griffin BA, McCaffrey DF, & Morral AR (2008). Inconsistencies in self-reported drug use by adolescents in substance abuse treatment: Implications for outcome and performance measurements. *Journal of substance abuse treatment*, 34(3), 347–355. [PubMed: 17614240]
- Harrison LD (1989). The Validity of Self-Reported Drug Use Among Arrestees Paper presented at the 41st meeting of the American Society of Criminology, Reno, NV.
- Harrison LD, & Hughes MS (1997). The Validity of Self-reported Drug Use: Improving the Accuracy of Survey Estimates. Retrieved from NIDA Research Monograph 167:
- Hausman JA, Abrevaya J, & Scott-Morton FM (1998). Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*, 87, 239–269.
- Krumpal I (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, 47(4), 2025–2047.
- Little RJ (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202.
- Livingston M, & Callinan S (2015). Underreporting in alcohol surveys: whose drinking is underestimated? *Journal of Studies on Alcohol and Drugs*, 76(1), 158–164. [PubMed: 25486405]
- Murphy SM, Friesner DL, & Rosenman R (2015). Opioid misuse among adolescents: new evidence from a misclassification analysis. *Applied health economics and health policy*, 13(2), 181–192. [PubMed: 25617182]
- Raghunathan TE, Lepkowski JM, Van Hoewyk J, & Solenberger P (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey methodology*, 27(1), 85–96.
- Substance Abuse and Mental Health Services Administration. (2014). Risk and Protective Factors Associated with Youth Marijuana Use. Retrieved from <https://www.samhsa.gov/capt/sites/default/files/resources/risk-protective-factors-marijuana-use.pdf>
- Tennekoon V, & Rosenman R (2013). Bias in measuring smoking behavior. *School of Economic Sciences Working Paper* 2013-10.
- Tennekoon V, & Rosenman R (2015). The pot calling the kettle black? A comparison of measures of current tobacco use. *Applied Economics*, 47(5), 431–448. [PubMed: 25587199]
- Tennekoon V, & Rosenman R (2016). Systematically misclassified binary dependent variables. *Communications in Statistics-Theory and Methods*, 45(9), 2538–2555. [PubMed: 27293307]
- Van Buuren S, Boshuizen HC, & Knook DL (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in medicine*, 18(6), 681–694. [PubMed: 10204197]
- Volkow ND, Baler RD, Compton WM, & Weiss SR (2014). Adverse health effects of marijuana use. *New England Journal of Medicine*, 370(23), 2219–2227. [PubMed: 24897085]
- Washington State Department of Health. (2014). Washington State Healthy Youth Survey 2012 Analytic Report. Retrieved from <http://www.doh.wa.gov/DataandStatisticalReports/DataSystems/HealthyYouthSurvey>
- Washington State Department of Health. (2016). Healthy Youth Survey Data Analysis and Technical Assistance Manual. Retrieved from <http://www.AskHYS.net>
- Williams RJ, & Nowatzki N (2005). Validity of adolescent self-report of substance use. *Substance Use & Misuse*, 40(3), 299–311.

Wright D, Gfroerer J, & Epstein J (1997). The use of external data sources and ratio estimation to improve estimates of hardcore drug use from the NHSDA. NIDA Res Monogr, 167, 477–497. [PubMed: 9243574]

Descriptive Statistics

Table 1.

	Raw Data		Imputed Data	
	n		n	
Used marijuana in prior 30 days	34,122	12%	34,122	12%
<i>Demographic and Socioeconomic Variables</i>				
Age, mean (SD)	25,168	15 (1.65)	34,817	15 (1.64)
Female	34,067	51%	35,202	51%
White/Caucasian	33,573	51%	34,653	50%
Mom has a college degree	24,414	36%	34,776	35%
Dad has a college degree	24,394	35%	34,775	34%
<i>School Bonds</i>				
Grades, ^a mean (SD)	32,764	4.13 (.97)	35,157	4.12 (.97)
<i>Family Bonds</i>				
Able to confide in parents	19,576	84%	35,071	83%
<i>Other Problem Behaviors</i>				
Consumed alcohol in the prior 30 days	34,083	14%	35,198	15%
Smoked cigarettes in the prior 30 days	34,068	6%	35,210	6%
Misused pain medication in the prior 30 days	25,285	4%	34,795	3%
Consumed <i>other illicit</i> substances in the prior 30 days	33,994	3%	35,188	3%
<i>Exposure to deviant social influences</i>				
Gang affiliation in prior 12 months	25,090	5%	34,802	5%
Friends okay with marijuana use, ^b mean (SD)	10,697	1.92 (1.14)	34,673	1.76 (1.07)
<i>Personality and Attitude</i>				
Feelings of major depression in prior 12 months	25,296	31%	34,799	30%
Seriously considered suicide in prior 12 months	25,239	18%	34,796	17%
Regular participation in supervised after-school activities, ^c mean (SD)	24,442	2.11 (.87)	34,787	2.11 (.87)
<i>School and Neighborhood Characteristics</i>				
School communicated with parents	20,799	35%	35,100	34%
Bullied in prior 30 days	33,892	25%	34,582	25%
Community laws favorable to drug use	20,932	30%	35,105	31%

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

	Raw Data		Imputed Data	
	n		n	
Ease of obtaining marijuana, ^d mean (SD)	20,389	1.86 (1.17)	35,023	1.92 (1.19)
Lived in urban area	34,122	91%	35,262	91%

^a 1="mostly Fs", 5="mostly As";

^b 1="very wrong", 4="not at all wrong";

^c 1=0 days, 3=3 or more days;

^d 1="very hard", 4="very easy,"

Table 2.**Probit Regression Results**

Variable	Estimate	Std. error	p-value
<i>Used Marijuana</i>			
Constant	−2.49	0.13	<0.001
<i>Demographic and Socioeconomic Variables</i>			
Age	0.05	0.01	<0.001
Female	−0.07	0.02	0.01
White/Caucasian	−0.07	0.02	0.01
Mom has a college degree	−0.04	0.03	0.23
Dad has a college degree	−0.01	0.03	0.80
<i>School Bonds</i>			
Grades	−0.15	0.01	<0.001
<i>Family Bonds</i>			
Able to confide in parents	0.02	0.03	0.55
<i>Other Problem Behaviors</i>			
Consumed alcohol in the prior 30 days	1.10	0.03	<0.001
Smoked cigarettes in the prior 30 days	0.65	0.04	<0.001
Misused pain medication in the prior 30 days	0.39	0.06	<0.001
Consumed <i>other illicit</i> substances in the prior 30 days	0.73	0.06	<0.001
<i>Exposure to deviant social influences</i>			
Gang affiliation in prior 12 months	0.45	0.04	<0.001
Friends okay with marijuana use	0.19	0.01	<0.001
<i>Personality and Attitude</i>			
Feelings of major depression in prior 12 months	0.06	0.03	0.05
Seriously considered suicide in prior 12 months	0.13	0.03	<0.001
Regular participation in supervised after-school activities	−0.03	0.01	0.02
<i>School and Neighborhood Characteristics</i>			
School communicated with parents	−0.03	0.03	0.33
Bullied in prior 30 days	−0.13	0.03	<0.001
Ease of obtaining marijuana	0.17	0.01	<0.001
Community laws favorable to drug use	−0.03	0.03	0.19
Lived in urban area	−0.03	0.04	0.46

Table 3.**Misclassification Model Regression Results**

Variable	Estimate	Std. error	p-value
<i>Equation 1: Used Marijuana</i>			
Constant	−2.81	0.25	<0.001
<i>Demographic and Socioeconomic Variables</i>			
Age	0.10	0.02	<0.001
Female	−0.09	0.05	0.05
White/Caucasian	−0.06	0.05	0.20
Mom has a college degree	−0.08	0.06	0.18
Dad has a college degree	−0.04	0.06	0.54
<i>School Bonds</i>			
Grades	−0.20	0.02	<0.001
<i>Family Bonds</i>			
Able to confide in parents	−0.11	0.06	0.06
<i>Other Problem Behaviors</i>			
Consumed alcohol in the prior 30 days	1.48	0.07	<0.001
Smoked cigarettes in the prior 30 days	1.23	0.15	<0.001
Misused pain medication in the prior 30 days	1.21	0.17	<0.001
Consumed <i>other illicit</i> substances in the prior 30 days	1.65	0.23	<0.001
<i>Exposure to deviant social influences</i>			
Gang affiliation in prior 12 months	0.65	0.08	<0.001
Friends okay with marijuana use	0.25	0.02	<0.001
<i>Personality and Attitude</i>			
Feelings of major depression in prior 12 months	0.12	0.04	0.003
Seriously considered suicide in prior 12 months	0.16	0.05	0.001
Regular participation in supervised after-school activities	−0.04	0.02	0.02
<i>School and Neighborhood Characteristics</i>			
School communicated with parents	−0.04	0.04	0.26
Bullied in prior 30 days	−0.17	0.04	<0.001
Ease of obtaining marijuana	0.21	0.02	<0.001
Community laws favorable to drug use	−0.04	0.04	0.24
Lived in urban area	−0.02	0.06	0.73
<i>Equation 2: Used marijuana, but reported otherwise</i>			
Constant	−0.67	0.36	0.07
Age	0.03	0.02	0.15
Female	0.02	0.06	0.76
White/Caucasian	0.07	0.06	0.22
Mom has a college degree	−0.08	0.08	0.31
Dad has a college degree	−0.02	0.08	0.80
Consumed alcohol in the prior 30 days	−0.25	0.09	0.01
Smoked cigarettes in the prior 30 days	−0.18	0.07	0.01

Variable	Estimate	Std. error	p-value
Misused pain medication in the prior 30 days	−0.07	0.07	0.38
Consumed <i>other illicit</i> substances in the prior 30 days	−0.52	0.08	<0.001
Grades	0.03	0.03	0.25
Able to confide in parents	−0.16	0.06	0.01
Equation 3: Did not use marijuana, but reported otherwise			
Constant	4.50	7.32	0.54
Age	−1.05	0.76	0.17
Female	−3.27	1.62	0.04
White/Caucasian	2.46	1.30	0.06
Mom has a college degree	−4.89	2.29	0.03
Dad has a college degree	2.33	1.68	0.16
Consumed alcohol in the prior 30 days	−2.98	2.90	0.30
Smoked cigarettes in the prior 30 days	6.15	2.57	0.02
Misused pain medication in the prior 30 days	0.51	1.42	0.72
Consumed <i>other illicit</i> substances in the prior 30 days	4.35	2.01	0.03
Able to confide in parents	−0.36	0.66	0.58
Friends okay with marijuana use	1.08	0.67	0.11

Table 4.

Prevalence Estimates

Variable	Estimate	Std. error	95% confidence interval	
Probit regression	12.3%	0.6%	11.1%	13.5%
Misclassification model	18.1%	1.4%	15.5%	20.8%

Table 5.**Differences Between Probit and Misclassification Model Results**

Variable	Difference	p-value
Constant	−0.32	0.21
Demographic and Socioeconomic Variables		
Age	0.05	0.03
Female	−0.02	0.37
White/Caucasian	0.01	0.39
Mom has a college degree	−0.04	0.33
Dad has a college degree	−0.03	0.36
School Bonds		
Grades	−0.05	0.03
Family Bonds		
Able to confide in parents	−0.13	0.06
Other Problem Behaviors		
Consumed alcohol in the prior 30 days	0.38	<0.001
Smoked cigarettes in the prior 30 days	0.58	<0.001
Misused pain medication in the prior 30 days	0.82	<0.001
Consumed other illicit substances in the prior 30 days	0.92	<0.001
Exposure to deviant social influences		
Gang affiliation in prior 12 months	0.2	0.03
Friends okay with marijuana use	0.06	0.01
Personality and Attitude		
Feelings of major depression in prior 12 months	0.06	0.19
Seriously considered suicide in prior 12 months	0.03	0.35
Regular participation in supervised after-school activities	−0.01	0.36
School and Neighborhood Characteristics		
School communicated with parents	−0.01	0.39
Bullied in prior 30 days	−0.04	0.29
Ease of obtaining marijuana	0.04	0.08
Community laws favorable to drug use	−0.01	0.39
Lived in urban area	0.01	0.40