

Original article

Prevalence and correlates of diabetes among criminal justice–involved individuals in the United States



Craig A. Rolling, PhD ^{a,*}, Michael G. Vaughn, PhD ^{b,c}, Dagmar Velez, MPH ^d,
Dylan B. Jackson, PhD ^e, Katherine J. Holzer, MSW ^b, Lisa Jaegers, PhD ^{b,f},
Brian B. Boutwell, PhD ^{a,b,g}

^a Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University, St. Louis, MO

^b School of Social Work, College for Public Health and Social Justice, Saint Louis University, St. Louis, MO

^c Graduate School of Social Work, Yonsei University, Seoul, Republic of Korea

^d Mercy Virtual Care Center, Chesterfield, MO

^e Department of Criminal Justice, The University of Texas at San Antonio, San Antonio

^f Department of Occupational Science and Occupational Therapy, Doisy College of Health Sciences, Saint Louis University, St. Louis, MO

^g Department of Family and Community Medicine, School of Medicine, Saint Louis University, St. Louis, MO

ARTICLE INFO

Article history:

Received 8 June 2018

Accepted 29 May 2019

Available online 7 June 2019

Keywords:

Diabetes

Probationers

Parolees

Machine learning

Regression trees

ABSTRACT

Purpose: Diabetes is one of the most prevalent and fastest-growing adverse health conditions in the United States and disproportionately affects those demographic and socioeconomic groups that are also more likely to be involved with the criminal justice (CJ) system. This study examines the prevalence and correlates of diabetes among CJ-involved individuals in the United States.

Methods: Using traditional statistical modeling and modern machine learning methods, data from the National Study on Drug Use and Health were analyzed to compare the correlates and predictive interactions of diabetes diagnosis among those respondents on probation and parole to a sample, matched by age and gender, who were not.

Results: Subjects involved in the CJ system were 15% more likely (1.66% vs. 1.44%, $P = .015$) to report a past-year diagnosis of diabetes than a sample of noninvolved individuals matched by age and sex, although this association was not statistically significant after adjusting for demographic and behavioral confounders. Similar trends in diabetes prevalence emerged for the non-CJ and CJ groups with regard to income, depression (OR of 2.38 and 1.65 for the CJ and non-CJ groups, respectively) and attainment of college education (OR of 0.64 and 0.30 for the CJ and non-CJ groups, respectively, compared with those with less than a high school education). Results also suggested that a generally high propensity toward risk taking had a negative effect on diabetes for the non-CJ group (OR 0.78; 95% CI 0.69–0.87), yet increased the odds of diabetes (OR 1.38; 95% CI 1.02–1.85) for the CJ group.

Conclusions: Involvement in the U.S. CJ system is correlated with a higher prevalence of diabetes and differing risk factors for diabetes diagnosis. Further research is necessary, however, to unpack the precise causal pathways that underlie the associational trends in the current analysis.

© 2019 Elsevier Inc. All rights reserved.

Introduction

The prevalence of diabetes has significantly increased among the U.S. adult population over the past few decades. Using cross-sectional data from the National Health and Nutrition Examination Survey (NHANES) between 1988–1994 and 2011–2012,

Menke et al. [1] estimate that the prevalence increased from 9.8% to 12.4%. An increase in prevalence was observed among males and females, each racial/ethnic group, and every education and income level; however, the most rapid increase occurred among African Americans and Mexican Americans. According to the CDC, African Americans and Hispanics now have significantly higher age-adjusted rates of diabetes prevalence than do non-Hispanic whites [2].

The long-term consequences associated with diabetes include cardiovascular disease, neuropathy, nephropathy, retinopathy, and foot damage [3]. These complications could be exacerbated for

* Corresponding author. Saint Louis University, 3545 Lafayette Avenue, St. Louis, MO 63104. Tel.: +1-314-977-8107.

E-mail address: craig.rolling@slu.edu (C.A. Rolling).

probationers and parolees, who often experience limited access to health care on release from formal incarceration [4]. Most studies examining the rate of diabetes among this population find that the prevalence estimates are similar to the U.S. general population [5–8]; however, previous studies adjusting for age indicate that the prevalence is likely higher among inmates.

For example, Wilper et al. [8] used data from the 2002 Survey of Inmates in Local Jails and the 2004 Survey of Inmates in State and Federal Correctional Facilities to analyze diabetes prevalence. Comparing their findings with the noninstitutionalized U.S. adult population data from the 2003–2004 NHANES indicated that the rate of diabetes among inmates was comparable with that of the U.S. noninstitutionalized population. However, once the results were age-standardized to the U.S. census, the prevalence of diabetes was found to be higher among inmates (11.1% of federal inmates, 10.1% of state inmates, and 8.1% of jail inmates) compared with the U.S. population (6.5%). Binswanger et al. [9] used the same data as well as the 2002–2004 NHANES data for comparison, finding that jail and prison inmates have a higher prevalence of diabetes compared with noninstitutionalized adults across three age groups, after adjusting for sex. On a smaller scale, Langevin et al. [10] investigated the prevalence of diabetes among 915 male offenders using medical records from 1964 to 2006, finding that approximately 15% of male offenders were diagnosed with diabetes, compared with 6.2% in the general population.

Although previous studies on the rate of diabetes in the criminal justice (CJ) population have focused primarily on those currently incarcerated [5,8,9,11–13], few studies have examined the rate of diabetes among probationers and parolees despite major barriers to health care for this population. Primary obstacles to care for this population include lack of insurance, inadequate coordination between the CJ system and public health organizations, and housing and employment struggles that may cause individuals to neglect their health [4]. However, the epidemiology and etiology of diabetes specifically among probationers and parolees is largely unknown.

The present study

There have been few studies of chronic health conditions such as diabetes conducted among CJ-involved populations and even fewer that have examined the correlates of these conditions in nationally representative data sources. The objective of the present study is to examine whether there are differential correlates of diabetes among CJ-involved individuals compared with those in the general population. Chronic health conditions such as diabetes may be complicated by the extensive risk histories and psychiatric comorbidity found among justice-involved populations. In addition to using logistic regression analyses to uncover these patterns of correlates, we also use machine learning methods in the form of regression trees and random forests to further explore the structure of these relationships.

Methods

Sample and procedures

This study is based on public-use data collected between 2005 and 2014 as part of the National Survey on Drug Use and Health (NSDUH) [14]. Consistent with previous studies, NSDUH data were pooled across years to increase the analytic sample size (no respondents were interviewed in multiple years) and improve population estimates [15,16]. The NSDUH interview utilizes a computer-assisted interviewing methodology to increase the likelihood of valid respondent reports of substance abuse and other high-risk behaviors [14]. Additional details regarding the NSDUH design and procedures is available elsewhere [14].

The goal of the present study is to compare the prevalence and correlates of diabetes for those involved in the CJ system (CJ group) with those for a comparable population not involved in the CJ system (non-CJ group). A total of 20,019 respondents in the 2005–2014 samples responded that they had been on probation and/or parole within the past 12 months. Each of these individuals in the CJ group was matched to ten people in the non-CJ group with the same gender and age category. Thus, the non-CJ sample had $n = 200,190$. This ten-to-one matching was performed to increase statistical power while retaining comparability between the two groups (CJ and non-CJ) with respect to age and gender.

Measures

Measures used in the NSDUH have been refined over several decades, and their psychometric properties have been shown to possess adequate reliability [7,17,18].

Diabetes

Respondents were queried regarding whether a doctor or other medical professional had diagnosed them with diabetes in the past year. This condition was dichotomously measured as yes or no.

Criminal justice system involvement

Respondents were asked if they had been on probation or parole/supervised release at any time in the past 12 months. Although probationers and parolees both are classified as being involved with the CJ system for the purpose of this study, the level of CJ involvement is different for the two groups. While those on probation may have been exposed to jail for only a short time (or not at all), respondents on parole or supervised release had previously spent time in prison.

Sociodemographics

The following demographic variables were used: age (12–17, 18–25, 26–34, 35–49, and 50 or older), gender, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, and other), total annual family income (less than \$20,000, \$20,000 to \$49,999, \$50,000 to \$74,999, and \$75,000 or more), and metropolitan population density (classified as large, ≥ 1 million; small, less than 1 million; and nonmetropolitan).

Drug and alcohol use

Indicators of dependence on four common substances (nicotine, alcohol, marijuana, and pain medication) were included in our analysis. For nicotine, the nicotine dependence syndrome scale score was used, whereas dependence on the other substances was dichotomously measured. Previous studies examining substance use disorders using the NSDUH measures have shown acceptable reliability and validity [19–21]. Selling illegal drugs (yes or no) was used based on a respondent's past year report. An item measuring the ease with which respondents could obtain illicit drugs was also included as a dichotomous measure (fairly to very easy, or not easy).

Mental health and personality characteristics

Two major chronic mental health conditions, anxiety and depression, were assessed dichotomously in the NSDUH survey and used in our analysis. Risk propensity was measured by summing two items related to the frequency or enjoyment of participation in dangerous or risky behaviors, then categorizing into low, medium, or high propensity for such behavior.

Data analysis

To examine the correlates of diabetes in the CJ and non-CJ samples, we used three analysis methods: logistic regression, a traditional statistical technique; regression trees [22], an algorithmic machine learning procedure; and random forests [23], an ensemble of regression trees designed to improve their stability and predictive power. Diabetes diagnosis was the dependent variable in each analysis, and the same set of predictors was used in each procedure. Each statistical method was performed separately for the CJ-involved (CJ group) and those not involved in the CJ system (non-CJ group). In each case, the non-CJ individuals without diabetes were used as the reference group (nonresponders). Logistic regression, regression trees, and random forests each analyze associations differently and thus allow for unique insights when applied to this large sample. All analysis was done in R software [24].

We first used logistic regression to examine correlates of diabetes among sociodemographic factors, indicators of drug and alcohol use, and mental health and personality characteristics for the CJ and non-CJ groups. Regression trees were then applied to both groups using the same dependent and predictor variables. Because of the low overall incidence of the diabetes outcome, we used regression trees rather than classification trees. Trees were fit using the `rpart` function in the `rpart` R package [25] with a complexity parameter of 0.0005 and a minimum threshold of 50 observations in each terminal node. These parameters were changed from their default values to produce more fully developed trees for additional insights. Finally, random forests were built on both groups to more accurately assess the discriminative power of individual predictor variables in the regression tree framework [26]. Random forests and the resulting variable importance plots were produced using the `randomForest` R package [27].

Results

Logistic regression

Overall, individuals involved in the CJ system were 15% more likely (1.66% vs. 1.44%, $P = .015$) to report a past-year diagnosis of diabetes than a comparable population of noninvolved individuals matched by age and gender. Odds ratios from the logistic regressions for the two groups are shown in Table 1. There are similar directional trends in diabetes prevalence within the non-CJ and CJ groups with regard to income (lower OR for each increasing level of income), depression (OR of 2.38 and 1.65 for the CJ and non-CJ groups, respectively), and attainment of college education (OR of 0.64 and 0.30 for the CJ and non-CJ groups, respectively, compared with those with less than a high school education). However, the two sets of odds ratios also feature several differences. Perhaps most notably, a high-risk propensity had a negative effect on diabetes diagnosis for the non-CJ group (OR 0.78; 95% CI 0.69–0.87), but was associated with higher odds of diabetes (OR 1.38; 95% CI 1.02–1.85) for the CJ group. Nicotine and alcohol dependence also were associated with significant increases in diabetes risk for the CJ group, but had odds ratios less than 1 (although not statistically significant) for the non-CJ group.

Regression trees

The regression tree for the non-CJ group is presented in Figure 1. The tree produced a total of nine terminal nodes, in which the prevalence of diabetes diagnosis ranged from 0.6% (those under 26 years old) to 21.6% (Hispanic and non-Hispanic blacks 50 or older). The tree's initial split divides the sample into those who are younger than 50 years and those who are 50 years or older. There is

Table 1

OR and 95% CI for diabetes outcome for non-CJ-involved and CJ-involved individuals

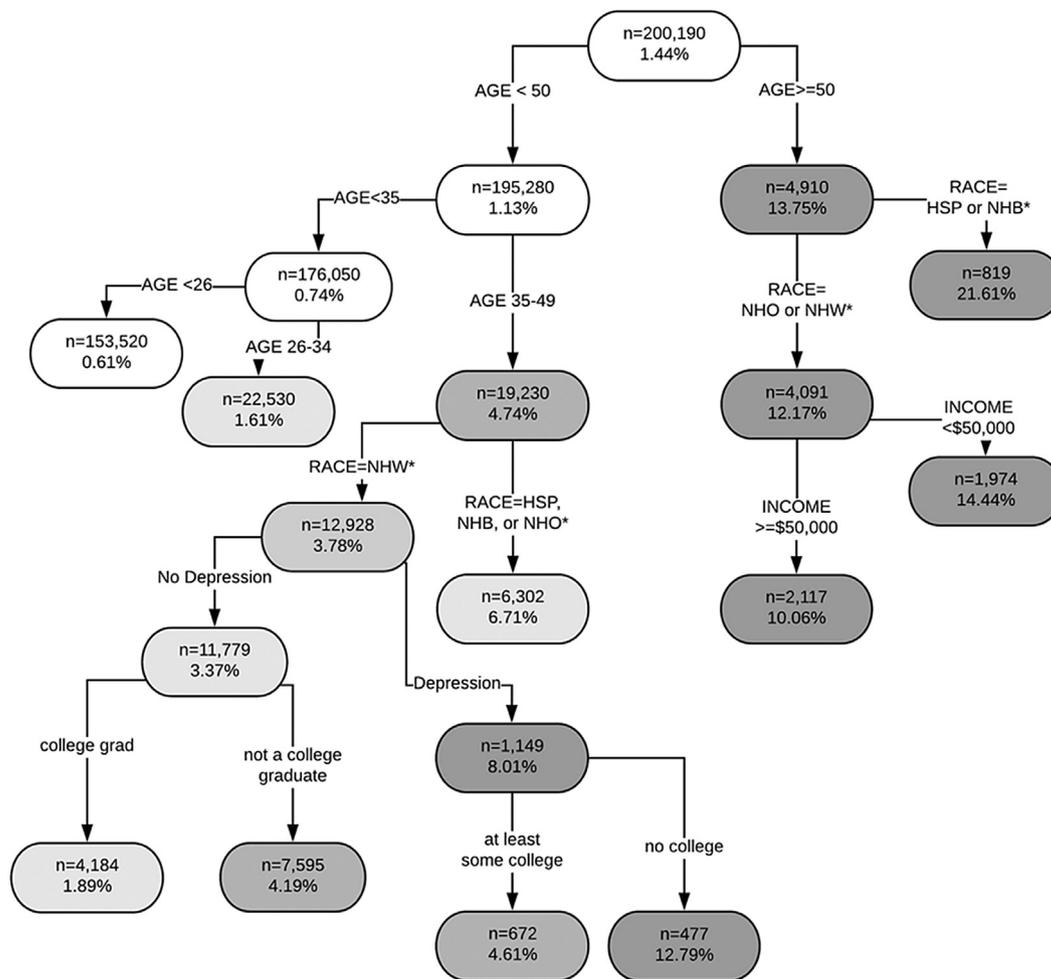
Variables	Non-CJ-involved		CJ-involved	
	OR	95% CI	OR	95% CI
Sociodemographic factors				
Age (ref = 18–25 y old)				
12–17 y old	0.55	0.42, 0.73	1.13	0.61, 2.05
26–34 y old	2.45	2.15, 2.79	1.85	1.19, 2.82
35–49 y old	7.64	6.88, 8.49	9.09	6.70, 12.4
50+ y old	24.90	22.1, 28.0	35.31	25.2, 49.4
Gender (male)	1.09	1.00, 1.18	1.37	1.08, 1.72
Race				
Hispanic	1.47	1.32, 1.64	1.79	1.28, 2.47
Non-Hispanic black	1.73	1.54, 1.93	2.59	1.90, 3.51
Non-Hispanic other	1.21	1.05, 1.39	2.35	1.65, 3.28
Income (ref ≤ \$20,000)				
\$20,000–\$49,000	0.87	0.79, 0.96	0.69	0.54, 0.89
\$50,000–\$74,999	0.79	0.69, 0.89	0.33	0.21, 0.49
>\$75,000	0.74	0.66, 0.84	0.30	0.20, 0.45
Education (ref = less than HS)				
High school graduate	0.98	0.88, 1.10	0.74	0.55, 0.99
Some college	0.92	0.81, 1.04	0.77	0.56, 1.07
College graduate	0.64	0.56, 0.74	0.30	0.18, 0.50
County type (ref = large metro)				
Small metro	0.88	0.80, 0.97	0.93	0.70, 1.25
Nonmetro	1.04	0.93, 1.17	1.40	1.03, 1.90
Drug and alcohol use				
Marijuana dependence (past yr)	0.73	0.50, 1.03	1.53	0.86, 2.56
Dependence on analgesics (past yr)	2.34	1.62, 3.28	1.85	0.88, 3.48
Nicotine dependence (NDSS score)	0.91	0.80, 1.04	2.41	1.84, 3.15
Alcohol dependence (past yr)	0.94	0.76, 1.14	1.89	1.30, 2.68
Sold illegal drugs	0.96	0.76, 1.06	1.08	0.88, 1.29
Drug access	1.15	1.01, 1.30	1.38	0.99, 1.90
Other health-related characteristics				
Depression (past year)	1.65	1.42, 1.91	2.38	1.69, 3.31
Anxiety (past year)	1.02	0.85, 1.22	1.44	0.97, 2.11
Risk Propensity (ref = low)				
Medium	0.90	0.80, 1.02	1.26	0.91, 1.72
High	0.78	0.69, 0.87	1.38	1.02, 1.85

Significant odds ratios are in bold.

a pronounced difference in prevalence in these two groups, with only 1.1% of those younger than 50 years having diabetes, compared with a prevalence of 13.7% among those 50 years and older. Subsequent splits identify predictors of diabetes within specific subgroups. For example, among those in the non-CJ group who are 50 years or over, black and Hispanic individuals were almost twice as likely to have diabetes (21.6%) as those of other races (12.2%). Deeper in the tree, we see that whether or not an individual has depression creates a large split among those who are ages 35–49 years and white: 8.0% of those who were depressed had diabetes, compared with only 3.4% who did not indicate depression.

Figure 2 shows the regression tree analysis results for the CJ group. The starting node, with a sample size of 197,639, is the sum of individuals in the non-CJ group with no diabetes and individuals in the CJ group with diabetes. The proportions shown in Figure 2 are the number of people in the CJ group with diabetes, divided by that number plus the number in the non-CJ group without diabetes. The tree produced 11 terminal nodes in which the prevalence of diabetes and CJ involvement ranged from 0.1% (for those under 35) to 17.2% (for respondents 50 or older with depression and nicotine dependence). As with the non-CJ tree results, splitting the individuals by age category of less than 50 versus 50 or older was the first partition from the initial sample. However, comparing the trees in Figures 1 and 2, we see that county type, nicotine dependence, alcohol dependence, and risk propensity all are used as partitioning variables in Figure 2 (CJ tree) but not in Figure 1 (non-CJ tree), indicating that different covariates are useful for identifying higher-prevalence groups in the CJ versus non-CJ samples.

Regression Tree for Diabetes in Non-CJ Group



*NHO=NonHispanic Other, NHW=NonHispanic White, HSP=Hispanic, NHB=NonHispanic Black

Fig. 1. Regression tree for diabetes diagnosis in non-CJ-involved sample.

Random forest variable importance plots

Random forest tree ensembles were run for the non-CJ and CJ groups to analyze the relative importance of diabetes correlates in the regression tree framework. Figure 3 is a side-by-side variable importance plot that shows the scaled mean decrease in node impurity (measured by Gini index) that results from splitting on the variable, averaged over all trees in the forest. Put simply, variables with a higher importance measure more frequently produce tree partitions that result in more homogeneous nodes. The importance measures were scaled separately for each forest so that the minimum and maximum importance values for each forest had values of 0 and 1, respectively.

The variable importance plot shows some similarities, but also some differences, between the variable importance measures in the two groups. Although age produced the most helpful splits in the non-CJ forest, it was far from the most important factor by this measure in the CJ forest. This is consistent with our findings that diabetes is more prevalent in the CJ group and appears more frequently at a younger age in this group.

County type had the highest importance measure for the CJ forest, indicating that place of residence is an important determinant of diabetes diagnosis and involvement in the CJ system. Race,

education, and easy access to hard drugs also had higher importance measures in the CJ group than in the non-CJ group.

Supplemental analysis

We conducted further analysis of diabetes as it relates to age, race/ethnicity, and the type of CJ exposure. To determine whether CJ exposure was independently associated with diabetes diagnosis in the presence of the other covariates, we combined the CJ and non-CJ groups and included CJ exposure as an independent variable with the other covariates in a single logistic regression with diabetes diagnosis as the dependent variable (Table 2). In this model, the adjusted coefficient of CJ exposure was not significant ($P = .67$). About one quarter of those in our CJ sample were on parole as opposed to probation only, and parolees possess a higher unadjusted rate of diabetes diagnosis (2.04%) compared with those on probation only (1.52%; P -value for difference = .014). However, the distinction between probation and parole was no longer significant ($P = .88$) after accounting for the other covariates in the CJ logistic regression model.

Diabetes diagnoses clearly become more common with age, and we examined the association between diabetes and CJ status stratified by age group, to determine whether there was evidence of effect modification by age.

Regression Tree for Diabetes in CJ Group

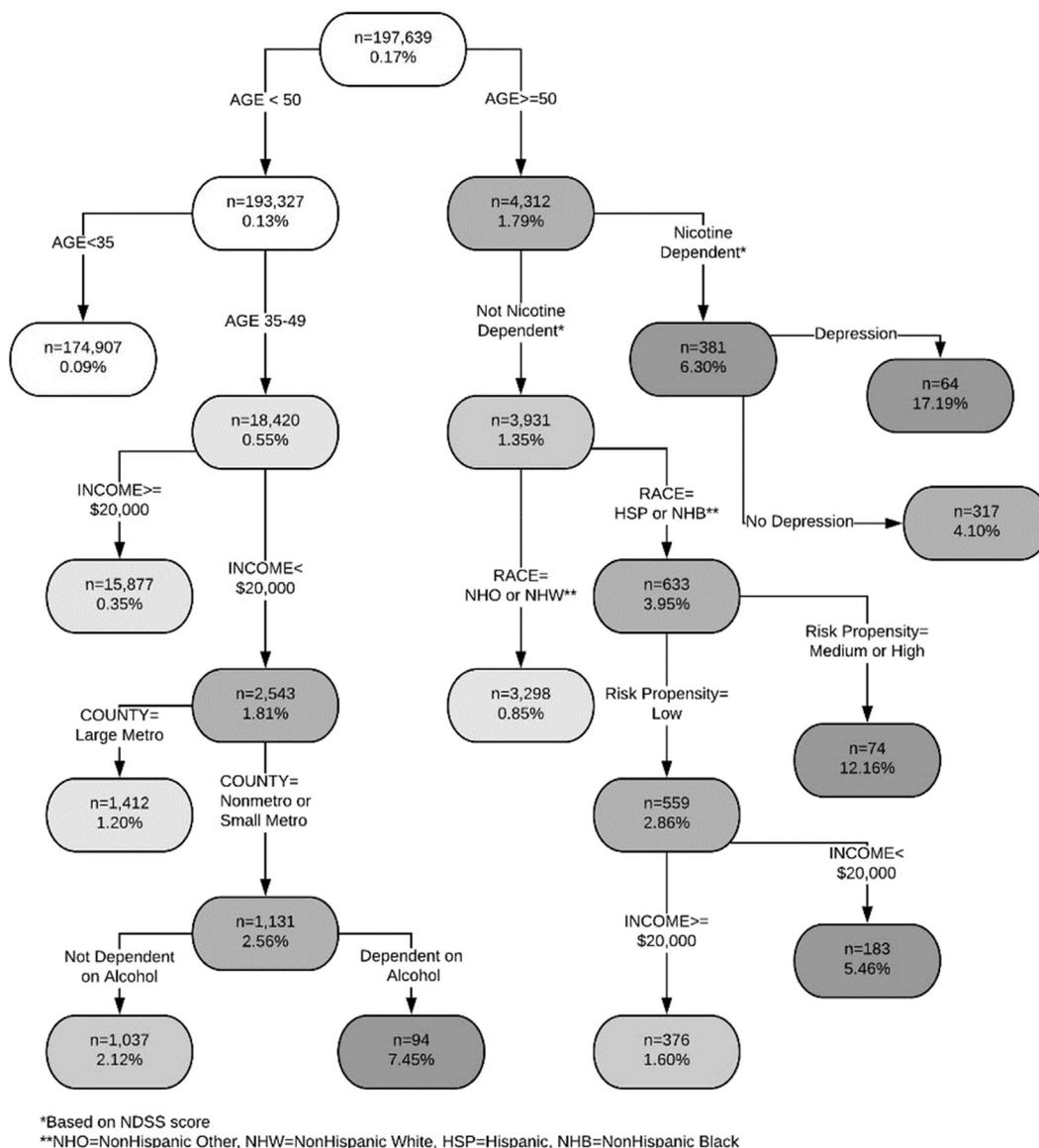


Fig. 2. Regression tree for diabetes diagnosis and CJ involvement.

There is no monotonic trend of effect modification by age, although Table 3 provides some evidence that the association between CJ involvement and diabetes is strongest for the youngest and the oldest age groups under study. Likewise, Table 3 provides no clear evidence of effect modification by race/ethnicity, although the association between CJ involvement and diabetes may be stronger for whites and those classified as “other” (mostly Asians and those of more than one race).

Discussion

CJ-involved persons are vulnerable to a host of health risks, including sexually transmitted diseases, hepatitis, pancreatitis, and various severe health limitations [7,28]. Even so, research has generally suggested that diabetes, a growing health problem in the United States and globally [1], may not be significantly more common among the justice-involved [7], despite age-adjusted analyses suggesting that diabetes may in fact be more prevalent

among jail and prison inmates than the U.S. population [8]. This body of research, however, has yet to explore the etiology and rate of diabetes among probationers and parolees, despite the challenges that these populations face in obtaining community health care and coordinate services that promote long-term health. The present study explores the correlates of diabetes among probationers and parolees, including risk propensity and psychiatric comorbidities, using a large, nationally representative data source. The study yielded a number of key findings.

First, logistic regression results indicated that CJ-involved individuals were an estimated 15% more likely to report a diagnosis of diabetes than non-CJ individuals matched on age and gender. However, the association between CJ involvement and diabetes diagnosis was not significant after adjusting for additional demographic and behavioral confounders. This suggests that the observed association is a confounded, rather than a causal, relationship. Logistic regression models also suggested that a high-risk propensity, male gender, dropping out of high school, and

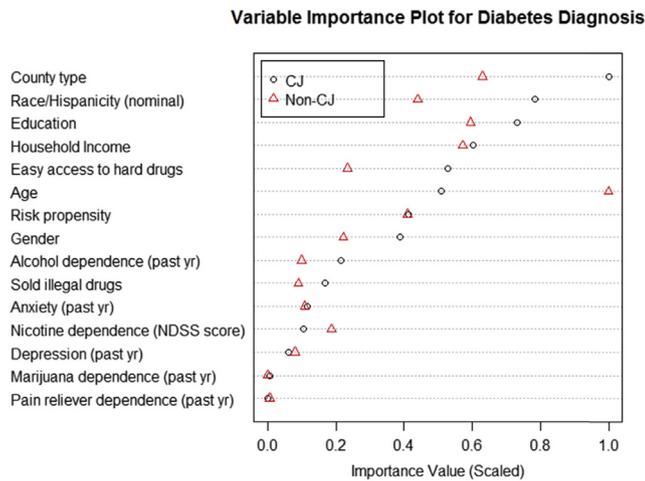


Fig. 3. Random forest side-by-side variable importance plot.

dependence on nicotine and alcohol were significantly and positively predictive of a diagnosis of diabetes among the CJ-involved, but were not significantly predictive of diabetes among the non-CJ-involved participants.

In addition, regression and classification trees were incorporated into the present study. Similar machine learning approaches have recently increased in public health more generally, and because of the large number of observations in the NSDUH data, these methods

Table 2
OR and 95% CI for diabetes outcome, non-CJ and CJ individuals combined

Variables	Entire sample	
	OR	95% CI
CJ involvement (probation or parole)	1.02	0.91, 1.16
Sociodemographic factors		
Age (ref = 18–25 y old)		
12–17 y old	0.51	0.40, 0.65
26–34 y old	2.45	2.16, 2.77
35–49 y old	7.20	6.51, 7.97
50+ y old	24.3	21.7, 27.2
Gender (male)	1.12	1.03, 1.21
Race		
Hispanic	1.23	1.11, 1.37
Non-Hispanic black	1.43	1.29, 1.60
Non-Hispanic other	1.37	1.21, 1.55
Income (ref ≤ \$20,000)		
\$20,000–\$49,000	0.94	0.86, 1.04
\$50,000–\$74,999	0.86	0.76, 0.97
>\$75,000	0.74	0.66, 0.83
Education (ref = less than HS)		
High school graduate	0.89	0.80, 1.00
Some college	0.85	0.76, 0.95
College graduate	0.61	0.53, 0.69
County type (ref = large metro)		
Small metro	0.93	0.85, 1.03
Nonmetro	1.13	1.02, 1.26
Drug and alcohol use		
Marijuana dependence (past yr)	0.84	0.62, 1.14
Dependence on analgesics (past yr)	1.21	0.83, 1.75
Nicotine dependence (NDSS score)	0.97	0.86, 1.09
Alcohol dependence (past yr)	0.81	0.67, 0.97
Sold illegal drugs	0.95	0.87, 1.04
Drug access	1.17	1.04, 1.32
Other health-related characteristics		
Depression (past year)	1.62	1.41, 1.85
Anxiety (past year)	1.13	0.96, 1.33
Risk propensity (ref = low)		
Medium	0.90	0.80, 1.01
High	0.74	0.66, 0.82

Significant odds ratios are in bold.

Table 3
Diabetes prevalence for non-CJ and CJ individuals by age and by race/ethnicity

Category	Non-CJ-involved	CJ-involved	OR (95% CI)
Age (% of CJ sample)			
12–17 y old (29%)	0.44%	0.73%	1.65 (1.17, 2.26)
18–25 y old (48%)	0.71%	0.86%	1.22 (0.96, 1.52)
26–34 y old (11%)	1.61%	1.29%	0.80 (0.54, 1.15)
35–49 y old (10%)	4.74%	5.30%	1.13 (0.91, 1.38)
50+ y old (2%)	13.7%	15.7%	1.17 (0.90, 1.50)
Race/ethnicity (% of CJ sample)			
Hispanic (20%)	1.58%	1.50%	0.95 (0.72, 1.23)
Non-Hispanic black (18%)	1.96%	1.97%	1.01 (0.78, 1.29)
Non-Hispanic other (9%)	1.27%	2.35%	1.87 (1.34, 2.57)
Non-Hispanic white (53%)	1.32%	1.49%	1.13 (0.96, 1.33)

are well-placed to shed insight on the population of interest in the present study [29]. In particular, these trees implicitly model interactions (i.e., subgroup-specific relationships) and thus may lead to insights beyond what can be learned from a logistic regression alone [30]. Regression tree results revealed that regardless of whether participants were CJ-involved or not, the prevalence of diabetes was high among older participants (i.e., 50 years and older), particularly black and Hispanic participants. Among the non-CJ-involved, depression also appeared to be linked to an increased risk of diabetes among middle-aged, white subjects. In the CJ-involved group, among those aged 50 years or older, nicotine dependence coupled with depression elevated the prevalence of diabetes substantially (from 1.79% to 17.19%). Among the CJ-involved who were 50 years and older but not dependent on nicotine, being Hispanic or black with a medium- to high-risk propensity also was associated with notable increases in the prevalence of diabetes.

Overall, the results suggest some meaningful differences in the correlates of diabetes among CJ-involved and non-CJ-involved individuals, which in turn may have important implications for how this growing medical problem might be addressed among different subpopulations defined by a history of CJ involvement. Generally speaking, the results suggest a need to more closely align the research agendas for scholars in both the public health and CJ disciplines. As research in and around this topic continues to amass, policy initiatives have an opportunity to arise and further galvanize the efforts to prevent and treat the development of diabetes among at risk subsets of the population who often have barriers to affordable, accessible health care.

What should not be understated, of course, is the associational nature of our study which precludes, at this point, making strong and direct policy recommendations rooted in some deep sense of whether CJ involvement is causing specific poor health outcomes. What does seem reasonable given the pattern of findings, however, might be targeted efforts to encourage adherence to medications on release from jail or prison, an initiative aided by the fact that key common generic drugs (such as metformin for diabetes) are often freely, or cheaply, accessible. Moreover, if adherence among probationers and parolees were made a priority among probation and parole officials, it could represent a key financial investment in the health and well-being of the estimated 4.5 million [31] U.S. probationers and parolees, operating ultimately to save millions of dollars across the health care sector otherwise stemming from complications common in untreated diabetes and psychiatric illness.

Despite the contribution that the present study makes to our understanding of the correlates and prevalence of diabetes among an understudied subset of justice-involved individuals, it is not without its limitations. First, the diabetes diagnosis is 1) likely dependent on the extent to which an individual has received health services that would detect diabetes and 2) lacking in terms of the severity of the symptomatology being manifested with the

diagnosis. Persons exposed to the CJ system may have greater screening or access to care than those with similar demographics but not in the CJ system. Prisons are required to provide medical care to all inmates, and the Federal Bureau of Prisons recommends screening inmates for diabetes. However, the frequency of screening varies widely across correctional facilities, and the extent to which greater access to care explains the increased prevalence of diabetes in the CJ group cannot be determined from the data. Relatedly, someone may have diabetes or prediabetes and not yet be aware, which is not captured in this measure. Second, data lack the necessary temporal ordering to make a strong causal inference with respect to the structure of correlates identified. In conclusion, our results—while constrained in regards to causal inference—offer a window onto various correlates of diabetes in both individuals netted in the CJ system, as well as those who have not been. While there was some overlap in risk factors across groups, some correlates seemed to be unique to CJ-involved participants. If these results ultimately replicate, they suggest that public health campaigns and policy aimed at improving health in the U.S. population may need to be tailored to fit specific risk profiles for CJ-involved and non-CJ-involved individuals.

Acknowledgment

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- [1] Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and trends in diabetes among adults in the United States, 1988–2012. *JAMA* 2015;314(10):1021–9.
- [2] Centers for Disease Control and Prevention. National diabetes statistics report. Estimates of diabetes and its burden in the United States. <https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf>. [Accessed 5 July 2019].
- [3] Mayo Foundation for Medical Education, Research. Diabetes. <https://www.mayoclinic.org/diseases-conditions/diabetes/symptoms-causes/syc-20371444>. [Accessed 22 March 2018].
- [4] Rich JD, Chandler R, Williams BA, Dumont D, Wang EA, Taxman FS, et al. How health care reform can transform the health of criminal justice-involved individuals. *Health Aff (Millwood)* 2014;33(3):462–7.
- [5] Baillargeon J, Black SA, Pulvino J, Dunn K. The disease profile of Texas prison inmates. *Ann Epidemiol* 2000;10(2):74–80.
- [6] Vaughn MG, DeLisi M, Beaver KM, Perron BE, Abdon A. Toward a criminal justice epidemiology: behavioral and physical health of probationers and parolees in the United States. *J Criminal Justice* 2012;40(3):165–73.
- [7] Vaughn MG, Salas-Wright CP, DeLisi M, Piquero AR. Health associations of drug-involved and criminal-justice-involved adults in the United States. *Crim Justice Behav* 2014;41(3):318–36.
- [8] Wilper AP, Woolhandler S, Boyd JW, Lasser KE, McCormick D, Bor DH, et al. The health and health care of U.S. prisoners: results of a nationwide survey. *Am J Public Health* 2009;99(4):666–72.
- [9] Binswanger IA, Krueger PM, Steiner JF. Prevalence of chronic medical conditions among jail and prison inmates in the United States compared with the general population. *J Epidemiol Community Health* 2009;63(11):912–9.
- [10] Langevin R, Langevin M, Curnoe S, Bain J. The prevalence of diabetes among sexual and violent offenders and its co-occurrence with cognitive impairment, mania, psychotic symptoms and aggressive behavior. *Int J Prisoner Health* 2008;4(2):83–95.
- [11] D'Souza RM, Butler T, Petrovsky N. Assessment of cardiovascular disease risk factors and diabetes mellitus in Australian prisons: is the prisoner population unhealthier than the rest of the Australian population? *Aust N Z J Public Health* 2005;29(4):318–23.
- [12] Falter RG. Selected predictors of health services needs of inmates over age 50. *J Correctional Health Care* 1999;6(2):149–75.
- [13] Harzke AJ, Baillargeon JG, Pruitt SL, Pulvino JS, Paar DP, Kelley MF. Prevalence of chronic medical conditions among inmates in the Texas prison system. *J Urban Health* 2010;87(3):486–503.
- [14] U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. National Survey on Drug Use and Health 2002–2014 (NSDUH-2002-2014-DS0001) 2019. <https://datafiles.samhsa.gov/>.
- [15] Nakawaki B, Crano WD. Predicting adolescents' persistence, non-persistence, and recent onset of nonmedical use of opioids and stimulants. *Addict Behav* 2012;37(6):716–21.
- [16] Hedden S, Gfroerer J, Barker P, Smith S, Pemberton MR, Saavedra LM, et al. Comparison of NSDUH mental health data and methods with other data sources. In CBHSQ Data Review. Substance Abuse and Mental Health Services Administration (US). 2012.
- [17] Chromy JR, Feder M, Gfroerer J, Hirsch E, Kennet J, Morton KB. Reliability of key measures in the National Survey on Drug Use and Health (HHS Publication No. SMA 09-4425, Methodology Series M8). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies; 2010.
- [18] Kennet J, Painter D, Hunter SR, Granger RA, Bowman KR. Assessing the reliability of key measures in the National Survey on Drug Use and Health using test-retest methodology. Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies; 2006.
- [19] Fearn NE, Vaughn MG, Nelson EJ, Salas-Wright CP, DeLisi M, Qian Z. Trends and correlates of substance use disorders among probationers and parolees in the United States 2002–2014. *Drug Alcohol Depend* 2016;167:128–39.
- [20] Jordan BK, Karg RS, Batts KR, Epstein JF, Wiessen C. A clinical validation of the National Survey on Drug Use and Health assessment of substance use disorders. *Addict Behav* 2008;33(6):782–98.
- [21] Grucza RA, Abbacchi AM, Przybeck TR, Gfroerer JC. Discrepancies in estimates of prevalence and correlates of substance use and disorders between two national surveys. *Addiction* 2007;102(4):623–9.
- [22] Breiman L, Friedman J, Olshen R, Stone C. Classification and regression trees. Wadsworth: Chapman and Hall/CRC; 1984.
- [23] Breiman L. Random forests. *Machine Learn* 2001;45(1):5–32.
- [24] R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018. <https://www.R-project.org/>. [Accessed 5 July 2019].
- [25] Therneau T, Atkinson B, Ripley B. rpart: Recursive Partitioning and Regression Trees. R package version 4.1–13. 2015.
- [26] Archer KJ, Kimes RV. Empirical characterization of random forest variable importance measures. *Comput Stat Data Anal* 2008;52(4):2249–60.
- [27] Liaw A, Wiener M. Classification and regression by randomForest. *R News* 2002;2(3):18–22.
- [28] Schnittker J, John A. Enduring stigma: the long-term effects of incarceration on health. *J Health Soc Behav* 2007;48(2):115–30.
- [29] Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med* 2003;26(3):172–81.
- [30] Buchner F, Wasem J, Schillo S. Regression trees identify relevant interactions: Can this improve the predictive performance of risk adjustment? *Health Econ* 2017;26(1):74–85.
- [31] U.S. Bureau of Justice Statistics. Probation and Parole in the United States, 2016. NCJ 251148. <https://www.bjs.gov/index.cfm?ty=pbdetail&iid=6188>. [Accessed 26 April 2018].