

Using Random Forest to examine factors associated with unprotected sexual behavior among Black Americans post-release from incarceration in New York City

Running head: post-release sexual behavior

Tawandra L. Rowell-Cunsolo, PhD; Yue Long, MA; Betsy Szeto, BS; Rahma Mkuu, MPH; and Nabila El-Bassel, PhD

Tawandra L. Rowell-Cunsolo, PhD\* (corresponding author)

Assistant Professor of Social Welfare Science

Columbia University

School of Nursing

617 West 168th Street

New York, NY 10032

212-305-8611 (Phone)

212-543-6003 (Fax)

[tr2301@cumc.columbia.edu](mailto:tr2301@cumc.columbia.edu)

Yue Long, MA

Columbia University

Quantitative Methods in the Social Sciences

Betsy Szeto

Columbia University

Mailman School of Public Health

Rahma Mkuu, MPH

Texas A & M University

Nabila El-Bassel, PhD

Columbia University

Social Intervention Group

#### Acknowledgements

This work was supported by the National Institute on Drug Abuse under Grant number K01DA036411 (PI: Rowell-Cunsolo).

## Abstract

**Introduction:** Black Americans are overrepresented among incarcerated individuals and those infected with human immunodeficiency virus (HIV), which is largely driven by sexual behavior. Because HIV risk behaviors reportedly increase post-incarceration, we assessed the extent to which they are engaging in risk behavior, which is essential for developing the most effective corrections and community-based prevention interventions.

**Methods:** One hundred and seventy-one formerly incarcerated individuals were interviewed about their engagement in unprotected sexual behavior post-incarceration. Participants were recruited from agencies in the metropolitan New York City area. We conducted a log-binomial regression and random forest to assess factors associated with unprotected sexual behavior.

**Results:** Most of our participants (64%) reportedly engaged in unprotected sexual behavior post-release from incarceration. According to our regression results, illicit drug use was a predictor of unprotected sexual behavior ( $RR=1.31$ , 95%CI [1.02, 1.74],  $p < .05$ ), especially cocaine usage ( $RR=1.34$ , 95%CI [1.08, 1.67],  $p < .05$ ). Our random forest results demonstrated that total time in jail and cocaine usage were among the most important variables associated with unprotected sexual behavior.

**Conclusions:** Formerly incarcerated Black Americans are at an elevated risk for HIV post-incarceration. Harm reduction approaches such as drug treatment and pre-exposure prophylaxis (PrEP) should be considered for this population.

## Introduction

Black Americans are incarcerated at disproportionate rates in the United States, which currently leads the world in the number of citizens incarcerated per capita (The Sentencing Project, 2014). Recent statistics show that Black men are eight times more likely than White men to be incarcerated, and Black women are three times more likely than White women to experience incarceration (A. Carson, 2014).

In addition to their disproportionate incarceration, according to the Centers for Disease Control and Prevention (CDC), nationally, human immunodeficiency virus (HIV) prevalence rates among Black Americans remain disproportionately high. The Black population in the U.S. currently represents the segment of the U.S. population where HIV is steadily increasing and for which the HIV burden is most substantial (Centers for Disease Control and Prevention, 2015). Black Americans comprise 12 percent of the U.S. population, but 46 percent of new HIV infections in 2013 (Centers for Disease Control and Prevention, 2015).

A growing body of research has shown that incarceration is associated with an elevated risk for HIV transmission for Black Americans (Dauria et al., 2015; Epperson, El-Bassel, Gilbert, Orellana, & Chang, 2008; Khan, Behrend, Adimora, Weir, Tisdale, et al., 2011; Khan et al., 2009; Krebs, 2006). Incarceration has been linked to dissolution of primary intimate relationships (Khan, Behrend, Adimora, Weir, White, et al., 2011). Because incarceration removes such a substantial number of eligible Black men from their communities, it creates an imbalance in the male-to-female sex ratio, limiting the availability of sexual partners (Johnson & Raphael, 2009), which promotes relationship instability (Khan et al., 2008; Knittel, Snow, Riolo, Griffith, & Morenoff, 2015). This scenario places Black Americans at an elevated risk for HIV infection, even when their sexual behavior is not any riskier than individuals from other racial groups (Hallfors, Iritani, Miller, & Bauer, 2007).

Unprotected sexual behavior is a substantial driver of HIV among Black Americans (Centers for Disease Control and Prevention, 2015). Because incarceration is disruptive to sexual partnerships, the period immediately following incarceration is particularly prone to HIV vulnerability. This period may be chaotic for formerly incarcerated individuals attempt to rebuild their lives (Clear, Rose, & Ryder, 2001; Holtfreter & Watanaporn, 2014), and may include an increase in unprotected sexual behavior, transactional sexual activities, and the initiation of multiple (or concurrent) sexual partnerships, which are all substantial contributors to HIV infection (Adams et al., 2011; Adimora & Schoenbach, 2005; Adimora et al., 2004; Grieb, Davey-Rothwell, & Latkin, 2012; Khan, Behrend, Adimora, Weir, White, et al., 2011; Thomas, Levandowski, Isler, Torrone, & Wilson, 2008).

Recent data suggest that the number of prisoners being released currently may exceed the number being admitted into correctional facilities (E. A. Carson & Golinelli, 2013). Approximately 600,000 individuals are released from prisons each year; a significant proportion are Black Americans (E. A. Carson & Golinelli, 2013), and many have a history of drug dependency (Mumola & Karberg, 2006), both of which are significant predictors of HIV seropositivity (Hagan, Perlman, & Des Jarlais, 2011; McQuillan et al., 2006).

Presumably, the longer an individual is incarcerated, the more fractured and distant his relationships with sexual partners become. Although history of incarceration has been recognized as a risk factor for HIV transmission, more information about the risk behaviors of formerly incarcerated individuals is needed to devise programs that adequately respond to their risk profile and prevention needs. Given the aforementioned risks associated with incarceration among Black Americans, the primary purpose of this study was threefold: 1) to examine the extent to which incarceration duration is associated with unprotected sexual behavior among Black Americans recently released from incarceration; 2) to identify additional factors (in addition to incarceration duration) that may be associated with unprotected sexual

behavior among formerly incarcerated Black Americans; and 3) to assess the utility of using random forest to establish factors associated with unprotected sexual behavior post-incarceration among our sample.

## Methods

### Participants and Recruitment

This cross-sectional study was conducted from January 2014 through August 2015. Individuals who met the following inclusion criteria were eligible for participation: 1) released from prison or jail; 2) living in the metropolitan New York City area; and 3) above the age of 18 years old. This sample was recruited from community-based agencies that provide services to ex-offenders in the New York City metropolitan area, including those offering substance abuse treatment centers transitional housing organizations and homeless shelter, harm reduction agencies and programs, and employment agencies. Recruitment flyers and advertisements with basic information about the study were posted at agencies throughout the metropolitan area, and given to agency staff to distribute. The flyers contained a phone number for individuals to call if they were interested in participating in the study. Research staff also gave presentations at orientation meetings and educational seminars for individuals recently released from incarceration. Finally, research participants were also encouraged, and compensated, for recruiting other individuals into the study.

Participants were interviewed by trained research assistants. During the informed consent process participants were reminded that their participation was anonymous, no information would be shared with any criminal justice agencies or representatives, and that a Certificate of Confidentiality was issued to further protect their rights. Study protocols were approved by the Columbia University Medical Center Institutional Review Board.

## Measures

Demographics. Data on socio-demographics such as gender, age, race and ethnicity, sexual orientation, and marital status were collected. Because of low frequencies for some categories, race (Black and non-Black), marital status (single and married), and sexual orientation (heterosexual and non-heterosexual) were collapsed into two categories. Participants were also asked to disclose whether they had medical insurance and insurance type.

Criminal history. We assessed the number of times individuals had been incarcerated as an adult, and type of offense(s) committed. Participants were also asked whether they were previously and currently on probation or parole, total incarceration duration in jail and prison time, and time since release from incarceration. Total jail and prison time were coded as continuous variables measured in days.

Illicit drug use. Participants were asked to report any illicit drug use post-release from prison. They were also asked to provide the types of drugs used. For drug use categories, a code of 0 was assigned if they had not used a specific drug post-incarceration, and a 1 if they reported using that specific drug.

Sexual behavior. Participants were asked whether they engaged in unprotected sexual behavior since they had been released from incarceration. Other sexual behavioral measures include the number of male and female sexual partners post-release from incarceration.

## Data Analytical Plan

Descriptive statistics were used to examine the distribution of the sample. We conducted chi-square tests to determine whether there were any differences between participants who reportedly engaged in unprotected sexual behavior post-release from incarceration versus those who did not. Fisher's exact test was used for variables with expected cell counts of less than five. We also used t-tests to compare the means of continuous variables. We screened for statistically significant variables at the .10 level and

calculated their risk ratio using log-binomial regression. Lastly, we entered statistically significant variables at the .20 level during bivariate analyses into random forest model to predict unprotected sexual behavior, and selected variables that were the most effective in identifying unprotected sexual behavior. All data analytical procedures were performed using R statistical software.

In addition to our regression analysis, we used random forest (RF) procedures to assess factors associated with unprotected sexual behavior. Random forest is an algorithm for classification developed by Breiman that uses an ensemble of classification trees to predict outcomes and to rank the importance of variables in establishing predictions (Boulesteix, Janitza, Kruppa, & Konig, 2012; Breiman, 2001). In RF, each of the classification trees is built using a bootstrap sample of the data, and at each split the candidate set of variables is a random subset of the variables. For each tree, two-thirds of the sample is bootstrapped for training, and the remaining one-third is referred to as the out-of-bag samples (OOB). The bootstrap results are aggregated (bagging) to decrease ensemble variance, and the random variable selection contributes to the low correlation of individual trees. As a result, the algorithm yields an ensemble that can achieve both low bias and low variance (Breiman, 2001; Diaz-Uriarte & de Andres, 2006; Kuhn & Johnson, 2013).

Random forest has several characteristics that make it ideal for our dataset. Random forest reportedly has great predictive performance (Kuhn & Johnson, 2013), and the advantage of quantifying the impact of predictors in the ensemble by calculating variable importance. To quantify the impact of predictors of unprotected sexual behavior, we used the method introduced by Breiman to calculate variable importance in random forest (Breiman, 2001), which calculates differences in prediction accuracy on OOB samples before and after permuting each predictor variable. Since the trees are trained independently on bootstrapped samples, and the forests are independently trained on each of the training subsets during data

resampling, we averaged the local scores and standardized it by its standard deviation to obtain the raw importance score for each predictor. Figure 1 illustrates this approach for each forest.

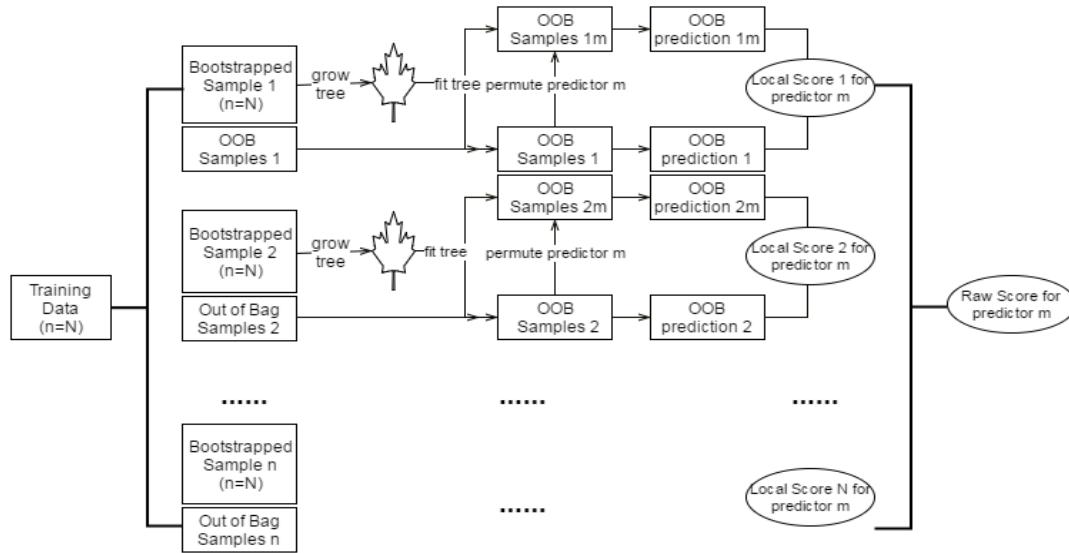


Figure 1: Diagram of Variable Importance Calculation Process for Predictor m in Random Forest

Because the outcome in our dataset (unprotected sexual behavior) was imbalanced with a larger proportion of positive samples (64% Yes vs. 36% No), we utilized *post hoc* sampling approaches to attenuate the effects of class imbalance during model training. Previous research suggests that, in general, for low-dimensional data such as ours, down-sampling techniques are more effective than oversampling techniques (Hulse, M., & Napolitano, 2007). Thus, we adopted two down-sampling schemes to deal with class imbalance in our study: down sampling inside of data resampling (DOWN) and balanced random forest with down sampling (BRF). The details of the BRF and DOWN algorithm are described elsewhere (Chen, Liaw, & Breiman, 2004; Hulse et al., 2007; Kuhn, 2016). Because drug use after release was categorized as heroin use, cocaine use and marijuana use, we substituted the summary variable with three dummies in the analysis, which enabled us to maintain the full dataset.

To examine our models' performance, we used repeated 10-fold cross-validation, recommended by Kuhn and Johnson (2013), which reportedly reduces bias and variance without substantially increasing computational costs. For each repetition of the 10-fold cross validation on the full dataset ( $n=171$ ,  $p=8$ ), we trained RF on the 9-fold training subsets to predict one of the 50 resampling subsets. The prediction was compared with actual class outcomes and evaluated using measurements including the receiver operating characteristic (ROC), sensitivity, specificity, accuracy, Kappa, precision and f1-score. We calculated quartile points (25%, 50%, 75%), min, max, and mean for each measurement.

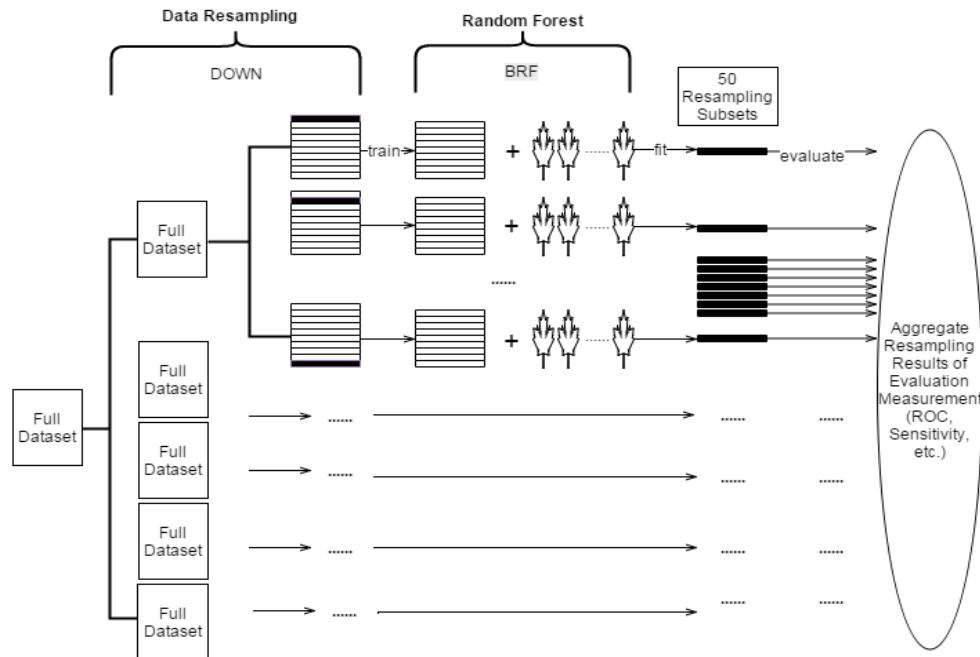


Figure 2: Diagram of Analytical Plan with Repeated 10-fold CV and Subsampling for Class Imbalance

We considered different parameter configurations for *mtry* (number of predictors considered at each split) following the suggestion of numerous scholars (Breiman, 2003; Kuhn & Johnson, 2013; Liaw & Wiener, 2002). We set *ntree* (number of trees for the forest) at 1500 and *nodesize* (the minimum size of terminal nodes) at 1. As Figure 2 demonstrates, DOWN was conducted during data resampling, whereas

BRF was grown directly on balanced bootstrap samples. We compared their results with those obtained without any subsampling (RF). With different subsampling methods, the size of OOB samples and local score from the classification tree differs, resulting in variability in the size of the score, regardless of the similarity of their ranking. In recognition of this scenario, we also scaled the raw score by dividing by the highest score.

## Results

The vast majority of our 171 participants were Black Americans (97%; n=165). Our sample was largely male (82%) and heterosexual (80%). The age range was 19 to 66 years old ( $M=45.0$ ,  $SD=10.1$ ). On average, participants had been released from incarceration within the previous 10 months. The mean for incarceration duration was 3.5 years (range: 5 days to 32 years). Although most participants reported drug offenses (89%), participants were also incarcerated for committing violent (19%) and property offenses (17%). Of those who reportedly used illicit drugs after being released, the most commonly substances used were crack and/or powder cocaine (59%), marijuana (55%), and heroin (31%).

Unprotected sexual behavior post-incarceration was reported by most participants (64.3%). The average number of sexual partners for these participants was 2.37 ( $SD=3.14$ ). On average, women reported 1.89 ( $SD=2.08$ ) sexual partners post-release from incarceration while men reported 2.47 ( $SD=3.32$ ) sexual partners post-release. For women, the average number of male sexual partners was 2.41 ( $SD=2.38$ ). While 10.6 percent of the men who engaged in unprotected sexual behavior reportedly had sex with men, on average, the number of females partners for these men was 3.37 ( $SD=4.41$ ). On average, the men who reportedly had sex with other men had 6.06 ( $SD=9.10$ ) male sexual partners.

As shown in Table 1, at the bivariate level, participants who reportedly used illicit drugs after release from incarceration were more likely than those who did not used drugs to have unprotected sex

(70% vs. 53%,  $p=.03$ ). The odds of engaging in unprotected sexual behavior were higher among participants who used marijuana ( $p=0.07$ ), heroin ( $p=0.10$ ), and cocaine (0.01) compared to individuals who did not use those substances after being released from incarceration. When the means for participant age, incarceration duration and time since release from incarceration were examined (Table 2), total time spent in jail was the only statistically significant predictor of unprotected sexual behavior. As Table 2 demonstrates, participants who spent more time in jail during their lifetime (in years) had a lower odds of having unprotected sex ( $M = 3.1$ ,  $SD= 3.6$ ;  $p=0.05$ ). Total time spent in prison was not statistically significant. As summarized in Table 3, the univariate risk ratio of engaging in unprotected sex were significantly higher among participants who reportedly used illicit drugs post-release from incarceration ( $RR=1.31$ ,  $95\%CI [1.02, 1.74]$ ,  $p <.05$ ), especially cocaine users ( $RR=1.34$ ,  $95\%CI [1.08, 1.67]$ ,  $p <.05$ ).

Figure 3 presents the mean ROC over 50 resampling subsets for each value of  $mtry$  different across sub-sampling schemes. We observed that random forest with no subsampling (RF) and balanced random forest (BRF) had stable performance across parameter values, whereas random forest with inside down sampling (DOWN) did not. The mean ROC was roughly between 60% and 63% for all sampling schemes across  $mtry$ , confirming claims that in RF, parameter tuning did not have much impact on model performance (Kuhn & Johnson, 2013), especially with low-dimensional datasets.

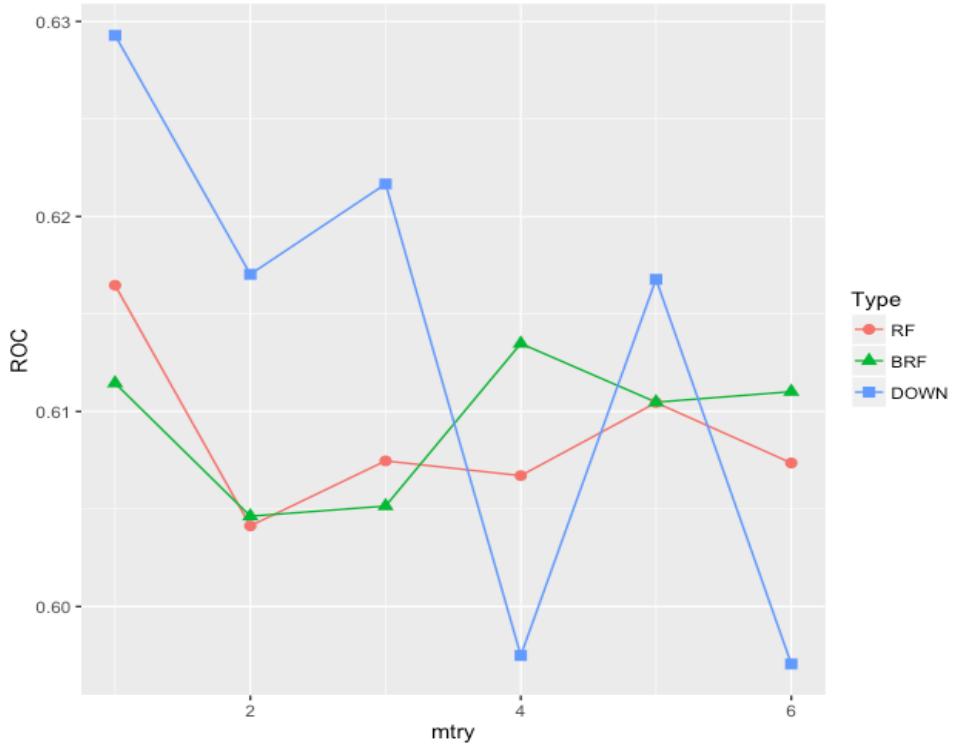


Figure 3. Mean ROC on Resampling Subsets over Tuning Values for  $mtry$

Table 4 presents evaluation results on the resampling subsets using the optimal tuning parameter value (based on mean ROC). The optimal  $mtry$  for both RF and DOWN was 1, while the optimal  $mtry$  for BRF was 4. Mean ROC was above 60% across all sub-sampling methods, where it was the highest by a small margin using DOWN (62.93%). Sensitivity was high with RF (97.82%) and BRF (95.82%), but it was reduced to 54.36% with DOWN. However, specificity was very low for both RF and BRF, but increased to 59.10% for DOWN. F1 score was high for both RF (77.53%) and BRF (76.89%), and average with DOWN (60.64%).

Table 5 presents the variable importance table calculated on out-of-bag samples. Numbers in the brackets represent raw scores. The results demonstrated that time since release from prison (TSRP) and total time in jail (TTJ) were ranked among the top three variables across all sub-sampling methods. Cocaine use post-incarceration was the most important *discrete* variable in the predictor set; cocaine use

was constantly ranked the forth place across sub-sampling methods. Heroin use was identified as the least important predictor.

## Discussion

The aim of this study was to explore predictors of unprotected sexual behavior among formerly incarcerated individuals, primarily Black Americans. Comparable to prior research in this area (Morrow, 2009; Seal, Eldridge, Kacanek, Binson, & MacGowan, 2007; Stephenson et al., 2006), a large proportion of our study sample reportedly engaged in unprotected sexual behavior relatively quickly post-release from incarceration. Illicit drug use, in particular, was a predictor of unprotected sexual behavior (Meade et al., 2014), as was observed in a previous study exploring condom use before incarceration (Braithwaite & Stephens, 2005). Cocaine users, who represented 37 percent of our sample, were especially more likely than non-cocaine users to reportedly engage in unprotected sexual behavior. Given the well-documented associations between drug use and risky sexual behavior (Vosburgh, Mansergh, Sullivan, & Purcell, 2012), particularly cocaine usage (Harrell, Mancha, Petras, Trenz, & Latimer, 2012; Hatfield, Horvath, Jacoby, & Rosser, 2009; Maranda, Han, & Rainone, 2004), the importance of this relationship cannot be underestimated.

For formerly incarcerated individuals, drug use is routinely associated with recidivism (Gray, Fields, & Maxwell, 2001; Prendergast, 2009), especially during the year following release from incarceration; the vast majority are re-incarcerated within a few years (Durose, Cooper, & Snyder, 2014). Hence, preventive interventions and educational programs targeting this population should also include strategies for reducing substance use. Participants in our study reported a high number of previous incarcerations, and although it was not associated with unprotected sexual behavior, it does suggest that periods of incarceration may be in opportune moment to provide risk reduction counseling and/or educational interventions for this population (Underhill, Dumont, & Operario, 2014). Participants who

spent more time in jail were less likely to report unprotected sexual behavior after release from incarceration, aligning with other study findings comparing risk behaviors pre- and post-incarceration (Adams & Kendall, 2011). This may indicate that jail time may serve as a protective factor for formerly incarcerated individuals (Hearn, Whitehead, Khan, & Latimer, 2015).

Our results also demonstrated that random forest is a rigorous and effective analytical strategy for examining predictors of HIV risk behaviors among vulnerable populations. The ability to rank the importance of variables in predicting HIV risk behaviors can be used to inform targeted prevention programs for formerly incarcerated individuals. Our analysis confirmed previous research suggesting that down sampling demonstrates superior performance for low-dimensional data (Hulse et al., 2007). There was some overlap between the results from our regression and RF models.

This study does have a number of limitations that should be acknowledged. The information collected in this study is highly sensitive, and may have been underreported by our participants. Although information about unprotected sexual behavior was collected, data on the type of sexual acts and their frequency were not collected, which may mask the magnitude of the risk behaviors in which our participants were engaged. Also, since our participants are self-selected and our convenience sample approach is vulnerable to selection bias, our findings should not be generalized to other formerly incarcerated individuals. Despite these limitations, our study findings highlight the importance of engaging formerly incarcerated Black Americans in HIV risk reduction programs and interventions. These data can be used to inform targeted prevention strategies for individuals who have experienced incarceration.

## Conclusions

Discharge planning programs that include HIV prevention and educational information designed to promote healthy behaviors among those who are being released from incarceration are warranted.

Formerly incarcerated individuals with a history of illicit drug use, in particular, may be in need of targeted programs designed to reduce the risks associated with their drug use. Because illicit drug use is associated with risky sexual behavior, such services and programs should be accessible prior to release from incarceration. Since prison-based substance use programs are inconsistently offered to the prison population (National Institute on Drug Abuse, 2012), correctional officials should seek to strengthen and expand their drug abuse programming to ensure that effective programs are implemented in a manner that is respective of the human rights of incarcerated individuals with substance use problems.

Prevention strategies for formerly incarcerated Black Americans should include information on HIV transmission modes, the importance of healthy sexual decision-making, and drug treatment. Specifically, harm reduction approaches such as evidence-based drug treatment and administration of pre-exposure prophylaxis (PrEP) should be considered. Implementation in the correctional institution and/or post-release community for the period immediately following incarceration should be explored. Providing this segment of the population with the most effective tools to reduce their risk of infection is crucial to reducing the burden of HIV among vulnerable populations.

## References

- Adams, J., Nowels, C., Corsi, K., Long, J., Steiner, J. F., & Binswanger, I. A. (2011). HIV Risk After Release From Prison: A Qualitative Study of Former Inmates. *Jaids-Journal of Acquired Immune Deficiency Syndromes*, 57(5), 429-434. doi: 10.1097/QAI.0b013e31821e9f41
- Adimora, A. A., & Schoenbach, V. J. (2005). Social context, sexual networks, and racial disparities in rates of sexually transmitted infections. *J Infect Dis*, 191 Suppl 1, S115-122. doi: 10.1086/425280
- Adimora, A. A., Schoenbach, V. J., Martinson, F., Donaldson, K. H., Stancil, T. R., & Fullilove, R. E. (2004). Concurrent sexual partnerships among African Americans in the rural south. *Ann Epidemiol*, 14(3), 155-160. doi: 10.1016/s1047-2797(03)00129-7
- Boulesteix, A. L., Janitza, S., Kruppa, J., & Konig, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery*, 2(6), 493-507. doi: 10.1002/widm.1072
- Braithwaite, R., & Stephens, T. (2005). Use of protective barriers and unprotected sex among adult male prison inmates prior to incarceration. *International Journal of Std & Aids*, 16(3), 224-226. doi: Doi 10.1258/0956462053420112
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. doi: Doi 10.1023/A:1010933404324
- Breiman, L. (2003). Manual on setting up, using, and understanding random forests. Berkeley, CA.
- Carson, A. (2014). *Prisoners in 2013*. (NCJ 247282). Washington, DC: U.S. Department of Justice.
- Carson, E. A., & Golinelli, D. (2013). *Prisoners in 2012: Trends in admissions and releases, 1991-2012*. (NCJ 243920). Washington, DC.
- Centers for Disease Control and Prevention. (2015). *HIV among African Americans*. Atlanta, GA.

- Chen, C., Liaw, A., & Breiman, L. (2004). Using random forest to learn imbalanced data. *Discovery*(1999), 1-12.
- Clear, T. R., Rose, D. R., & Ryder, J. A. (2001). Incarceration and the community: The problem of removing and returning offenders. *Crime & Delinquency*, 47(3), 335-351. doi: Doi 10.1177/0011128701047003003
- Dauria, E. F., Oakley, L., Arriola, K. J., Elifson, K., Wingood, G., & Cooper, H. L. F. (2015). Collateral consequences: implications of male incarceration rates, imbalanced sex ratios and partner availability for heterosexual Black women. *Culture Health & Sexuality*, 17(10), 1190-1206. doi: 10.1080/13691058.2015.1045035
- Diaz-Uriarte, R., & de Andres, S. A. (2006). Gene selection and classification of microarray data using random forest. *Bmc Bioinformatics*, 7. doi: Artn 3  
10.1186/1471-2105-7-3
- Durose, M. R., Cooper, A. D., & Snyder, H. N. . (2014). *Recidivism of prisoners released in 30 states in 2005: Patterns from 2005 to 2010*. Washington, DC.
- Epperson, M., El-Bassel, N., Gilbert, L., Orellana, E. R., & Chang, M. (2008). Increased HIV risk associated with criminal justice involvement among men on methadone. *AIDS Behav*, 12(1), 51-57. doi: 10.1007/s10461-007-9298-4
- Gray, M. K., Fields, M., & Maxwell, S. R. (2001). Examining probation violations: who, what, and when. *Crime & Delinquency*, 47(4), 537-557.
- Grieb, S. M., Davey-Rothwell, M., & Latkin, C. A. (2012). Concurrent sexual partnerships among urban African American high-risk women with main sex partners. *AIDS Behav*, 16(2), 323-333. doi: 10.1007/s10461-011-9954-6

- Hagan, H., Perlman, D. C., & Des Jarlais, D. C. (2011). Sexual risk and HIV infection among drug users in New York City: a pilot study. *Subst Use Misuse*, 46(2-3), 201-207. doi: 10.3109/10826084.2011.521470
- Hallfors, D. D., Iritani, B. J., Miller, W. C., & Bauer, D. J. (2007). Sexual and drug behavior patterns and HIV and STD racial disparities: the need for new directions. *Am J Public Health*, 97(1), 125-132. doi: 10.2105/ajph.2005.075747
- Harrell, P. T., Mancha, B. E., Petras, H., Trenz, R. C., & Latimer, W. W. (2012). Latent classes of heroin and cocaine users predict unique HIV/HCV risk factors. *Drug and Alcohol Dependence*, 122(3), 220-227. doi: 10.1016/j.drugalcdep.2011.10.001
- Hatfield, L. A., Horvath, K. J., Jacoby, S. M., & Rosser, B. R. S. (2009). Comparison of Substance Use and Risky Sexual Behavior Among a Diverse Sample of Urban, HIV-Positive Men Who Have Sex with Men. *Journal of Addictive Diseases*, 28(3), 208-218. doi: 10.1080/10550880903014726
- Hearn, L. E., Whitehead, N. E., Khan, M. R., & Latimer, W. W. (2015). Time Since Release from Incarceration and HIV Risk Behaviors Among Women: The Potential Protective Role of Committed Partners During Re-entry. *Aids and Behavior*, 19(6), 1070-1077. doi: 10.1007/s10461-014-0886-9
- Holtfreter, K., & Watanaporn, K. A. (2014). THE TRANSITION FROM PRISON TO COMMUNITY INITIATIVE An Examination of Gender Responsiveness for Female Offender Reentry. *Criminal Justice and Behavior*, 41(1), 41-57. doi: 10.1177/0093854813504406
- Hulse, J. V., M., Khoshgoftaar T., & Napolitano, A. (2007). *Experimental perspectives on learning from imbalanced data*. Paper presented at the 24th international conference on Machine learning., Corvallis, Oregon: Oregon State University.

- Johnson, R. C., & Raphael, S. (2009). The Effects of Male Incarceration Dynamics on Acquired Immune Deficiency Syndrome Infection Rates among African American Women and Men. *Journal of Law & Economics*, 52(2), 251-293.
- Khan, M. R., Behrend, L., Adimora, A. A., Weir, S. S., Tisdale, C., & Wohl, D. A. (2011). Dissolution of primary intimate relationships during incarceration and associations with post-release STI/HIV risk behavior in a Southeastern city. *Sex Transm Dis*, 38(1), 43-47. doi: 10.1097/OLQ.0b013e3181e969d0
- Khan, M. R., Behrend, L., Adimora, A. A., Weir, S. S., White, B. L., & Wohl, D. A. (2011). Dissolution of primary intimate relationships during incarceration and implications for post-release HIV transmission. *Journal of Urban Health*, 88(2), 365-375.
- Khan, M. R., Doherty, I. A., Schoenbach, V. J., Taylor, E. M., Epperson, M. W., & Adimora, A. A. (2009). Incarceration and high-risk sex partnerships among men in the United States. *J Urban Health*, 86(4), 584-601. doi: 10.1007/s11524-009-9348-5
- Khan, M. R., Miller, W. C., Schoenbach, V. J., Weir, S. S., Kaufman, J. S., Wohl, D. A., & Adimora, A. A. (2008). Timing and duration of incarceration and high-risk sexual partnerships among African Americans in North Carolina. *Annals of Epidemiology*, 18(5), 403-410. doi: 10.1016/j.annepidem.2007.12.003
- Knittel, A. K., Snow, R. C., Riolo, R. L., Griffith, D. M., & Morenoff, J. (2015). Modeling the community-level effects of male incarceration on the sexual partnerships of men and women. *Social Science & Medicine*, 147, 270-279.
- Krebs, Christopher P. (2006). INMATE FACTORS ASSOCIATED WITH HIV TRANSMISSION IN PRISON\*. *Criminology & Public Policy*, 5(1), 113-135.

- Kuhn, M. (2016). *caret: Classification and Regression Training. R package version 6.0-68*. Retrieved from <https://CRAN.R-project.org/package=caret>
- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* New York: Springer.
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest.
- Maranda, M. J., Han, C. L., & Rainone, G. A. (2004). Crack cocaine and sex. *Journal of Psychoactive Drugs*, 36(3), 315-322.
- McQuillan, G. M., Kruszon-Moran, D., Kottiri, B. J., Kamimoto, L. A., Lam, L., Cowart, M. F., . . . Spira, T. J. (2006). Prevalence of HIV in the US household population: the National Health and Nutrition Examination Surveys, 1988 to 2002. *J Acquir Immune Defic Syndr*, 41(5), 651-656. doi: 10.1097/01.qai.0000194235.31078.f6
- Meade, C. S., Bevilacqua, L. A., Moore, E. D., Griffin, M. L., Gardin, J. G., Potter, J. S., . . . Weiss, R. D. (2014). Concurrent Substance Abuse Is Associated with Sexual Risk Behavior among Adults Seeking Treatment for Prescription Opioid Dependence. *American Journal on Addictions*, 23(1), 27-33. doi: 10.1111/j.1521-0391.2013.12057.x
- Morrow, K. M. (2009). HIV, STD, and hepatitis risk behaviors of young men before and after incarceration. *AIDS care*, 21(2), 235-243. doi: 10.1080/09540120802017586
- Mumola, Christopher J., & Karberg, Jennifer C. (2006). Drug Use and Dependence, State and Federal Prisoners, 2004 (pp. 12p-12p). Washington, DC: Office of Justice Programs, Bureau of Justice Statistics.
- National Institute on Drug Abuse. (2012). *Principles of Drug Abuse Treatment for Criminal Justice Populations: A Research-Based Guide*. (NIH publication No. 11-5316). Bethesda, MD: National Institutes of Health, National Institute on Drug Abuse Retrieved from [www.drugabuse.gov/publications/principles-drug-abuse-treatment-criminal-justice-populations](http://www.drugabuse.gov/publications/principles-drug-abuse-treatment-criminal-justice-populations).

- Prendergast, M. L. (2009). Interventions to promote successful re-entry among drug-abusing parolees. *Addiction Science & Clinical Practice*, 5(1), 4-13.
- Seal, D. W., Eldridge, G. D., Kacanek, D., Binson, D., & MacGowan, R. J. (2007). A longitudinal, qualitative analysis of the context of substance use and sexual behavior among 18- to 29-year-old men after their release from prison. *Social Science & Medicine*, 65(11), 2394-2406. doi: <http://dx.doi.org/10.1016/j.socscimed.2007.06.014>
- Speed, T. (2003). *Statistical analysis of gene expression microarray data*. Boca Raton, FL: CRC Press.
- Stephenson, B. L., Wohl, D. A., McKaig, R., Golin, C. E., Shain, L., Adamian, M., . . . Kaplan, A. H. (2006). Sexual behaviours of HIV-seropositive men and women following release from prison. *Int J STD AIDS*, 17(2), 103-108. doi: 10.1258/095646206775455775
- The Sentencing Project. (2014). Trends in U.S. Corrections. [http://sentencingproject.org/doc/publications/inc\\_Trends\\_in\\_Corrections\\_Fact\\_sheet.pdf](http://sentencingproject.org/doc/publications/inc_Trends_in_Corrections_Fact_sheet.pdf)
- Thomas, J. C., Levandowski, B. A., Isler, M. R., Torrone, E., & Wilson, G. (2008). Incarceration and sexually transmitted infections: a neighborhood perspective. *J Urban Health*, 85(1), 90-99. doi: 10.1007/s11524-007-9231-1
- Underhill, K., Dumont, D., & Operario, D. (2014). HIV Prevention for Adults With Criminal Justice Involvement: A Systematic Review of HIV Risk-Reduction Interventions in Incarceration and Community Settings. *American Journal of Public Health*, 104(11), E27-E53. doi: 10.2105/Ajph.2014.302152
- Vosburgh, H. W., Mansergh, G., Sullivan, P. S., & Purcell, D. W. (2012). A Review of the Literature on Event-Level Substance Use and Sexual Risk Behavior Among Men Who Have Sex with Men. *Aids and Behavior*, 16(6), 1394-1410. doi: 10.1007/s10461-011-0131-8

