

## Evaluating substance abuse treatment programs within the US: reasons for program non-completion for dropouts and transfers

Sarah Kurihara, Immanuel Abdi, Utkarsh Baranwal, Vrinda Mittal

### Introduction

Within the United States, substance abuse results in an economic burden of more than 400 billion dollars. Within this, tens of billions of dollars are spent every year in the context of substance abuse treatment programs (Meara, 2005). A large proportion of the individuals who enter substance abuse programs do not end up staying in their program until completion. Due to the continued substance abuse among individuals who leave these programs, the high proportion of individuals who leave substance abuse programs results in inefficiencies in the capital spending in substance abuse treatment programs, while also taking up capacity from other individuals who have a higher chance of completing a treatment program.

Therefore, we are primarily interested in understanding the reasons for program non-completion so that recommendations could be made to improve the completion rate for treatment facilities across the United States. When considering the reasons for program non-completion, two main cases stand out as having government, policy-related, or systematic reasons for occurrence: Dropouts and Transfers. We seek to understand for both events: who is dropping out, who is transferring and why. The main goal would be to have individuals enter the substance abuse system and successfully complete their program on the first try. Therefore, we also look to understand what the top-performing treatment facilities are doing differently and compare them to the bottom five treatment facilities to see if we can understand the reasons for program failure.

### Data Wrangling and Cleaning

#### Data Organization

To ensure that the data met the assumption of independence required for hypothesis testing and did not bias the models, we only considered observations without a previous treatment episode in the Treatments dataset. This removed the observations where the individual had been in the substance abuse system before. Despite removing these observations, the sample size is still sufficient for data analysis purposes with over 1 million observations.

Group	NOPRIOR	COMPLETED	Description	N
A	0	1	Successful 1st attempt	464,316
B	0	0	Unsuccessful 1st attempt	858,825
C	1	1	Successful 2nd+ attempt	932,056
D	1	0	Unsuccessful 2nd+ attempt	1,438,106
E	-9	0 or 1	Unknown attempt number	355,833

**Table 1.** To avoid issues with dependent samples, Groups C, D, and E were removed from the analysis.

**Data Cleaning**

The data was primarily categorical, so it was converted to factors when it was imported. The data was then re-encoded using the data dictionary provided. The Treatments dataset ranges from 2017 to 2020, but all years were kept in the dataset because there was no significant difference in completion rates across the years.

**Feature Engineering**

The original Treatments dataset contains 81 columns. Exploratory data analysis (EDA) drastically reduced the number of features used in modeling and exploratory data analysis in order to reduce collinearity between variables. For example, GENDER was included, but PREG was not since these variables are highly correlated with one another (i.e., only women can be pregnant). Similarly, DSMCRIT was included but PSYPROB was not since both variables speak to alcohol vs. mental health disorders.

**Exploratory Data Analysis****Initial EDA**

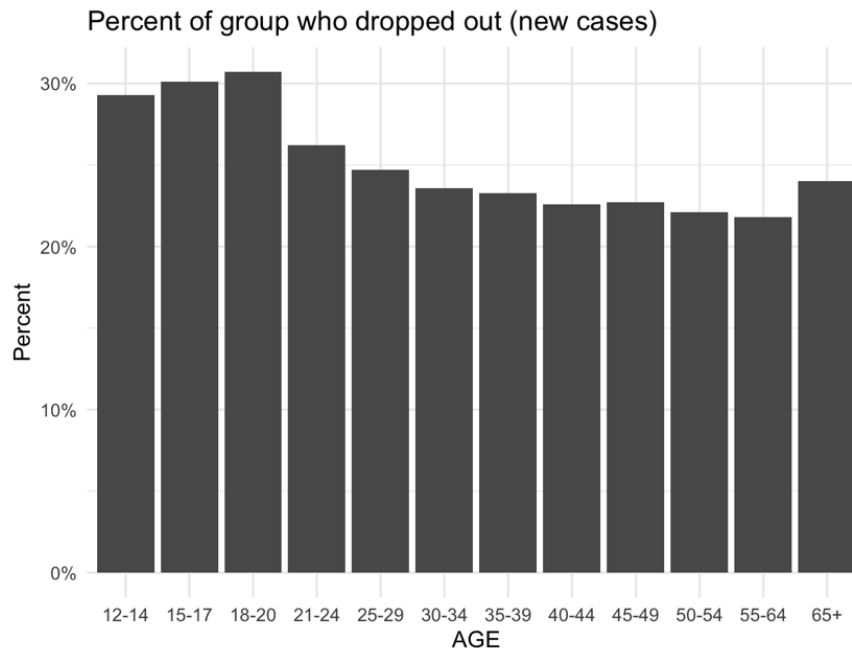
64.91% of all cases are not completed, with the majority of the non-completed cases due to transfers (30.8%) and dropouts (24.2%). Terminations and incarcerations occur at a much lower rate (5% and 1.2%), respectively.

	New Cases		Repeat Cases	
Reason	Count	Percentage	Count	Percentage
Completed	464,316	35.1%	932,056	39.3%
Dropped Out	320,427	24.2%	604,845	25.5%
Terminated	66,563	5%	149,679	6.3%
Transferred	407,641	30.8%	534,496	22.6%
Incarcerated	15,223	1.2%	41,630	1.8%
Death	2,186	0.2%	6,703	0.3%
Other	46,785	3.5%	100,753	4.3%
<b>TOTAL</b>	1,323,141		2,370,162	

**Table 2. Total counts for new cases and repeat cases**

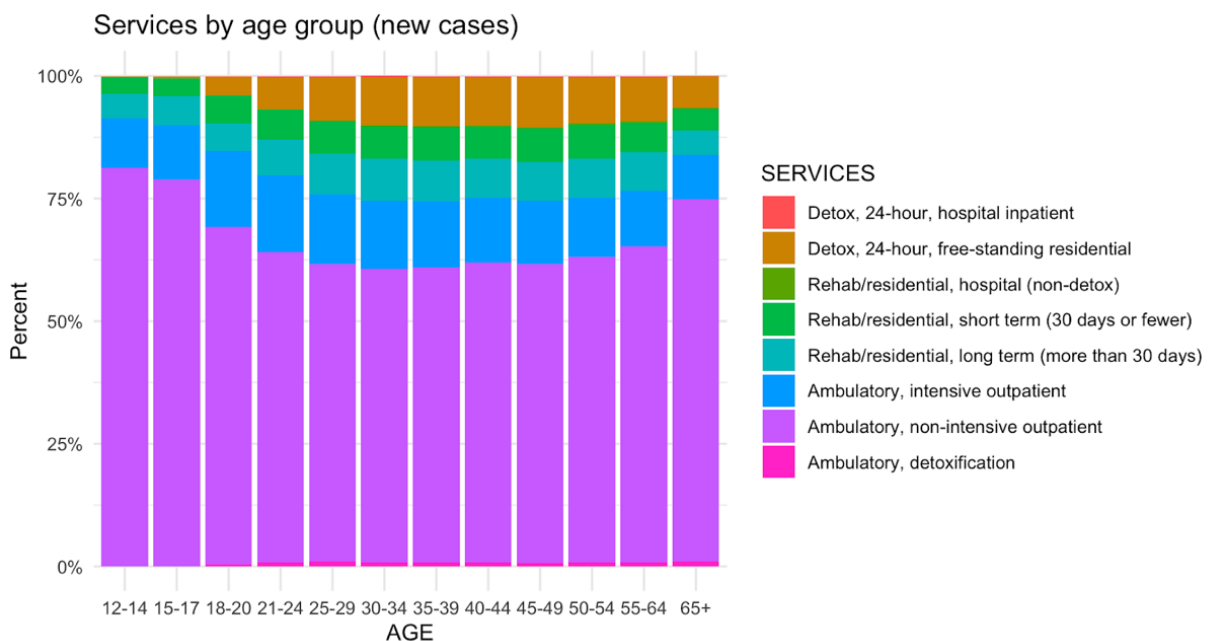
**Drop-Outs**

Demographic features were explored to understand individuals who dropped out of treatment. Controlling for age, we found that 32% of teens (ages 12-20) are dropping out of treatment, a rate that is higher than any other age group. A 2-sample proportion test was run to investigate the hypothesis that teens have an equal drop-out rate compared to the next generation (ages 21-34). A p-value of 0.001 indicates that we reject the null hypothesis that dropout rates for teens are equal and conclude that dropout rates are greater for teens than for all other generations.

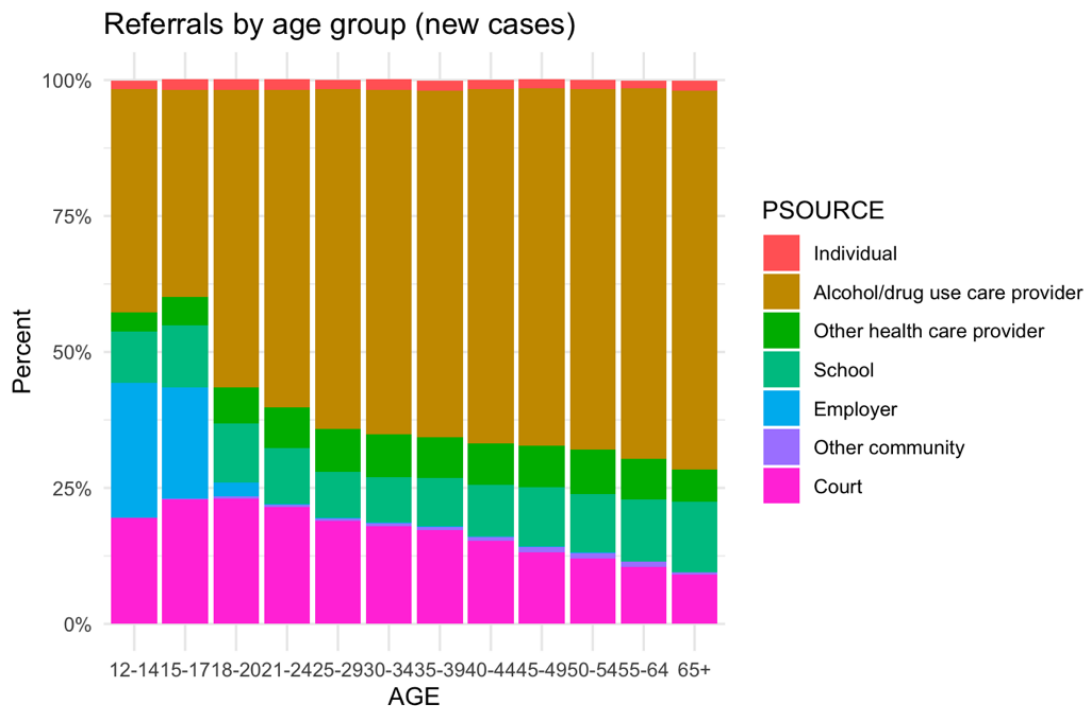


**Figure 1. Dropout Rate by Age:** Teenagers have higher rates of dropout than older populations ( $p=0.001$ )

To understand what programs these teens are involved in, we dove deeper into the service type, the primary substance used, and the primary referral source. We found that the majority of teens (>50%) were admitted to ambulatory, non-Intensive outpatient programs and were primarily referred to these programs by their employers. Alternatively, adults are enrolled in more severe 24-hour, detox programs and are typically referred by their alcohol or drug care provider.

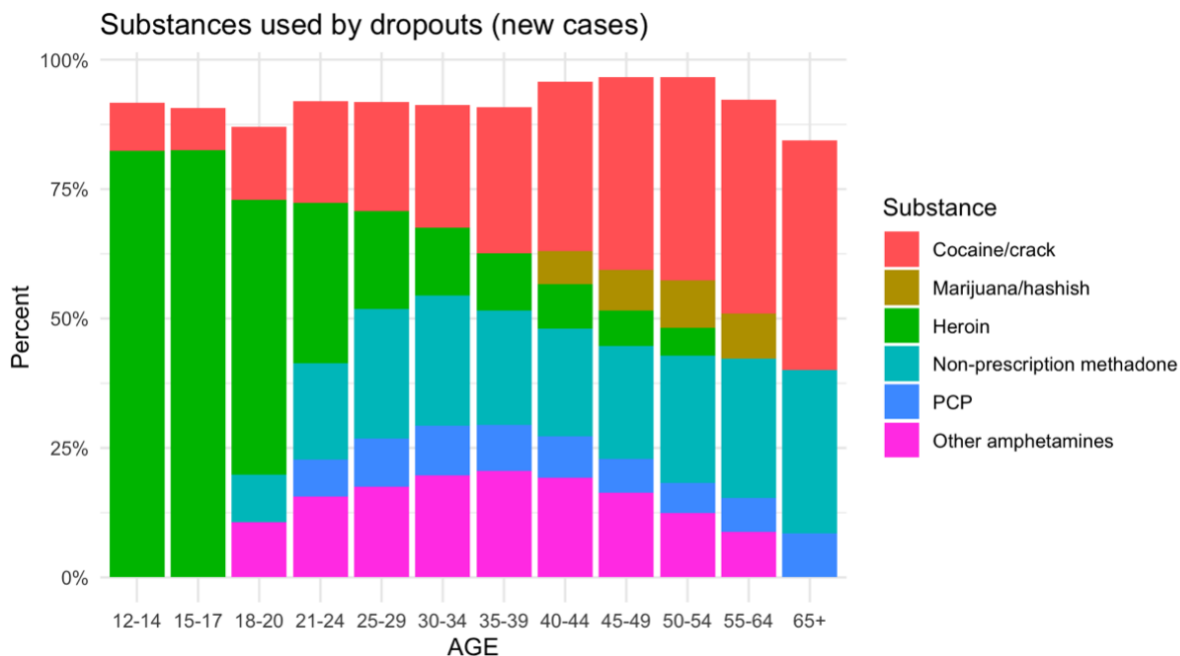


**Figure 2. Services by Age:** Teens attend low-level care facilities (Ambulatory, non-intensive outpatient) at greater rates, whereas adults attend high-level care facilities (detox, 24-hour, hospital inpatient, and free-standing residential)



**Figure 3. Primary Referral Source by Age:** Teens are typically referred to treatment facilities by their employer, whereas adults are referred to treatment facilities by the alcohol/drug care provider

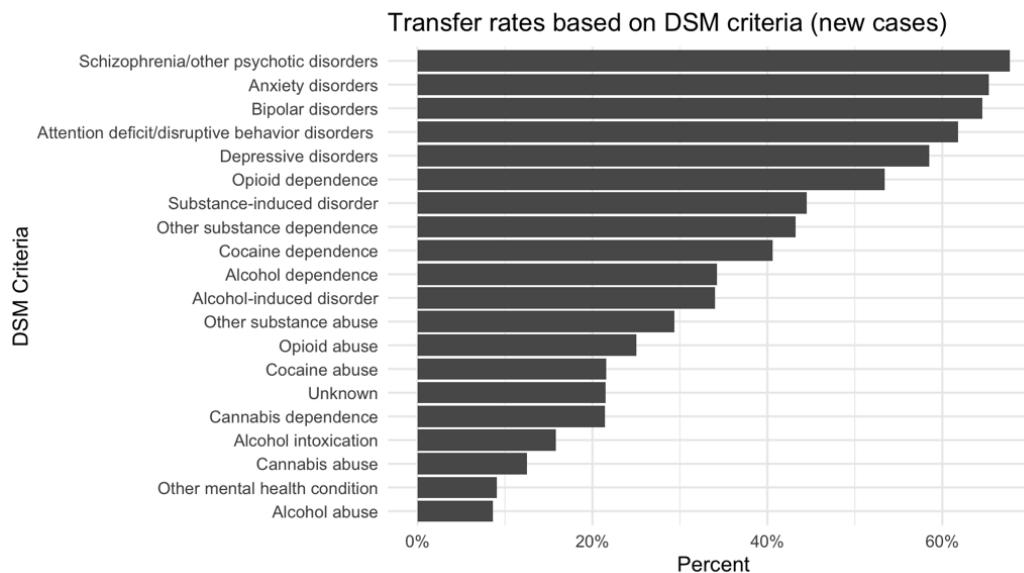
Lastly, we explored the primary substance used by these demographic groups and found that teens (12-20) are predominantly using heroin (25-75%), whereas adults are using non-prescription methadone and cocaine/crack and methadone.



**Figure 4. Substances Used by Age:** Teenagers are more likely to use heroin than adults, whereas adults typically use opiates and other synthetic drugs.

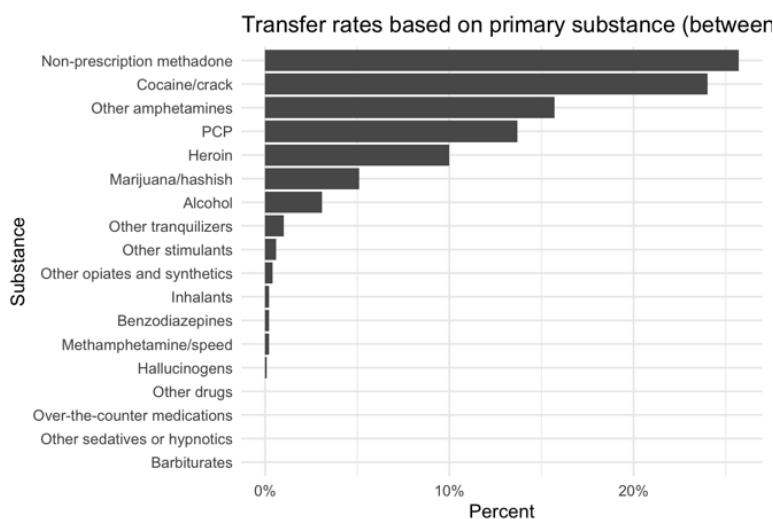
## Transfers

Over 60% of Schizophrenia, Anxiety, Bipolar, and Attention Deficit disorders cases are transferred. This highlights the significant impact that mental health has on completing substance abuse treatment programs.

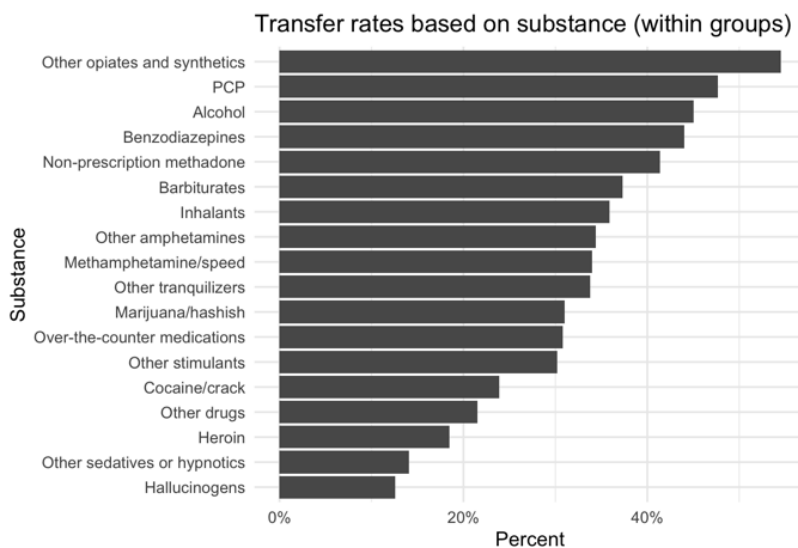


**Figure 5. Transfer Rates per DSM Criteria:** Mental health disorders make up the majority of transfer-related discharges.

25.7% of cases that result in a transfer are treating Methadone addiction and 24% of cases that result in a transfer are treating Cocaine/crack. Among the top substances with the highest transfer rates include Opiates/Synthetics (54.5%), PCP (47.7%), and Alcohol (45%). This indicates the need for more treatment facilities specific to Methadone and Cocaine, while users who use Opiates, PCP, and Alcohol need alternative forms of treatment to support their specific substance.

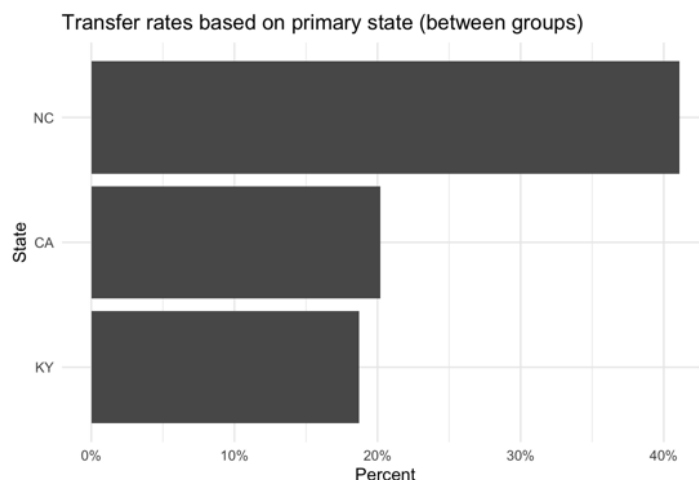


**Figure 6a: Top 3 Substances Used in Transfer Cases (Between Analysis):** Methadone (n=103,410, 25.7%), Cocaine/crack (n=96798, 24%), Amphetamines (n=63228, 15.7%).

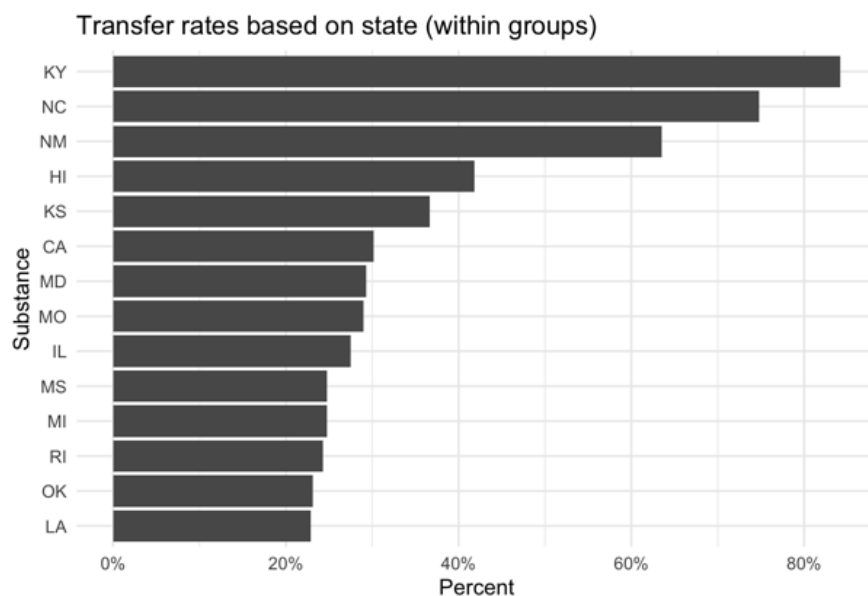


**Figure 6b: Top 3 Substances Used in Transfer Cases (Within Analysis):** Opiates/Synthetics (n=1700, 54.5%), PCP (n=55195, 47.7%), Alcohol (n=12448, 45%)

Similarly, to evaluate the impact of geography on transfer rates, the same analysis was run using the State FIPS codes. For new cases, nearly 41% of all transfers occur from North Carolina, indicating that triage is not effectively placing individuals in this state. Looking within a state, Kentucky has a 74.8% rate of transfer. This indicates the need for additional support and funding for treatment centers in Kentucky. Both Kentucky and North Carolina appear as states with high transfer rates, regardless of looking between states or within these groups. Therefore, more resources should be provided to these states to facilitate better treatment options for individuals.



**Figure 7a: Top 3 States Involved in Transfer Cases for New Cases:** North Carolina (n=167363, 41.1%), California (n=82140, 20.2%), Kentucky (n=76074, 18.7%)



**Figure 7b: Top 3 States Involved in Transfer Cases for New Cases and Repeat Cases:** Kentucky (n=167363, 74.8%), North Carolina (n=167363, 74.8%), New Mexico (n=61, 63.5%)

## Modeling

### *Predicting outcomes of substance abuse treatment*

Analysis was executed elucidating which characteristics are strong indicators for treatment completion, patient transfer, and patient dropout. We construct three supervised random forest learning models to focus on each of these 3 separate cases. These models focused on the top five states with the highest treatment completion rates. The models are constructed with the premise that if we can identify the factors that determine treatment completion, dropout, and getting transferred based on data from states with high completion rates, they will lead to insights into why these states are succeeding. Demographic, economic, program treatment, substance abuse, and previous arrest information were all used as features within the usage of training the Random Forest models.

These models show that length of stay, services required at the treatment facility, age, and substance being used were all the most important features towards determining completion. In terms of transfers, length of stay, age services required at the treatment facility, and their frequency of attending self-help were the most important determinants. For predicting a dropout, length of stay, frequency of attending self-help, mental illness, and primary substance of abuse were all the most important determinants. Lastly, we found that specific substances, such as methamphetamines and marijuana, lower self-help attendance rate, and race leads to higher patient transfer rates.

Model Prediction Category	Treatment Completion	Patient Transfer	Patient Drop Out
Training Accuracy	100%	100%	100%
Testing Accuracy	76.8%	86.3%	80%

**Table 5. Evaluation of performance within each model for predicting treatment completion, drop out, and getting transferred to a different facility.**

## **Conclusion**

Within this analysis, we analyzed the causes of dropouts and transfers within substance abuse programs in the United States. We discovered that there are various regions, treatment-specific, biological, as well as substance-type characteristics that influence whether an individual will complete a substance abuse program. Firstly, initial statistical analysis has indicated that age is a strong predictor of completion, with people aged 12-20 having the highest rate of treatment program dropout. Additionally, mental abuse disorders such as schizophrenia, anxiety, bipolar disorder, and attention deficit disorder yield high rates of facility transfer. The type of substance that an individual uses also results in differences transferring to different facilities. Methadone, opiates, PCP, and alcohol all result in common transfers.

We also executed an analysis using a supervised machine learning approach to determine whether we can predict if an individual will finish, transfer, or drop out of a treatment program. These models show that length of stay, services required at the treatment facility, age, and substance being used were all the most important features towards determining completion. In terms of transfers, length of stay, age services required at the treatment facility, and their frequency of attending self-help were the most important determinants. For predicting a dropout, length of stay, frequency of attending self-help, mental illness, and primary substance of abuse were all the most important determinants. Lastly, we found that specific substances, such as methamphetamines and marijuana, lower self-help attendance rate, and race leads to higher patient transfer rates. We hope this analysis provides information for treatment centers to better evaluate patients and increases the rates of individuals who successfully complete their substance abuse treatment programs within the first attempt.

## **References**

Meara E, Frank RG. Spending on substance abuse treatment: how much is enough? *Addiction*. 2005 Sep;100(9):1240-8. doi: 10.1111/j.1360-0443.2005.01227.x. PMID: 16128713; PMCID: PMC1402649.