



Original article

## Predictors of Adolescents' First Episode of Homelessness Following Substance Use Treatment


 Graham T. DiGuseppi, Sc.M.<sup>a,\*</sup>, Jordan P. Davis, M.S.W., Ph.D.<sup>a</sup>, Daniel Leightley, M.Sc., Ph.D.<sup>b</sup>, and Eric Rice, M.A., Ph.D.<sup>a</sup>
<sup>a</sup>Suzanne Dworak-Peck School of Social Work, University of Southern California, Los Angeles, California

<sup>b</sup>King's Centre for Military Health Research, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

Article history: Received July 10, 2019; Accepted November 23, 2019

Keywords: Substance use; Victimization; Homelessness; Machine learning

---

 See Related Editorial on p.381
 

---

### A B S T R A C T

**Purpose:** A growing body of research has identified correlates (i.e., predictors) of youth homelessness. However, such risk and protective factors have not been identified for youth receiving substance use treatment. Using characteristics collected at treatment intake, the present study sought to identify predictors of youths' first episode of homelessness during the 12 months after substance use treatment entry.

**Methods:** Data come from a longitudinal study of adolescents (n = 17,911; aged 12–17 years) receiving substance use treatment throughout the U.S. Participants completed surveys at intake and at 3, 6, and 12 months later. Logistic regression and Lasso machine learning regression were used to predict participants' first episode of homelessness in the 12 months after treatment intake.

**Results:** After excluding adolescents reporting previous experiences of homelessness, 5.0% of adolescents reported their first episode of homelessness over the 12 months after treatment intake. The results from logistic and lasso models were generally consistent. Final models revealed that adolescents who were older, male, reported more victimization experiences, mental health problems, family problems, deviant peer relationships, and substance use problems (more treatment episodes and illicit drug dependence) were more likely to report experiencing homelessness. Hispanic/Latino adolescents were less likely to experience homelessness, compared with white adolescents.

**Conclusions:** The results point to the important risk and protective factors that can be assessed at treatment entry to identify adolescents at greater risk of experiencing their first episode of homelessness.

© 2020 Society for Adolescent Health and Medicine. All rights reserved.

### IMPLICATIONS AND CONTRIBUTION

Adolescents with specific demographic characteristics and adverse life experiences may be at high risk of experiencing homelessness while receiving substance use treatment. The results of this study may help treatment providers identify adolescents who may benefit from homelessness prevention efforts while in treatment.

Nationwide, nearly 4,000 unaccompanied adolescents aged <18 years experienced homelessness on a single night in 2018 [1]. However, point-in-time counts are known to underestimate

the magnitude and number of homeless youth in the U.S. [2,3]. A recent nationally representative survey suggests that the past-year household prevalence of homelessness among adolescents aged 13–17 years may be as high as 4.3% [3]. These recent estimates highlight the need to understand the correlates of youth homelessness among general and at-risk populations to inform appropriate targeting of prevention and intervention services.

**Conflicts of interest:** The authors have no conflict of interest to declare.

\* Address correspondence to: Graham T. DiGuseppi, Sc.M., Dworak-Peck School of Social Work, University of Southern California, 669 W. 34th St, Los Angeles, CA 90089.

E-mail address: [diguisep@usc.edu](mailto:diguisep@usc.edu) (G.T. DiGuseppi).

Although research on youth homelessness is growing, the prevalence and correlates of homelessness among adolescents receiving substance use treatment have not been well established. In 2018, it was estimated that 159,000 adolescents (0.6%) received some form of substance use treatment in the previous year [4]. Adolescents in treatment for substance use disorders generally experience many of the same psychosocial problems faced by homeless youth, including childhood maltreatment and dysfunctional family environments [5,6], school problems and delinquency [7], and poor mental health [8]. Given the high prevalence of substance use among adolescents and young adults experiencing homelessness [9–12], problematic substance use (e.g., using any substance that is causing intrapersonal or interpersonal distress) itself may be a risk factor for homelessness, either independently or in combination with other psychosocial problems. Identifying salient risk factors for homelessness among youth leaving substance use disorder treatment may inform interventions to prevent homelessness and further escalation of substance use problems.

Predictors of youth homelessness among the population-based samples in the U.S. have been identified in several studies [3,13,14]. Most recently, using a nationally representative sample of adolescents and young adults, Morton et al. [3] found that demographic factors, such as being a young parent, African American, Hispanic, lesbian, gay, bisexual, or transgender, and not completing high school, were independently associated with higher risk of homelessness in the previous 12 months. Other studies using data from the National Longitudinal Study of Adolescent Health (Add Health) and the National Longitudinal Study on Youth-97 (NLSY97) have identified risk and protective factors associated with prospective runaway and homelessness experiences. In one study using Add Health data, lower family relationship quality, school adjustment problems, and victimization experiences during adolescence were prospectively associated with homelessness in young adulthood [13]. A second Add Health study identified childhood risk factors (poor family functioning, socioeconomic disadvantage, and separation from parents or caregivers) and current risk factors (socioeconomic difficulties, mental health, and drug addiction problems) associated with lifetime experiences of homelessness reported by young adults [14]. Using NLSY97 data, multiple runaway episodes, poor school performance, nontraditional family structure, and parental work limitations because of health problems were significant risk factors for homelessness in young adulthood [15]. In contrast, permissive parenting and Hispanic ethnicity were found to protect against later homelessness [15]. Relevant to the present study, risk factors associated with adolescents' first runaway experiences over a 2-year period have included lower socioeconomic status, female gender, neighborhood and personal victimization, delinquency, and school suspensions [16]. African American race, Hispanic ethnicity, and parental monitoring predicted fewer runaway episodes at follow-up.

The risk and protective factors identified in these studies are consistent with theoretical and empirical research conducted with smaller samples of homeless youth [17–19]. The Risk Amplification Model (RAM) highlights the importance of family conflict as a precipitant of homelessness experiences (being kicked out of home, running away, or institutionalization) [19]. According to the RAM, runaway or homelessness experiences can

amplify these early negative life events, placing adolescents at greater risk for a variety of negative outcomes and risk behaviors learned via association with deviant peers [20,21].

Grounded in an ecological developmental framework [17], the present study sought to identify predictors of the first episode of homelessness (broadly defined) in the 12 months after substance use treatment entry. Although prior work has identified relatively consistent risk and protective factors of homelessness among nationally representative samples, the present study is the first to examine these factors among adolescents in substance use disorder treatment. In selecting important factors, we grouped predictor variables assessed at treatment entry into different socioecological domains. Specifically, we sought to understand the utility of predictor variables from the individual, familial, peer, and treatment domains in predicting adolescents' first episode of homelessness. Our analytic approach used traditional statistical methods (binary logistic regression) as well as machine learning (ML) Lasso regression. ML approaches have been shown to be ideal for predicting outcomes using large data sets [22] and have only recently been applied to investigate mental health outcomes [23].

## Methods

### Data source and participants

The current sample consists of adolescents ( $n = 20,069$ , aged 12–17 years) entering treatment from 2002 to 2012 at 192 Substance Abuse and Mental Health Services Association funded substance use treatment clinics throughout the U.S. Treatment sites used the Global Appraisal of Individual Needs (GAIN), a comprehensive biopsychosocial assessment administered via computer by trained program staff at treatment intake and at 3, 6, 9, and 12 months postintake. Informed consent to provide deidentified data was obtained from individuals receiving treatment under the supervision of each site's institutional review board, and pooled data were obtained from the GAIN coordinating center, Chestnut Health Systems [24]. Participants were excluded from analysis if they reported any lifetime or past 90-day experience of homelessness at intake ( $n = 2,158$ ), resulting in an analytic sample of 17,911 adolescents (Table 1). Participants were referred to treatment from a variety of sources, most frequently juvenile justice (50.2%), personal/self-referral (18.7%), family members (10.3%), or school/job sources (8.2%). Levels of care at intake were most commonly outpatient (70.1%), intensive outpatient (9.7%), early intervention (5.0%), and post-residential continuing care (3.5%). Sites reported that adolescents received a variety of treatments, including adolescent community reinforcement approach (38.6%), motivational enhancement therapy combined with cognitive behavioral therapy (37.9%), case management, 12-step facilitation, other "treatment as usual" approaches (9.4%), and other evidence-based treatments (8.2%).

### Measures

All variables were assessed at the time of intake, and the main outcome variable was assessed at 3, 6, and 12 months post-intake (Table 1). GAIN measures and scales have demonstrated to be reliable and predictive across numerous adolescent studies. Detailed psychometric information can be found in the GAIN administration guide [25].

**Table 1**  
Sample characteristics at treatment intake (n = 17,911 adolescents)

Variable	Description	Values	n or M	% or SD	% Missing
Age	Age at intake in years	Count (12–17)	15.56	1.21	.00
Male	Male gender identity	Dichotomous (yes/no)	13,426	74.96%	.03
Race/ethnicity	Racial/ethnicity in seven groups; multiracial if identified as more than one racial group	Nominal			.03
	African-American		2,873	16.04%	
	White		6,442	35.97%	
	Hispanic		5,475	30.57%	
	Asian		138	.77%	
	Native American/Alaskan		256	1.43%	
	Native				
	Multiracial		2,584	14.43%	
	Other		137	.76%	
LGBQ	Lesbian, gay, bisexual, or questioning/curious identity	Dichotomous (yes/no)	251	1.40%	74.91
Recent school problems	From the GAIN training problem scale: expelled from school or high levels of school-related problems (past 12 months), or missing 5 + days of school, or being in trouble/suspended (past 90 days)	Dichotomous (yes/no)	10,349	57.78%	1.26
Financial Problems Scale	Number of problems related to budgeting, insufficient income to pay bills, borrowing money, arguing about money, and using emergency public services (food bank, soup kitchen) in past 12 months	Count (0–10)	.17	.74	60.19
Juvenile justice involvement	Involvement in the juvenile justice system in 13 different ways (awaiting trial, on probation/parole, etc.)	Dichotomous (yes/no)	13,107	73.18%	.21
Disability compensation	Receipt of supplemental security income, social security disability insurance, or other disability compensation	Dichotomous (yes/no)	326	1.82%	3.72
General Victimization Scale	Number of lifetime victimization experiences (physical, emotional, sexual) plus the number of traumagenic factors involved (duration and type of experiences, relationship to perpetrator) and worries about victimization experience(s) happening again.	Count (0–15)	2.70	2.92	.09
Mental health					
Depressive/mood disorder	Past 12-month mood disorder based on DSM-IV symptom counts	Dichotomous (yes/no)	5,323	29.72%	.06
Generalized anxiety disorder	Past 12-month generalized anxiety disorder based on DSM-IV symptom counts	Dichotomous (yes/no)	1,565	8.74%	.13
Traumatic stress disorder	Past 12-month high traumatic stress based on DSM-IV symptom counts	Dichotomous (yes/no)	3,584	20.01%	.16
Suicide problems	Suicidal ideation or a plan, means, or making a suicide attempt in the past 12 months	Dichotomous (yes/no)	1,635	9.13%	.16
Conduct disorder	Antisocial behaviors (bullying, physical fights, theft, etc.) at least twice in the past 12 months	Dichotomous (yes/no)	8,022	44.79%	.23
Family					
CPS/foster care involvement	Self-reported in custody of the state, foster care, or other group home or child care institution	Dichotomous (yes/no)	1,746	9.75%	.60
Family problems	Trouble at home or with family for any reason in past 90 days	Dichotomous (yes/no)	4,104	22.91%	1.21
History of substance use	Blood relatives' history of problems with alcohol or drug use	Dichotomous (yes/no)	12,325	68.81%	1.64
History of mental illness	Blood relatives' history of emotional, mental, or psychological problems	Dichotomous (yes/no)	5,971	33.34%	3.98
Parenthood	Has one or more children	Dichotomous (yes/no)	731	4.08%	.00
Social					
Social Risk Index	Sum of people currently socialize with who were employed, involved in illegal activity, substance use, physical and verbal altercations, etc. in the past 90 days. Reverse score number of people employed in substance use treatment or recovery.	Count (0–28)	13.16	4.44	4.17
General Social Support Index	Number of social support sources (counselors, family, friends, and colleagues from work or school)	Count (0–9)	6.47	2.51	
Substance use treatment					
No insurance	Not covered by any type of insurance, court, or health program	Dichotomous (yes/no)	1,574	8.79%	62.53
Prior treatment episodes	Number of prior substance abuse treatment episodes	Ordinal (none, once, twice, three, four, or five or more episodes)	.49	.95	.29
Alcohol use disorder	Self-reported lifetime alcohol dependence (DSM-IV symptoms)	Dichotomous (yes/no)	3,025	16.89%	13.56
Marijuana use disorder	Self-reported lifetime marijuana dependence (DSM-IV symptoms)	Dichotomous (yes/no)	5,868	32.76%	12.42
Other drug use disorder	Self-reported lifetime other (not alcohol or marijuana) drug dependence (DSM-IV symptoms)	Dichotomous (yes/no)	5,320	29.70%	6.98
Self-help meeting attendance	Number of days attending self-help group meetings, i.e., AA/NA/CA or Social Recovery in the past 90 days	Count (0–90)	1.65	7.46	.43
Homelessness	One or more day(s) of homelessness in the past 90 days at 3, 6, or 12-month follow-up	Dichotomous (yes/no)	887	4.95%	.46

AA = alcoholics anonymous; CA = cocaine anonymous; CPS = child protective services; GAIN = Global Assessment of Individual Needs; NA = narcotics anonymous; SD = standard deviation.

**Individual domain.** Individual-level predictors include demographics, sexual orientation (lesbian, gay, bisexual, and queer/questioning [LGBQ] status), recent school-related problems, juvenile justice involvement, financial problems, disability compensation (supplemental security income, social security disability insurance, or other compensation benefits), victimization, and mental health. Financial problems were assessed using the financial problems scale [26]. Lifetime victimization was assessed using the General Victimization Scale [27]. Past year depression, generalized anxiety disorder, and traumatic stress disorder are indicated by clinical cutoff scores based on DSM-IV symptom counts. Suicide problems were indicated if participants reported having thoughts of suicide, a plan, means, or making a suicide attempt in the past 12 months. Conduct disorder was indicated by antisocial behaviors (bullying, physical fights, theft, etc.) at least twice in the past 12 months.

**Family domain.** Predictor variables in the family domain included lifetime child protective services/foster care involvement, recent (past 90-day) family problems, family history of substance use, family history of mental illness, and parenthood (having one or more children).

**Social domain.** Social support was assessed using the General Social Support Index [28], and the social risk was assessed using the Social Risk Index [25].

**Treatment domain.** Treatment variables included insurance coverage, number of prior substance abuse treatment episodes, self-help meeting attendance, and three variables reflecting lifetime substance dependence (DSM-IV criteria) for alcohol, marijuana, and other drug(s).

**First episode of homelessness.** This binary outcome variable was derived from participant responses to the question, “During the past 90 days, on how many days have you been homeless or had to stay with someone else to avoid being homeless?” A dummy variable was created such that participants who reported at least one day of homelessness at the 3-, 6-, or 12-month follow-up interviews were coded as 1, and participants who reported 0 days of homelessness at the 3, 6, and 12-month follow-up interviews were coded as 0. Nine-month follow-up data were not used, given high amounts of missing data (primarily because of funding restrictions or inconsistent data collection practices across sites).

#### Data analysis

**Logistic regression.** Associations between predictor variables and the first episode of homelessness were first examined using bivariate logistic regression models (Supplementary Table 1). Predictor variables with significant bivariate relations ( $p < .10$ ) to the first episode of homelessness were selected for inclusion in multivariate models [29]. Bivariate correlations among the predictor variables were also run to confirm that no multicollinearity was present ( $r < .90$ ) [29]. In the multivariate logistic regression models, a sequential/hierarchical model building approach was used, such that groups of predictor variables within each domain were entered in steps (i.e., individual, familial, peers, and treatment). All predictor variables were entered simultaneously in a final model. All models included age, gender, and race/ethnicity.

**Machine learning.** In addition to traditional logistic regression models, we also used an ML approach to help determine which features (i.e., independent variables) were most important in predicting our outcome of interest. That is, rather than interpreting a potentially large number of variables that predict the first episode of homelessness, we used ML to rank order our predictors from our final model. Our ML approach used the Sk-Learn module in Python [30] to run Lasso regression models to identify variables that contribute to adolescents' first episode of homelessness. Lasso models use an embedded method for identifying variables with the largest predictive power (i.e., contribute the most to the prediction of the outcome variable) [31]. In this approach, we use Lasso regularization, where if a feature is determined to be irrelevant, the Lasso model penalizes the coefficient, setting it to zero. Features with a coefficient of zero are removed from the model; remaining features are iteratively tested to determine positive and negative contribution to model outcome.

**Missing data.** High amounts of data were missing for five predictor variables assessed at intake: LGBQ identity (74.9% missing), financial problems (60.2% missing), insurance coverage (62.5% missing), alcohol dependence (13.6% missing), and marijuana dependence (12.4% missing). This is likely because of lack of knowledge on the part of the participant, survey skip patterns, and items being added in newer versions of the GAIN. Follow-up rates were 88.7%, 78.5%, and 45.4% at the 3-, 6-, and 12-month follow-ups, respectively. However, all adolescents completed at least one follow-up, and only 83 (.5%) were missing homelessness data over all three follow-ups. Those with missing homelessness data at all follow-ups (compared with those with any) were more likely to be African American, high school graduates, unemployed, have lower social risk, have more substance use treatment episodes, and attend more self-help meetings (all  $ps < .05$ ). Our analyses handled cases with missing data using a full-information maximum likelihood estimator in Mplus version 8 (Muthén & Muthén, 1998–2017).

#### Results

Over the 12 months after treatment intake, 887 (5.0%) adolescents reported experiencing homelessness on at least one day. Results of bivariate logistic regression models can be found in the Supplementary Table 1.

#### Multivariate logistic regression

In Model 1, individual predictors associated with experiencing higher odds of homelessness were age, male gender, victimization, depression, and conduct disorder (Table 2). Hispanic ethnicity was associated with a lower odds of experiencing homelessness. In Model 2, family factors associated with experiencing higher odds of homelessness were child protective services/foster care involvement, recent family problems, family history of substance use, and family history of mental illness. In Model 3, the social risk was significantly associated with higher odds of experiencing homelessness. In Model 4, treatment factors associated with greater odds of homelessness were the number of prior treatment episodes, lifetime marijuana dependence, and lifetime illicit drug use dependence. In Model 5 (full model), age, male gender, victimization, conduct disorder, family problems, family history of substance use, social risk, number of

**Table 2**

Multivariate logistic regression models predicting first episode of homelessness (adolescents; n = 17,906)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
	Individual	Familial	Social	Treatment	Full Model
Age	<b>1.12 (1.06–1.19)</b>	<b>1.14 (1.07–1.21)</b>	<b>1.10 (1.04–1.17)</b>	1.06 (1.00–1.13)	<b>1.09 (1.03–1.16)</b>
Male	<b>1.23 (1.05–1.44)</b>	<b>1.42 (1.23–1.64)</b>	<b>1.61 (1.39–1.86)</b>	<b>1.54 (1.33–1.78)</b>	<b>1.25 (1.07–1.47)</b>
African-American	.86 (.69–1.06)	.89 (.72–1.11)	<b>.74 (.60–.91)</b>	.89 (.71–1.10)	.97 (.78–1.21)
Other	1.04 (.87–1.25)	1.13 (.94–1.35)	1.08 (.90–1.29)	1.13 (.94–1.34)	1.07 (.89–1.28)
Asian	.49 (.18–1.34)	.65 (.24–1.75)	.47 (.17–1.29)	.52 (.19–1.42)	.61 (.22–1.67)
Hispanic	<b>.59 (.49–.71)</b>	<b>.68 (.57–.82)</b>	<b>.56 (.46–.67)</b>	<b>.60 (.50–.72)</b>	<b>.64 (.53–.77)</b>
LGBQ	1.50 (.93–2.45)				1.41 (.86–2.30)
Recent school problems	1.11 (.96–1.29)				1.11 (.94–1.30)
Financial problems	1.11 (1.00–1.24)				1.10 (.98–1.23)
Victimization	<b>1.08 (1.06–1.11)</b>				<b>1.06 (1.03–1.09)</b>
Depression	<b>1.29 (1.08–1.55)</b>				1.15 (.96–1.38)
Generalized anxiety disorder	.96 (.77–1.19)				.94 (.76–1.18)
Traumatic stress disorder	1.07 (.89–1.28)				1.02 (.85–1.23)
Suicide problems	1.21 (.98–1.49)				1.21 (.98–1.50)
Conduct disorder	<b>1.41 (1.21–1.65)</b>				<b>1.18 (1.01–1.39)</b>
CPS/Foster care		<b>1.40 (1.15–1.70)</b>			1.22 (1.00–1.50)
Family problems		<b>1.54 (1.32–1.79)</b>			<b>1.23 (1.05–1.45)</b>
Family history of substance use		<b>1.67 (1.38–2.02)</b>			<b>1.35 (1.11–1.65)</b>
Family history of mental illness		<b>1.38 (1.18–1.60)</b>			1.11 (.94–1.30)
Social risk			<b>1.06 (1.05–1.08)</b>		<b>1.02 (1.01–1.04)</b>
No insurance				.91 (.66–1.25)	1.07 (.77–1.47)
Num. prior treatment episodes				<b>1.19 (1.12–1.27)</b>	<b>1.15 (1.08–1.23)</b>
AUD				1.19 (.99–1.44)	1.06 (.87–1.28)
MUD				<b>1.34 (1.13–1.59)</b>	1.19 (1.00–1.42)
ODD				<b>1.78 (1.53–2.08)</b>	<b>1.25 (1.05–1.48)</b>
Number of self-help meetings				1.00 (.99–1.00)	.99 (.99–1.00)

Odds ratios significant at the  $p < .05$  level are highlighted in bold.

AUD = alcohol use disorder; CPS = child protective services; LGBQ = lesbian, gay, bisexual, and queer/questioning; MUD = marijuana use disorder; ODD = other drug use disorder.

prior treatment episodes, and lifetime illicit drug dependence were associated with higher odds of experiencing homelessness. Hispanic ethnicity was associated with reduced odds of experiencing homelessness in the full model.

### ML regression

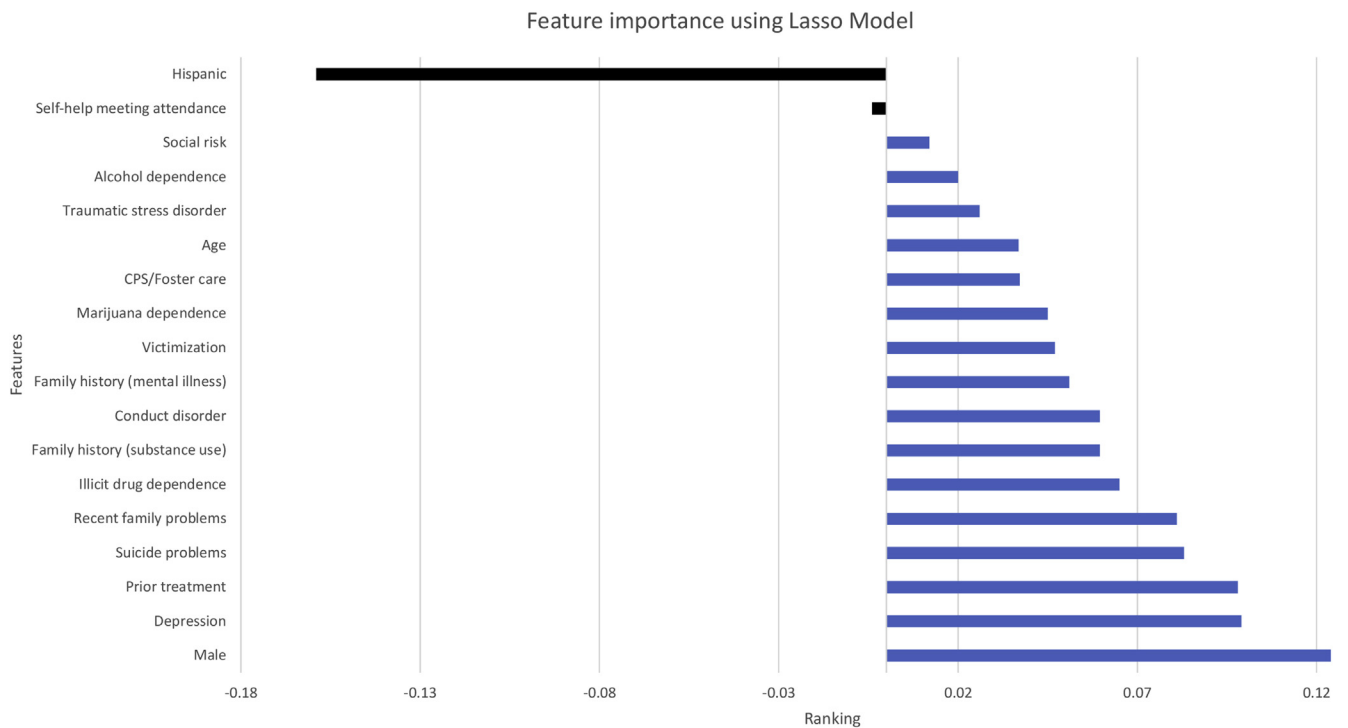
Figure 1 depicts the results of the adolescent ML model. The top five features that had the greatest contribution to the outcome were (in order of importance according to Lasso model coefficients) male gender, mood disorder (depression), number of prior treatment episodes, suicide problems, and weekly family problems. Hispanic race and number of self-help meetings negatively contributed to the outcome.

### Discussion

Adolescent homelessness is a serious problem in the U.S., affecting approximately 660,000 households each year [3]. Homelessness experiences such as being kicked out or running away from home may disproportionately affect adolescents with substance use problems, but the correlates of youth homelessness have yet to be investigated in this population. The present study set out to address this by identifying theoretical and empirical risk and protective factors from the homeless youth literature and test their association with adolescents' self-reported first episode of homelessness in the 12 months after substance use treatment entry. We used traditional regression and ML as two complementary approaches to identify risk and protective factors that contributed to this outcome.

Results revealed that the prevalence of homelessness among adolescents in substance use treatment may be substantially higher than in the general population. Just under 11% of adolescents reported experiencing homelessness before treatment entry. With these participants excluded from further analysis, 5.0% of the remaining sample reported experiencing their first episode of homelessness in the 12 months after treatment entry. This is slightly higher than the 3.0% incidence rate of first-time homelessness among adolescents in U.S. households [3], and the total number of adolescents reporting their first episode of homelessness in this study is roughly one fifth the 2019 point-in-time count nationwide [1]. The high prevalence of homelessness among this population is likely a result of the pervasive psychosocial adversities experienced by these youth, underscoring the need to understand the correlates of first episode of homelessness in this population.

Our logistic regression models revealed that at the individual level, older age and male gender were consistent risk factors across models. Hispanic adolescents, however, were less likely to report homelessness experiences. These characteristics were important in logistic and ML models. In fact, male gender and Hispanic ethnicity were the most important risk and protective factors, respectively, in the ML model. Our finding that Hispanic ethnicity is protective against adolescents' runaway or homelessness experiences is consistent with previous research with adolescent [15,16] and child welfare-involved samples [32]. This protective effect may be a result of familism and greater social capital among Latino families [33]. In contrast with previous research, LGBQ identity, black/African-American race, and recent school problems were not associated with participants' first episode of homelessness



**Figure 1.** Feature importance using the Lasso method (expressed in coefficients). Features that contribute positively and negatively to the outcome are displayed. Irrelevant features are not displayed.

[3,13,34,35]. These inconsistencies may be because of the inclusion of other important factors in our models or possibly may be a result of excluding adolescents with previous homelessness experiences from our analyses. The proportion of adolescents with recent school-related problems in our sample was relatively high, potentially limiting our ability to detect a significant relationship with the first episode of homelessness. This may also explain the nonsignificant relationship between juvenile justice involvement and the first episode of homelessness in our study. Logistic models showed that victimization, depression, and conduct disorder were significant individual risk factors as well, although depression did not remain significant in the full logistic regression model. In contrast, depression was the second most important feature contributing to homelessness in the ML model. Together, this suggests that victimization experiences (e.g., physical, emotional, or sexual abuse experienced in one's lifetime), depression symptoms, and conduct disorder (e.g., delinquent behaviors committed in the past year) are especially salient in adolescents' first homelessness episode after treatment intake.

Consistent with the RAM [19], all the variables in the familial logistic regression model were significant risk factors for the first episode of homelessness, but only family problems (e.g., "getting into trouble at home") and family history of substance abuse remained significant in the full logistic regression model. All family-related variables demonstrated importance in the ML model as well; among these, family problems emerged as the most important feature. This comports well with previous research highlighting the high rates of homelessness among youth in the foster care system [14] and youth with general family problems or poor family relationship quality [13,14]. The

present study adds to the literature that dysfunctional family environments contribute to experiences of youth homelessness within the context of substance use disorder treatment. We should note that contrary to recent research [3], parenthood was not associated with adolescents' homelessness experiences at the bivariate level, possibly because of the low number of parents in this sample.

Current results also suggest that adolescents' social risk (i.e., affiliations with deviant peers) may independently predict an increased risk of homelessness. Although social risk was an important feature in the ML model, it had a weak contribution to the outcome. This makes sense in light of the RAM, which tends to accentuate the negative influence of deviant peers on adolescent outcomes *after* adolescents become homeless and accumulate socialization experiences on the street [19]. It is easy to imagine though, how such negative relationships may coincide with strained family relationships, precipitating adolescents' *first* episode of homelessness.

Perhaps one of the most unique contributions of this article is the inclusion of treatment-related variables relevant to this adolescent population. Namely, the number of prior treatment episodes and lifetime illicit drug dependence (other than marijuana) were significant risk factors for homelessness in the final logistic regression model, and these characteristics demonstrated the importance for the ML regression model as well. This is consistent with previous research indicating problematic drug use (but not alcohol or gambling problems) as a significant risk factor for homelessness among young adults [14], and another population-based study finding that drug use was a stronger predictor of the first episode of homelessness among adults than alcohol use [36]. More extensive treatment histories and abuse or

dependence of illicit drugs (i.e., methamphetamines, opiates, etc.) may indicate more severe substance use issues [37], leading to significant problems in other life domains (i.e., housing stability). The ML model also revealed that more self-help group attendance may be protective against experiencing first episode of homeless, although this variable showed a weak contribution to the outcome. These treatment/substance use characteristics may be particularly important for identifying at-risk adolescents during treatment intake.

In addition to using typical logistic regression models, this study showed that ML regression may complement our understanding of how these risk and protective factors contribute to adolescents' homelessness experiences. Namely, ML regression offered an alternative method for ranking the predictive power of adolescent characteristics at treatment intake. Future work can explore how incorporating other ML techniques (e.g., decision trees) may contribute to our understanding of problems such as youth homelessness and help to identify those in greatest need of targeted services. Often, however, large data sets are needed to do ML, and these data-driven approaches will likely need to be complemented by relevant behavioral and social science theories [38].

There are several limitations to this study. First, our data were derived via participant self-report. Although we did our best to exclude any individuals who reported previous homelessness experiences at treatment intake, we cannot be certain that participants' experience of homelessness in the 12 months after treatment intake was their first episode of homelessness. Second, our dichotomous outcome does not offer detailed information about participants' homelessness experiences. For example, information regarding the timing, duration, reasons behind participants' homelessness (e.g., a brief runaway episode vs. more entrenched homelessness), or whether adolescents were accompanied or unaccompanied by parents or guardians was not examined. Future research may examine how risk and protective factors are associated with these aspects of adolescents' homelessness experiences during and after substance use treatment. As with any study relying on secondary data analysis, our selection of predictor variables was limited to what was available. Of note is the lack of a valid measure of household income and high amounts of missingness for some variables. Furthermore, it was beyond the scope of this article to determine the temporal or reciprocal relationships between risk and protective factors. Future research will be needed to determine how risk and protective factors may interact in complex ways across socioecological domains. Finally, we did not consider any additional ML algorithms to compare the predictive performance of the Lasso model. Future work should assess how highly features correlate between ML algorithms.

Results may inform homelessness prevention efforts within the context of substance use treatment settings. One strategy may be to focus efforts toward identifying individuals at treatment intake who may be at greater risk of experiencing their first episode of homelessness. Treatment providers may wish to explore providing evidence-based treatments such as family-based therapy, which has consistently been shown to reduce mental health and substance use problems among youth experiencing homelessness [39,40]. However, more rigorous clinical trials are needed to identify effective interventions for this vulnerable population and their impact on housing stability. Adolescents at greatest risk may be identified

by their demographic characteristics and a constellation of risk factors: those with poor mental health, family problems, victimization, problematic drug use, and prior treatment histories at treatment entry are at higher risk for subsequent homelessness.

### Acknowledgments

The authors wish to thank the research participants and the Global Appraisal of Individual Needs Coordinating Center for making this research possible.

### Funding sources

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

### Supplementary Data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.jadohealth.2019.11.312>.

### References

- [1] Henry M, Watt R, Mahathey A, et al. The 2019 Annual Homeless Assessment Report (AHAR) to Congress. U.S. Department of Housing and Urban Development; 2020.
- [2] Auerswald C, Lin J, Petry L, et al. Hidden in plain sight: An assessment of youth inclusion in point-in-time counts of California's unsheltered homeless population. The Homeless Hub; 2013.
- [3] Morton MH, Dworsky A, Matjasko JL, et al. Prevalence and correlates of youth homelessness in the United States. *J Adolesc Health* 2018;62:14–21.
- [4] Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the United States: Results from the 2018 National Survey on Drug Use and Health (HHS Publication No. PEP19-5068, NSDUH Series H-54). Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2019. Available at: <https://www.samhsa.gov/data/>. Accessed January 28, 2020.
- [5] Dube SR, Miller JW, Brown DW, et al. Adverse childhood experiences and the association with ever using alcohol and initiating alcohol use during adolescence. *J Adolesc Health* 2006;38:444.e1–444.e10.
- [6] Dube SR, Felitti VJ, Dong M, et al. Childhood abuse, neglect, and household dysfunction and the risk of illicit drug use: The adverse childhood experiences study. *Pediatrics* 2003;111:564–72.
- [7] Wang M-T, Fredricks JA. The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child Dev* 2014;85:722–37.
- [8] Storr CL, Pacek LR, Martins SS. Substance use disorders and adolescent psychopathology. *Public Health Rev* 2012;34:10.
- [9] Fazel S, Khosla V, Doll H, et al. The prevalence of mental disorders among the homeless in Western countries: Systematic review and meta-regression analysis. *Plos Med* 2008;5:e225.
- [10] Greene JM, Ennett ST, Ringwalt CL. Substance use among runaway and homeless youth in three national samples. *Am J Public Health* 1997;87:229–35.
- [11] Hodgson KJ, Shelton KH, van den Bree MBM, et al. Psychopathology in young people experiencing homelessness: A systematic Review. *Am J Public Health* 2013;103:e24–37.
- [12] Johnson KD, Whitbeck LB, Hoyt DR. Substance abuse disorders among homeless and runaway adolescents. *J Drug Issues* 2005;35:799–816.
- [13] van den Bree MBM, Shelton K, Bonner A, et al. A longitudinal population-based study of factors in adolescence predicting homelessness in young adulthood. *J Adolesc Health* 2009;45:571–8.
- [14] Shelton K, Taylor P, Bonner A, et al. Risk factors for homelessness: Evidence from a population-based study. *Psychiatr Serv* 2009;60:465–72.
- [15] Sznajder-Murray B, Jang B, Slesnick N, et al. Longitudinal predictors of homelessness: Findings from the National Longitudinal Survey of Youth-97. *J Youth Stud* 2015;18:1015–34.
- [16] Tyler KA, Bersani BE. A longitudinal study of early adolescent precursors to running away. *J Early Adolesc* 2008;28:230–51.
- [17] Haber MG, Toro PA. Homelessness among families, children, and adolescents: An ecological-developmental perspective. *Clin Child Fam Psychol Rev* 2004;7:123–64.

- [18] Rosenthal D, Mallett S, Myers P. Why do homeless young people leave home? *Aust N Z J Public Health* 2006;30:281–5.
- [19] Whitbeck LB, Hoyt DR, Yoder KA. A risk-amplification model of victimization and depressive symptoms among runaway and homeless adolescents. *Am J Community Psychol* 1999;27:273–96.
- [20] Tyler KA, Hoyt DR, Whitbeck LB. The effects of early sexual abuse on later sexual victimization among female homeless and runaway adolescents. *J Interpers Violence* 2000;15:235–50.
- [21] Bender K, Brown SM, Thompson SJ, et al. Multiple victimizations before and after leaving home associated with PTSD, depression, and substance use disorder among homeless youth. *Child Maltreat* 2015;20:115–24.
- [22] Crown WH. Potential application of machine learning in health outcomes research and some statistical cautions. *Value Health* 2015;18:137–40.
- [23] Leightley D, Williamson V, Darby J, et al. Identifying probable post-traumatic stress disorder: Applying supervised machine learning to data from a UK military cohort. *J Ment Health* 2019;28:34–41.
- [24] Ives M, Funk R, Belur V, et al. Accessing pooled GAIN and EBTx data. Available at: <https://chestnut.app.box.com/v/PooledData>. Accessed April 21, 2019.
- [25] Dennis ML, White M, Titus JC, et al. Global appraisal of individual needs: Administration guide for the GAIN and related measures. Normal, IL: Chestnut Health Systems; 2008.
- [26] Conrad KJ, Matters MD, Luchins DJ, et al. Development of a Money Mismanagement Measure and cross-validation due to suspected range restriction. *J Appl Meas* 2006;7:206–24.
- [27] Titus JC, Dennis ML, White WL, et al. Gender differences in victimization severity and outcomes among adolescents treated for substance abuse. *Child Maltreat* 2003;8:19–35.
- [28] Garner BR, Godley SH, Funk RR. Predictors of early therapeutic alliance among adolescents in substance abuse treatment. *J Psychoactive Drugs* 2008;40:55–65.
- [29] Hosmer DW, Lemeshow S. Applied logistic regression. 2nd ed. New York, NY: John Wiley & Sons, Inc.; 2000.
- [30] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. *J Mach Learn Res* 2011;12:2825–30.
- [31] Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B* 1996;58:267–88.
- [32] Schneiderman JU, Kennedy AK, Granger TA, et al. Predictors and correlates of unstable housing experiences among a child welfare-involved sample. *J Public Child Welf* 2019;13:1–17.
- [33] Martin MJ, Conger RD, Robins RW. Family stress processes and drug and alcohol use by Mexican American adolescents. *Dev Psychol* 2019;55:170–83.
- [34] Keuroghlian AS, Shtasel D, Bassuk EL. Out on the street: A public health and policy agenda for lesbian, gay, bisexual, and transgender youth who are homeless. *Am J Orthopsychiatry* 2014;84:66–72.
- [35] Van Leeuwen JM, Hopfer C, Hooks S, et al. A snapshot of substance abuse among homeless and runaway youth in Denver, Colorado. *J Community Health* 2004;29:217–29.
- [36] Thompson RG, Wall MM, Greenstein E, et al. Substance-use disorders and poverty as prospective predictors of first-time homelessness in the United States. *Am J Public Health* 2013;103:S282–8.
- [37] Anglin MD, Hser Y-I, Grella CE. Drug addiction and treatment careers among clients in the drug abuse treatment outcome study (DATOS). *Psychol Addict Behav* 1997;11:308–23.
- [38] Yarkoni T, Westfall J. Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspect Psychol Sci* 2017;12:1100–22.
- [39] Xiang X. A review of interventions for substance use among homeless youth. *Res Soc Work Pract* 2013;23:34–45.
- [40] Coren E, Hossain R, Pardo Pardo J, Bakker B. Interventions for promoting reintegration and reducing harmful behaviour and lifestyles in street-connected children and young people. *Cochrane Database Syst Rev* 2016;CD009823.