AIDS and Behavior (2022) 26:3279–3288
https://doi.org/10.1007/s10461-022-03672-y

ORIGINAL PAPER

Multicenter Development and Validation of a Model for Predicting Retention in Care Among People with HIV

Jessica P. Ridgway[1] · Aswathy Ajith[2] · Eleanor E. Friedman[1] · Michael J. Mugavero[3] · Mari M. Kitahata[4] · Heidi M. Crane[4] · Richard D. Moore[5] · Allison Webel[6] · Edward R. Cachay[7] · Katerina A. Christopoulos[8] · Kenneth H. Mayer[9] · Sonia Napravnik[10] · Anoop Mayampurath[11]

Accepted: 23 March 2022 / Published online: 8 April 2022

Abstract
Predictive analytics can be used to identify people with HIV currently retained in care who are at risk for future disengagement from care, allowing for prioritization of retention interventions. We utilized machine learning methods to develop predictive models of retention in care, defined as no more than a 12 month gap between HIV care appointments in the Center for AIDS Research Network of Integrated Clinical Systems (CNICS) cohort. Data were split longitudinally into derivation and validation cohorts. We created logistic regression (LR), random forest (RF), and gradient boosted machine (XGB) models within a discrete-time survival analysis framework and compared their performance to a baseline model that included only demographics, viral suppression, and retention history. 21,267 Patients with 507,687 visits from 2007 to 2018 were included. The LR model outperformed the baseline model (AUC 0.68 [0.67–0.70] vs. 0.60 [0.59–0.62], $P < 0.001$). RF and XGB models had similar performance to the LR model. Top features in the LR model included retention history, age, and viral suppression.

Keywords Retention in care · Predictive analytics · Machine learning

Introduction

In the US, only half of people with HIV (PWH) are engaged in regular medical care [1]. PWH not retained in care have worse health outcomes and are more likely to transmit HIV than PWH engaged in regular care [2–4]. Indeed, PWH who are diagnosed with HIV but not engaged in medical care account for the largest proportion of HIV transmission events in the US [5]. Thus, improving retention in care is a key focus of the national plan for eliminating HIV in the US [6].

Most public health efforts to improve retention in care among PWH have focused on re-engaging PWH who have lapsed in care, albeit with limited success. For example, Data to Care programs facilitate data sharing between public health departments and HIV providers to identify PWH not retained in care to inform reengagement efforts [7]. Further, while these efforts to reengage PWH who have already

✉ Jessica P. Ridgway
Jessica.ridgway@uchospitals.edu

1 Department of Medicine, University of Chicago, 5841 S Maryland Ave, MC 5065, Chicago, IL 60637, USA

2 Center for Research Informatics, University of Chicago, Chicago, IL, USA

3 Department of Medicine, University of Alabama, Birmingham, AL, USA

4 Department of Medicine, University of Washington, Seattle, WA, USA

5 Department of Medicine, Johns Hopkins University, Baltimore, MD, USA

6 Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, OH, USA

7 Department of Medicine, University of California San Diego, La Jolla, CA, USA

8 Department of Medicine, University of California San Francisco, San Francisco, CA, USA

9 Fenway Health, Boston, MA, USA

10 Department of Medicine, University of North Carolina, Chapel Hill, NC, USA

11 Department of Pediatrics, University of Chicago, Chicago, IL, USA

lapsed in care are important, they are time and resource-intensive [8]. Therefore, preventing patients from disengaging from care in the first place may be a more effective strategy for improving retention in care among PWH.

Predictive analytics can be used to identify PWH currently retained in care who are at risk for future disengagement from care. A risk prediction model capable of quantifying a specific patient's risk for future disengagement from care based on their unique characteristics could be utilized to prioritize retention resources for patients who would most benefit from them. Recently, Ramachandran et al. developed a machine learning predictive model of retention in HIV care using electronic medical record (EMR) data [9]. However, their model was limited by the use of data from a single HIV care clinic representing a relatively homogeneous urban population, potentially limiting generalizability to other settings. Therefore, the current study aimed to develop and validate a predictive model of retention in care using large-scale clinical data from eight diverse clinical HIV care sites from across the US.

Methods

Data Sources and Study Population

The Centers for AIDS Research (CFAR) Network of Integrated Clinical Systems (CNICS) is a prospective observational cohort study of adult PWH who receive primary HIV care at one of eight CFAR affiliated medical centers (Case Western Reserve University, University of Alabama Birmingham, University of California San Francisco, University of North Carolina, University of Washington, University of California San Diego, Fenway Community Health Center of Harvard University, and Johns Hopkins University) [10, 11]. Methods of data collection in CNICS have been previously reported [10]. Briefly, comprehensive clinical data collected through EMRs and institutional data systems undergo rigorous data quality assessment and are harmonized in a central data repository that is updated quarterly.

CNICS participants with at least one HIV care visit between January 1, 2007 and June 26, 2018 were included in the study cohort. If a participant died during the study period, they were censored on their date of death, and their data for the year prior was not included in the model. We examined demographic characteristics, medical appointment attendance information, clinic site, diagnoses, laboratory results, and medications. All data were deidentified by CNICS prior to transfer to researchers. CNICS research has been approved by the Institutional Review Boards at each site and this study was approved by the University of Chicago Biological and Sciences Division Institutional Review Board.

Outcomes

We utilized two measures of retention in care as the outcomes of interest. Our primary outcome measure was a 12-month gap, in which a patient was considered retained in care if no more than 365 days elapsed between HIV care encounters [12]. An HIV care encounter was defined as a clinic appointment attended with an HIV provider. We also considered a secondary measure of retention, the National HIV/AIDS Strategy (NHAS) outcome, in which a patient was considered retained in care if the patient attended at least 2 HIV care visits greater than 90 days apart within a 365-day period [13]. Characteristics of patients with and without the outcome were compared using descriptive statistics (t-tests for continuous variables, and chi-squared tests for categorical variables).

Predictor Variables

Our models were trained on 177 variables including patient demographic characteristics, past history of retention (either 12-month gap or NHAS depending on the modelled outcome), HIV transmission risk factor, insurance information, laboratory results (HIV viral load, CD4 count, sexually transmitted infection test results), diagnoses (cardiovascular disease, diabetes, hypertension, mental health disorders, pulmonary disease, renal disorders, AIDS defining illness), substance use based on ICD9/10 codes (methamphetamines/crystal, alcohol, hallucinogens, inhalants, marijuana, illicit opioids, sedative hypnotics and anxiolytics, cocaine/crack, and tobacco smoking), and antiretroviral therapy.

While CNICS performs chart review as well as other adjudication and validation procedures to verify certain diagnoses in participants' medical records [14], we included all diagnoses based on diagnosis codes (e.g., ICD9/10 codes) including those that were unverified. We included all diagnoses because our goal was to create a model that is replicable and generalizable in other settings that do not necessarily have the resources to perform chart reviews or other adjudication/validation procedures to verify diagnoses. The complete list of variables is provided in Supplementary Table 1.

Model Development

We split the data longitudinally into independent derivation (years 2007–2014) and validation (years 2015–2018) cohorts to develop the predictive models. We chose to split the data longitudinally rather than at random to more accurately reflect the way the model would be utilized in a clinical setting. We utilized a discrete-time survival analysis framework where data from the derivation cohort were discretized

into one-year intervals with the last recorded information chosen as representative for that interval. All models were then optimized to predict retention outcomes over a one-year period from each visit. We constructed a baseline logistic regression (LR) model (Baseline) that utilized past visit history (retained/not-retained in the past year according to the outcome measure of interest), viral suppression in the past year, and patient characteristics (age, birth sex, gender, race, HIV risk factor, years since first HIV care encounter) as inputs to predict retention in care. We derived several machine learning models using all 177 variables (Supplementary Table 1), including: LR, random forest (RF), and gradient boosted machine (XGB) models and compared the performance of these models with each other and with the baseline model. RF models are tree-based machine learning models that build upon a collection of decision trees, while XGB models are based on the gradient boosted decision tree algorithm that involves sequentially building decision trees that correct errors made by prior trees. Hyperparameter optimization for all models were performed using fivefold cross validation.

Model Performance

The primary measure of evaluating model performance was area under the receiver operating characteristic curve (AUC). Model AUCs were compared using DeLong's method [15]. We further calculated sensitivity and specificity at various thresholds of predicted retention in care. We also measured variable importance and created variable importance plots for our best-performing models for each outcome. Finally, we conducted a sensitivity analysis for comparing final model performance across each site. All analyses were performed using R version 3.5.2 (R Project for Statistical Computing), with two-sided $P < 0.001$ values denoting statistical significance to account for the large sample size of persons in our study population.

Results

Study Population

The study cohort consisted of a total of 21,267 patients with 507,687 visits from 2007 to 2018, of which 14,644 (69.0%) were consistently retained in care using the 12-month gap retention outcome. Those who were retained per the 12-month gap definition were more likely to be older (mean age of 40.4 years vs. 36.8 years, $P < 0.001$) and less likely to be Black (37.6% vs. 41.0%, $P < 0.001$). They were more likely to be Hispanic (13.8% vs. 12.5%, $P < 0.013$) and were also more likely to be men who had sex with men (57.9% vs. 53.8%, $P < 0.001$). Based on the NHAS outcome definition of retention, 9137 (43.0%) were retained in HIV care throughout the duration of the study period. The patients retained per this definition were more likely Black (40.3% vs. 37.4%, $P < 0.001$), slightly older (40.1 years vs. 38.7 years, $P < 0.001$) and more likely to be non-Hispanic (88.0% vs. 85.6% $P < 0.001$). Table 1a and 1b compare the characteristics between patients consistently retained and not consistently retained for each definition.

Model Performance

On a test set of 14,819 patients in the validation data set, the LR model outperformed the baseline in predicting patient retention using the 12-month gap definition (AUC 0.68 [0.67–0.70] vs. 0.60 [0.59–0.62], $P < 0.001$, see Table 2). Model performance in predicting the 12-month gap outcome of retention decreased when extending from the LR model to the RF (RF AUC 0.64 [0.63–0.66] vs. LR AUC 0.68 [0.67–0.70], $P < 0.001$). Additionally, model performance did not significantly improve when extending from the LR model to the XGB model (XGB AUC 0.69 [0.68–0.69] vs. LR AUC 0.68 [0.67–0.70]). Further, adjusting for multicollinearity among variables in the LR model by incorporating regularization did not improve the performance of the model significantly (regularized LR AUC 0.69 [0.68–0.70] vs. LR AUC 0.68 [0.67–0.70]). When considering the NHAS-outcome definition of retention in care, the LR model slightly outperformed the baseline (LR AUC 0.59 [0.59–0.60] vs. baseline AUC 0.58 [0.58–0.59], $P < 0.001$). The RF model improved performance minimally, but not significantly, over the LR model (RF AUC 0.60 [0.59–0.60] vs. LR AUC 0.59 [0.59–0.60], $P = 0.005$), while the XGB model did not perform better than the LR model (XGB AUC 0.59 [0.58–0.60] vs. LR AUC 0.59 [0.59–0.60], $P = 0.70$).

We further compared the performance of the LR models against the baseline models at various thresholds for both definitions of retention in care (Table 3). For the 12-month gap outcome, at a sensitivity of 73% the LR model had a higher specificity (51% vs. 42%) in predicting retention in care. Similarly, at a specificity of around 65%, the LR model had a higher sensitivity (63% vs. 45%) than the baseline model. The positive and negative predictive values of the LR model at these thresholds were 99% and 2%, respectively. When considering the NHAS definition of retention in care, we observe that at a sensitivity of 80%, the LR model had similar specificity (32% vs. 31%) compared to baseline. At this threshold, the LR model had positive and negative predictive values of 98% and 12%, respectively. At a specificity of 70%, the LR model also had a similar sensitivity (41% vs. 40%) to the baseline model, a positive predictive value of 99% and a negative predictive value of 2%.

Figures 1 and 2 depict the ten most important features, as measured using the absolute value of the t-statistic, for

Table 1 Characteristics of patients that were retained and not retained in care throughout the study period as per (a) the 12-month gap and (b) NHAS definition of retention in care

(a)

	Retained during study period (N = 14,644)	Not retained during study period (N = 6623)	P-value
Birth sex, n (%)			
Male	12,048 (82.3%)	5400 (81.5%)	0.146
Female	2596 (17.7%)	1222 (18.5%)	
Intersex	0 (0.0%)	1 (0.0%)	
Present sex, n (%)			
Male	11,895 (81.2%)	5310 (80.2%)	0.074
Female	2749 (18.8%)	1313 (19.8%)	
Race, n (%)			
African American/Black	5501(37.6%)	2717 (41.0%)	< 0.001
White	7666 (52.3%)	3248 (49.0%)	
Other	1477 (10.1%)	658 (9.9%)	
Ethnicity, n (%)			
Hispanic/Latino	2018 (13.8%)	829 (12.5%)	0.013
Non-Hispanic/Latino	12,626 (86.2%)	5794 (87.5%)	
Age			
Mean (SD)	40.4 (10.7)	36.8 (9.79)	< 0.001
CDC HIV transmission category, n (%)			
Heterosexual contact	3583 (24.5%)	1757 (26.5%)	< 0.001
MSM	8483 (57.9%)	3560 (53.8%)	
MSM/IDU	886 (6.1%)	504 (7.6%)	
IDU	1169 (8.0%)	583 (8.8%)	
Other	217 (1.5%)	104 (1.6%)	
Unknown	306 (2.1%)	115 (1.7%)	

(b)

	Retained during study period (N = 9137)	Not retained during study period (N = 12,130)	P-value
Birth sex, n (%)			
Male	7498 (82.1%)	9950 (82.0%)	0.685
Female	1639 (17.9%)	2179 (18.0%)	
Intersex	0 (0.0%)	1 (0.0%)	
Present sex, n (%)			
Male	7417 (81.2%)	9788 (80.7%)	0.385
Female	1720 (18.8%)	2342 (19.3%)	
Race, n (%)			
African American/Black	3685 (40.3%)	4533 (37.4%)	< 0.001
White	4580 (50.1%)	6334 (52.2%)	
Other	872 (9.5%)	1263 (10.4%)	
Ethnicity, n (%)			
Hispanic/Latino	1101 (12.0%)	1746 (14.4%)	< 0.001
Non-Hispanic/Latino	8036 (88.0%)	10,384 (85.6%)	
Age			
Mean (SD)	40.1 (10.9)	38.7 (10.3)	< 0.001
CDC HIV transmission category, n (%)			
Heterosexual contact	2401 (26.3%)	2939 (24.2%)	< 0.001
MSM	5347 (58.5%)	6686 (55.2%)	
MSM/IDU	461 (5.0%)	929 (7.7%)	

Table 1 (continued)

(b)

	Retained during study period (N = 9137)	Not retained during study period (N = 12,130)	P-value
IDU	606 (6.6%)	1146 (9.5%)	
Other	130 (1.4%)	191 (1.6%)	
Unknown	192 (2.1%)	229 (1.9%)	

CDC Centers for Disease Control and Prevention, *SD* standard deviation, *MSM* men who have sex with men, *IDU* injection drug use

Table 2 AUC for 12-month gap and NHAS definitions of retention in care, for different statistical models

Model	AUC for 12-month gap measure of retention in care	AUC for NHAS-outcome measure of retention in care
Baseline	0.60 (0.59–0.62)	0.58 (0.58–0.59)
LR	0.68 (0.67–0.70)	0.59 (0.59–0.60)
RF	0.64 (0.63–0.66)	0.60 (0.59–0.60)
XGB	0.69 (0.68–0.71)	0.59 (0.58–0.60)

AUC under the receiver operating characteristic curve, *NHAS* National HIV/AIDS Strategy, *LR* logistic regression, *RF* random forest, *XGB* gradient boosted machine

the LR model for both definitions of retention in care. These variables are grouped into four categories: demographic characteristics, laboratory results, diagnoses, and past visit history. Supplementary Table 2 (Table S2) shows the odds ratios for each of these variables in both models. As can be seen, age, retention in the prior year, viral suppression over the previous year, number of prior viral load tests, and a diagnosis of dyslipidemia were important and significantly associated with retention in the future, regardless of the retention outcome measure. Other variables that were important for prediction and increased the likelihood of being retained according to the 12-month gap outcome included the number of prior CD4 tests and the number of STI tests in the last year. A negative syphilis test in the last 6 months was associated with decreased odds of retention for the 12-month gap outcome. Insurance type was also an important predictor of retention for both models. For the 12-month gap outcome, patients with Medicaid or Ryan White had decreased odds of retention. For the NHAS outcome, patients with private insurance or Medicare had increased odds of retention, whereas those who utilized Ryan White had decreased odds of retention. Finally, the final LR model performance across each site is depicted in Supplementary Table 3. Most sites demonstrated equivalence in model performance for both outcomes, with overlapping AUCs and 95% CI.

Discussion

In this study we developed a predictive model for retention in care among patients with HIV using a multicenter dataset. Our LR model performed with better discrimination and accuracy than a standard baseline model based only on demographic characteristics, viral suppression, and past retention history. Notably, there was no significant improvement in model performance with advanced machine learning algorithms such as RF and gradient boosted machines compared to LR. While our models may not currently be accurate enough for clinical deployment, they are interpretable and with further refinement have the potential to be utilized in clinical HIV care to allow providers to prioritize retention interventions for patients most at risk for future disengagement from care.

While others have developed machine learning models to predict missed appointments in other fields such as oncology and primary care [16, 17], there have been few machine learning models developed to predict engagement in HIV care. Ramachandran et al. developed several machine learning models to predict retention in HIV care at a single urban HIV care center [9]. They found that a RF model was more accurate for predicting retention in care than LR or decision tree models. However, as the model was only developed using data from a single center, it may lack generalizability. The current study is unique in that it includes data from diverse HIV care sites across the US (both geographically and demographically) and so may be more broadly generalizable to the population of PWH across the US.

The lack of improvement in model performance in extending from LR to advanced machine learning models may be attributed to the categorical nature of most of our variables. The structured nature of the dataset could prevent machine learning methods from utilizing non-linear trends and complicated interactions to improve performance. In addition, it is worth noting that not only did the LR model have the best performance, LR models are also interpretable and easily implementable as a point-of-care tool to predict retention in care.

Our study indicates that one of the most important features for predicting retention in care is patient history of

Table 3 Sensitivity and specificity of different cut-offs for the LR model in comparison to the baseline model for predicting retention in care as defined by 12-month gap and NHAS-outcome measures

Outcome measure for retention in care	Model cutoff	Sensitivity (%, 95% CI)	Specificity (%, 95% CI)
Twelve-month gap measure	LR model (predicted probability × 100)		
	≥ 80	93 (93, 94)	19 (16, 21)
	≥ 86	84 (84, 84)	36 (34, 38)
	≥ 89	73 (73, 74)	51 (48, 53)
	≥ 91	63 (62, 63)	64 (62, 67)
	≥ 92	56 (55, 56)	71 (69, 74)
	≥ 94	38 (38, 39)	85 (83, 87)
	≥ 95	29 (28, 29)	90 (89, 92)
	Baseline model (predicted probability × 100)		
	≥ 87	90 (89, 90)	21 (19, 23)
	≥ 88.8	80 (80, 80)	33 (31, 35)
	≥ 90.3	73 (73, 73)	42 (40, 45)
	≥ 91.5	64 (64, 65)	52 (50, 55)
	≥ 92.2	51 (51, 51)	60 (59, 63)
	≥ 92.5	45 (45, 46)	65 (63, 68)
	≥ 93.1	33 (33, 34)	77 (75, 79)
NHAS-outcome measure	LR model (predicted probability × 100)		
	≥ 82	80 (80, 80)	32 (32, 33)
	≥ 84	70 (69, 70)	44 (43, 45)
	≥ 86	60 (60, 60)	53 (52, 54)
	≥ 87	56 (56, 57)	57 (56, 57)
	≥ 89	41 (41, 42)	70 (69, 71)
	≥ 91	28 (28,29)	80 (79, 81)
	Baseline model (predicted probability × 100)		
	≥ 84.2	80 (80, 81)	31 (30, 32)
	≥ 86.3	69 (68, 69)	44 (43,45)
	≥ 87.5	60 (60, 60)	53 (52, 54)
	≥ 88.5	49 (49, 50)	62 (61, 61)
	≥ 89.4	40 (39, 40)	70 (69, 71)
	≥ 90	34 (34, 34)	74 (73, 75)

Model cut-offs are represented as predicted probabilities × 100

LR logistic regression, *CI* confidence interval, *NHAS* National HIV/AIDS Strategy

retention in care in the previous year. This finding is consistent with prior studies among PWH that an individual patient's prior patterns of engagement in care are predictive of future engagement in care [9, 18]. Several different definition of retention in care exist, with some definitions related to patterns of attended HIV care visits, and others related to missed visits [12, 19]. All of these retention definitions are strongly correlated with one another and with HIV viral suppression [12]. Our models used two outcomes related to attended visits, 12-month gap and NHAS. In our study, more patients were classified as retained using the 12-month gap definition of retention than NHAS, likely because the 12-month gap definition is less restrictive, only requiring one visit within a year vs. two visits in a year as required by the NHAS definition. Of note, Pence et al. developed a LR model to predict a different retention measure, missed HIV care visits, using CNICS data [18]. Their study similarly found that missed visits in the prior year was the most important predictor of future missed visits.

In addition to prior retention history, features related to laboratory tests, demographics, and comorbidities also ranked highly in our models. HIV viral suppression in the last year was an important predictor in the models for both outcome measures, and last CD4 value was an important predictor in the NHAS model. Others have also found that viral suppression and CD4 count are associated with retention in care among PWH [20–23]. Other laboratory test-related features that ranked highly in the models were more

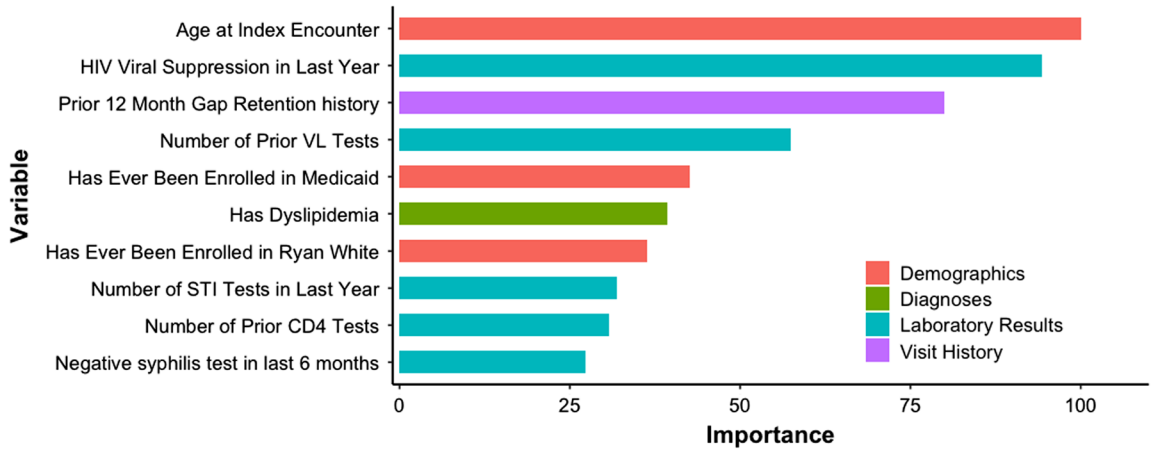

Fig. 1 Variable importance for 12-month gap outcome of retention in care. Figure depicts the ten most important features, as measured using the absolute value of the t-statistic, for the logistic regression model for the 12-month gap outcome of retention in care. The following variables were positively associated with retention in care: older age, HIV viral suppression in the last year, retention in the previous year, higher number of prior viral load tests, dyslipidemia, increased number of STI tests in the last year, higher number of prior CD4 tests. The following variables were negatively associated with retention in care: Medicaid, Ryan White insurance, and negative syphilis test in the last 6 months

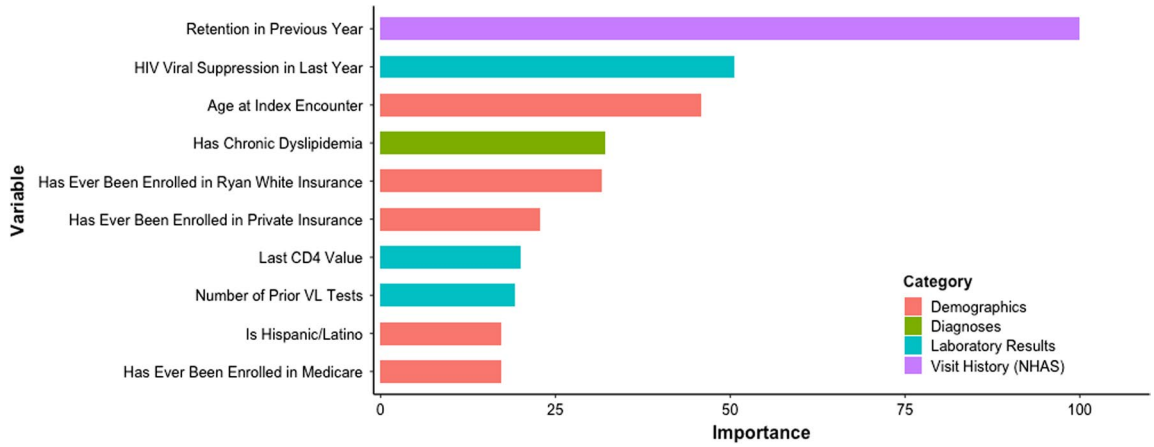

Fig. 2 Variable Importance for NHAS outcome of retention in care. Figure depicts the ten most important features, as measured using the absolute value of the t-statistic, for the logistic regression model for both the NHAS outcome of retention in care. The following variables were positively associated with retention in care: retention in the previous year, HIV viral suppression in the last year, older age, dyslipidemia, private insurance, higher number of prior viral load tests, Medicare insurance. The following variables were negatively associated with retention in care: Ryan White insurance, Hispanic/Latino ethnicity. Last CD4 value did not have a consistent directional association with retention in care

representative of patterns of healthcare utilization rather than actual test results, such as number of prior HIV viral load tests, number of prior CD4 count tests, and number STI tests in the last year. Somewhat counterintuitively, patients who tested negative for syphilis in the prior 6 months had lower odds of retention than those who did not have a syphilis test or those who tested positive. This could be because persons tested for syphilis have a higher risk for lapsing in care due to behavioral risk factors, but those that test positive are retained in both HIV and STI care in order to receive treatment for syphilis. Important demographic features in the models included age and ethnicity. This is consistent with findings from previous studies that have found older age is associated with higher retention in care among PWH [22, 24]. In the NHAS retention model, Hispanic individuals were less likely to be retained in care. Other studies have found mixed results in terms of the association between ethnicity and retention in care [25–27]. In both of our models, features related to insurance status also ranked highly. Others have also found that insurance status is associated with retained in care [28–30]. Chronic comorbidities like dyslipidemia were also important features for predicting

retention in care. Giordano et al. similarly found that PWH with chronic comorbidities were more likely to be retained in care than those without chronic comorbidities [20]. These individuals may be more likely to be retained in care because they access healthcare for their other comorbidities in addition to HIV. We examined these comorbidities both individually and as a count of total comorbid conditions. When we examined a count of comorbid conditions our results indicated no clinical difference between those retained and those not retained for both retention outcomes (12-month gap, median of 2 IQR (1–3) comorbidities for those retained and not retained; NHAS outcome, median of 2 IQR (1–3) comorbidities for those retained and not retained). Of note, when developing our models, we included diagnoses based on ICD9/10 codes which may lack sensitivity and specificity for identifying patients with medical comorbidities. As a result, we may not necessarily be identifying patients who actually have each diagnosis, e.g. dyslipidemia, but rather patients who received an ICD9/10 code for dyslipidemia. This use of ICD9/10 codes is consistent with our goal in using machine learning to develop a predictive model using all available data, rather than to accurately confirm clinical characteristics associated with an outcome.

While we found several other demographic differences between PWH who were retained throughout the study period vs. those not retained during the study period, i.e. race, sex, HIV risk factor, these factors were not among the most important features in the model. Of note, while behavioral health characteristics such as depression and substance use disorder have been strongly associated with retention in care in other studies [18, 31–33], these were not among the most important features in our models. Behavioral health disorders may still impact retention in care in the study population, but were not among the features that most strongly predicted future retention. It is important to note that we identified behavioral health disorders based on ICD9/10 codes and not based on patient reported outcomes. Indeed, one prior CNICS study that incorporated patient reported outcomes to create an Index of Engagement had an AUC of 0.69 for predicting suboptimal retention in care in the year after administration [34].

With further refinement, our predictive models of retention in care could be utilized by HIV clinics to prioritize retention interventions for patients most at risk for future disengagement from care. Currently, most HIV care clinics provide retention resources based on provider referrals or clinical intuition. Some clinics have developed systems for risk stratifying patients most at risk for disengagement from care. For example, the Data for Care Alabama Program uses risk stratification based on patients' number of prior missed clinic visits in order to prioritize patients for retention interventions [35]. Our model could potentially provide a more accurate method to identify patients at risk for lapsing in care. The scores retuned by our model indicate the likelihood of poor retention on a continuous scale. Depending on resources, a clinic can set thresholds for identifying patients at risk for poor retention at different cutoffs for predicted probabilities, with associated tradeoffs between sensitivity and positive predictive value. A clinic with limited resources could choose to use a model cut point that flags a relatively lower number of patients as at risk for non-retention and in need of retention resources. Alternatively, a clinic with more retention resources could choose to implement the model with a cut point with a higher number of patients flagged as in need of retention resources.

Our study has several limitations. In measuring the outcomes of retention in care, we were unable to account for patients who transferred care outside of CNICS, including those who were incarcerated or entered a skilled nursing facility. Therefore, participants who transferred care may have been inaccurately categorized as not retained. While our model performed better than the baseline model, the accuracy of the best performing model was fair with an AUC of 0.68. The model included data from structured EMR fields. It is possible that the inclusion of other variables that reflect social and structural disparities (e.g., education level, housing status, employment) or patient-reported outcomes could improve model performance. For example, patient-reported internalized stigma has been associated with retention in the CNICS cohort [36]. The addition of natural language processing of unstructured clinical notes [37, 38] or inclusion of other data sources, e.g., geospatial data, social media data, or other open source data [39, 40], could also potentially improve the models' accuracy. In addition, while the HIV care sites included in the study represented sites with geographic diversity across the US, the majority of sites are academic medical centers, and so the results may not be generalizable to patients seen in community-based clinics. We did not adjust for site in our models because our goal was to create broadly generalizable models that can be used at other hospitals. Although most sites demonstrated equivalence in model performance, there were a few sites that had a lower performance. Thus, our models would need to be validated at specific sites prior to deployment. We also chose not to adjust for calendar year so that our models would be generalizable in the future, but there may have been temporal trends during the study period that could confound our results. For example, patterns of follow up for HIV care evolved over the study period with stable patients being seen less frequently in more recent years, which could have impacted model performance. Finally, our study is retrospective and may be prone to confounding. The performance of the model in a prospective setting must be analyzed before implementation.

In conclusion, we developed a predictive model of retention in care among PWH using EMR data from eight HIV

care clinics across the US. We compared various models utilizing different machine learning methods, and found that a LR model had the best performance. While the model may not be accurate enough for clinical deployment in its current state, with further refinement such a model could be utilized in HIV care clinics to risk stratify patients for retention interventions. Future work should evaluate the performance of predictive models utilizing natural language processing of clinical notes or additional data elements. In addition, prospective evaluation of the clinical utility of such models in real world practice is needed.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10461-022-03672-y.

Author Contributions JR, EF, AA, and AM contributed to the study conception and design. MM, MK, HC, RM, AW, EC, KC, HM, and SN contributed to data collection. Analysis was performed by EF, AA, and AM. The first draft of the manuscript was written by JR and all authors contributed to critical revisions of the manuscript.

Funding This work was supported by the NIH-funded CFAR Network of Integrated Clinical Systems (R24 AI067039) and the CFARs at University of Alabama at Birmingham (P30 AI027767), University of Washington (P30 AI027757), University of California San Diego (P30 AI036214), University of California San Francisco (P30 AI027763), Case Western Reserve University (P30 AI036219), Johns Hopkins University (P30 AI094189, U01 DA036935), Fenway Health/Harvard University (P30 AI060354), and University of North Carolina Chapel Hill (P30 AI50410). This work was also supported by NIH Grant 1K23MH121190-01.

Data Availability Data are available upon request through CNICS.

Code Availability Code is available upon reasonable request.

Declarations

Conflict of interest Edward Cachay has received research grants paid to UC Reagents from Gilead Sciences and Merck & Co., Inc., and has been on an advisory board for Gilead Sciences. Jessica Ridgway reports serving as an expert witness on a legal case for Gilead Sciences.

Ethical Approval The study was approved by the University of Chicago Institutional Review Board.

Informed Consent Informed consent was obtained for all individuals involved in the study.

Consent for Publication N/A.

References

1. Centers for Disease Control and Prevention. Selected national HIV prevention and care outcomes. Centers for Disease Control and Prevention. https://www.cdc.gov/hiv/pdf/library/slidesets/cdc-hiv-prevention-and-care-outcomes-2018.pdf. Accessed 16 Oct 2020.
2. Mugavero MJ, Lin HY, Willig JH, et al. Missed visits and mortality among patients establishing initial outpatient HIV treatment. Clin Infect Dis. 2009;48(2):248–56.
3. Park WB, Choe PG, Kim SH, et al. One-year adherence to clinic visits after highly active antiretroviral therapy: a predictor of clinical progress in HIV patients. J Intern Med. 2007;261(3):268–75.
4. Skarbinski J, Rosenberg E, Paz-Bailey G, et al. Human immunodeficiency virus transmission at each step of the care continuum in the United States. JAMA Intern Med. 2015;175(4):588–96.
5. Centers for Disease Control and Prevention. Ending the HIV epidemic: HIV treatment is prevention. Vital Signs, March 2019. https://www.cdc.gov/vitalsigns/end-hiv/. Accessed 16 Oct 2020.
6. Department of Health and Human Services. Ending the HIV epidemic: plan for America. https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview. Accessed 16 Oct 2020.
7. Sweeney P, DiNenno EA, Flores SA, et al. HIV data to care-using public health data to improve HIV care and prevention. J Acquir Immune Defic Syndr. 2019;82(Suppl 1):S1–5.
8. Chang EJ, Fleming M, Nunez A, Dombrowski JC. Predictors of successful HIV care re-engagement among persons poorly engaged in HIV care. AIDS Behav. 2019;23(9):2490–7.
9. Ramachandran A, Kumar A, Koenig H, et al. Predictive analytics for retention in care in an urban HIV clinic. Sci Rep. 2020;10(1):6421.
10. Kitahata MM, Rodriguez B, Haubrich R, et al. Cohort profile: the Centers for AIDS Research Network of Integrated Clinical Systems. Int J Epidemiol. 2008;37(5):948–55.
11. Edwards JK, Cole SR, Westreich D, et al. Loss to clinic and five-year mortality among HIV-infected antiretroviral therapy initiators. PLoS ONE. 2014;9(7):e102305.
12. Mugavero MJ, Westfall AO, Zinski A, et al. Measuring retention in HIV care: the elusive gold standard. J Acquir Immune Defic Syndr. 2012;61(5):574–80.
13. National HIV/AIDS Strategy for the United States. Updated to 2020. https://www.hiv.gov/sites/default/files/nhas-2020-action-plan.pdf. Accessed 13 Aug 2021.
14. Crane HM, Heckbert SR, Drozd DR, et al. Lessons learned from the design and implementation of myocardial infarction adjudication tailored for HIV clinical cohorts. Am J Epidemiol. 2014;179(8):996–1005.
15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–45.
16. Percac-Lima S, Cronin PR, Ryan DP, Chabner BA, Daly EA, Kimball AB. Patient navigation based on predictive modeling decreases no-show rates in cancer care. Cancer. 2015;121(10):1662–70.
17. Huang Y, Hanauer DA. Patient no-show predictive model development using multiple data sources for an effective overbooking approach. Appl Clin Inform. 2014;5(3):836–60.
18. Pence BW, Bengtson AM, Boswell S, et al. Who will show? Predicting missed visits among patients in routine HIV primary care in the United States. AIDS Behav. 2019;23(2):418–26.
19. Batey DS, Kay ES, Westfall AO, et al. Are missed- and kept-visit measures capturing different aspects of retention in HIV primary care? AIDS Care. 2020;32(1):98–103.
20. Giordano TP, Hartman C, Gifford AL, Backus LI, Morgan RO. Predictors of retention in HIV care among a national cohort of US Veterans. HIV Clin Trials. 2009;10(5):299–305.
21. Horstmann E, Brown J, Islam F, Buck J, Agins BD. Retaining HIV-infected patients in care: where are we? Where do we go from here? Clin Infect Dis. 2010;50(5):752–61.
22. Judd RT, Friedman EE, Schmitt J, Ridgway JP. Association between patient-reported barriers and HIV clinic appointment

attendance: a prospective cohort study. AIDS Care. 2021. https://doi.org/10.1080/09540121.2021.1906401.
23. Wawrzyniak AJ, Rodriguez AE, Falcon AE, et al. Association of individual and systemic barriers to optimal medical care in people living with HIV/AIDS in Miami-Dade County. J Acquir Immune Defic Syndr. 2015;69(Suppl 1):S63-72.
24. Israelski D, Gore-Felton C, Power R, Wood MJ, Koopman C. Sociodemographic characteristics associated with medical appointment adherence among HIV-seropositive patients seeking treatment in a county outpatient facility. Prev Med. 2001;33(5):470–5.
25. Sheehan DM, Fennie KP, Mauck DE, Maddox LM, Lieb S, Trepka MJ. Retention in HIV care and viral suppression: individual- and neighborhood-level predictors of racial/ethnic differences, Florida, 2015. AIDS Patient Care STDS. 2017;31(4):167–75.
26. Fleishman JA, Yehia BR, Moore RD, Korthuis PT, Gebo KA, HIVR Network. Establishment, retention, and loss to follow-up in outpatient HIV care. J Acquir Immune Defic Syndr. 2012;60(3):249–59.
27. Myers K, Li T, Baum M, Ibanez G, Fennie K. The individual, interactive, and syndemic effect of substance use, depression, education, and ethnicity on retention in HIV care. Int J STD AIDS. 2021;32(2):184–93.
28. Palacio H, Shiboski CH, Yelin EH, Hessol NA, Greenblatt RM. Access to and utilization of primary care services among HIV-infected women. J Acquir Immune Defic Syndr. 1999;21(4):293–300.
29. Kay ES, Edmonds A, Ludema C, et al. Health insurance and AIDS Drug Assistance Program (ADAP) increases retention in care among women living with HIV in the United States. AIDS Care. 2020;33(8):1–8.
30. Kay ES, Batey DS, Mugavero MJ. The Ryan White HIV/AIDS Program: Supplementary Service Provision Post-Affordable Care Act. AIDS Patient Care STDS. 2018;32(7):265–71.
31. Cunningham CO, Buck J, Shaw FM, Spiegel LS, Heo M, Agins BD. Factors associated with returning to HIV care after a gap in care in New York State. J Acquir Immune Defic Syndr. 2014;66(4):419–27.
32. Pecoraro A, Royer-Malvestuto C, Rosenwasser B, et al. Factors contributing to dropping out from and returning to HIV treatment in an inner city primary care HIV clinic in the United States. AIDS Care. 2013;25(11):1399–406.
33. Zuniga JA, Yoo-Jeong M, Dai T, Guo Y, Waldrop-Valverde D. The role of depression in retention in care for persons living with HIV. AIDS Patient Care STDS. 2016;30(1):34–8.
34. Christopoulos KA, Neilands TB, Koester KA, et al. The HIV index: using a patient-reported outcome on engagement in HIV care to explain sub-optimal retention in care and virologic control. Clin Infect Dis. 2020. https://doi.org/10.1093/cid/ciaa1892.
35. Sohail M, Rastegar J, Long D, et al. Data for Care (D4C) Alabama: clinic-wide risk stratification with enhanced personal contacts for retention in HIV care via the Alabama Quality Management Group. J Acquir Immune Defic Syndr. 2019;82(Suppl 3):S192–8.
36. Pearson CA, Johnson MO, Neilands TB, et al. Internalized HIV stigma predicts suboptimal retention in care among people living with HIV in the United States. AIDS Patient Care STDS. 2021;35(5):188–93.
37. Oliwa T, Furner B, Schmitt J, Schneider J, Ridgway JP. Development of a predictive model for retention in HIV care using natural language processing of clinical notes. J Am Med Inform Assoc. 2021;28(1):104–12.
38. Ridgway JP, Uvin A, Schmitt J, et al. Natural language processing of clinical notes to identify mental illness and substance use among people living with HIV: retrospective cohort study. JMIR Med Inform. 2021;9(3):e23456.
39. Ridgway JP, Lee A, Devlin S, Kerman J, Mayampurath A. Machine learning and clinical informatics for improving HIV care continuum outcomes. Curr HIV/AIDS Rep. 2021;18(3):229–36.
40. Olatosi B, Zhang J, Weissman S, Hu J, Haider MR, Li X. Using big data analytics to improve HIV medical care utilisation in South Carolina: a study protocol. BMJ Open. 2019;9(7):e027688.