https://dspace.org/

A Risk Prediction Model to Identify People Living with HIV Who are Highrisk for Disengagement from Care after HIV Diagnosis in South Africa

D Ibrahima; W Hallee; M McNairy; I Hari; E et al ...

SPRINGER/PLENUM PUBLISHERS
https://hdl.handle.net/10539/43219
WOS

Downloaded from a DSpace repository. https://dspace.org/

AIDS and Behavior (2024) 28:3362–3372
https://doi.org/10.1007/s10461-024-04430-y

ORIGINAL PAPER

A Risk Prediction Model to Identify People Living with HIV Who are High-risk for Disengagement from Care after HIV Diagnosis in South Africa

Dieye Ibrahima[1] · Wong Hallee[1] · McNairy Margeret[2] · Iyer Hari[3] · Tshabalala Gugulethu[4] · Fata Amanda[5] · Bor Jacob[6] · Koenig Serena P[5,7] · Otwombe Kennedy[4,8] · Katz Ingrid T[5,7]

Accepted: 21 June 2024 / Published online: 10 July 2024
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract
The provision of ART in South Africa has transformed the HIV epidemic, resulting in an increase in life expectancy by over 10 years. Despite this, nearly 2 million people living with HIV are not on treatment. The objective of this study was to develop and externally validate a practical risk assessment tool to identify people with HIV (PWH) at highest risk for attrition from care after testing. A machine learning model incorporating clinical and psychosocial factors was developed in a primary cohort of 498 PWH. LASSO regression analysis was used to optimize variable selection. Multivariable logistic regression analysis was applied to build a model using 80% of the primary cohort as a training dataset and validated using the remaining 20% of the primary cohort and data from an independent cohort of 96 participants. The risk score was developed using the Sullivan and D'Agostino point based method. Of 498 participants with mean age 35.7 years, 192 (38%) did not initiate ART after diagnosis. Controlling for site, factors associated with non-engagement in care included being < 35 years, feeling abandoned by God, maladaptive coping strategies using alcohol or other drugs, no difficulty concentrating, and having high levels of confidence in one's ability to handle personal challenges. An effective risk score can enable clinicians and implementers to focus on tailoring care for those most in need of ongoing support. Further research should focus on potential strategies to enhance the generalizability and evaluate the implementation of the proposed risk prediction model in HIV treatment programs.

Keywords South Africa · Engagement in care · HIV · Risk score

Resumen
La provisión de TAR (Terapia Antirretroviral) en Sudáfrica ha transformado la epidemia del VIH, resultando en un aumento de la esperanza de vida de más de 10 años. Los últimos objetivos de tratamiento del VIH se sitúan en 94-75-92, con brechas notables después de las pruebas. El objetivo de este estudio fue desarrollar y validar externamente una herramienta práctica de evaluación de riesgos para identificar a las personas con VIH (PVH) con mayor riesgo de deserción

✉ Katz Ingrid T
IKatz2@bwh.harvard.edu

Otwombe Kennedy
otwombek@phru.co.za

1 Analysis Group Inc., Boston, MA, USA

2 Division of General Internal Medicine, Centre for Global Health, Weill Cornell Medicine, NY, USA

3 Section of Cancer Epidemiology and Health Outcomes, Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, USA

4 Perinatal HIV Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

5 Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA

6 Department of Global Health and Epidemiology, School of Public Health, Boston University, Boston, MA, USA

7 Harvard Medical School, Boston, MA, USA

8 School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

del cuidado después de las pruebas. Se desarrolló un modelo sencillo de aprendizaje automático que incorpora factores clínicos y psicosociales en una cohorte primaria de 498 PVH. Se utilizó el análisis de regresión LASSO para optimizar la selección de variables. Se aplicó un análisis de regresión logística multivariable para construir un modelo usando el 80% de la cohorte primaria como conjunto de datos de entrenamiento y validado usando el 20% restante de la cohorte primaria y datos de una cohorte independiente de 96 participantes. El puntaje de riesgo se desarrolló utilizando el método basado en puntos de Sullivan y D'Agostino. De los 498 participantes con una edad media de 35,7 años, 192 (38%) no iniciaron TAR después del diagnóstico. Controlando por sitio, los factores asociados con la no participación en el cuidado incluyeron tener menos de 35 años, sentirse abandonado por Dios, estrategias de afrontamiento desadaptativas usando alcohol u otras drogas, no tener dificultades para concentrarse y tener altos niveles de confianza en la capacidad de manejar desafíos personales. Un puntaje de riesgo efectivo puede permitir a los clínicos y ejecutores enfocarse en personalizar el cuidado para aquellos que más necesitan apoyo continuo. Las investigaciones futuras deben centrarse en estrategias potenciales para mejorar la generalización y evaluar la implementación del modelo de predicción de riesgo propuesto en los programas de tratamiento del VIH.

Palabras Clave Sudáfrica · Participación en el cuidado · VIH · Puntuación de riesgo

Introduction

Global HIV treatment initiatives have focused on increasing access to antiretroviral therapy (ART), with the goal of creating an "AIDS-free generation" [1,2]. In South Africa, the provision and widespread availability of ART has transformed the HIV epidemic over the past twenty years, resulting in an increase in life expectancy by over 10 years [3,4]. South Africa has provided ART for over 5.8 million people with HIV (PWH) as of 2023 [5]. While this reflects a significant uptake in treatment over the past decade, one-quarter of the nearly 7.8 million PWH in South Africa are still not on ART. This gap confirms the growing evidence that treatment availability alone is insufficient to curb the epidemic [6–8].

Previous research in South Africa has indicated that barriers to engaging and remaining in care after HIV testing have been a significant obstacle in achieving epidemic control [9–16]. Evidence from South Africa revealed ART-eligible PWH who did not engage in care after testing often reported "feeling healthy," [17] and believed that the perceived risks [18] and social costs of initiating ART outweighed its known benefits. This finding suggests optimizing ART initiation for PWH in South Africa will require identifying who is at highest risk for disengaging from care and providing targeted support to offset those risks [19,20].

A validated prediction tool that identifies PWH at highest risk for non-engagement could appropriately and efficiently tailor interventions to improve uptake of treatment [21]. Such a tool could be incorporated into a comprehensive testing strategy, allowing immediate triage of PWH who are most likely to defer the next phase of care and provide appropriate management. Previous work has identified that clinical criteria such as weight or immune status is not accurate in identifying who is high-risk for attrition after testing; [22] likely because psychosocial factors are more predictive of risk yet are not part of routine screening. Previous work in South Africa has primarily focused on the association between care engagement and clinical, sociodemographic, household, and health system factors [23–25]. While more recent models have incorporated the role of a range of socio-structural, socio-cognitive, and psychosocial factors on engagement in care, [26–29] they have not been externally validated or translated into practical tools that help identify PWH with highest risk of non-engagement in low-resource settings.

The objective of this study was to develop and externally validate a comprehensive prediction tool that used data from two cohorts of PWH in South Africa that included psychosocial variables to predict risk of treatment non-initiation after testing.

Methods

Study Setting and Population

Data used for this analysis were obtained from a longitudinal cohort of 500 PWH, enrolled between June 2014 and June of 2015, who were newly diagnosed with HIV and eligible for ART through public health facilities in South Africa. ART-eligibility in South Africa shifted during enrollment for this study. Specifically, before January 1, 2015, PWH were eligible for treatment if they had a CD4 count $\leq$350 cells/μL, an advanced clinical diagnosis including WHO Stage III/IV HIV illness, or who had a CD4 count between 350 and 500 cells/μL and were "ready and motivated to start." [30] After January 1, 2015, the CD4 count threshold changed to $\leq$500 cells/μL with the same clinical criteria. Universal ART was introduced in South Africa in September 2016.

Three hundred of these participants were enrolled at a Department of Health Clinic in Gugulethu, Cape Town, and the other 200 were enrolled in clinics within Soweto, Johannesburg. Surveys were administered at baseline, 3-, and 6-months following study enrollment. Participants were prospectively followed for nine months to track medical outcomes and rates of ART initiation. Laboratory data were extracted from the National Health Laboratory Service (NHLS), [31] which is the largest diagnostic pathology service in South Africa, supporting the national and provincial health departments in the delivery of healthcare. Further details on the study design and methods of the primary study are described elsewhere [32]. A second cohort enrolled 100 participants newly diagnosed with HIV in Soweto between January 2017 and November 2018. Participants were followed for six months to track medical outcomes and rates of ART initiation.

Study Measures

Outcome

ART initiation over a six-month period was studied as a binary outcome, determined using clinical indicators associated with care engagement (e.g., initial measurement of creatinine and/or viral load test at 6 months). Initiation at six months post-test was selected in the parent study given that there can be a significant lag between diagnosis to ART initiation, and multiple other studies measure ART initiation within 6 months [33,34]. Participants were considered to have initiated ART based on the date of a test of kidney function (creatinine clearance) reported to NHLS. Creatine clearance is routinely performed prior to initiating tenofovir, which is part of a standard, first-line treatment regimen in South Africa. This method of recording ART initiation in NHLS has been validated as an accurate way to impute treatment initiation for PWH receiving public sector healthcare in South Africa [35].

Candidate Predictors

Variables were included as potential predictors of ART non-initiation based on expert knowledge and literature review [13,17,36]. These variables include clinical factors, demographics, and psychosocial measures.

Clinical Factors and Demographics

CD4 cell count at the time of testing, obtained either through blood draw or finger prick, served as a clinical measure. Sociodemographic information included site of testing, sex, age, marital status, children, and employment status, all measured at baseline.

Psychosocial Variables

Four key psychosocial domains were examined: perceived stress, mental health, social support and coping strategies.

1) Mental health was assessed using the Patient Health Questionnaire-9 (PHQ-9) [37]. This questionnaire consists of nine items, such as the frequency with which poor appetite or overeating is experienced. Items are rated on a scale from 0 (Not at all) to 3 (Nearly every day), reflecting the frequency of symptoms.
2) Perceived stress was measured using the Perceived Stress Scale (PSS)N[38], which comprises 10 items. Questions include the level of confidence individuals have in handling personal problems, with responses ranging from 0 (never) to 4 (very often).
3) Social support was evaluated using the Multidimensional Scale of Perceived Social Support. Support from significant others, family, and friends was measured, with four items per category. Responses ranged from 1 (very strongly disagree) to 7 (very strongly agree) [39].
4) Coping strategies were analyzed using two components. The first component involved the use of the Brief Cope scale, which includes subscales for self-distraction, denial, substance use, disengagement, venting, and self-blame. Each subscale contains two items, rated from 1 (not at all) to 4 (a lot) [40]. The second component pertained to religious coping, assessed through the 14-item Brief RCOPE. This measure includes both positive (e.g., “seeking God’s love and care”) and negative aspects (e.g., “wondering whether God has abandoned me”), with seven items for each, rated from 1 (not at all) to 4 (a great deal) [41].

The final list of candidate predictors was selected from all potential predictors through discussions with clinical experts, literature review, and preliminary univariate analyses.

Statistical Analyses

Baseline Characteristics

Baseline characteristics were summarized among ART initiators and non-initiators using means and standard deviations (SD) for continuous variables and frequency counts and percentages for categorical variables. The Wilcoxon rank sum test was used to compare continuous variables between initiators and non-initiators. For categorical

variables, Pearson's chi-square test was utilized. Data were presented for both the derivation and external datasets.

All participants with non-missing data on the measures involved in the analyses were included. To create a risk-scoring tool that allows for predictions to be made using basic arithmetic on paper, we dichotomized all psychosocial variables. This approach, guided by expert input, also addresses small sample sizes in certain response categories and ensures the tool's applicability in settings with limited resources.

Training and Testing Samples

A random sample of 80% of the primary cohort ($N=398$) was used as a training dataset (training sample) to build the predictive model, and the remaining 20% (hold-out sample; $N=100$) was held out for testing. The model was further validated using the independent secondary cohort (external sample).

The holdout sample, drawn from a period when ART initiation was advised for CD4 counts below 500 cells/µL, contrasts with the external sample's broader range of CD4 counts, reflecting the updated guidelines for initiating ART regardless of CD4 count. To address the variability in CD4 counts, we rescaled these values by dividing by 100, ensuring a consistent scale across both the external and holdout samples.

Variable Selection and Multivariable Logistic Regression

Variable selection was optimized using Least Absolute Shrinkage and Selection Operator (LASSO). This method selects a subset of predictors that minimize prediction error by imposing a constraint on the model parameters that causes regression coefficients for some variables to shrink toward zero. Variables with a regression coefficient equal to zero after the shrinkage process are excluded from the final model while variables with nonzero regression coefficient are most strongly associated with the response variable. A coefficient profile plot was produced against values of the shrinkage penalty (lambda). To determine the optimal penalty, the partial likelihood deviance (binomial deviance) curve was also plotted against values of lambda, and dotted vertical lines were drawn at the optimal values using the minimum criteria and the 1 standard error of the minimum criteria (the 1-SE criteria). Using this method, the minimal subset of predictors leading to a good performance was selected based on a 10-fold cross-validation using the training sample.

Multivariable logistic regression analysis was applied to build a predictive model based on the predictors selected from the LASSO using the training data set. Model calibration was assessed using the Hosmer-Lemeshow goodness of fit (GOF) test. For external validation, model performance was tested in the hold-out and external samples using the area under the curve (AUC) C-statistic. The performance of the model was compared to the performance of a base model based on demographic and clinical variables. We also assessed statistical significance of the coefficient at the 5% and 10% level.

Simplified Risk Score Tool

To provide care centers with a quantitative tool to predict individual probability of non-initiation, a simple tool for computing risk scores was derived from the multivariable logistic model using the methods of Sullivan and D'Agostino [29]. Points for each significant predictor are summed up to create a simple table indicating how risk changes for each risk score. Predicted risks for different point totals were reported, and the correlation between the point-based risk scores and the logistic model-based risk scores was assessed. All statistical analysis was conducted using R statistical programming language version 4.2.1 [42].

Results

Baseline Characteristics

Of the 500 participants enrolled in the derivation dataset, 498 (99.6%) were included in the follow-up analysis. Of these, 192 (38.6%) did not initiate ART over the six-month follow-up period. The two participants excluded from the follow-up had incomplete data: one did not provide answers to questions regarding social support, and the other was excluded due to a missing date of birth. Baseline characteristics are summarized in Table 1. Participants who were less likely to start treatment were slightly younger (mean [SD] age: 34.2 [9.0] vs. 36.7 [9.1] years, $t=34{,}184$, $p<0.01$) and were more likely to reside in Soweto vs. Gugulethu (49% vs. 34%, $t=11.67$, $p<0.01$), compared to those who engaged in care. Other participant demographics and CD4 count were similar between the two groups.

Socio-behavioral measures revealed associations between non-engagement in care and various factors. Among those who did not initiate treatment, 38.5% expressed confidence in their ability to handle a personal problem, compared to 28.8% among those who did initiate treatment ($t=4.71$, $p<0.05$). Similarly, feeling abandoned by God was reported by 30.7% of non-initiators versus 17.3% of initiators ($t=11.41$, $p<0.001$). Use of maladaptive coping

Table 1 Baseline characteristics in the derivation cohort by treatment status

	Treatment Initiation			
	Yes, $N=306^1$	***No***, $N=192^1$	**Test statistic**[2]	**p-value**
Age	36.65 (9.11)	34.21 (8.98)	34,184	0.002
Female	166 / 306 (54%)	94 / 192 (49%)	1.12	0.3
Lives in Soweto	103 / 306 (34%)	95 / 192 (49%)	11.67	< 0.001
Married	49 / 306 (16%)	25 / 192 (13%)	0.62	0.4
Has children	247 / 306 (81%)	152 / 192 (79%)	0.09	0.7
Employed	147 / 306 (48%)	90 / 192 (47%)	0.03	0.8
Completed high school	80 / 306 (26%)	40 / 192 (21%)	1.54	0.2
CD4 count/100	2.43 (1.38)	2.40 (1.28)	29839.5	0.8
Confident about ability[a]	88 / 306 (29%)	74 / 192 (39%)	4.71	0.023
No appetite problems[b]	259 / 306 (85%)	172 / 192 (90%)	2.07	0.12
Feels abandoned by God[c]	53 / 306 (17%)	59 / 192 (31%)	11.41	< 0.001
Copes using drugs[d]	18 / 306 (5.9%)	30 / 192 (16%)	11.76	< 0.001
No concentration problems[e]	272 / 306 (89%)	182 / 192 (95%)	4.40	0.024
No problems facing barriers[f]	245 / 306 (80%)	167 / 192 (87%)	3.48	0.047
Thinks no one cares[g]	17 / 306 (5.6%)	16 / 192 (8.3%)	1.06	0.2

[1] Mean (SD); n / N (%)

[2] Wilcoxon rank sum test; Pearson's Chi-squared test

Notes: [a] "In the last month, how often have you felt confident about your ability to handle your personal problems?" – Based on Item 5 of the Perceived Stress Scale (PSS). Dichotomization: "fairly/very often" vs. other responses

[b] "Over the last two weeks, how often have you been bothered by poor appetite or over-eating" – Based on Item 5 of the Patient Health Questionnaire (PHQ-9). Dichotomization: Other responses vs. "nearly every day"

[c] "Wondered whether God had abandoned me" - From the Negative Religious Coping subscale of the Brief RCOPE. Dichotomization: "a great deal/quite a bit" vs. other responses

[d] "I've been using alcohol or other drugs to make myself feel better" - From the Substance Use subscale of the Brief Cope scale. Dichotomization: "a lot/a medium amount" vs. other responses

[e] "Over the last two weeks, how often have you been bothered by trouble concentrating on things, such as reading the newspaper or watching television" – Based on Item 7 of the Patient Health Questionnaire (PHQ-9). Dichotomization: "not at all/several days" vs. other responses

[f] "In the last month, how often have you felt difficulties were piling up so high that you could not overcome them" – Based on Item 10 of the Perceived Stress Scale (PSS). Dichotomization: "never/almost never/sometimes" vs. other responses

[g] "There is a special person in my life who cares about my feelings" - from the Significant Others subscale of the Multidimensional Scale of Perceived Social Support. Dichotomization: "disagree/disagree strongly/very strongly/don't know" vs. other responses

techniques, such as alcohol or other drugs, was observed in 15.6% of non-initiators compared to 5.9% of initiators ($t=11.76$, $p<0.001$). Additionally, 94.8% of non-initiators reported no challenges with concentration, against 88.9% of initiators ($t=4.40$, $p<0.05$). Other socio-behavioral characteristics did not significantly differ between the two groups.

Training and Testing Samples

After random sampling in a ratio of 4:1, 398 and 100 participants were included in the training and hold-out test sets, respectively. Participant characteristics were similar in these two cohorts. The external cohort included 96 participants with follow-up information. With an ART non-initiation rate of 71%, participants in this external cohort were significantly different from those in the primary cohort in characteristics such as location, education, and psychological scales (Table 2).

Variable Selection and Multivariable Logistic Regression

Six potential predictors were selected based on data from the training set (Fig. 1) and had nonzero coefficients in the LASSO regression model. Controlling for site, these predictors included age, feelings towards God, coping strategies, ability to concentrate, and confidence in one's ability to handle personal challenges. The results of the logistic regression analysis among these selected variables using the training data set are given in Table 3. Controlling for recruitment location, factors associated with non-engagement in care included being < 35 years old, feeling abandoned by God, maladaptive coping strategies using alcohol or other drugs, not having concentration problems, and having high levels of confidence in one's ability to handle person challenges. Based on p-values in the final logistic regression model, all predictors showed significant statistical differences at the 10% level, and four of them at the 5% level. The logistic model had a good fit (GOF $t=3.58$, p-value

Table 2 Characteristics by type of sample

	Treatment Initiation		
	Training sample[1], *N* = 398	***Hold-out sample***[1], *N* = 100	***External sample***[1], *N* = 89
Age	35.52 (9.14)	36.47 (9.07)	37.53 (10.48)
Female	206 / 398 (52%)	54 / 100 (54%)	24 / 89 (27%)
Lives in Soweto	150 / 398 (38%)	48 / 100 (48%)	89 / 89 (100%)
Married	52 / 398 (13%)	22 / 100 (22%)	17 / 89 (19%)
Has children	320 / 398 (80%)	79 / 100 (79%)	77 / 89 (87%)
Employed	187 / 398 (47%)	50 / 100 (50%)	43 / 89 (48%)
Completed high school	96 / 398 (24%)	24 / 100 (24%)	48 / 89 (54%)
CD4 count/100	2.40 (1.34)	2.46 (1.34)	4.02 (3.13)
Confident about ability	131 / 398 (33%)	31 / 100 (31%)	33 / 89 (37%)
No appetite problems	347 / 398 (87%)	84 / 100 (84%)	69 / 89 (78%)
Feels abandoned by God	88 / 398 (22%)	24 / 100 (24%)	31 / 89 (35%)
Copes using drugs	38 / 398 (9.5%)	10 / 100 (10%)	16 / 89 (18%)
No concentration problems	364 / 398 (91%)	90 / 100 (90%)	71 / 89 (80%)
No problems facing barriers	333 / 398 (84%)	79 / 100 (79%)	66 / 89 (74%)
Thinks no one cares	31 / 398 (7.8%)	2 / 100 (2.0%)	16 / 89 (18%)
Defaulted treatment	151 / 398 (38%)	41 / 100 (41%)	63 / 89 (71%)

[1] Mean (SD); n / N (%)

of 0.89). C-statistic values of 0.74 and 0.60 in the hold-out sample and the external test sample, respectively (Fig. 2). These statistics show a better performance of the model compared to a base model with only demographics and clinical variables (Fig. 3). These results are consistent with previous work showing that clinical measures alone may not accurately identify people at high-risk for attrition after testing and emphasize the importance of including psychosocial factors [9–11,13,17,19].

Simplified Risk Score Tool

Following the methods of Sullivan and D'Agostino, a point-based score that predicts an individual's risk of defaulting treatment during the six months follow-up after baseline was developed. Only predictors with a significant association with treatment non-initiation (Table 3) were used to derive the points. Table 4 describes the scores of participants in the validation dataset. Converting the results from the model into a point-base score, risk levels ranged from low risk −4 to very high risk 14 (3–86% risk, respectively). Based on total points, participants were categorized as low risk (0–10%), medium risk (12–25%), high risk (30–60%), and very high risk (65–86%). Among the external test cohort 5.6% were low risk, 37.1% medium risk, 50.6% high risk, and 6.7% very high risk. The correlation between the point-based risk scores and the multivariable logistic model-based risk scores in the external sample was 0.99, which suggests the simplified tool accurately reflects predictions from the final, validated model.

Discussion

In this prospective study of PWH in South Africa, we developed and validated a risk assessment tool to identify those at highest risk for not engaging in care following HIV testing. We found that including socio-behavioral factors at time of HIV testing improves prediction accuracy as compared to only demographic and clinical measures. Important socio-behavioral predictors of treatment non-initiation include maladaptive coping strategies, specifically substance use and negative religious coping, and high levels of confidence in one's ability to overcome challenges, without facing challenges to one's ability to focus or concentrate.

Predictors from our model have been individually associated with non-engagement in care among PWH. Maladaptive coping, including avoidant coping, has been widely found to have a negative impact on HIV treatment adherence. [43–45] A qualitative study examining internalized stigma, ART uptake, and viral load suppression found that participants who demonstrated avoidant coping at three months had a lower likelihood of ART uptake at six months following diagnosis. [26]

In the case of avoidant coping through substance use, prior data has shown the risk of treatment non-initiation for PWH can lead to poor decision-making regarding medication use, and ultimately leads to poor HIV clinical outcomes. [46–50] Our study also found that feeling abandoned by God, a marker of negative religious coping, was a risk factor. Negative religious coping has been associated with high psychological distress and low quality of life in other settings such as advanced cancer care [51,52].

We also found that high levels of confidence in one's ability to overcome challenges was a predictor of treatment non-initiation. To our knowledge, the association between overconfidence, misplaced confidence, or inappropriate confidence and HIV treatment initiation or adherence has not been studied. However, behavioral research has consistently found that people tend to have unrealistic beliefs about their own knowledge and abilities [53]. Overconfidence refers to when belief in one's ability exceeds reality

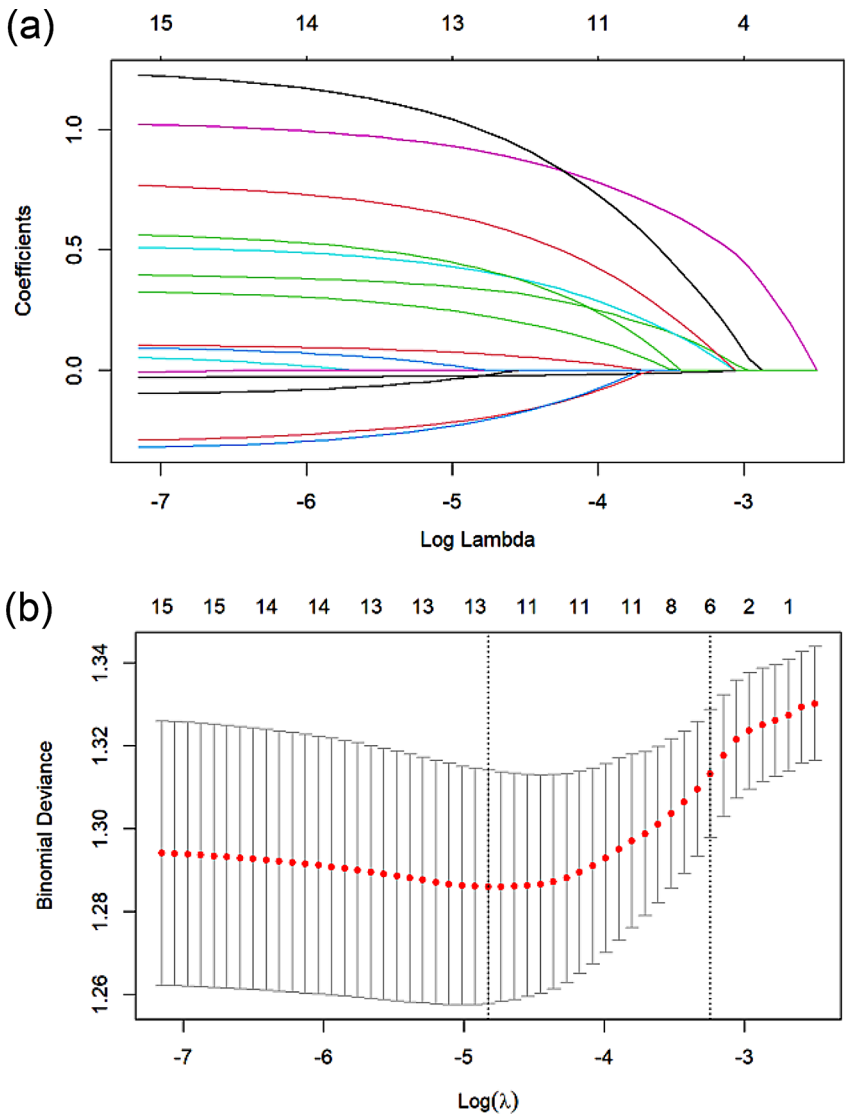

Fig. 1 Variable selection by LASSO binary logistic regression model. A coefficient profile plot was produced against the log(lambda) sequence (**a**). By verifying the optimal parameter (lambda) in the LASSO model, the partial likelihood deviance (binomial deviance) curve was plotted versus log(lambda), and dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 standard error of the minimum criteria (the 1-SE criteria) (**b**). Six variables with nonzero coefficients were selected by the lambda based on 1 standard error criteria

Table 3 Multivariable logistic model results based on features from the LASSO analysis

Intercept and variables	Coefficients	Odds ratios	Confidence intervals	*P*-value
Intercept	-1.571	0.208	(0.054, 0.730)	< 0.05
Age	-0.029	0.971	(0.948, 0.994)	< 0.05
Lives in Soweto	0.419	1.520	(0.940, 2.457)	0.087
Feels abandoned by God	0.507	1.660	(0.956, 2.888)	0.072
Copes using drugs	1.023	2.782	(1.326, 6.043)	< 0.05
No concentration problems	1.194	3.300	(1.370, 9.008)	< 0.05
No problems facing barriers	0.731	2.077	(1.111, 4.036)	< 0.05

[54]. For our sample, participants may have placed more confidence in their ability to start treatment than was realistic. More research is needed to determine the association between overconfidence or misplaced confidence and risk of HIV treatment non-initiation.

In our analysis, CD4 count was not found to be associated with ART non-initiation. Prior research has had mixed results in relation to the impact of CD4 on engagement in care. Some studies have found an association of a lower CD4 count with a higher rate of ART initiation due to a greater need for ART [26,32,33]. However, other studies have suggested that patients with high CD4 counts are more likely to start ART, as they may have fewer barriers to overcome to initiate treatment [55].

The period immediately following diagnosis is critical to ensure timely engagement and long-term retention in care and the findings of this study have potentially major implications for how clinicians identify and respond to patients

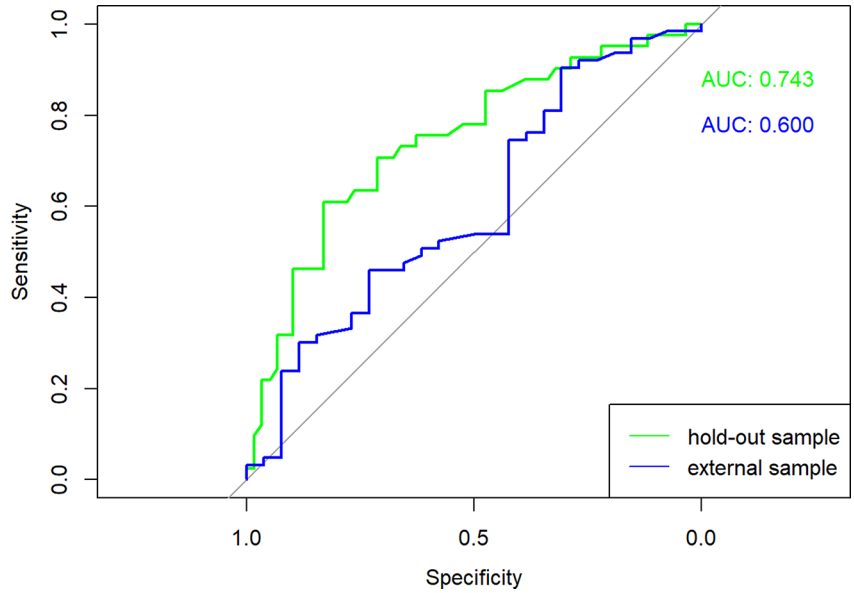

Fig. 2 Receiver operating characteristic curve (ROC) validation of the final prediction model. This figure reflects the out-of-sample performance of the logistic model including the six variables selected using LASSO

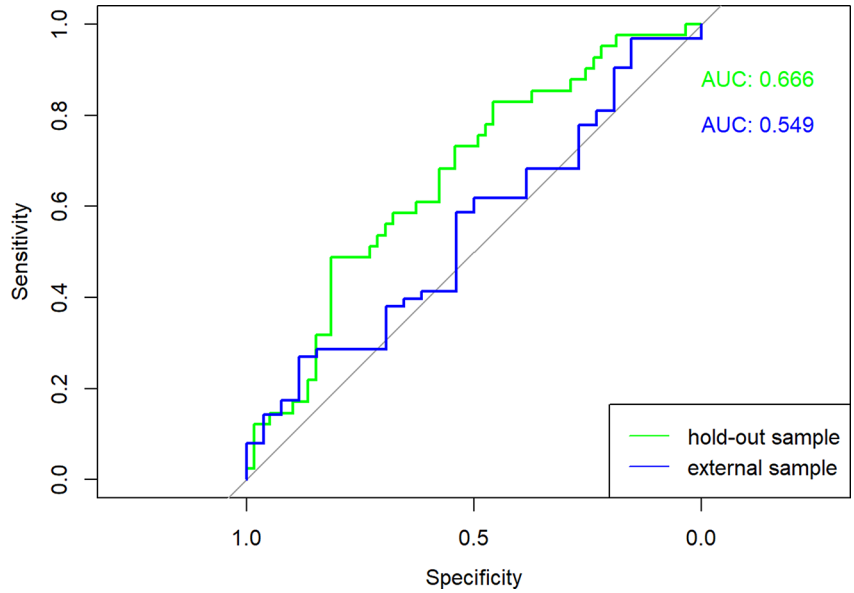

Fig. 3 Receiver operating characteristic curve (ROC) validation of the base prediction model. This figure reflects the out-of-sample performance of the base logistic model including all demographic and clinical measures: age, gender, marital status, employment status, education, living location, children, and CD4 count

at highest risk of non-engagement. Previous research has shown that up to 30% of PWH do not start treatment following testing [56,57]. There is a critical need to identify this population at the point of testing and provide the necessary support to ensure their engagement in care [58]. The proposed risk score enables clinicians to quickly assess individual PWH's risk of disengaging from care using a simple point-based system, particularly critical in the Test and Treat era. Using a risk score is advantageous because it is easy to interpret, and predictions can be calculated using simple arithmetic, without a computer or calculator. The predictions from the proposed risk score system are highly correlated with the multivariate logistic model, indicating a risk score can be used to make predictions without sacrificing accuracy.

An important limitation of the study is that it only followed participants for six months. While timely engagement in care is important in health outcomes, our study does not distinguish between individuals who entered care after six months and those who never engaged in care. Other studies have used time-to-event analyses that continuously follow participants for longer periods of time [23,25,56]. However, the longer treatment initiation is delayed, the less likely it might be to occur [12]. Therefore, if the goal is to increase early rates of treatment initiation, appropriate interventions should be enacted immediately after diagnosis. Another important limitation is that these data are from 2014 to 2015, a time when CD4 cutoffs were utilized to determine treatment eligibility in South Africa. Currently, ART is universally recommended for all PWH, with same-day or rapid

Table 4 Simplified risk score for ART default (**A**) and corresponding risk associated with each point total (**B**)

A.

Risk factor	Categories	Points
Age	18–30	0
	30–40	-1
	40–50	-2
	50–60	-3
	>= 60	-4
Copes using drugs[a]	0	0
	1	4
Feels abandoned by God[b]	0	0
	1	2
Lives in Soweto	0	0
	1	1
No concentration problems[c]	0	0
	1	4
No problems facing difficulties[d]	0	0
	1	3

B.

Point total	**Risk**	**N (%) in the external test set**
-4–0	< 10%	5 (5.6%)
1–4	12–25%	33 (37.1%)
5–9	30–60%	45 (50.6%)
10–14	65–86%	6 (6.7%)

Notes: [a] "I've been using alcohol or other drugs to make myself feel better" - From the Substance Use subscale of the Brief Cope scale. Dichotomization: "a lot/a medium amount" vs. other responses

[b] "Wondered whether God had abandoned me" - From the Negative Religious Coping subscale of the Brief RCOPE. Dichotomization: "a great deal/quite a bit" vs. other responses

[c] "Over the last two weeks, how often have you been bothered by trouble concentrating on things, such as reading the newspaper or watching television" – Based on Item 7 of the Patient Health Questionnaire (PHQ-9). Dichotomization: "not at all/several days" vs. other responses

[d] "In the last month, how often have you felt difficulties were piling up so high that you could not overcome them" – Based on Item 10 of the Perceived Stress Scale (PSS). Dichotomization: "never/almost never/sometimes" vs. other responses

ART initiation as soon as possible after HIV diagnosis [59]. Future studies with larger samples performed in the context of the current treatment options can further refine our understanding of barriers to care after testing.

Another important limitation of our study is its limited generalizability. A significant predictor was the site of the study (Soweto), suggesting that location may be a proxy for other crucial factors affecting treatment initiation. This raises questions about the applicability of our risk score tool beyond the specific populations studied. While the tool demonstrates good predictive ability within our study's context, its utility in other settings, even within South Africa, is uncertain. This limitation is significant and suggests the need for future research to explore more generalizable predictors, such as broader socioeconomic factors and aspects of healthcare provision, which vary geographically.

This study also had several strengths. By utilizing measures of psychosocial characteristics in the risk model, we were able to show the importance of these assessments to improve predictions of HIV-related health decision-making. In addition, we validated our predictive model using two sets of datasets, which showed a good overall model performance.

Our research serves as a proof of concept that a comprehensive approach, encompassing socio-behavioral factors along with demographics and clinical characteristics, is essential for predicting treatment non-initiation among PWH. This approach highlights the need for future studies, especially under universal ART recommendations, to develop more broadly applicable predictive models and inform early intervention strategies to enhance treatment initiation.

Conclusion

To realize the UNAIDS 95:95:95 goal in South Africa, there is a crucial need to identify subpopulations at highest risk for targeted strategies to support ART initiation and linkage to care after testing. Our risk score tool combines measures of demographics with socio-behavioral characteristics in simple machine learning models to provide a practical strategy to for ensuring those at highest risk for non-initiation are identified early in the care cascade, allowing for focused individualized counselling and risk assessment. Further research should focus on potential strategies for further evaluation in the current context of HIV treatment programs in South Africa and potential implementation strategies in using a risk score in clinical settings.

Funding This work was funded by NIH grant K23 MH097667.

Declarations

Ethical Approval Ethical approval for this study was provided by the Human Subjects Ethics Committees of the University of Witwatersrand, University of Cape Town, the Gauteng Department of Health and Partners Healthcare. Informed consent was obtained from all individual participants included in the study.

Competing Interests The authors declare that they have no competing interests.

References

1. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med. 2011;365(6):493–505.
2. Davies E. Hillary Clinton looks to an AIDS free generation. BMJ: Br Med J (Online) 2012; 345.
3. Bor J, Herbst AJ, Newell M-L, Bärnighausen T. Increases in adult life expectancy in rural South Africa: valuing the scale-up of HIV treatment. Science. 2013;339(6122):961–5.
4. Joint United Nations Programme on HIV/AIDS (UNAIDS). South Africa: Country Profile. 2021. http://www.unaids.org/en/regionscountries/countries/southafrica.
5. United Nations General Assembly. Despite remarkable achievement in tackling HIV/AIDS. Disease remains 'a Global Crisis, causing a death every minute', Senior Official tells General Assembly. SEVENTY-SEVENTH SESSION, 78TH MEETING (AM), GA/12509. United Nations General Assembly; 2023.
6. Heffron R, Ngure K, Mugo N, et al. Willingness of Kenyan HIV-1 serodiscordant couples to use antiretroviral-based HIV-1 prevention strategies. JAIDS J Acquir Immune Defic Syndr. 2012;61(1):116–9.
7. Kranzer K, Zeinecker J, Ginsberg P, et al. Linkage to HIV care and antiretroviral therapy in Cape Town, South Africa. PLoS ONE. 2010;5(11):e13801.
8. Nglazi MD, Lawn SD, Kaplan R, et al. Changes in programmatic outcomes during 7 years of scale-up at a community-based antiretroviral treatment service in South Africa. J Acquir Immune Defic Syndr. 2011;56(1):e1.
9. Bor J, Chiu C, Ahmed S, et al. Failure to initiate HIV treatment in patients with high CD4 counts: evidence from demographic surveillance in rural South Africa. Trop Med Int Health. 2018;23(2):206–20.
10. Curran K, Ngure K, Shell-Duncan B, et al. If I am given antiretrovirals I will think I am nearing the grave': Kenyan HIV serodiscordant couples' attitudes regarding early initiation of antiretroviral therapy. AIDS. 2014;28(2):227–33.
11. Geng EH, Bwana MB, Muyindike W, et al. Failure to initiate antiretroviral therapy, loss to follow-up and mortality among HIV-infected patients during the pre-ART period in Uganda. J Acquir Immune Defic Syndr. 2013;63(2):e64–71.
12. Hoffmann CJ, Lewis JJ, Dowdy DW, et al. Mortality associated with delays between clinic entry and ART initiation in resource-limited settings: results of a transition-state model. J Acquir Immune Defic Syndr. 2013;63(1):105–11.
13. Katz IT, Essien T, Marinda ET, et al. Antiretroviral therapy refusal among newly diagnosed HIV-infected adults. AIDS. 2011;25(17):2177–81.
14. Mujugira A, Celum C, Thomas KK, et al. Delay of antiretroviral therapy initiation is common in East African HIV-infected individuals in serodiscordant partnerships. J Acquir Immune Defic Syndr. 2014;66(4):436–42.
15. Nachega JB, Uthman OA, del Rio C, Mugavero MJ, Rees H, Mills EJ. Addressing the Achilles' heel in the HIV care continuum for the success of a test-and-treat strategy to achieve an AIDS-free generation. Clin Infect Dis. 2014;59(Suppl 1):S21–7.
16. Stinson K, Myer L. Barriers to initiating antiretroviral therapy during pregnancy: a qualitative study of women attending services in Cape Town, South Africa. Afr J AIDS Res. 2012;11(1):65–73.
17. Katz IT, Dietrich J, Tshabalala G, et al. Understanding treatment refusal among adults presenting for HIV-testing in Soweto, South Africa: a qualitative study. AIDS Behav. 2015;19(4):704–14.
18. Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. Science. 1974;185(4157):1124–31.
19. Hontelez JA, Newell M-L, Bland RM, Munnelly K, Lessells RJ, Bärnighausen T. Human resources needs for universal access to antiretroviral therapy in South Africa: a time and motion study. Hum Resour Health. 2012;10:39.
20. Van Damme W, Kober K, Kegels G. Scaling-up antiretroviral treatment in southern African countries with human resource shortage: how will health systems adapt? Soc Sci Med. 2008;66(10):2108–21.
21. McNairy ML, Abrams EJ, Rabkin M, El-Sadr WM. Clinical decision tools are needed to identify HIV-positive patients at high risk for poor outcomes after initiation of antiretroviral therapy. PLoS Med. 2017;14(4):e1002278.
22. McNairy ML, Jannat-Khah D, Pape JW, et al. Predicting death and lost to follow-up among adults initiating antiretroviral therapy in resource-limited settings: derivation and external validation of a risk score in Haiti. PLoS ONE. 2018;13(8):e0201945.
23. Boyer S, Iwuji C, Gosset A, et al. Factors associated with antiretroviral treatment initiation amongst HIV-positive individuals linked to care within a universal test and treat programme: early findings of the ANRS 12249 TasP trial in rural South Africa. AIDS Care. 2016;28(Suppl 3):39–51.
24. Plazy M, Newell ML, Orne-Gliemann J, Naidu K, Dabis F, Dray-Spira R. Barriers to antiretroviral treatment initiation in rural KwaZulu-Natal, South Africa. HIV Med. 2015;16(9):521–32.
25. Sabapathy K, Mubekapi-Musadaidzwa C, Mulubwa C, et al. Predictors of timely linkage-to-ART within universal test and treat in the HPTN 071 (PopART) trial in Zambia and South Africa: findings from a nested case-control study. J Int AIDS Soc. 2017;20(4):e25037.
26. Earnshaw VA, Bogart LM, Laurenceau J-P, et al. Internalized HIV stigma, ART initiation and HIV-1 RNA suppression in South Africa: exploring avoidant coping as a longitudinal mediator. J Int AIDS Soc. 2018;21(10):e25198–e.
27. Sy KTL, Tariq S, Ramjee G, et al. Predictors of antiretroviral therapy initiation in eThekwini (Durban), South Africa: findings from a prospective cohort study. PLoS ONE. 2021;16(2):e0246744–e.
28. Filiatreau LM, Pettifor A, Edwards JK, et al. Associations between key psychosocial stressors and viral suppression and Retention in Care among Youth with HIV in Rural South Africa. AIDS Behav. 2021;25(8):2358–68.
29. Laurenzi CA, du Toit S, Ameyan W, et al. Psychosocial interventions for improving engagement in care and health and behavioural outcomes for adolescents and young people living with HIV: a systematic review and meta-analysis. J Int AIDS Soc. 2021;24(8):e25741.
30. Meintjes G, Black J, Conradie F, et al. Adult antiretroviral therapy guidelines 2014: guideline. South Afr J HIV Med. 2014;15(4):121–43.
31. National Health Laboratory Service (NHLS). 2023. https://www.nhls.ac.za/.
32. Katz IT, Bogart LM, Dietrich JJ, et al. Understanding the role of resilience resources, antiretroviral therapy initiation, and HIV-1 RNA suppression among people living with HIV in South Africa: a prospective cohort study. AIDS. 2019;33(Suppl 1):S71–9.
33. Bor J, Ahmed S, Fox MP, et al. Effect of eliminating CD4-count thresholds on HIV treatment initiation in South Africa: an empirical modeling study. PLoS ONE. 2017;12(6):e0178249.
34. Fox MP, Larson B, Rosen S. Defining retention and attrition in pre-antiretroviral HIV care: proposals based on experience in Africa. 2012; 17(10): 1235–44.
35. Maskew M, Bor J, Hendrickson C, et al. Imputing HIV treatment start dates from routine laboratory data in South Africa: a validation study. BMC Health Serv Res. 2017;17(1):1–9.
36. Alvarez-Uria G, Pakam R, Midde M, Naik PK. Predictors of delayed antiretroviral therapy initiation, mortality, and loss to followup in HIV infected patients eligible for HIV treatment: data

from an HIV cohort study in India. *Biomed Res Int* 2013; 2013: 849042-.
37. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–13.
38. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav. 1983;24(4):385–96.
39. Zimet GD, Powell SS, Farley GK, Werkman S, Berkoff KA. Psychometric characteristics of the multidimensional scale of perceived social support. J Pers Assess. 1990;55(3–4):610–7.
40. Carver CS. You want to measure coping but your protocol'too long: consider the brief cope. Int J Behav Med. 1997;4(1):92–100.
41. Pargament KI, Koenig HG, Perez LM. The many methods of religious coping: Development and initial validation of the RCOPE. J Clin Psychol. 2000;56(4):519–43.
42. RStudio, RStudio. Integrated Development for R. Boston, MA; 2020.
43. Katz IT, Ryu AE, Onuegbu AG, et al. Impact of HIV-related stigma on treatment adherence: systematic review and meta-synthesis. J Int AIDS Soc. 2013;16:18640.
44. Lyimo RA, Stutterheim SE, Hospers HJ, de Glee T, van der Ven A, de Bruin M. Stigma, disclosure, coping, and medication adherence among people living with HIV/AIDS in Northern Tanzania. AIDS Patient Care STDs. 2014;28(2):98–105.
45. Turan B, Hatcher AM, Weiser SD, Johnson MO, Rice WS, Turan JM. Framing mechanisms linking HIV-related stigma, adherence to treatment, and health outcomes. Am J Public Health. 2017;107(6):863–9.
46. Braithwaite RS, Bryant KJ. Influence of alcohol consumption on adherence to and toxicity of antiretroviral therapy and survival. Alcohol Res Health. 2010;33(3):280–7.
47. Detsis M, Tsioutis C, Karageorgos SA, Sideroglou T, Hatzakis A, Mylonakis E. Factors Associated with HIV Testing and HIV Treatment Adherence: a systematic review. Curr Pharm Des. 2017;23(18):2568–78.
48. Gonzalez A, Barinas J, O'Cleirigh C. Substance use: impact on adherence and HIV Medical Treatment. Curr HIV/AIDS Rep. 2011;8(4):223–34.
49. Li L, Luo S, Lan CW, et al. Alcohol Use, HIV Treatment Adherence, and sexual risk among people with a history of Injecting Drug Use in Vietnam. AIDS Behav. 2017;21(Suppl 2):167–73.
50. Sileo KM, Kizito W, Wanyenze RK, et al. Substance use and its effect on antiretroviral treatment adherence among male fisherfolk living with HIV/AIDS in Uganda. PLoS ONE. 2019;14(6):e0216892.
51. Ano GG, Vasconcelles EB. Religious coping and psychological adjustment to stress: a meta-analysis. J Clin Psychol. 2005;61(4):461–80.
52. Tarakeshwar N, Vanderwerker LC, Paulk E, Pearce MJ, Kasl SV, Prigerson HG. Religious coping is associated with the quality of life of patients with advanced cancer. J Palliat Med. 2006;9(3):646–57. https://doi.org/10.1089/jpm.2006.9.646.
53. Parker AM, Stone ER. Identifying the effects of unjustified confidence versus overconfidence: lessons learned from two analytic methods. J Behav Decis Mak. 2014;27(2):134–45.
54. Mellers B, Fincher K, Drummond C, Bigony M. Chapter 1 - Surprise: a belief or an emotion? In: Pammi VSC, Srinivasan N, editors. Progress in Brain Research. Elsevier; 2013. pp. 3–19.
55. Clouse K, Pettifor A, Maskew M, et al. Initiating antiretroviral therapy when presenting with higher CD4 cell counts results in reduced loss to follow-up in a resource-limited setting. Aids. 2013;27(4):645–50.
56. Seeley J, Bond V, Yang B, et al. Understanding the time needed to Link to care and start ART in seven HPTN 071 (PopART) Study communities in Zambia and South Africa. AIDS Behav. 2019;23(4):929–46.
57. Koenig SP, Dorvil N, Dévieux JG, et al. Same-day HIV testing with initiation of antiretroviral therapy versus standard care for persons living with HIV: a randomized unblinded trial. PLoS Med. 2017;14(7):e1002357.
58. Maughan-Brown B, Harrison A, Galárraga O, et al. Factors affecting linkage to HIV care and ART initiation following referral for ART by a mobile health clinic in South Africa: evidence from a multimethod study. J Behav Med. 2019;42:883–97.
59. Republic of South Africa. Implementation of the universal test and treat strategy for HIV positive patients and differentiated care for stable patients 2016. In: Health Do, editor.; 2016.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.