Published in final edited form as:
Stat Med. 2018 January 30; 37(2): 302–319. doi:10.1002/sim.7502.

A state transition framework for patient-level modeling of engagement and retention in HIV care using longitudinal cohort data†

Hana Lee[1,*], **Joseph W Hogan**[1,2], **Becky L Genberg**[3], **Xiaotian K Wu**[1], **Beverly S Musick**[4], **Ann Mwangi**[2,5], and **Paula Braitstein**[2,5,6,7,8]

[1]Department of Biostatistics, Brown University,121 S. Main Street, Providence, RI 02912, U.S.A

[2]Academic Model Providing Access to Healthcare (AMPATH), Eldoret, Kenya

[3]Department of Health Services, Policy & Practice, Brown University, Rhode Island, U.S.A

[4]Division of Biostatistics, School of Medicine, Indiana University, Indiana, U.S.A

[5]College of Health Sciences, School of Medicine, Moi University, Eldoret, Kenya

[6]Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

[7]Fairbanks School of Public Health, Indiana University, Indiana, U.S.A

[8]Regenstrief Institute, Indiana, U.S.A

Abstract

The HIV care cascade is a conceptual model used to outline the benchmarks that reflects effectiveness of HIV care in the whole HIV care continuum. The models can be used to identify barriers contributing to poor outcomes along each benchmark in the cascade such as disengagement from care or death. Recently the HIV care cascade has been widely applied to monitor progress towards HIV prevention and care goals in an attempt to develop strategies to improve health outcomes along the care continuum. Yet there are challenges in quantifying successes and gaps in HIV care using the cascade models that are partly due to the lack of analytic approaches. The availability of large cohort data presents an opportunity to develop a coherent statistical framework for analysis of the HIV care cascade. Motivated by data from the Academic Model Providing Access to Healthcare (AMPATH), which has provided HIV care to nearly 200,000 individuals in western Kenya since 2001, we developed a state transition framework that can characterize patient-level movements through the multiple stages of the HIV care cascade. We describe how to transform large observational data into an analyzable format. We then illustrate the state transition framework via multistate modeling to quantify dynamics in retention aspects of care. The proposed modeling approach identifies the transition probabilities of moving through each stage in the care cascade. In addition this approach allows regression-based estimation to characterize effects of (time-varying) predictors of within and between state transitions such as retention, disengagement, re-entry into care, transfer-out, and mortality.

†hlee@stat.brown.edu

*Correspondence to: Hana Lee, Department of Biostatistics, Brown University, 121 S. Main Street, Providence, RI 02912, U.S.A.

Keywords

HIV care cascade; HIV care continuum; Healthcare system evaluation; Multistate models; State transition representation

1. Introduction

The HIV care cascade (or continuum) is a conceptual model describing key benchmarks that people living with HIV (PLWH) must pass through to maximize benefits of antiviral therapy (ART). In most formulations, the optimal pathway consists of (1) HIV diagnosis through testing, (2) linkage to care, (3) engagement and retention in care, (4) initiation of ART through retention, and (5) sustained suppression of viral load. The conceptual model provides a useful framework for defining and evaluating the benchmarks that measure the effectiveness of HIV care, and for developing strategies to improve HIV outcomes for PLWH [1–4]. The HIV care cascade has become a framework for monitoring progress and identifying HIV care needs in the US since the release of the National HIV/AIDS Strategy in 2010. Furthermore, the HIV care cascade is used globally as a monitoring rubric to evaluate the performance of HIV/AIDS health system management; the World Health Organization (WHO) has emphasized the cascade model as the central assessment metric for HIV care programs [5]. The UNAIDS recently announced a new global target based on steps (1), (4), and (5) in the cascade: by the year 2020, 90 percent of PLWH should be diagnosed and know their status, 90 percent of those diagnosed on antiviral therapy, and 90 percent of those on therapy have viral suppression

Quantitative analyses such as macro level summaries of proportion meeting specific benchmarks, models that examine predictors of engagement in each stage or progression through cascade stages, can provide important information needed to intervene to minimize the negative outcomes and optimize HIV care and treatment efforts to break the cycle of HIV transmission and morbidity. Despite the global acceptance and utility of the HIV care cascade as a conceptual model, our empirical understanding about patient flow through the continuum is still limited, in part because the statistical methods for analyzing cascade data do not have a unified framework. Broadly speaking, there are three main modes of summarizing data related to the care cascade. *Macro-level analyses* rely on characterizing targeted aspects of the cascade by presenting aggregated data summaries (e.g., number and/or proportion of patients) in each stage of the cascade at certain time points or across time periods [2, 6, 7]. By looking at numbers or proportions of PLWH at each stage, one can readily identify ‘leaks’ or stages where improvements are needed. *Risk-factor and regression analyses* use individual-level data to identify or evaluate the effect of patient- or program-level factors associated with reaching specific benchmarks such as linkage, retention, and ART initiation [1, 8–11]. For this type of analysis, data are sometimes aggregated across time period to define outcome, or time to event outcomes are considered. A third mode of analysis uses *simulation techniques* based on an underlying model of progression through the cascade. The mathematical model is specified in advance, and uses inputs from multiple data sources to inform values or ranges of values for the parameters. Parameter values are calibrated using extant data on outcomes of interest, and simulation from the calibrated model are used to obtain predictions of outcomes of interest under different scenarios. This approach has been used to represent complex versions of the cascade and to evaluate impact of different intervention strategies or policies [12–16].

The increasing availability of large-scale patient level cohort data is providing new opportunities for data-driven analyses of the care cascade. For example, the Academic Model Providing Access to Healthcare (AMPATH), based in Eldoret, Kenya, provides HIV and primary care to nearly 200,000 individuals in western Kenya and maintains an electronic health record (EHR) known as the AMPATH Medical Record System (AMRS) [17]. Twice each year, raw data from the AMRS are formatted into a research-grade database that can be used to investigate various clinical and epidemiologic questions related to patient care. Data from the AMRS form a significant component of the NIH-funded International Epidemiology Databases to Evaluate AIDS (IeDEA) for the East African region. The IeDEA consortium, which aggregates patient-level data in multiple regions around the world, contains patient-level data for 1.7 million PLWH globally (iedea.org). Other large-scale cohorts include the CFAR Network of Integrated Clinical Systems (CNICS) [18], EuroCoord [19], Veterans Aging Cohort Study (VACS) [20], to name just a few.

Although the approaches described above are commonly used to analyze cascade data, they do not always take full advantage of the information available in longitudinal patient-level data. Macro-level summaries provide a useful program- or community-level snapshot, but frequently aggregate data over multiple time periods or collapse patient-level longitudinal data into single summaries (such as ‘engaged in care’). This approach can overlook phenomena such as cyclic patient behavior such as coming in and out of care over time [12]. Some regression approaches use only partial longitudinal information, or collapse longitudinal patient data, which precludes examination of or adjustment for important temporal trends such as time, period or cohort effects. Consequently, many analyses reported in the literature tend to capture cross-sectional snapshots of patient behavior [21–26]. Mathematical models, which are more complex, rely on data summaries from different sources of inputs in order to generate simulated outcomes. This raises the question of whether the models represent a specific population of interest. Although mathematical models can handle complexity that cannot be modeled with sparse, micro-level cohort data, the availability of large cohorts of patients with substantial longitudinal information on clinical outcomes presents an opportunity to develop a coherent and portable statistical framework for modeling progression through the care cascade, and for characterizing the role of individual-level factors associated with meeting (or not meeting) key milestones.

However, this ‘big data’ opportunity presents specific challenges: Patient-level cohort data are observational in nature. Data are recorded at irregular time points. Some covariates may be sparsely measured over time, and are often not available from all cohort members. Defining state of care based on raw patient data may not be straightforward, especially for those who have incomplete follow-up or long gaps between clinic visits. These features can make it difficult to even capture a cross-sectional snapshot of the cohort behavior as well as to estimate a temporal trend. These challenges require us to address operationalization of the cascade, data preparation, and specification of the model that can capture the cyclic patient behavior in longitudinal data.

To address some of these challenges, we propose using multistate transition models. We illustrate our approach through analysis of data from 92,215 individuals enrolled in HIV care in AMPATH, the largest HIV care program in Kenya and one of the largest in sub-Saharan Africa. We develop a model for engagement and retention in care, which illustrates many of the key advantages of the multistate modeling approach, such as handling cyclic behavior, modeling time and period effects, and handling competing risks such as death, disengagement, and transfer out of care. The multistate model provides a natural way to incorporate both patient- and program-level covariates and provides a natural framework for extensions that accommodate more states.

The rest of the paper is organized as follows: Section 2 provides details about the AMPATH cohort and provides motivation for adopting the state transition framework. In Section 3, we provide a review of multistate models and make connections to our current application. Section 4 provides details of the model specification for the AMPATH cohort, including how the raw data on engagement in care are translated into discrete states at each time point. Our data analysis appears in Section 5, and a discussion of future directions and potential extensions of the model in Section 6.

2. Data Source and Analysis Goals

AMPATH is a partnership between Moi University College of Health Sciences, Moi Teaching and Referral Hospital in Kenya and a consortium of North American institutions, and has provided HIV care to nearly 200,000 individuals in western Kenya since 2001. Including over 60 HIV/AIDS clinics, AMPATH is one of the largest HIV/AIDS care programs in sub-Saharan Africa. Clinical visit information from individuals who are engaged in medical care within AMPATH is recorded through an electronic medical records database, the AMPATH Medical Record System (AMRS). The AMRS uses an implementation of OpenMRS (wws.openmrs.org), a web-based open source electronic medical record platform that aims to build and manage health systems in the developing countries. Information from AMRS has been used to monitor and evaluate comprehensive intervention and treatment programs in western Kenya [13, 17, 27, 28]

The overall goal of our analysis is to use patient level data to model the part of the HIV care cascade in AMPATH that relates to engagement and retention in care. We utilize a multistate transition model that characterizes individual-level membership in these discrete states as a function of time: engaged in care, disengaged from care, transferred out of care, or deceased. Individuals may pass back and forth between engaged and disengaged, while transferred out and deceased are absorbing states. Duration of disengagement is classified as short, medium or long term, as described below. Figure 1 is a graphical representation of the states and possible transitions.

Given a specific operationalization of the care cascade – or in our case, the process of engagement in care – a key challenge is translation of patient-level data into well-defined stages in the cascade. In the AMPATH data, as with all electronic health records, frequency of observation times are heterogeneous within and between individuals, and key clinical information (e.g., CD4 counts) can be measured sporadically and at irregular times. We use

AMPATH patient monitoring guidelines to set equally spaced time intervals to align patient-level data and ascertain membership in each phase over time. In general, AMPATH patients are monitored every 3 months when on ART, and every 6 months if not on ART, so that all patients are expected to re-visit a clinic at six-month intervals regardless of their ART status. We therefore construct a dataset whereby state membership is ascertained and recorded at 200-day intervals following enrollment (every 6 months plus a grace period). Information on time-varying covariates such as CD4 count and treatment status are treated in a similar way. The algorithm used to convert individual patient records into state membership outcomes is described in more detail in Section 4.1.

The multistate model enables several types of summaries and analyses. We can generate summaries of the proportion in each state and rates of transition between states as a function of time for examining temporal trends. The model naturally incorporates the commonly-observed cyclic pattern of temporary disengagement followed by re-engagement in care [29]. State transitions can be modeled as a function of covariates using multinomial regression for repeated measures, enabling examination of the effects of individual- and program-level factors for evaluation and prediction. Treating state membership using multinomial regression has the added advantage of handling competing risks for terminal events such as death and transfer out of care.

Given the focus on engagement and retention, we use ART initiation as a covariate rather than a state to examine its effect on outcomes over time. More broadly, we examine individual-level characteristics such as age, gender, CD4 counts, and calendar year that predict passage from one stage of the cascade to another, illustrating how the model can be used to identify potential factors associated with negative outcomes such as disengagement and death.

3. Multistate process and Markov models

A multistate process is a stochastic process that represents movement through a series of discrete states over time. Let $\{S(t) : t \geq 0\}$ be a multistate process with a finite state space $\mathscr{S} = \{1, 2, \dots , L\}$. Multistate processes can be characterized in terms of probability of transition from state k at time t to state l at time $t + u$, for $u > 0$, given by

$$p_{kl}(t,t+u)=\mathrm{pr}\{S(t+u)=l|S(t)=k\}, \quad k,l \in \mathscr{S},\ t \geq 0,\ u>0. \quad (1)$$

For continuous-time processes, a transition intensity function

$$q_{kl}(t)=\lim_{\Delta t \to 0}\frac{p_{kl}(t,t+\Delta t)}{\Delta t}, \quad (2)$$

describes instantaneous risk of transitioning from state k to l at time t, and can be elaborated by including dependence on covariates or on prior states. For discrete-time processes where the time increment u is a fixed constant, $S(t)$ is observed at times t_0, t_1, t_2, …, where $t_0 = 0$ and $t_j = uj$ for $j = 1, 2, …$. We can rewrite (1) in terms of transition probabilities

$$p_{kl}(t_{j'}, t_j) = \mathrm{pr}(S_j = l | S_{j'} = k), \quad k, l \in \mathscr{S}, \, t_{j'} < t_j,$$

where $S_j = S(t_j)$. Let $\{\mathbf{X}(t) : t \geq 0\}$ denote a (possibly multivariate) covariate process, and let $\mathscr{F}_{t^-}$ denote the accumulated history of $S(t)$ and $\mathbf{X}(t)$ up to but not including time t. Some components of $\mathbf{X}(t)$ may be time invariant (e.g., baseline covariates and attributes such as gender and age at enrollment). In this article, we will model one-step transitions between states using a discrete-time model having a first-order dependence structure such that

$$\begin{aligned} p_{kl}(t_j, t_{j-1} | \mathscr{F}_{t^-}) &= \mathrm{pr}(S_j = l | S_{j-1} = k, \mathscr{F}_{t_j}) \\ &= \mathrm{pr}(S_j = l | S_{j-1} = k, \mathbf{X}_j), \end{aligned} \tag{3}$$

where, using a slight abuse of notation, $\mathbf{X}_j = \mathbf{X}(t_j^-)$ is the information in $\mathbf{X}(t)$ available up to but not including t_j. In words, this assumption implies first-order dependence in state transitions conditionally on covariate information $\mathbf{X}_j$.

An early accounting of large sample theory and inference procedures for multistate models can be found in Albert [30]. Continuous time formulations can also be represented in terms of counting processes; see Kalbfleisch and Lawless [31] and Andersen et al. [32] for a comprehensive treatment. A variety of applications are described in Putter et al. [33] and Therneau and Grambsch [34]. Discrete-time models can be formulated in terms of log-linear models [35], which is closer to the approach we take here.

Multistate transition models have broad application and have been widely applied in biomedical research, particularly for describing disease progression [36–39] and evolution [40]. In HIV/AIDS, they have been used to model disease progression [41–44] and to characterize trajectories of associated markers of progression [45, 46]. Another important application concerns the evolution of antiretroviral drug resistance [47, 48]. Discrete time models also have been used to characterize and predict HIV/AIDS epidemics [49, 50].

Applications of multistate modeling to the HIV care cascade are relatively new but can be expected to increase. For example, Yehia et al. [51] used multinomial models to characterize transitions between states defined in terms of retention in care and viral suppression; however, their analysis considered only one-year transitions during a brief time period. Nosyk et al. [29] use a recurrent event model to characterize retention on antiviral therapy. The emerging importance of the HIV care cascade as a framework for monitoring patient- and program-level outcomes, combined with the increasing availability of large-scale cohort data and data from electronic health records, provides a natural and important opportunity for new applications of multistate modeling. The model presented here is intended to describe a basic framework and to provide a basis for formulating more elaborate and complex statistical approaches to representing the full care cascade.

4. Multistate model for engagement in care

4.1. Defining states

Let $i = 1, \ldots, n$ index individuals in a cohort enrolled at time $t = 0$ and followed until a fixed date on which the database is closed. Because of staggered enrollment times, the maximum possible follow up time for individual i is t_i^*. Following Figure 2, we discretize the time axis into intervals $[0, t_1]$, $(t_1, t_2]$, $(t_2, t_3]$, … having equal length $u = t_j - t_{j-1}$, where $t_0 = 0$ is enrollment time.

Patient-level data related to engagement in care can be represented in terms of three distinct counting processes. For individual i and for $t \geq 0$, let $N_i^V(t)$ denote number of visits, and let $N_i^D(t)$ and $N_i^T(t)$ denote zero-one counting processes for death and transfer out of care, respectively. Let $Y_i(t) = I\{t \leq t_i^*, N_i^T(t) = N_i^D(t) = 0\}$ denote an 'at-risk' variable that indicates whether an individual is eligible to return for clinic visits. Information in the counting processes is used to generate a discrete-time multistate process $\{S_i(t) : t = 0, t_1, t_2, \ldots, t_{J_i}\}$, where $S_i(t) \in \{1, 2, \ldots, L\}$ is state membership at time t and $J_i = \max_j\{t_j \leq t_i^*\}$ is the number of intervals $(t_{j-1}, t_j]$ such that $t_j \leq t_i^*$. Alternatively we can write $\{S_{ij} : j = 0, 1, 2, \ldots, J_i\}$, where $S_{ij} = S_i(t_j)$.

For the AMPATH data, counting process information is converted into the discrete-time process as follows:

$$S_{ij} = \begin{cases} 1 & \text{if} \quad Y_i(t_j)\int_{(t_{j-1},t_j]} dN_i^V(u) > 0 & \text{(engaged)} \\ 2 & \text{if} \quad Y_i(t_j)\int_{(t_{j-1},t_j]} dN_i^V(u) = 0 \text{ and } S_{i,j-1} = 1 & \text{(disengaged for 1 period)} \\ 3 & \text{if} \quad Y_i(t_j)\int_{(t_{j-1},t_j]} dN_i^V(u) = 0 \text{ and } S_{i,j-1} = 2 & \text{(disengaged for 2 periods)} \\ 4 & \text{if} \quad Y_i(t_j)\int_{(t_{j-1},t_j]} dN_i^V(u) = 0 \text{ and } S_{i,j-1} \in \{3, 4\} & \text{(disengaged for 3+ periods)} \\ 5 & \text{if} \quad N_i^T(t_j) = 1 & \text{(transferred out)} \\ 6 & \text{if} \quad N_i^D(t_j) = 1 & \text{(deceased)} \end{cases} \tag{4}$$

Put simply, S_{ij} denotes state membership at the end of interval $(t_{j-1}, t_j]$. All cohort members are engaged at enrollment to care so that $S_{i0} = 1$ for all i. For those with $Y_i(t_j) = 1$ (not dead or transferred out), $S_{ij} = 1$ (engaged) if one or more visits occur within the interval and $S_{ij} =$ 2, 3, or 4 (disengaged) if not. States 2, 3 and 4 represent durations of disengagement from care, measured in intervals. In principle the number of 'disengaged' states can be increased. Preliminary analyses of the AMPATH cohort indicates that the probability of returning to care after 3 or more intervals without a visit is less than .02, motivating our choice to use only 3 categories; in principle, the number of categories can be greater. Note that $S_{ij} = 4$ is not an absorbing state: individuals can return to care even after being disengaged for 3 or more intervals.

States 5 (transfer out) and 6 (death) are absorbing states. In the AMPATH EMR, a transfer indicator (yes/no) is available but transfer date is not. Therefore transfer is assumed to occur in the first interval following the most recent engagement in care. For those who died, state membership becomes deceased from the date of death.

Translation of counting processes into discrete-time states is illustrated in Figure 2. Each panel depicts data from a single patient starting from enrollment at $t = 0$ and extending through database closure (August 24, 2016). Grey vertical lines represent patient visits; at each visit, $dN_i^V(t)=1$. State membership is represented along the bottom of Figure 2. Database closure can induce right censoring of the last time interval. In our analysis we assume this administrative censoring is non-informative; hence we do not include information from these incompletely observed intervals.

In addition to information about state of engagement, we assume there is information available on an r-dimensional covariate process $\{\mathbf{X}_{ij} : j = 0, 1, 2, \dots , J_i\}$, where $\mathbf{X}_{ij}=\mathbf{X}_i(t_j^-)$ is the most recently observed value of $\mathbf{X}_i(t)$ at the instant prior to t_j. Referring again to Figure 2, observed CD4 counts are plotted as a function of time, and treatment status is represented using solid or dashed vertical line for each visit. Some of the covariates, such as gender and age at enrollment, are time-independent.

4.2. State transitions

An important quantity describing change over time is the state transition probability $p_{jkl} = \text{pr}(S_j = l / S_{j-1} = k)$, which forms the basis of several types of analyses. In the absence of covariates, the collection of transition probabilities can be used to form the $L \times L$ transition matrix $\mathbf{P}_j$, where element (k, l) is p_{jkl}. An annotated version of $\mathbf{P}_j$ corresponding to the AMPATH engagement states in (4) is given in Table 2. Note that the sum of each row of $\mathbf{P}_j$ is 1.

From a programmatic point of view, time-specific or aggregated summaries of $\mathbf{P}_j$ can be useful in identifying ‘leaks’ in the process of retention in care. Specific elements of the transition matrix can be plotted as a function of time, providing visual representation to identify periods of high risk where unfavorable outcomes such as patient disengagement from care or death are more pronounced. Moreover, under certain assumptions, the transition matrix can be used to estimate marginal probability of state membership as a function of time. To illustrate, let $\boldsymbol{\pi}_{jl} = \text{pr}(S_j = l)$ and let $\boldsymbol{\pi}_j = (\boldsymbol{\pi}_{j1}, \boldsymbol{\pi}_{j2}, \dots , \boldsymbol{\pi}_{jL})^{\mathrm{T}}$ denote the vector of state membership probabilities at time j. Under the assumption of first-order Markov dependence, and in the absence of covariates, marginal state membership probabilities are calculated as $\boldsymbol{\pi}_j=\boldsymbol{\pi}_1^{\mathrm{T}}\prod_{m=2}^{j}\mathbf{P}_m$.

4.3. Regression models for state transitions

In the discrete time domain, state membership S_{ij} at each time point t_j follows multinomial distribution. Recall that $S_{ij} \in \{1, \dots , L\}$. The general multinomial model can be elaborated as follows to incorporate the effect of covariates on transition dynamics:

$$S_{ij}|S_{i,j-1}{=}k, \mathbf{X}_{ij}{=}\mathbf{x}, Y_i(t_{j-1}){=}1{\sim}\text{Mult}\{\mathbf{p}_{ijk}(\mathbf{x})\},\ j{=}1,\ldots,J_i; k{=}1,\ldots,L, \quad (5)$$

where $\mathbf{p}_{ijk}(\mathbf{x}) = (p_{ijk1}(\mathbf{x}), \ldots, p_{ijkL}(\mathbf{x}))^T$ is an L-vector having elements $p_{ijkl}(\mathbf{x}) = \text{pr}(S_{ij} = l / S_{i,j-1} = k, \mathbf{X}_{ij} = \mathbf{x})$, with $\sum_{l=1}^{L} p_{ijkl}(\mathbf{x}){=}1$. Having absorbing states or order restrictions, as in Table 2, constrains a subset of the transition probabilities to be zero. In practice the domain of k in model (5) is limited to non-absorbing states, a constraint that is easily imposed in practice by conditioning on $Y_i(t_{j-1}) = 1$ and restricting sample at t_j to those still in the risk set. Although this model assumes first-order dependence as in (3), more general dependence structures are possible and are discussed further in our data analysis in Section 5.

The specification in (5) lends itself directly to formulation of regression models for longitudinal multinomial data [35, 52]. We adopt the commonly-used loglinear specification that models rate of transition from one state to another relative to a reference state, and captures covariate effects in terms of log relative rate ratios (RRR). Suppressing subscript i, a fully general specification is

$$\log\left\{\frac{p_{jkl}(\mathbf{x}_j)}{p_{jk1}(\mathbf{x}_j)}\right\}{=}g_{jkl}(\mathbf{x}_j) \quad (6)$$

where, for $j = 1, \ldots, \max_i\{J_i\}$, $k = 1, \ldots, L$ and $l = 2, \ldots, L$, the coefficient $g_{jkl}(\mathbf{x})$ is a function that represents the effect of $\mathbf{x}_j$ on log rate of the transition from state $k \to l$ relative to $k \to 1$. In practice it will typically be necessary to impose structure on g in order to fit the model, such as assuming g has a known functional form, that the form of g is constant over time, or that g is indexed by a finite-dimensional parameter vector.

Subscripting g_{jkl} by k implicitly indicates that prior state S_{j-1} is part of the regression function, and that regression effects are state-specific. For example, looking at transitions from $k = 1$ (engaged) at t_{j-1} to $l = 2$ (disengaged) at t_j as a function of a scalar, time invariant covariate x, we can write

$$\log\left\{\frac{p_{j12}(x)}{p_{j11}(x)}\right\}{=}I(S_{j-1}{=}1)(\alpha_{j12}{+}\beta_{j12}x),$$

where α_{j12} is an intercept term and β_{j12} is the log RRR characterizing the effect of x on transition rate $p_{j12}(x)$ (from engaged to disengaged) relative to rate $p_{j11}(x)$ (from engaged to engaged). When g is linear in a regression coefficient β, as in this example, the relative risk ratio $\exp(\beta)$ may provide more natural interpretation about the effect of the covariate.

In addition to the effect of clinical or socio-economic covariates on the relative transitions, we can also estimate various types of temporal variations such as *time, period,* or *cohort effects* that are often of interest in epidemiologic studies. Briefly, the time effect refers to the effect of follow-up time, the period effect is the effect of calendar time, and the cohort effect represents the effect of enrollment time/date. It is known that these three effects cannot be evaluated simultaneously, even with longitudinal data. However, any two of them can be captured within the regression function by appropriately parameterizing the design matrix. In Section 5, we estimate period effects using a regression spline and time effects using indicator variables for interval number *j*.

4.4. State membership prediction

Recall that in the absence of covariates, the matrix $\mathbf{P}_j$ contains the one-step transition probabilities $\{p_{jkl}\}$ as listed in Table 2. Elements of a matrix $\mathbf{P}_j(\mathbf{x})$ having covariate-dependent transition probabilities $\{p_{jkl}(\mathbf{x})\}$ are calculated via

$$p_{jk1}(\mathbf{x})=\frac{1}{1+\sum_{m=2}^{L}\exp\{g_{jkm}(\mathbf{x})\}}, \quad l=1;$$
$$p_{jkl}(\mathbf{x})=\frac{\exp\{g_{jkl}(\mathbf{x})\}}{1+\sum_{m=2}^{L}\exp\{g_{jkm}(\mathbf{x})\}}, \quad l=2,\ldots,L.$$

For settings where covariates are exogenous, marginal state membership probabilities as a function of $\mathbf{x}_j$ can be calculated as described in Section 4.2. Specifically, let $\bar{\mathbf{x}}_j = (\mathbf{x}_1, \ldots, \mathbf{x}_j)$ denote the longitudinal history of $\{\mathbf{x}_j\}$, and let $\boldsymbol{\pi}_j(\bar{\mathbf{x}}_j)$ represent the $L \times 1$ vector of (covariate-specific) marginal state probabilities $\boldsymbol{\pi}_{jl}(\bar{\mathbf{x}}_j) = \mathrm{pr}(S_j = l / \bar{\mathbf{x}}_j)$. Then $\boldsymbol{\pi}_j(\overline{\mathbf{x}}_j)=\boldsymbol{\pi}_1^{\mathrm{T}}(\mathbf{x}_1)\prod_{m=2}^{j}\mathbf{P}_m(\mathbf{x}_m)$. Averaging over the distribution of $\bar{\mathbf{x}}_j$ yields predicted distribution across states for the population. When covariates are endogenous, as is the case with CD4 count and similar disease markers, predicting state membership probabilities requires a model for the joint distribution of the covariate and state transition processes.

4.5. Estimation and model diagnostics

The time-specific transition probabilities listed in Table 2 can be estimated using those in the risk set at each time t_j,

$$\hat{p}_{jkl}=\frac{\sum_i Y_i(t_{j-1})I\{S_{ij}=l, S_{ij-1}=k\}}{\sum_i Y_i(t_{j-1})I\{S_{ij-1}=k\}}$$

Under the assumption of first-order dependence, these can subsequently be used to estimate state membership probabilities as described in Section 4.2.

Estimation and inference about covariate effects requires simplifications of the regression functions $g_{jkl}(\mathbf{x})$. A natural simplification, which we adopt in our example, is to assume the functional form of g is known, is constant over time, and is indexed by an unknown vector of regression parameters. Specifically, we consider versions of model (6) that can be written as

$$\log\left\{\frac{p_{jkl}(\mathbf{x}_j)}{p_{jk1}(\mathbf{x}_j)}\right\}=\alpha_{jkl}+\mathbf{x}_j\boldsymbol{\beta}_{kl}+h_{kl}(d_j;\boldsymbol{\gamma}_{kl}) \qquad (7)$$

where, for $k \rightarrow l$ transitions, α_{jkl} is a time-specific intercept and $\boldsymbol{\beta}_{kl}$ is a vector of regression coefficients having the same dimension as $\mathbf{x}_j$. The term $h_{kl}(d_j ; \boldsymbol{\gamma}_{kl})$ captures a transition-specific *period* effect: d_j is calendar date corresponding to interval endpoint t_j (which varies by individual) and each $h_{kl}(; \boldsymbol{\gamma}_{kl})$ is a smooth function parameterized by a finite-dimensional parameter $\boldsymbol{\gamma}_{kl}$. In practice h can specified using regression splines or other low-rank smoothing techniques. In our application we specify h using thin plate regression splines [53] and implement estimation using the `mgcv` package in `R`. In principle, estimation and inference can be carried out using any statistical software package that fits multinomial regression models.

Consistent estimates of model parameters in (7) can be obtained by using maximum likelihood to fit a multinomial regression to the longitudinal observations, treating them as if they are independent. Consistency in this case requires (i) length of follow up is non-informative in the sense that it is unrelated to state transition rate at each time t_j, conditionally on $\mathbf{x}_j$; and (ii) the number of individuals increases in such a way that the ratio of individuals to intervals is a constant. This second condition permits estimation of time-specific intercepts even though the number of time points may increase with the number of individuals. Asymptotic normality of the parameter estimates relies on mild regularity conditions in [54] along with some conditions on working covariance structure described in [55]. Estimation of standard errors must acknowledge within-subject correlation. Either robust standard errors [56] or bootstrap resampling (within individuals as the sampling unit) can then be used.

Model fit can be assessed by comparing, at each time j and for those with $Y_i(t_{j-1}) = 1$, the observed and predicted proportion in each state (i.e., marginal state membership probability). Chi-square type goodness-of-fit tests applied to large datasets may be prone to generating statistically significant discrepancies; visual comparisons may be more useful for identifying meaningful differences between observed and fitted values. Examples are provided in our analysis of AMPATH data.

Finally, the assumption of first-order Markov dependence is a strong one. As a practical matter, it is may not be feasible to test all possible violations of first-order dependence. However, one straightforward approach — used in our analysis — is to fit model (6) with S_{j-2} as a covariate, possibly interacted with elements of $\mathbf{X}_j$, and test for violations of first-order Markov dependence using a Wald test for the added covariates (with standard errors calculated as indicated above).

5. Application to AMPATH Data

5.1. Overview

We illustrate application of the multistate modeling approach using data from AMPATH as described in Section 2. Our analysis uses data on 92,215 HIV infected adults aged 18 years or greater who enrolled in AMPATH supported clinics between June 2008 and August 2016. The time axis originates at enrollment in care ($t_0 = 0$; baseline). Characteristics at enrollment and data on follow-up information are shown in Table 1. The database closure date is August 24th, 2016, which yields a maximum of $\max_i\{J_i\} = 15$ intervals for ascertaining state membership.

We operationalized the 6-state engagement process as described in Section 2 using intervals of length $u = 200$ days. We used the mapping in (4) to define, for each individual, state membership S_{ij} at times $\{t_0, t_1, t_2, \ldots\} = \{0, 200, 400, \ldots\}$. We also incorporated covariate information on age and gender at enrollment, and on CD4 count and treatment status, both of which vary over time. Figure 2 depicts raw data on visit times, treatment status and CD4 count for four different patients, and illustrates how information on visit times (vertical lines) is converted into state membership (color-coded bars along horizontal axis). Figure 3 shows the number of individuals with available data as a function of potential length of follow-up time.

Our analysis is designed to characterize state transitions appearing in Table 2, and has several components. First, we summarize state transition rates in $\mathbf{P}_j$ as a function of time since enrollment and overall (aggregating over time). Next, we use multinomial regression to estimate the effects of age, gender, time-varying CD4 count and time-varying treatment status on state transitions over time. We also use the model to characterize period effects (calendar time) and effect of time since enrollment. Finally, we apply goodness of fit tests and investigate possible violations of the first-order Markov assumption.

5.2. Summarizing state transition rates and temporal trends

Table 3 provides a summary of state transition rates using all data aggregated over time. Entries in the upper left indicate the cyclic nature of engagement in care: among those engaged at a specific time, rate of continued engagement is .85 and rate of disengagement is .13. Among those disengaged for one period of time, about 11 percent return to care.

Transitions related to the disengaged states illustrate the deleterious effects of being disengaged in care for more than 200 days. Among those disengaged for a single interval (Disengaged 1), 88 percent will remain disengaged for a second interval (Disengaged 2); subsequent to that, 94 percent will become disengaged for 3 or more intervals (Disengaged 3^+). Hence, among those who become disengaged for one interval, the probability of remaining disengaged for 3 or more intervals is $(.94)(.88) \approx .83$.

Table 3 also indicates that the per-interval mortality rate is slightly over .02 (aggregating values in the last column). It must be noted that mortality estimates are based only on deaths that are confirmed and recorded in the AMRS, and that these estimates have a potentially substantial downward bias [57–59]. Another potential source of mortality rate under-estimation is reporting lag because death registration data is not linked directly to AMRS.

Figure 4 shows temporal trends in the state transition rates, and can be viewed as a visualization of the entries of the transition matrix $\mathbf{P}_j$ as a function of interval endpoint times t_j. The plot is rendered in the logit scale to make temporal trends and fluctuations more apparent. The plot of transitions from the engaged state ($S = 1$) suggest that the most critical period for retention in care is the first interval after enrollment, where we see a sharp decrease in re-engagement following enrollment. Referring to transitions from 'disengaged 1' ($S = 2$, upper right panel), we see that those who are disengaged in this first interval are additionally at higher risk for remaining disengaged in the second interval. Thereafter, retention and re-engagement rates (green lines) tend to remain steady or even increase slightly, except for those who have been disengaged for more than 3 intervals ($S = 4$, lower right panel). Rates for mortality and transfer-out are relatively low and remain roughly constant over time.

Overall, plots for transitions from various disengagement status imply that disengaged patients who did not come back in the first to the second follow up years are at high risk of having a long gap to come back for care and ultimately become lost. These plots simultaneously identify when and where the greatest gaps in AMPATH care exist. The gaps are attributable to high continued disengagement rate in the early follow up period, which further leads to a high chance of having a long gap and becoming lost-to-follow-up in the following years.

5.3. Regression modeling

We use a version of the regression model given in (7) for our analysis in this section. Specifically, our model has the form

$$\log\left\{\frac{p_{jkl}(\mathbf{x}_{ij},d_{ij})}{p_{jk1}(\mathbf{x}_{ij},d_{ij})}\right\}=\alpha_{jkl}+\mathbf{x}_j\boldsymbol{\beta}_{kl}+h_{kl}(d_{ij};\boldsymbol{\gamma}_{kl}),\qquad k=1,\ldots,4;l=2,\ldots,6.$$

The domain of k indicates that only those transitions $k \to l$ from states $k = 1, \ldots, 4$ are being modeled. Further restrictions apply as they relate to modeling transitions from the disengaged state (Table 2); in particular the model only applies to transition rates that are not deterministically 0 or 1. The variable d_{ij} captures period effect, and represents the number of days elapsed from January 1, 2008 until the (individual-specific) date associated with time t_j. The period effect is captured by the smooth function h_{kl}, which is specific to each $k \to l$ transition but does not vary with enrollment time. As indicated previously, we use thin plate regression splines to specify h_{kl}.

The covariate matrix $\mathbf{X}_{ij}$ is set up as follows. Let $\mathbf{X}_{i0} = (I\{\text{Age}_i \geq 35\}, \text{Male}_i)$ denote the 1×2 indicator vector for age and gender at baseline ($\text{Male}_i = 1$ if male, 0 if female). We define a 3-level nominal variable for time-varying CD4 count based on the most recently observed value of CD4 as of t_j, and a two-level variable for treatment based on whether or not an individual is on antiretroviral therapy at t_j (1 if yes, 0 if no). The CD4 variable takes value 1 if the most recently observed CD4 count is less than or equal to 350; value 2 if over 350, and value 0 if CD4 has not been measured between t_0 and t_j. We then create a 1×5 vector of indicators to reflect the full interaction between these two, using $\{\mathrm{CD4} < 350, \mathrm{ARV}^-\}$ as the reference category. Appended to $\mathbf{X}_{i0}$, this generates a 1×7 covariate vector $\mathbf{X}_{ij}$ for the risk set $\{i : Y_i(t_j) = 1\}$ at each time t_j, so that the coefficient vector $\boldsymbol{\beta}_{kl}$ has dimension 7×1. Table 4 and 5 show the effect of covariates in $\mathbf{X}_j$ on state transition probabilities in terms of relative rate ratios (RRR).

Table 4 shows covariate effects for transitions from the engaged state ($S = 1$). To illustrate interpretations, consider the effect of age on transition from engaged to disengaged, for which RRR = .64. The RRR represents a comparison, between those with Age ≥ 35 and Age <35, of the rate ratio p_{12}/p_{11}, where the transition rate engaged → disengaged is in the numerator and engaged → engaged in the denominator (i.e., transition *to* engaged is the reference transition). Using a slight abuse of notation, we can see that the RRR is actually a ratio of ratios by writing

$$\mathrm{RRR}(\mathrm{Age} \geq 35{:}\mathrm{Age}{<}35) = \frac{p_{12}(\mathrm{Age} \geq 35)/p_{11}(\mathrm{Age} \geq 35)}{p_{12}(\mathrm{Age}{<}35)/p_{11}(\mathrm{Age}{<}35)} = .64.$$

In short, the RRR here indicates that older individuals are less likely to disengage in care, once engaged. Reading across the first row, we see that older individuals also are less likely to transfer from care (RRR = .57) and are at higher risk for mortality (RRR = 1.14), once engaged in care. The gender effect indicates that men are at higher risk for becoming disengaged and for mortality. Note that the RRR for any individual covariate is conditional on other covariates included in the model. Hence the RRR shown in Tables 4 and 5 are conditional on the other covariates listed as well as calendar time.

An examination of the time-varying covariates in Table 4 allows comparison of groups defined by their CD4-treatment profile. Looking first at those with CD4 <350, we see those on ARV have lower rates of disengagement, transfer and mortality, compared to those not on ARV (the reference group). The effect of treatment within the other CD4 categories can be seen by comparing RRR within category. For example, among those with CD4 ≥ 350, the RRR for becoming disengaged is .12 while on ARV and .31 if not, suggesting lower overall risk of disengagement while on ARV. The effect of ARV is more pronounced if information on CD4 is absent. Though not shown here, confidence intervals for these within-category comparisons can be obtained by changing the reference category or by direct computation from the estimated variance-covariance matrix of the model parameters.

Table 5 summarizes covariate effects for transitions from disengaged states. Among the key findings: (i) once disengaged from care for one interval, men and those over 35 are less likely to remain disengaged (or, more likely to return to care), but the effect dissipates once disengagement lasts for two or more intervals; (ii) within each CD4 category, those on treatment are less likely to remain disengaged and tend to have lower rates of mortality.

Figure 5 shows estimated calendar-year effects in terms of shifts in log relative risk of disengagement and death among those who are engaged in care. The left panel indicates that rate of disengagement from care increased between 2008 and 2014, but leveled out thereafter; the right panel indicates a substantial reduction in observed mortality starting around 2013. It is possible that these trends are tied to programmatic changes such as the introduction of viral load monitoring in 2014. However attributing the trends to specific causes requires a more comprehensive analysis and the curves should not be over-interpreted. It is likely that the sharp decrease in observed mortality among those engaged in care is at least partially attributable to reporting lags. Similar trends are seen in the period effect for transitions from disengaged state to mortality (not shown). For those effects, incorporating information from contact tracing and double sampling could be used to partially correct this potential bias, as indicated above [57–59].

5.4. Model assessment

To assess potential lack of fit, we constructed plots of observed and fitted state membership probabilities for each interval with follow-up data available. Results are presented in Figure 1 in Appendix. With small exceptions, the model shows good agreement with the observed data.

A key assumption in our model is first-order Markov assumption. In our implementation, we assume state membership at t_j depends only on state and covariate information at t_{j-1}, and conditionally on that information, is independent of observed-data history at or before t_{j-2}. To assess the effects of potential violations of this assumption, we re-fit the model of transitions from *engaged*, adding information about S_{j-2} as a categorical covariate. We examined a second order dependence for the engagement model only, because disengagement state at t_j is always determined by second order state membership S_{j-2}. Results shown in Table 1 in Appendix indicate that there is evidence of second order dependence in transitions from engaged to other states. The results imply that those who engaged in two consecutive intervals are less likely to disengage from care, transfer-out, and die, compared to those who missed visits and return to care. However, second order dependence did not change the substance of our findings and the estimated effect of covariates on transition rates are very robust between two multistate models with and without adjusting for S_{j-2}.

Using second order dependence model, we re-generated the plots of observed and fitted state membership probabilities. Figure 2 in Appendix shows that inclusion of second-order term seems to improve the model fit for early disengagement (day 200 in short disengagement) and transfer-out. Although the impact of second-order term does not seem substantial, it would be important to account for higher order dependence in model for predictive inference about state membership.

6. Discussion

This paper describes the use of a multistate modeling framework for the HIV care cascade. With the growing and widespread availability of large-scale individual-level cohort data, our work addresses the need for statistical modeling approaches that can be used to take full advantage of the rich information about longitudinal individual-level outcomes in these on these large datasets for both programmatic and research purposes [21].

We focus here on the process of engagement and retention in care, and use a discrete-time multistate model of longitudinal transitions between states that characterize engagement, duration of disengagement, transfer out of care, and mortality. State membership is defined at pre-specified interval endpoints that correspond to a maximum expected lag time between visits for individuals engaged in HIV care, where the lag time is set by the user.

The model works on the assumption that care engagement status will be ascertained at discrete points in time, starting from enrollment in care. This confers some advantages: it enables transparent translation of irregular data from electronic health records into an analyzable format; it allows users to apply, with some specialization, statistical software for fitting multinomial models for longitudinal data; it enables prediction, evaluation of covariate effects, and estimation of state membership probabilities at fixed points in time. We have given suggestions on how to translate data on multiple counting processes into a discrete-time state transition process, and illustrated how to use the model by analyzing data from over 90,000 individuals from AMPATH. We have also shown how to use the model to deal with cyclic engagement in care, long-term disengagement, and competing risks of terminal events such as transfer from care and death. With some modification, the model can be extended to handle other categories that are important in the care cascade.

The model does rely on some key assumptions. A basic assumption is first-order Markov dependence. We used it in our example, but we also showed how it can be relaxed. A second assumption is that the expected visit frequency in our definition of ‘engaged in care’ applies to all individuals at all time points, when in reality visit frequencies can be dependent on clinical characteristics and treatment status.

The existing limitations of our model, combined with the broader scope of analytic needs related to the HIV care cascade, point to several potential avenues of further development. First, the definition of ‘engaged in care’ can be elaborated to accommodate expected time of return visit; more generally, a model structure that accommodates patient visits as a continuous-time process can be considered. Both of these extensions would trade simplicity for complexity, and it will be important to understand the potential benefits of increasing model complexity as it pertains to fulfilling the larger analytic goals related to summarizing the cascade.

Second, we are working on models that increase the number of states needs to be expanded to accommodate the important benchmarks related to initiating antiviral treatment and having viral load suppression. AMPATH has recently instituted annual viral load monitoring for most of its clients, which makes this extension more practically feasible. Third, as most data on the care cascade is observational, there is a need to develop a framework for causal inference or, more generally, handling exogenous covariates. We are currently working on using inverse probability weighting and g estimation, which build naturally on the longitudinal regression framework, to assess causal effects for key cascade-related outcomes. In addition, we are working on incorporating variable selection tools.

Fourth, variable selection and the ability to handle more covariates is clearly important. Data from cohort studies typically collect comprehensive information about risk factors, diagnosis, and lab test results. We are working to incorporate formal approaches to variable selection to allow users to identify specific determinants of transition dynamics along the care cascade, and thereby make better use of existing resources for interventions.

Finally, misclassification of deaths and other outcomes is a common problem in state membership ascertainment. Incorporation of data from contact tracing and double sampling can lead to significant improvements in estimation. A major priority in the next step of model development is incorporation of external information derived by tracing those who are disengaged for an extended period of time.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Associate Editor and two reviewers for the careful reading of our paper and all of the helpful comments that greatly improved the presentation of the article. The authors were partially supported by NIH grants R01 AI1083441, and P30 AI042853 (Providence-Boston Center for AIDS Research), and NIMH grant K01 MH099966. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Ulett KB, Willig JH, Lin HY, Routman JS, Abroms S, Allison J, Chatham A, Raper JL, Saag MS, Mugavero MJ. The therapeutic implications of timely linkage and early retention in HIV care. AIDS Patient Care and STDs. 2009; 23(1):41–49. [PubMed: 19055408]
2. Gardner EM, McLees MP, Steiner JF, del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. Clinical Infectious Diseases. 2011; 52(6):793–800. [PubMed: 21367734]
3. McNairy ML, El-Sadr WM. The HIV care continuum: No partial credit given. AIDS. 2012; 26(14): 1735–1738. [PubMed: 22614888]
4. Kranzer K, Govindasamy D, Ford N, Johnston V, Lawn SD. Quantifying and addressing losses along the continuum of care for people living with HIV infection in sub-Saharan Africa: A systematic review. Journal of the International AIDS Society. 2012; 15(2)
5. Meeting report on framework for metrics to support effective treatment as prevention; 2–3 April 2012; Geneva, Switzerland. 2012. http://www.who.int/iris/handle/10665/75387
6. [Accessed: 2017-02-28] HIV Care Continuum for the United States and Puerto Rico. 2014. http://aidsetc.org/resource/hiv-care-continuum-united-states-and-puerto-rico
7. Nosyk B, Montaner JS, Colley G, Lima VD, Chan K, Heath K, Yip B, Samji H, Gilbert M, Barrios R, et al. The cascade of HIV care in British Columbia, Canada, 1996–2011: A population-based retrospective cohort study. The Lancet Infectious Diseases. 2014; 14(1):40–49. [PubMed: 24076277]
8. Kim MH, Ahmed S, Buck WC, Preidis GA, Hosseinipour MC, Bhalakia A, Nanthuru D, Kazembe PN, Chimbwandira F, Giordano TP, et al. The Tingathe programme: A pilot intervention using community health workers to create a continuum of care in the prevention of mother to child transmission of HIV (PMTCT) cascade of services in Malawi. Journal of the International AIDS Society. 2012; 15(4)
9. Medley A, Ackers M, Amolloh M, Owuor P, Muttai H, Audi B, Sewe M, Laserson K. Early uptake of HIV clinical care after testing HIV-positive during home-based testing and counseling in western Kenya. AIDS and Behavior. 2013; 17(1):224–234. [PubMed: 23076720]

10. Mugavero MJ, Amico KR, Horn T, Thompson MA. The state of engagement in HIV care in the United States: From cascade to continuum to control. Clinical Infectious Diseases. 2013; 57(8): 1164–1171. [PubMed: 23797289]
11. Boyer S, Iwuji C, Gosset A, Protopopescu C, Okesola N, Plazy M, Spire B, Orne-Gliemann J, McGrath N, Pillay D, 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(sup3):39–51. [PubMed: 27421051]
12. Hallett TB, Eaton JW. A side door into care cascade for HIV-infected patients? Journal of Acquired Immune Deficiency Syndromes. 2013; 63:S228–S232. [PubMed: 23764640]
13. Olney JJ, Braitstein P, Eaton JW, Sang E, Nyambura M, Kimaiyo S, McRobie E, Hogan JW, Hallett TB. Evaluating strategies to improve HIV care outcomes in Kenya: A modelling study. The Lancet HIV. 2016; 3(12):e592–e600. [PubMed: 27771231]
14. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: A mathematical model. The Lancet. 2009; 373(9657):48–57.
15. Hontelez J, De Vlas SJ, Tanser F, Bakker R, Bärnighausen T, Newell ML, Baltussen R, Lurie MN. The impact of the new WHO antiretroviral treatment guidelines on HIV epidemic dynamics and cost in South Africa. PLoS ONE. 2011; 6(7):e21919. [PubMed: 21799755]
16. Ryan GW, Bloom EW, Lowsky DJ, Linthicum MT, Juday T, Rosenblatt L, Kulkarni S, Goldman DP, Sayles JN. Data-driven decision-making tools to improve public resource allocation for care and prevention of HIV/AIDS. Health Affairs. 2014; 33(3):410–417. [PubMed: 24590938]
17. Tierney WM, Rotich JK, Hannan TJ, Siika AM, Biondich PG, Mamlin BW, Nyandiko WM, Kimaiyo S, Wools-Kaloustian K, Sidle JE, et al. The AMPATH Medical Record System: Creating, implementing, and sustaining an electronic medical record system to support HIV/AIDS care in western Kenya. Studies in Health Technology and Informatics. 2007; 129(1):372. [PubMed: 17911742]
18. Kitahata MM, Rodriguez B, Haubrich R, Boswell S, Mathews WC, Lederman MM, Lober WB, Van Rompaey SE, Crane HM, Moore RD, Bertram M, Kahn JO, Saag MS. Cohort profile: the Centers for AIDS Research Network of Integrated Clinical Systems. International Journal of Epidemiology. 2008; 37:948–955. [PubMed: 18263650]
19. Gourlay A, Noori T, Pharris A, Axelsson M, Costagliola D, Cowan S, Croxford S, d'Arminio Monforte A, Del Amo J, Delpech V, Díaz A, Girardi E, Gunsenheimer-Bartmeyer B, Hernando V, Jose S, Leierer G, Nikolopoulos G, Obel N, Op de Coul E, Paraskeva D, Reiss P, Sabin C, Sasse A, Schmid D, Sonnerborg A, Spina A, Suligoi B, Supervie V, Touloumi G, Van Beckhoven D, van Sighem A, Vourli G, Zangerle R, Porter K. European HIV Continuum of Care Working Group. The human immunodeficiency virus continuum of care in European Union countries in 2013: Data and challenges. Clinical Infectious Diseases. 2017; 64:1644–1656. [PubMed: 28369283]
20. Justice AC, Dombrowski E, Conigliaro J, Fultz SL, Gibson D, Madenwald T, Goulet J, Simberkoff M, Butt AA, Rimland D, Rodriguez-Barradas MC, Gibert CL, Oursler KAK, Brown S, Leaf DA, Goetz MB, Bryant K. Veterans Aging Cohort Study (VACS): Overview and description. Medical Care. 2006; 44:S13–S24. [PubMed: 16849964]
21. Haber N, Pillay D, Porter K, Bärnighausen T. Constructing the cascade of HIV care: Methods for measurement. Current Opinion in HIV and AIDS. 2016; 11(1):102–108. [PubMed: 26545266]
22. Giordano TP, Hartman C, Gifford AL, Backus LI, Morgan RO. Predictors of retention in HIV care among a national cohort of US veterans. HIV Clinical Trials. 2009; 10(5):299–305. [PubMed: 19906622]
23. Marks G, Gardner LI, Craw J, Crepaz N. Entry and retention in medical care among HIV-diagnosed persons: A meta-analysis. AIDS. 2010; 24(17):2665–2678. [PubMed: 20841990]
24. Thompson MA, Mugavero MJ, Amico KR, Cargill VA, Chang LW, Gross R, Orrell C, Altice FL, Bangsberg DR, Bartlett JG, et al. Guidelines for improving entry into and retention in care and antiretroviral adherence for persons with HIV: Evidence-based recommendations from an International Association of Physicians in AIDS Care panel. Annals of Internal Medicine. 2012; 156(11):817–833. [PubMed: 22393036]

25. Mugavero MJ, Westfall AO, Cole SR, Geng EH, Crane HM, Kitahata MM, Mathews WC, Napravnik S, Eron JJ, Moore RD, et al. Beyond core indicators of retention in HIV care: Missed clinic visits are independently associated with all-cause mortality. Clinical Infectious Diseases. 2014; 59(10):1471–1479. [PubMed: 25091306]

26. Yehia BR, Stewart L, Momplaisir F, Mody A, Holtzman CW, Jacobs LM, Hines J, Mounzer K, Glanz K, Metlay JP, et al. Barriers and facilitators to patient retention in HIV care. BMC Infectious Diseases. 2015; 15(1):246. [PubMed: 26123158]

27. Braitstein P, Ayuo P, Mwangi A, Wools-Kaloustian K, Musick B, Siika A, Kimaiyo S. Sustainability of first-line antiretroviral regimens: Findings from a large HIV treatment program in western Kenya. Journal of Acquired Immune Deficiency Syndromes. 2010; 53(2):254–259. [PubMed: 19745752]

28. Genberg BL, Naanyu V, Wachira J, Hogan JW, Sang E, Nyambura M, Odawa M, Duefield C, Ndege S, Braitstein P. Linkage to and engagement in HIV care in western Kenya: An observational study using population-based estimates from home-based counselling and testing. The Lancet HIV. 2015; 2(1):e20–e26. [PubMed: 25621303]

29. Nosyk B, Lourenco L, Min JE, Shopin D, Lima VD, Montaner JSG. Characterizing retention in HAART as a recurrent event process: Insights into 'cascade churn'. AIDS. 2015; 29(13):1681–1689. [PubMed: 26372279]

30. Albert A. Estimating the infinitesimal generator of a continuous time, finite state Markov process. The Annals of Mathematical Statistics. 1962; 33(2):727–753.

31. Kalbfleisch J, Lawless JF. The analysis of panel data under a Markov assumption. Journal of the American Statistical Association. 1985; 80(392):863–871.

32. Andersen, PK., Borgan, Ø., Gill, RD. Statistical Models Based on Counting Processes. Springer-Verlag; New York: 1993.

33. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. Statistics in Medicine. 2007; 26:2389–2430. [PubMed: 17031868]

34. Therneau, TM., Grambsch, PM. Modeling Survival Data: Extending the Cox Model. New York: Springer; 2000.

35. Agresti, A. Categorical Data Analysis. 3. Wiley and Sons; 2012.

36. Kay R. AMarkov model for analysing cancer markers and disease states in survival studies. Biometrics. 1986; 42(4):855–865. [PubMed: 2434150]

37. Jackson CH, Sharples LD, Thompson SG, Duffy SW, Couto E. Multistate Markov models for disease progression with classification error. Journal of the Royal Statistical Society: Series D (The Statistician). 2003; 52(2):193–209.

38. Sweeting M, Farewell V, De Angelis D. Multi-state Markov models for disease progression in the presence of informative examination times: An application to hepatitis C. Statistics in Medicine. 2010; 29(11):1161–1174. [PubMed: 20437454]

39. Lange JM, Hubbard RA, Inoue LY, Minin VN. A joint model for multistate disease processes and random informative observation times, with applications to electronic medical records data. Biometrics. 2015; 71(1):90–101. [PubMed: 25319319]

40. Minin VN, Suchard MA. Counting labeled transitions in continuous-time Markov models of evolution. Journal of Mathematical Biology. 2008; 56(3):391–412. [PubMed: 17874105]

41. Longini IM, Clark WS, Byers RH, Ward JW, Darrow WW, Lemp GF, Hethcote HW. Statistical analysis of the stages of HIV infection using a Markov model. Statistics in Medicine. 1989; 8(7):831–843. [PubMed: 2772443]

42. Gentleman R, Lawless J, Lindsey J, Yan P. Multi-state Markov models for analysing incomplete disease history data with illustrations for HIV disease. Statistics in Medicine. 1994; 13(8):805–821. [PubMed: 7914028]

43. Goshu AT, Dessie ZG. Modelling progression of HIV/AIDS disease stages using semi-Markov processes. Journal of Data Science. 2013; 11(2):269–280. [PubMed: 26279666]

44. Nosyk B, Min J, Lima VD, Yip B, Hogg RS, Montaner JS. HIV-1 disease progression during highly active antiretroviral therapy: An application using population-level data in British Columbia: 1996–2011. Journal of Acquired Immune Deficiency Syndromes. 2013; 63(5):653. [PubMed: 24135777]

45. Satten GA, Longini IM Jr. Markov chains with measurement error: Estimating the 'true' course of a marker of the progression of human immunodeficiency virus disease. Journal of the Royal Statistical Society: Series C (Applied Statistics). 1996; 45(3):275–309.

46. Guihenneuc-Jouyaux C, Richardson S, Longini IM. Modeling markers of disease progression by a hidden Markov process: Application to characterizing CD4 cell decline. Biometrics. 2000; 56(3): 733–741. [PubMed: 10985209]

47. Foulkes A, DeGruttola V. Characterizing the progression of viral mutations over time. Journal of the American Statistical Association. 2003; 98:859–867.

48. Healy B, DeGruttola V. Hidden Markov models for settings with interval-censored transition times and uncertain time origin: Application to HIV genetic analyses. Biostatistics. 2007; 8(2):438. [PubMed: 16940036]

49. Nucita, A., Bernava, GM., Giglio, P., Peroni, M., Bartolo, M., Orlando, S., Marazzi, MC., Palombi, L. A Markov chain based model to predict HIV/AIDS epidemiological trends. International Conference on Model and Data Engineering; Springer; p. 225-236.

50. Lee S, Ko J, Tan X, Patel I, Balkrishnan R, Chang J. Markov chain modelling analysis of HIV/AIDS progression: A race-based forecast in the United States. Indian Journal of Pharmaceutical Sciences. 2014; 76(2):107. [PubMed: 24843183]

51. Yehia BR, Stephens-Shields AJ, Fleishman JA, Berry SA, Agwu AL, Metlay JP, Moore RD, Mathews WC, Nijhawan A, Rutstein R, et al. The HIV care continuum: Changes over time in retention in care and viral suppression. PLoS ONE. 2015; 10(6):e0129376. [PubMed: 26086089]

52. Li YP, Chan W. Analysis of longitudinal multinomial outcome data. Biometrical Journal. 2006; 48(2):319–326. [PubMed: 16708781]

53. Wood SN. Thin plate regression splines. Journal of the Royal Statistical Society, Series B. 2003; 65(1):95–114.

54. Liang K, Zeger S. Longitudinal Data Analysis Using Generalized Linear Models. Biometrika. 1986; 73(1):13–22.

55. Touloumis A, Agresti A, Kateri M. GEE for multinomial responses using a local odds ratios parameterization. Biometrics. 2013; 69(3):633–640. [PubMed: 23724948]

56. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics. 1986:121–130. [PubMed: 3719049]

57. Geng EH, Odeny TA, Lyamuya R, Nakiwogga-Muwanga A, Diero L, Bwana M, Braitstein P, Somi G, Kambugu A, Bukusi E, et al. Retention in care and patient-reported reasons for undocumented transfer or stopping care among HIV-infected patients on antiretroviral therapy in Eastern Africa: Application of a sampling-based approach. Clinical Infectious Diseases. 2015; 62(7):935–944. [PubMed: 26679625]

58. Geng EH, Odeny TA, Lyamuya RE, Nakiwogga-Muwanga A, Diero L, Bwana M, Muyindike W, Braitstein P, Somi GR, Kambugu A, et al. Estimation of mortality among HIV-infected people on antiretroviral treatment in east Africa: A sampling based approach in an observational, multisite, cohort study. The Lancet HIV. 2015; 2(3):e107–e116. [PubMed: 26424542]

59. Bakoyannis G, Yiannoutsos CT. Impact of and correction for outcome misclassification in cumulative incidence estimation. PloS one. 2015; 10(9):e0137454. [PubMed: 26331616]

7. Appendix

Appendix contains supplementary figures for Section 4 of the main text.

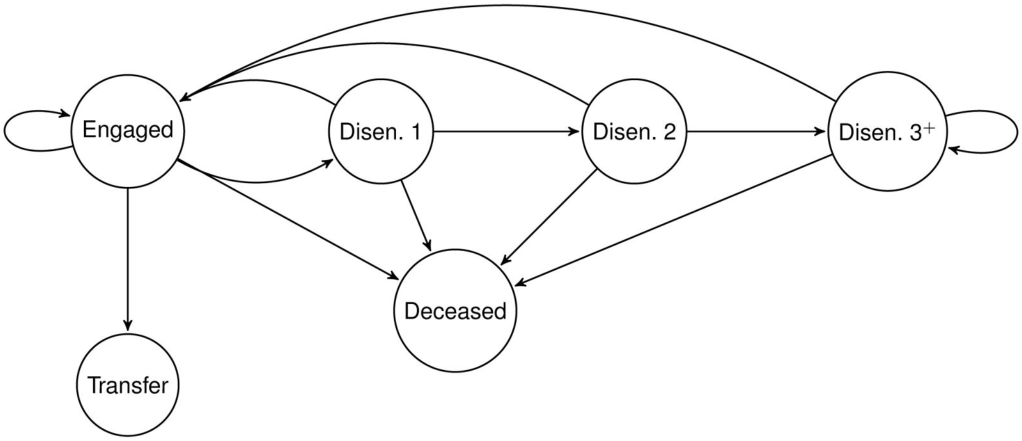

Figure 1.

The HIV care cascade using AMRS reflecting retention aspect of the cascade in care of AMPATH. At each state, patients are in the following states: `Engaged`: engaged in care, `Disen. 1`: disengaged from care for one interval (i.e., disengaged for a short-term), `Disen. 2`: disengaged from care for two consecutive intervals (i.e., disengaged for a moderate-term), `Disen. 3`$^+$: disengaged from care for more than two consecutive intervals (i.e., disengaged for a long-term), `Xfer`: transferred-out, and `Death`: deceased

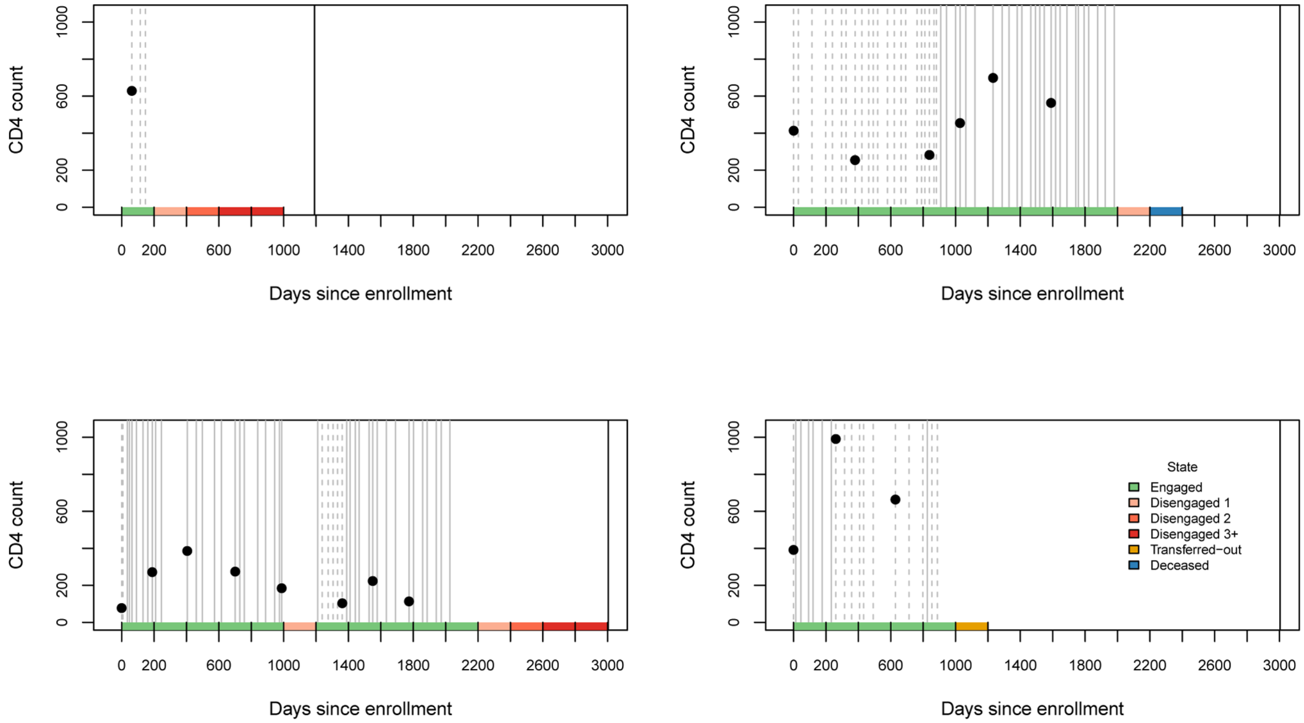

Figure 2.

Example of state ascertainment using data from four individuals in the AMPATH dataset. Grey vertical lines represent patient visits (solid if on treatment, dashed if not). The solid black vertical line is database closure date. CD4 counts are plotted as a function of time (black dot). Color-coded state membership depicted along the bottom of the graph.

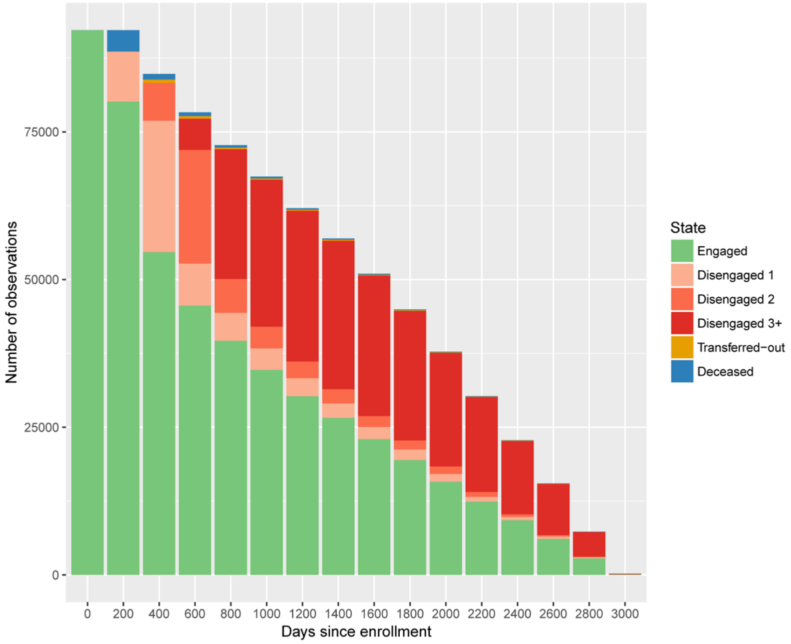

Figure 3.

Number of subjects in each state over time. At baseline, 92,215 unique individual records are available from AMPATH data. Six states are considered: engaged in care, disengaged from care for one interval (`Disengaged 1`), 3 = disengaged for two consecutive intervals (`Disengaged 2`), 4 = disengaged for more than two consecutive intervals (`Disengaged 3`$^+$), and 5 = transferred-out, 6=deceased. By definition, all individuals are engaged in care at baseline (day 0).

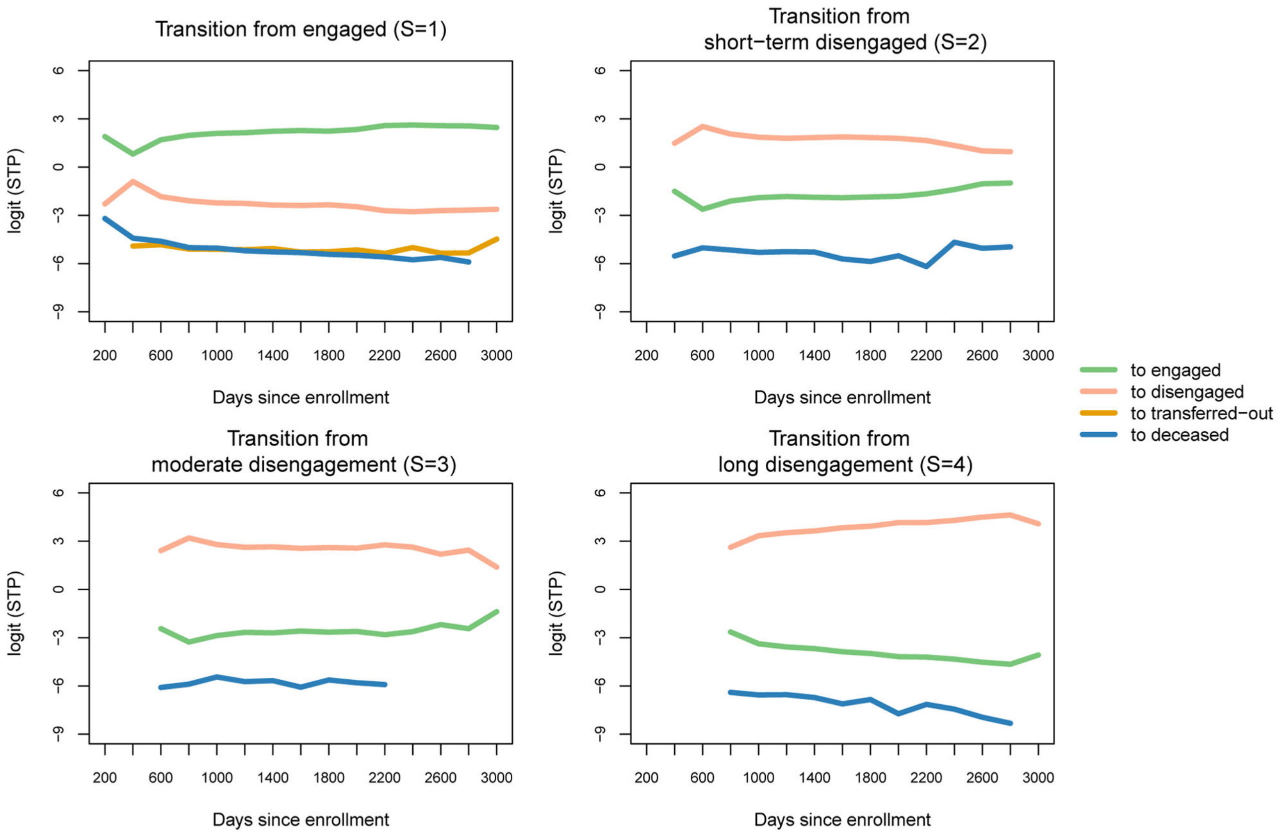

Figure 4.

Temporal trends in (unadjusted) state transition probabilities (STP) calculated at every 200 days. STP are presented in the logit scale to make temporal trends and fluctuations more apparent. All patients started from engagement in care at day 0, and thus no transitions from state 2 to other states were made at day 0. In each title, S=s (s=1,2,3,4) represents prior state membership. Transition to transfer-out was not allowed from state 2,3, and 4.

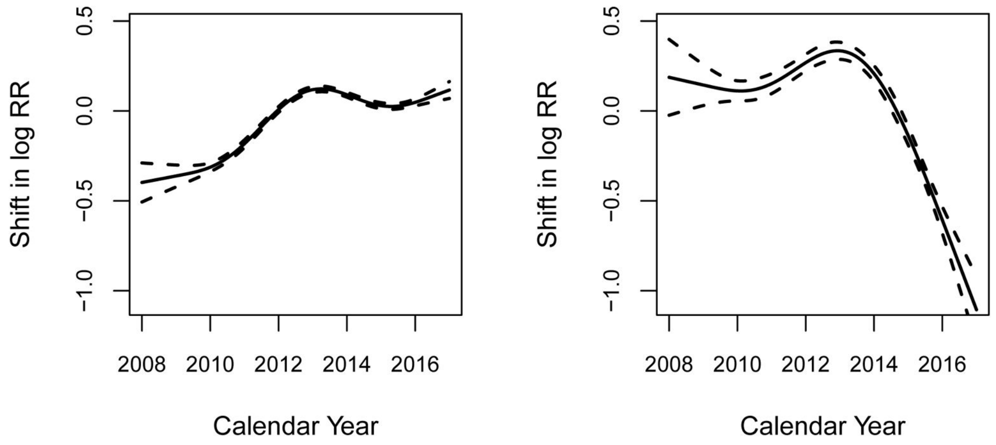

Figure 5.

Calendar year effects (i.e., period effects) on transition from engaged to disengaged (left panel) and transition from engaged to death (right panel). Splines were used to estimate the smoothed effect of calendar years.

Table 1

Baseline (i.e., at enrollment) characteristics of 92,215 HIV infected adults who enrolled in AMPATH between June 2008 and August 2016.

Variable	Number (%)	Median (IQR)
Male	30,900 (33.5)	
Age		35 (28, 43)
CD4[*]		247 (101, 439)
Taking ART	11,791 (12.8%)	
Year of enrollment		
2008	8,475 (9.2)	
2009	16,016 (17.4)	
2010	14,922 (16.2)	
2011	13,988 (15.2)	
2012	11,235 (12.2)	
2013	9,177 (10.0)	
2014	9,411 (10.2)	
2015	8,244 (8.9)	
2016	747 (0.8)	

[*] CD4 was measured and available for 31,535 (34%) patients at the time of enrollment to AMPATH care.

Table 2

Annotated transition matrix $\mathbf{P}_j$ corresponding to state definitions given in equation (4).

	State at t_j					
State at t_{j-1}	**Engaged**	**Diseng 1**	**Diseng 2**	**Diseng 3$^+$**	**Transfer-out**	**Death**
Engaged	p_{j11}	p_{j12}	0	0	p_{j15}	p_{j16}
Diseng 1	p_{j21}	0	p_{j23}	0	0	p_{j26}
Diseng 2	p_{j31}	0	0	p_{j34}	0	p_{j36}
Diseng 3$^+$	p_{j41}	0	0	p_{j44}	0	p_{j46}
Transfer-out	0	0	0	0	1	0
Death	0	0	0	0	0	1

Table 3

Mean state probabilities and state transition rates over follow-up period in the transition matrix format. `Diseng. 1, Diseng. 2`, and `Diseng. 3`$^+$ represents disengaged for a short-term, disengaged for a moderate-term, and disengaged for a long-term state, respectively.

	State at t_j					
State at t_{j-1}	**Engaged**	**Diseng. 1**	**Diseng. 2**	**Diseng. 3$^+$**	**Transfer-out**	**Death**
Engaged	.85	.13	0	0	.01	.01
Diseng. 1	.11	0	.88	0	0	.01
Diseng. 2	.05	0	0	.94	0	<.01
Diseng. 3$^+$	.02	0	0	.98	0	<.01
Transfer-out	0	0	0	0	1	0
Death	0	0	0	0	0	1

Table 4

Relative risk ratios (RRR) and 95% bootstrapped confidence intervals for effect of covariates on transitions from engaged in care ($S_{j-1} = 1$) to disengaged ($S_j = 2$), transfer-out ($S_j = 5$) or death ($S_j = 6$), relative to remaining engaged in care ($S_j = 1$); i.e., remaining engaged in care is the reference state.

State at t_{j-1}		**Engaged**	
State at t_j	**Disengaged**	**Transfer**	**Death**
Age ≥ 35	.64 (.63, .65)	.57 (.52, .62)	1.14 (1.08, 1.20)
Male	1.09 (1.07, 1.12)	.87 (.79, .95)	1.72 (1.63, 1.81)
CD4 < 350, ARV$^-$		Reference	
CD4 < 350, ARV$^+$	.16 (.16, .17)	.29 (.24, .34)	.47 (.44, .51)
CD4 ≥ 350, ARV$^-$	.31 (.30, .32)	.24 (.20, .30)	.11 (.10, .12)
CD4 ≥ 350, ARV$^+$	.12 (.11, .12)	.19 (.16, .23)	.13 (.12, .15)
No CD4, ARV$^-$	1.70 (1.62, 1.77)	1.18 (.90, 1.54)	.90 (.81, 1.00)
No CD4, ARV$^+$	.44 (.42, .46)	.55 (.45, .68)	.53 (.48, .59)

Table 5

Relative risk ratios (RRR) and 95% bootstrapped confidence intervals for effect of covariates on transitions from disengaged states to continued disengagement or death, relative to remaining engaged in care.

State at t_{j-1}	Disengaged 1		Disengaged 2		Disengaged 3^+	
State at t	**Disengaged 2**	**Death**	**Disengaged 3^+**	**Death**	**Disengaged 3^+**	**Death**
Age≥35	.88 (.83, .93)	1.32 (1.01, 1.72)	.95 (.87, 1.04)	1.42 (.95, 2.12)	.96 (.90, 1.03)	1.06 (.75, 1.49)
Male	.93 (.88, .99)	1.30 (1.02, 1.67)	.91 (.83, 1.01)	.96 (.66, 1.41)	.82 (.77, .88)	1.28 (.93, 1.77)
CD4<350, ARV$^-$	Reference		Reference		Reference	
CD4<350, ARV$^+$	.27 (.24, .30)	.71 (.49, 1.03)	.31 (.27, .37)	.77 (.44, 1.34)	.24 (.21, .27)	.50 (.33, .75)
CD4≥ 350, ARV$^-$	.54 (.47, .60)	.31 (.18, .52)	.74 (.62, .89)	.27 (.12, .63)	.69 (.60, .79)	.40 (.23, .70)
CD4≥ 350, ARV$^+$	.26 (.23, .29)	.29 (.17, .48)	.30 (.25, .36)	.14 (.06, .35)	.16 (.14, .19)	.10 (.05, .21)
No CD4, ARV$^-$	2.06 (1.74, 2.45)	1.45 (.83, 2.53)	2.05 (1.57, 2.66)	1.40 (.59, 3.35)	1.58 (1.30, 1.93)	.93 (.48, 1.80)
No CD4, ARV$^+$	.40 (.35, .45)	.42 (.24, .71)	.34 (.29, .41)	.42 (.19, .93)	.15 (.13, .17)	.16 (.08, .31)