Predicting individual-specific HIV survival functions - motivation, implementation, and potential applications.

Patrick Bullard

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2020

Committee:

Jared Baeten

Elizabeth Brown

Program Authorized to Offer Degree:

Epidemiology

University of Washington

Abstract

Predicting individual-specific HIV survival functions - motivation, implementation, and potential applications.

Patrick Bullard

Chair of the Supervisory Committee:

Jared Baeten MD PhD

Department of Epidemiology

Immense progress has been made in the development and delivery of HIV prevention interventions worldwide, and that is reflected in the trends of global HIV epidemic metrics. However, such a wide perspective overlooks relapses in several regions and numerous concentrated epidemics among certain subpopulations. Risk of HIV varies substantially across geographic, temporal, and social space. That, coupled with looming budget shortfalls, has motivated calls to optimize the delivery of HIV prevention interventions. One strategy is to provide interventions to those most at risk. Typically performed through community-level geographic prioritization or social stratification within an area, advances in data generation and analysis methods have enabled individual-specific risk assessments with rich interpretations. Among these are models that predict a patient-specific survival function. Here we fit and evaluate three such individual survival models, Accelerated Failure Time Weibull Regression (AFT_{weib}), Cox Proportional Hazards with the Kalbfleisch and Prentice estimator and Elastic Net regularization (Cox_{KPEN}), and Random Survival Forests (RSF) on data from the dapivirine ring ASPIRE trial. Evaluated on concordance, l1-loss, integrated Brier score, and visually, the fitted models demonstrate a moderate ability to rank survival times correctly but are poor predictors of individual survival time. Highly predictive features mirror those found previously. The visual nature of the prediction may enhance strategic risk communication. Future patient-specific HIV predictive models may perform extremely well. This would present risks. Data acquisition needs may be intensive and repetitive, with implications for feasibility and opportunity cost. Privacy, equity, and righteousness should be proactively and multidimensionally safeguarded.

Introduction

The suite of healthcare interventions on hand to check the global human immunodeficiency viruses (HIV) epidemic has expanded from frustratingly few 40 years ago, into the diverse and growing collection of effective measures deployed today[1,2]. Most notably, the dramatic uptake of simplified antiretroviral therapy (ART) regimens from an estimated 24% of candidates on treatment globally in 2010 to 62% in 2018 has been attributed with averting an estimated 1.3 million deaths per year in 2018 alone. What's more, beyond merely treating the disease, the non-infectivity accompanying successful ART has enabled the extremely productive treatment-as-prevention (TasP) modality[1]. Similarly, antiretrovirals used as pre-exposure prophylaxis (PrEP), targeted at individuals without HIV but at elevated exposure risk, offer another highly effective prevention option. The first PrEP agent to demonstrate efficacy was oral emtricitabine/tenofovir (TDF/FTC); with an estimated HIV prevention effect of >90%, adjusted for adherence[3]. TDF/FTC confers individual protection as well as population level benefits when deployed in conjunction with other HIV prevention methods[4,5,6]. Finally, novel PrEP formulations, including HIV microbicides delivered directly to the vaginal mucosa, exhibited partial efficacy in initial clinical trials and are currently under further evaluation[7,8].

Concomitant with the rollout of these interventions has been a global downtrend in HIV-associated mortality, from 1.2 million per year in 2010 to 770,000 per year in 2018[9]. Over the same time frame, the global incidence-to-prevalence ratio (IPR), a metric designed to capture the trajectory of the epidemic, has crept lower every year, from 6.6% to 4.6%, coming closer to the target ratio of 3%, the point at which, if sustained, the epidemic is likely to eventually subside[10,11]. However, these trends are tenuous and contingent on continued investment in the health service, social, and political aspects necessary to realize successful HIV care cascades: case ascertainment to treatment delivery to sustained viral suppression. However, the combined United States global spending on HIV (PEPFAR) has plateaued since 2010 at roughly $6.5 billion and potentially faces a sharp decline in 2020. The current administration is requesting only $4.9 billion for the fiscal year 2021[12], contributing to the flat funding of the crucial multilateral Global Fund since 2018. This is a concerning shortfall that accentuates, even more, the need to allocate resources efficiently. As summed up by UNAIDS in their refrain: "know your epidemic, know your response," the optimal HIV response depends on the specific characteristics of the epidemic in a given setting. Originally referring to an understanding of community-level features, this mantra should also be applied to individual-level risk. Ultimately, prevention resources should be invested in individuals at the highest risk of infection and risk in the HIV epidemic is fundamentally heterogeneous.

Parsing population variability in HIV infection might begin by looking across three broad aspects: person,

place, and time. Each of these dimensions can be described at various levels. The broadest aspect of heterogeneity is perhaps that which occurs across global geography. Most noticeable is the concentration of cases in East and South Africa. Despite progress in many areas, this region continues to be the hardest hit by HIV. Between 2010 and 2018, nearly one-half of the 17 million total incident infections occurred in East and South Africa[13]. And though it too is creeping lower, the incidence rate of HIV infection in this region remains over seven times the global average and nearly 26 times that of North America. Geographic heterogeneity persists to the national and sub-national levels as well. Finally, as high-resolution spatial data is making increasingly apparent, variability in predictors of HIV risk exist down to the most granular of spatial scales: from village to village and neighborhood to neighborhood[14].

Though to an extent inextricably linked with 'physical space', variability in risk exists across 'social space' as well. An HIV epidemic within a single area is often heuristically dichotomized as being general or concentrated. This designation reflects broad differences in the underlying transmission system. The epidemic is considered general when prevalence exceeds 1% within the large main-class social network or *general population* and is thus likely self-sustaining ($R_0 \geq 1$) through heterosexual contact alone. In contrast, concentrated epidemics exist where HIV prevalence in one or more defined sub-populations is 5% or greater. In such cases, these high prevalence sub-populations are termed key populations (KPs). The HIV transmission modes and comorbidities that are more relevant to KPs confer greater probability of infection per exposure (and often too, the *rates* of exposures are also higher) and such biological disparities are compounded, often acutely, by certain social determinants that manifest at all levels of social organization. The resulting crude relative risks for commonly recognized KPs - men who have sex with men, people who inject drugs, those engaged in transactional sex, and transgender people, are extreme. The estimated relative risks are 22, 22, 21, and 12, respectively[9,13]. If unaddressed, HIV prevalance can spiral out of control in KPs [ref]. Where both the conditions for a concentrated and general epidemic exist within the same area, the epidemic is said to be mixed. Characterizing HIV prevalence by social strata, or socio-medical network, reveals heterogeneity in risk across social space.

If there is a bright spot for priority populations, it may be the essentially uniform epistemic agreement of their outsized, unique needs and the steps being taken to meet them[15,16]. But what of general epidemics, how is infection risk distributed in this context? Here too it is variable and one broad aspect where that's seen in East and South Africa, where the epidemic is primarily generalized, is across gender. Women in a heterosexual and serodiscordant coupling are at roughly twice the infection risk per exposure, as would be a man. Coincident sexually transmitted infections (syphilis, chlamydia, gonorrhea, HPV, and herpes) push the risk to women higher still; these infections also increase the risk for men but, in most cases, to

a lesser degree. Additional unfortunate social and economic forces in this region put women, and young women in particular, at still greater risk of HIV infection. Adult women in this region experience roughly 1.5 times the number of new cases, as do adult men. For young women (15-24), that ratio is nearly 2.5[13]. The overwhelming majority of new HIV infections in women in East and South Africa occur during heterosexual contact[17,18]. Additional customized methods to counter this specific threat are necessary.

In 2018, roughly 75,000 people took PrEP at least once in East and South Africa, many as participants in implementation studies[19,20,21]. Though programs enrolled those at high risk, targeting female sex workers[22], serodiscordant couples[23], and young people[24] in high prevalence regions, hesitancy to initiate treatment was substantial[25]. Results consistently demonstrated the role of perceived personal risk in making the decision to initiate PrEP (in addition to concerns surrounding stigma, "pill-burden", and efficacy[26])[27,25], prompting calls for enhanced risk communication during clinic visits[20]. Randomized trials have demonstrated efficacy of individualized risk feedback in prompting health seeking behavior[28],[29,30]. Moreover, identifying individuals at the greatest risk for acquiring HIV would permit more efficient and cost-efficient PrEP services, in an era of cost containment. We hypothesized that *patient-specific survival models*, developed using advanced epidemiologic methods, could identify highest-risk individuals and provide intuitive visual summaries of their risk, allowing useful quantitative interpretations, and capturing heterogeneity of infection risk by person, place, and time. We evaluated three variants of so-called individual survival distribution (ISD) models using data from a large HIV incidence cohort of women in East and South Africa.

Methods

We fit three independent models on the ASPIRE data set that return patient specific survival curves: Weibull accelerated failure time, Cox regression with Kalbfleisch and Prentice baseline hazard and elastic net regularization, and random survival forests. Each model is evaluated on three quantitative performance metrics, concordance, L1 loss, and integrated Brier score.

The ASPIRE Study

The dapivirine vaginal ring was evaluated recently via a Phase III randomized double-blind placebo-controlled trial, ASPIRE[31] (and by a concurrent sister trial, The Ring Study[7]). ASIPRE enrolled 2629 women at 15 sites in Uganda, Malawi, Zimbabwe, and South Africa between August 2012 and June 2015. Participants were randomized to receive a placebo or 25 mg dapivirine laced vaginal ring at a 1:1 ratio, stratified by site. The trial protocol was reviewed and accepted by local ethics committees at each study site and each

participant gave written informed consent. The primary endpoint was incident HIV infection which was 27% lower in the dapivirine arm compared to placebo (95CI 1% to 46%) overall, with high-adherent subgroups seeing more benefit. Results of follow-up studies have similarly indicated safety and efficacy[32].

During ASPIRE follow-up, several case-report forms (CRFs) were filled out for each participant. The Baseline Behavior Assessment CRF includes variables that relate to sexual history (number of partners, sexual behavior) as well as yes/no responses regarding commonly held concerns with vaginal ring use (for example it becoming stuck or feeling uncomfortable). The Baseline Family Planning CRF captures information relating to a participant's contraceptive methods and number of prior pregnancies. The Demographic CRF contains information such as the participant's age, marital status, and ethnicity. The Demographic CRF also includes self-reported levels of alcohol and cigarette use. The Enrollment CRF includes administrative information relating to a participant entering the study including their randomization number, the date and time the prescription was assigned, and whether or not the ring (containing either dapivirine or placebo) was inserted. The Screening STI Test CRF reports the results of STI screening tests including those for *Trichomonas vaginalis*, *Neisseria gonorrhoeae*, and *Chlamydia trachomatis*. The last two CRFs are the MLR (Monthly Lab Report) and HCR (HIV Confirmatory Result) CRFs. The MLR CRF includes the results of several clinical tests performed monthly on trial participants including a pregnancy test as well as two rapid HIV tests. If either of the two HIV rapid tests were positive, confirmatory testing was performed (western blot and PCR), the results of which appear in the HCR CRF. All told, there are over 100 baseline variables and repeated observations on nearly 50 more.

Overview of Models

Both individually and in concert, these variables describe each patient in remarkable detail. Though the driving motivation to conduct ASPIRE was to rigorously evaluate the dapivirine vaginal ring via random treatment assignment, for this analysis we pooled both arms to create, effectively, an observational cohort study. We used the baseline covariates and the outcome HIV-free survival time to train three distinct survival models: Weibull accelerated failure time (AFT_{weib}), Cox regression with Kalbfleisch and Prentice baseline hazard and elastic net regularization (Cox_{KPEN}), and random survival forests (RSF). Each of these models is capable of returning a predicted survival function conditional on baseline covariate values.

General Survival Analysis

Survival analysis in general encompasses methods which attempt to describe, understand, and/or model, systems with time-to-event type responses. The ultimate objective of survival analysis is to characterize the

population survival time distribution. Though this distribution can be specified several ways[33], here we are interested in the survival function, which returns the probability of survival at a corresponding time. Figure 1 displays an example of a population survival function where survival times are drawn from a Weibull ($\lambda = 5, k = 1$) distribution.

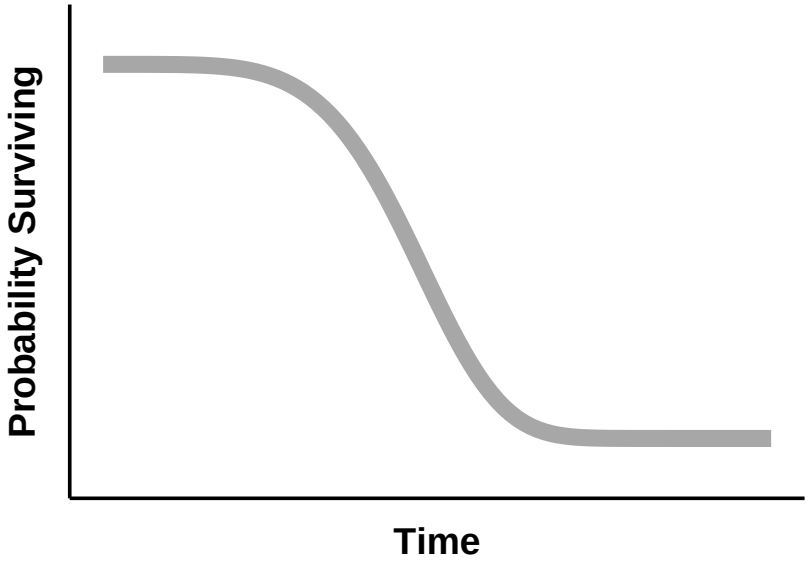

Figure 1. Parametric survival function; Weibull (..=5, k=1)

In nearly all biomedical contexts, attempts to directly construct the empiric analog of the population survival function, the sample survival function, runs into a problem. The response (in the case here, ASPIRE, the response is time until HIV infection) is not fully observed in all subjects. In fact, the response is only fully observed for the roughly 8% of participants in ASPIRE who became HIV infected over the course of the study. For the remaining 92%, the subjects who remained infection free, we know only their time-until-infection is *at least* as long as their study observation time. Such incomplete observations are collectively known as censoring. When the mechanism is subject-exit or study-termination before the event of interest, it's termed right-censoring. Figure 2 is a random sample of 100 ASPIRE participants illustrating the observational structure of the study including censoring.

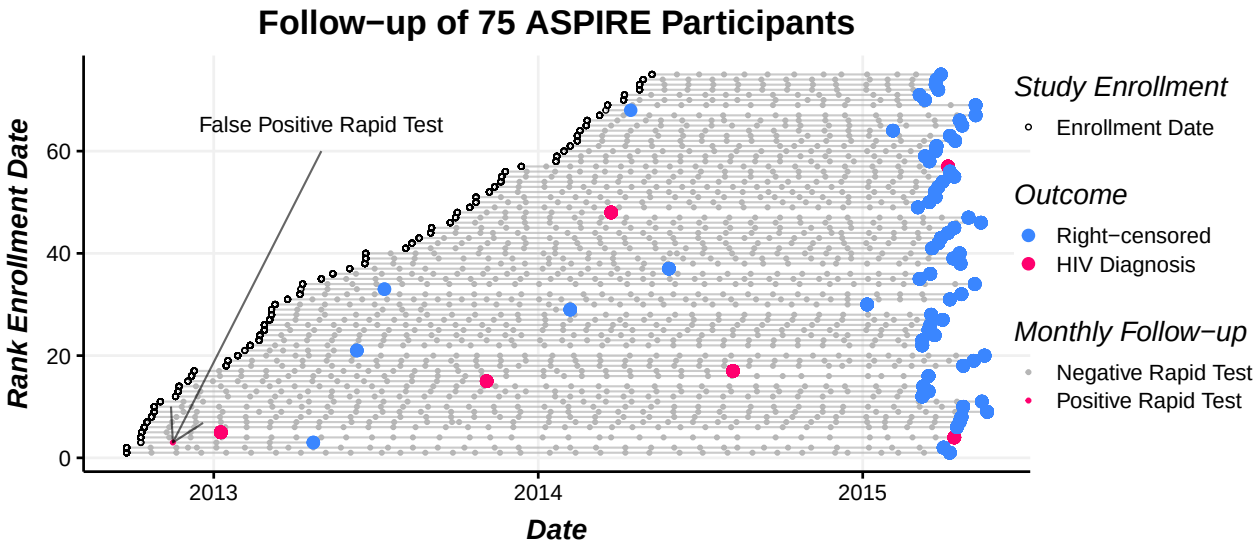

Figure 2. The enrollment and observational structure of the ASPIRE study

Assuming non-informative censoring (independence between censoring mechanism and event risk), various methods can provide unbiased estimates of the population survival time distribution. A common choice to estimate the survival function is the non-parametric Kaplan-Meier method. In the case of randomized controlled trials (clinical trials) the task is characterizing two survival time distributions: T|Drug and T|Placebo which are then compared for equivalence (generally via the log-rank test) to evaluate drug efficacy. To look deeper into any heterogeneity, further stratification can be performed, resulting in multiple estimated survival functions. Three or more estimated survival functions may be tested for equality via the Wilcoxon test. Figure 3 displays ASPIRE's Kaplan-Meier estimates following stratification on: nothing (overall survival), treatment arm assignment, and treatment arm assignment as well as study site country.

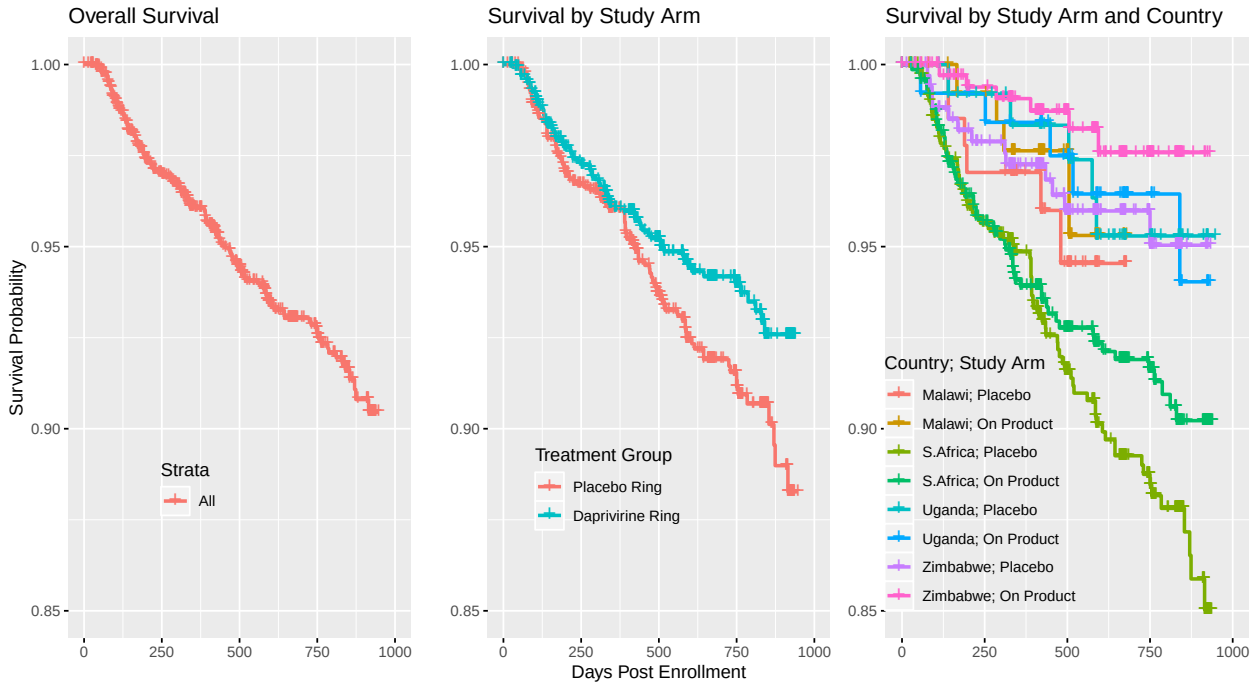

Figure 3. Kaplan−Meier survival estimates with increasing levels of stratification

Kaplan-Meier analysis is most useful for a smaller number of qualitative predictors. Further incorporation of an arbitrary number of continuous and qualitative predictors of survival time leads to a regression framework. Most commonly used in biomedical applications is the proportional hazard family of models where covariate values have a multiplicative effect on some shared baseline hazard function.

Weibull Accelerated Failure Time

AFT_{weib}

A fully parametric member of the proportional hazards model family is the Weibull accelerated failure time (AFT) model. AFT models can be conceptualized as modeling not the various survival times but the *speed* at which a subject approaches a *static* shared failure time with that speed being a function of covariates. Here we fit the AFT model specifying a Weibull distribution of survival times; AFT_{weib}. Weibull distributions typically take 2 parameters for survival analysis. The shape (λ) and scale (k) parameters $[W(\lambda, k)]$. AFT_{weib} assumes a survival time distribution of the form $T \sim W(exp(\beta X), \frac{1}{\sigma})$[6]. Using maximum likelihood to estimate betas and the sigma parameter allows this specification of the estimated survival function:

$$\hat{S}(t|X = x) = exp[-(\frac{t}{exp(\beta x)})^{1/\sigma}] \tag{1}$$

Cox Regression with estimated baseline hazard

Cox_{KPEN}

A common semi-parametric member of the proportional hazards family is the Cox model. An attractive feature of the Cox model is the common baseline hazard function does not need to be specified to fit or interpret beta values. Beta values fit with Cox's partial likelihood procedure and do not depend on the baseline hazard, which drops out of the calculation. Though that model form does have useful interpretations (for example, the relative magnitudes of Hazard Ratios delineate the expected order of failure), the fact the baseline hazard remains unknown precludes an estimation of the survival function. Several procedures have been proposed to specify the baseline hazard and thus allow an estimation of the survival function. Here we use the Kalbfleisch and Prentice method which itself is non parametric (it is analogous to the Kaplan-Meier method) and thus CoxKP remains semi-parametric. Another common estimator of the baseline hazard is the Breslow estimator[34]. Additionally, employed in this model is the elastic net extension to Cox regression. Most predictive models experience a trade-off between bias and variance. A common technique to navigate this trade-off is to include in the loss function (alongside the sum of squared prediction errors, in the case of linear regression), a term that captures the overall value of parameter coefficients. In this way, coefficient values are shrunk, which, to a point, reduces variance in excess of its cost in increased bias[35]. The weight given these additional loss terms is determined by a hyperparameter that is generally optimized (or "learned") via cross-validation. The sum of squared coefficient values and the sum of absolute coefficient values are incorporated in the loss functions of Lasso and Ridge regression respectively. The elastic net is the linear combination of Lasso and Ridge penalties, relatively weighted by a learned hyperparameter[36].

The Cox proportional hazards model:

$$h(t) = h_0(t)e^{(\beta^T X_{n*p})} \tag{2}$$

Relating baseline hazard to the survival function[37]:

$$S(t|X_i) = e^{-exp(\beta_{vec}X_{n_i*p})h_0()} = S_0(t)^{exp(\beta_{vec}X_{n_i*p})} \tag{3}$$

The KP method to estimate baseline survival $S_0(t)$ takes a form very similar to the Kaplan-Meier estimator (which is also known as the product limit estimator):

$$S_{0_{KP}}(t) = \prod_{i:t_i \leq t} \alpha_i^{\delta_i} \tag{4}$$

Where alpha is the baseline (covariates = 0) probability of continued survival at time = t and delta is censoring status at time = 1 (0=censored, 1 =uncensored). A simplifying assumption of no survival time ties (which is made here), results in this estimate of alpha:

$$\alpha_i = [1 - \frac{exp(\beta x_i)}{\sum_{t=t_i}^{t=min(t)} \beta x_t}]^{exp(-\beta x_i)} \tag{5}$$

Making the final KP estimator of the individual survival function:

$$\hat{S}(t|X = x) = \prod_{i:t_i \leq t} \alpha_i^{\delta_i}{}^{\beta x} \tag{6}$$

Random Survival Forests

RSF

The third model, Random Survival Forests (RSF), is fully non-parametric and predicated on the observation that many "weak learners" can combine into a much stronger one. RSF models are a relatively new addition to the survival analysis toolkit. They build off classic decision trees wherein feature space is recursively split to both maximize between node heterogeneity and minimize within node heterogeneity with respect to the outcome of interest. Development of this method to handle censored data and the incorporation of bootstrapping techniques have led to the ensemble type model RSF. Many trees (a forest) are generated on the training set. For each tree, only a random subset of features is available to split on. Consequently, though predictions of each tree are expected to have high variability relative to the others, the randomness causes its bias to be quite low. Observations associated with each terminal node on each tree are fed to the Nelson-Aalen (NA) estimator (an alternative to Kaplan-Meier)[38] and thus each terminal node is associated with a survival curve. Test set subjects are classified according to each random tree and the resulting set of NA survival curves are averaged to generate the ensemble prediction. This averaging step reduces variability. For detailed calculations see[38].

Evaluation Metrics

Multiple methods for evaluating model performance have been proposed and it is common to calculate several. In some cases, they may disagree. In that situation, the interpretations of each metric and how that relates to

the intended use for the model under construction are useful considerations. Briefly, *concordance* results from exhaustive pairwise comparisons between comparable subjects assessing if the order of failure implied by the *predicted* level of risk corresponds to the *observed* failure order of this pair. Concordance is the proportion of comparisons where concordance exists, giving it the interpretation: the probability the model is correct when it queries the order of failure of two subjects. Uncensored observations are comparable to all other uncensored observations; censored observations are comparable to uncensored observations whose outcome occurred before censoring[39]. *l1-loss* or mean absolute error, is the average absolute difference between predicted and observed survival times. Predicted survival is defined here as the median of the individual survival function To handle censored subjects where survival time is unobserved, the l1-loss algorithm makes a best guess approximation of their remaining survival time[40,41] and compares that to the median survival time. Because the estimate for remaining survival time is expected to be worse for those censored earlier (sooner) than those censored later, each censored subject's contribution to the L1-margin loss is weighted by its follow-up time. Finally, the *Brier Score* returns the mean squared difference between the predicted probability of failure and the actual observed status at a particular time. An extension to handle a continuous time interval is the *integrated Brier score*. The Brier score encompasses three features: uncertainty, calibration, and resolution[42]. The combination of uncertainty and resolution are roughly equivalent to concordance (leading some references to discuss a 2-component decomposition of the Brier score instead of 3) and calibration indicates the extent the modeled probabilities match the observed frequency of the outcome (i.e. a perfectly calibrated model is one where among 100 subjects predicted to have a 90% probability of the outcome, 90 in fact, do experience the outcome). Modifications to the Brier Score to handle right-censoring are described by Graf 1999[43].

Implementation

The baseline variables captured on the CRFs outlined above were joined into a single data frame, re-factored if necessary, and assessed for missingness. Repetitive or extraneous variables were dropped. Variables with missingness over 10% were dropped. Variables with missingness less than 10% were imputed with mean (numeric) or mode (categorical) values. Factor variables were one hot encoded. Following the suggestion by Griener et. al. 2018, each variable was individually assessed for association with incident HIV via Cox regression and excluded if insignificant at $\alpha = 0.1$.

Data cleaning, manipulation, and model fitting and evaluation were performed in RStudio using R version 3.5.1 (R Core Team 2019).

The AFT_{weib}, Cox_{KPEN}, and RSF models described above were fit withR packages: Survival, FastCox,

randomForestSRC and through adapting several R functions published by Haider et. al 2018[40]. The evaluation metrics described above: concordance, l1-loss, and integrated Brier score were computed via 5-fold cross validation. Means and standard deviations across validation sets reported. Test-set predicted individual survival functions were obtained for every subject via 5-fold cross validation.

Results

The ASPIRE cohort consisted of 2,629 HIV uninfected women between the ages of 18 and 45, enrolled at study sites in Malawi, South Africa, Uganda, and Zimbabwe (Table 1). Their mean age at enrollment was 27 years (sd, 6.2 years) and nearly half (45.6%) had completed secondary school. Just over half (54.9%) earned their own income and were not married (59.2%). Almost all participants reported having a primary sex partner (99.5%), though 3% reported that that person had changed in the previous three months. Nearly all (98.1%) reported at least one instance of vaginal sex in the last three months though this count varied considerably (min, 0; max, 99; mean, 26.5; sd, 24.7). Slightly under half (42.6%) reported using a condom (male, female, or both) in their last vaginal sexual encounter. Just 2.1% of participants reported an instance of anal sex in the previous three months.

These participants were followed for a total of 4,280 person-years (mean, 598 person-days; standard deviation, 241 person-days). Twelve women were lost to any follow-up and excluded from the analysis. Three others were retrospectively classified as HIV-positive at their time of enrollment and excluded as well. Throughout follow-up, 168 incident HIV infections were observed (incidence rate, 3.9 per 100 person-years).

Following data pre-processing, 85 baseline variables were joined into a single data frame. Of these, 76 had zero missingness. Of the nine that were incompletely observed, six related to STI screening had missingness over 10% and were dropped. Two variables had missingness less than 10%. The first, BFPlbct (number of live births) was actually, in fact, completely observed, NA was recorded in this field for women with zero pregnancies. The second variable with missingness, SSTph, (vaginal pH) had only one missing value, which was mean imputed.

Twenty-eight baseline features were associated with HIV infection at $\alpha = 0.1$ and passed to the model fitting and evaluation functions (Table 2).

Predicted survival curves specific to each ASPIRE subject were generated via cross-validation. Two plots of predicted individual survival functions are reported for each model. On the left are test set predicted survival functions and points indicating observed outcome for the same random sample of 26 ASPIRE subjects. On the right are curves of two random samples of ASPIRE subjects, 30 who contracted HIV and 30 who were

right-censored (did not acquire HIV during follow-up) with loess curves summarizing each of these strata.

Accelerated Failure Time

The accelerated failure time model fitting did not incorporate feature selection, and beta coefficients were estimated for all of the 28 variables that were separately associated with HIV-infection (Table 4). The most strongly predictive variables were infection with *Trichomonas vaginalis* and *Neisseria gonorrhoeae*, knowledge of primary sex partner's circumcision status, and completed secondary school or attended University.

For a qualitative visual summary of the predictive form of the AFT_{weib} model output, predicted survival curves are displayed for a random sample of 26 ASPIRE participants (Figure 5 left). Their interpretation is the probability (y-axis) that subject remains HIV-uninfected at any selected time-point (x-axis). To compare the model prediction with actual observation, overlaid on each predicted survival curve is the observed outcome for that subject, when they were either diagnosed with HIV or right-censored.

The AFT_{weib} model is fully parametric, and this is apparent in the constrained functional form of the predicated curves. Furthermore, the proportional hazards assumption guarantees these curves will not cross each other, allowing for a stable rank ordering of predicted risk. The AFT_{weib} model is constrained to a hazard function that is either monotonically increasing or monotonically decreasing. Here that hazard is monotonically decreasing, most evident in the slightly steeper initial slopes in predicted high-risk subjects. Among the random sample of 26 ASPIRE participants, one became HIV-infected. The AFT_{weib} model predicted the steepest survival curve for this subject.

To more directly compare predictions made for subjects who did vs. did acquire HIV during follow-up, the right-hand panel of Figure 5 displays an outcome balanced sample of predicted survival curves along with loess summaries for each. The loess summary of predicted curves stratified by observed outcome (Figure 5 right) are close. However, the loess summary of the HIV-infected strata is steeper, qualitatively indicating this model will, more often than not, correctly rank two individuals in terms of HIV risk.

Three quantitative performance metrics were estimated for the AFT_{weib} model via 5-fold cross-validation (Table 3). First, the probability that the AFT_{weib} model is correct when it queries the order of failure of two subjects, concordance, was estimated at 69% (sd, 0.06). Second, the average absolute difference between the median of the survival curve and the observed failure time among those who experienced the outcome or a "best guess" for those who did not, l1-loss, is 7371 days (sd, 849 days). Third, the integrated Brier score, is a composite that encompasses uncertainty, calibration, and resolution (see methods for a more detailed interpretation), and was estimated to be 0.047 (sd, 0.002).

Table 1: Baseline characteristics of the ASPIRE cohort

Characteristic	Summary
Age - yr.	
mean	27.23
SD	6.18
Study site - n (%)	
Malawi: Blantyre	130 (4.94)
Malawi: Lilongwe	142 (5.4)
South Africa: Botha's Hill	180 (6.85)
South Africa: Emavundleni Cent	166 (6.31)
South Africa: eThekwini	244 (9.28)
South Africa: Isipingo	117 (4.45)
South Africa: RK Khan	150 (5.71)
South Africa: Tongaat	103 (3.92)
South Africa: Umkomaas	103 (3.92)
South Africa: Verulam	150 (5.71)
South Africa: WHRI	213 (8.1)
Uganda: MU-JHU	253 (9.62)
Zimbabwe: Seke South	224 (8.52)
Zimbabwe: Spilhaus	230 (8.75)
Zimbabwe: Zengeza	224 (8.52)
Level of education - n (%)	
No schooling	23 (0.87)
Primary school, not complete	239 (9.09)
Primary school, complete	142 (5.4)
Secondary school, not complete	1026 (39.03)
Secondary school, complete	1044 (39.71)
Attended college or university	155 (5.9)
Earns own income - n (%)	
Earns own income (No)	1442 (54.85)
Earns own income (Yes)	1187 (45.15)
Marital Status - n (%)	
Unmarried	1555 (59.15)
Married	1074 (40.85)
Partner's ring awareness - n (%)	
Partner aware of ring use (Yes)	1680 (63.9)
Partner aware of ring use (No)	932 (35.45)
Partner aware of ring use (Not Sure)	4 (0.15)
No Primary Sex Partner	13 (0.49)
Two or more sex partners - n (%)	
Two or more sex partners (No)	2190 (83.3)
Two or more sex partners (Yes)	439 (16.7)
Instances of vaginal sex in an average month	
Mean	26.52
SD	24.66
Condom used w/ last sex act - n (%)	
Condom used w/ last sex act (No)	1120 (42.6)
Condom used w/ last sex act (Yes)	1509 (57.4)
Anal sex in past 3 months - n (%)	
Anal sex in past 3 months (No)	2575 (97.95)
Anal sex in past 3 months (Yes)	54 (2.05)
Treatment assignment - n (%)	
Placebo Ring	1316 (50.06)
Dapivirine Ring	1313 (49.94)

Table 2: Baseline features passed to model fitting functions

Baseline Characteristic	Hazard Ratio	p-value
Age at enrollment is greater or equal to 25 years	0.4802	0.0000
Instances of vaginal sex in an average month	0.9788	0.0000
Condom not always used during vaginal sex	0.5949	0.0090
No vaginal sex in the past week	1.4811	0.0276
Instances of vaginal sex in the last week	0.8822	0.0018
Primary sex partner (PSP) is uncircumcised	0.4010	0.0024
PSP is circumcised	0.5425	0.0412
Subject does not know HIV status of PSP	0.7149	0.0385
At least one specific concern with ring use	1.3515	0.0566
PSP is aware of subject's study participation	0.6932	0.0265
PSP is aware of vaginal ring use	0.7308	0.0449
Number of prior pregnancies	0.6959	0.0000
Number of cigarettes smoked daily	1.0355	0.0531
Attended college or university	0.4968	0.0980
Completed secondary school	0.6963	0.0354
Self-employed	0.5649	0.1071
Participant traveled 1-2 hours to clinic	0.5931	0.0426
Participant earns their own income	0.7714	0.0996
Participant is married	0.2864	0.0000
Participant received dapivirine vaginal ring	0.7282	0.0423
Severity of preexisting conditions	1.0538	0.0467
Study site is in South Africa	1.8908	0.0540
Serum creatinine mg/dL	0.2997	0.0885
Eosinophils cells/mm3	1.0006	0.0002
Platelets X1000/mm3	1.0027	0.0061
Positive for chlamydia	2.0811	0.0001
Positive for gonorrhea	2.7876	0.0001
Positive for trichomoniasis	1.9428	0.0044

Table 3: Evaluation of Accelerated Failure Time

Metric	Mean	SD
Concordance	0.6932	0.0607
L1 Loss	7371.0000	849.0000
Integrated Brier	0.0472	0.0023

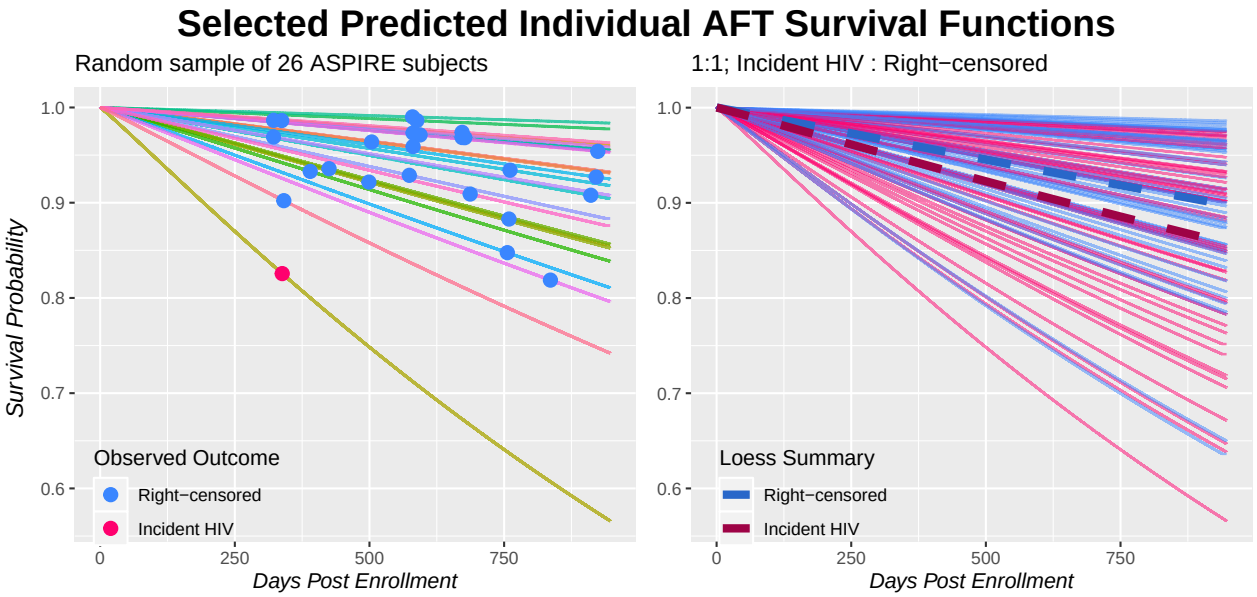

Figure 5.

Cox Regression with Kalbfleisch and Prentice baseline hazard and Elastic Net regularization

The Cox_{KPEN} model incorporates feature selection through the action of the Lasso component of Elastic Net regularization, which causes coefficient values of unselected variables to equal precisely zero. In this case, the Cox_{KPEN} model fit on the entire ASPIRE dataset with optimized hyperparameters only dropped one variable, "primary sex partner is circumcised," leaving 27 variables in the model (Table 6). Infections with *Trichomonas vaginalis*, *Chlamydia trachomatis*, and *Neisseria gonorrhoeae* were highly predictive as were married status and attending College or University.

Figure 6 summarizes Cox_{KPEN} predictions for the same sample of ASPIRE participants depicted in Figure 5 As with the AFT_{weib} model, Cox_{KPEN} makes the parametric proportional hazards assumption, again ensuring predicted survival curves do not cross. This, combined with the non-parametric estimation of

Table 4: Fitted AFT beta values arranged by absolute value

Baseline Characteristic	beta	exp.beta
Attended college or university	1.1520	3.1645
Positive for gonorrhea	-0.7331	0.4804
Completed secondary school	0.6311	1.8797
PSP is circumcised	0.6227	1.8640
Primary sex partner (PSP) is uncircumcised	0.5760	1.7789
Positive for trichomoniasis	-0.4929	0.6109
Serum creatinine mg/dL	-0.4681	0.6262
Participant is married	0.4173	1.5179
Participant received dapivirine vaginal ring	0.3893	1.4759
PSP is aware of subject's study participation	0.3284	1.3887
Participant traveled 1-2 hours to clinic	0.3141	1.3690
Positive for chlamydia	-0.2505	0.7784
Number of prior pregnancies	0.2226	1.2493
No vaginal sex in the past week	-0.1839	0.8320
Subject does not know HIV status of PSP	0.1604	1.1740
Condom not always used during vaginal sex	0.1555	1.1682
Self-employed	0.1261	1.1344
PSP is aware of vaginal ring use	-0.1158	0.8907
At least one specific concern with ring use	0.1051	1.1108
Age at enrollment is greater or equal to 25 years	0.0881	1.0921
Participant earns their own income	0.0741	1.0769
Number of cigarettes smoked daily	-0.0508	0.9505
Study site is in South Africa	-0.0280	0.9724
Severity of preexisting conditions	-0.0166	0.9835
Instances of vaginal sex in an average month	0.0092	1.0092
Platelets X1000/mm3	-0.0017	0.9983
Instances of vaginal sex in the last week	0.0008	1.0008
Eosinophils cells/mm3	-0.0005	0.9995

Table 5: Evaluation of Elastic Net CoxKP

Metric	Mean	SD
Concordance	0.6677	0.0656
L1 Loss	6136.0000	638.0000
Integrated Brier	0.0459	0.0011

baseline hazard, gives rise to the functional form of Cox_{KPEN} predictions, a common non-parametric survival function weighted by individual factors. Compared to the AFT_{weib} model, the Cox_{KPEN} predicted less divergence in risk among the same random sample of 26 ASPIRE participants as well as a moderately different rank ordering of risk among those subjects (Figure 6 left). Examining again the one subject who acquired HIV in this sample, the Cox_{KPEN} model predicts that they are at the third-highest risk, not the obviously high risk predicted by the AFT_{weib} model.

The loess summaries of predicted curves stratified by observed outcome (Figure 6 right) are quite close, but again qualitatively indicate this model will, more often than not, correctly rank two individuals in terms of HIV risk.

Three quantitative performance metrics were estimated for the Cox_{KPEN} via 5-fold cross-validation (Table 5). First, the probability the Cox_{KPEN} model is correct when it queries the order of failure of two subjects, concordance, is 66.8% (sd, 0.067). Second, the average absolute difference between the median of the survival curve and the observed failure time among those who experienced the outcome or a "best guess" for those who did not, l1-loss, is 6136 days (sd, 638 days). Third, the integrated Brier score encompassing uncertainty, calibration, and resolution is 0.046 (sd, 0.0011).

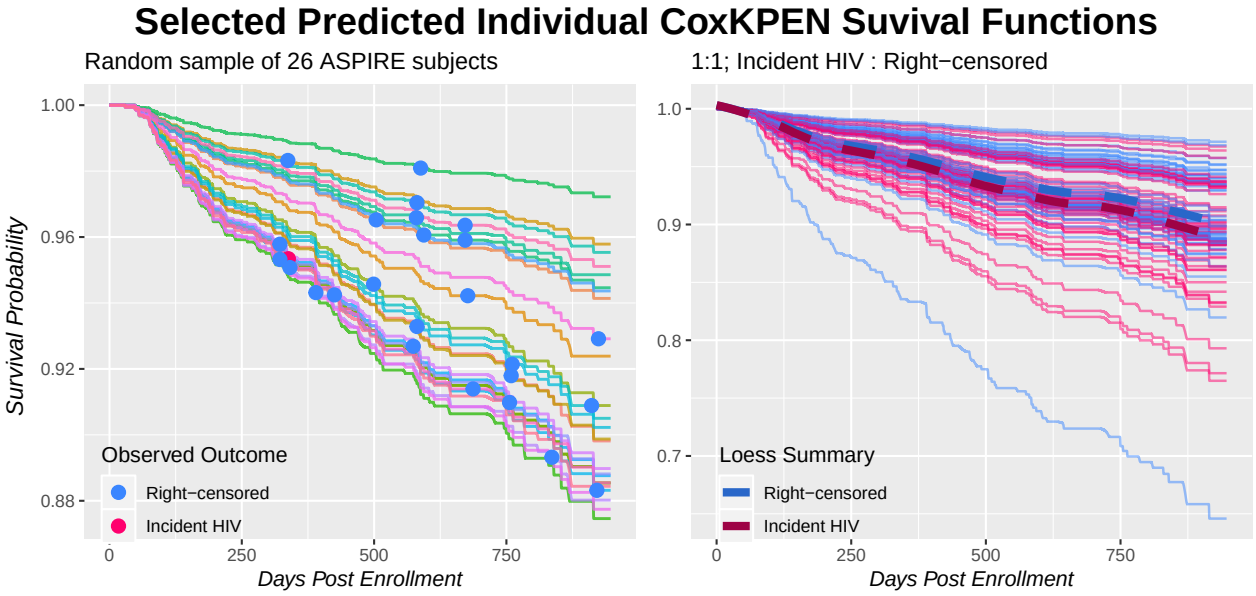

Figure 6.

Table 6: Fitted CoxKPEN beta values arranged by absolute value

Baseline Characteristic	beta	exp.beta
Positive for gonorrhea	0.2784	1.3210
Positive for trichomoniasis	0.1596	1.1730
Participant is married	-0.1405	0.8689
Positive for chlamydia	0.1359	1.1456
Attended college or university	-0.1172	0.8894
Age at enrollment is greater or equal to 25 years	-0.1029	0.9022
Study site is in South Africa	0.0996	1.1047
No vaginal sex in the past week	0.0947	1.0993
Condom not always used during vaginal sex	-0.0752	0.9276
Self-employed	-0.0675	0.9347
Participant traveled 1-2 hours to clinic	-0.0641	0.9379
Participant received dapivirine vaginal ring	-0.0633	0.9387
Completed secondary school	-0.0592	0.9425
PSP is aware of subject's study participation	-0.0517	0.9496
Serum creatinine mg/dL	-0.0414	0.9594
Number of prior pregnancies	-0.0391	0.9617
Primary sex partner (PSP) is uncircumcised	-0.0379	0.9628
Subject does not know HIV status of PSP	-0.0349	0.9657
PSP is aware of vaginal ring use	-0.0248	0.9755
Participant earns their own income	-0.0179	0.9823
At least one specific concern with ring use	0.0084	1.0084
Number of cigarettes smoked daily	0.0082	1.0082
Severity of preexisting conditions	0.0059	1.0059
Instances of vaginal sex in the last week	-0.0042	0.9958
Instances of vaginal sex in an average month	-0.0020	0.9980
Platelets X1000/mm3	0.0004	1.0004
Eosinophils cells/mm3	0.0002	1.0002
PSP is circumcised	0.0000	1.0000

Random Survival Forests

In contrast to the parametric AFT_{weib} and semi-parametric Cox_{KPEN} models, the RSF model is nonparametric, allowing its predicted survival curves to vary widely in form (Figure 7). These curves may cross, which in effect allows subject risk rankings to reorder across time, a degree of freedom not accessible to the AFT_{weib} or Cox_{KPEN} models. RSF functions are still constrained to be monotonically decreasing.

At 750 days, the range of survival probabilities predicted by the RSF model is between that of the AFT_{weib} and Cox_{KPEN} models (Figure 7 left), however the RSF model predicted the largest maximum probability of the three at that timepoint. The RSF model assigned the highest risk to only subject to acquire HIV, for most of the prediction time interval. A noticeable difference between the RSF versus AFT_{weib} and Cox_{KPEN} curves is that many of the RSF curves predict an early sharp drop in survival probability in higher-risk subjects. This trend is also apparent in the relatively steep initial slope of the loess summaries (Figure 7 right) followed by a leveling off. The two loess summaries exhibit the most divergence between roughly 300 and 800 days, suggesting this model is most able to correctly rank two individuals in terms of HIV risk in this time range.

Three quantitative performance metrics were estimated for the RSF via 5-fold cross-validation (Table 7). First, the probability the RSF model is correct when it queries the order of failure of two subjects, concordance, is 65.4% (sd, 0.046). Second, the average absolute difference between the median of the survival curve and the observed failure time among those who experienced the outcome or a "best guess" for those who did not, l1-loss, is 8048 days (sd, 641 days). Third, the cross-validated integrated Brier score encompassing uncertainty, calibration, and resolution is 0.0473 (sd, 0.002). Feature selection was not implemented in this RSF model. RSF models do not have interpretable variable coefficients.

Table 7: Evaluation of Random Survival Forest

Metric	Mean	SD
Concordance	0.6544	0.0460
L1 Loss	8048.0000	641.0000
Integrated Brier	0.0473	0.0019

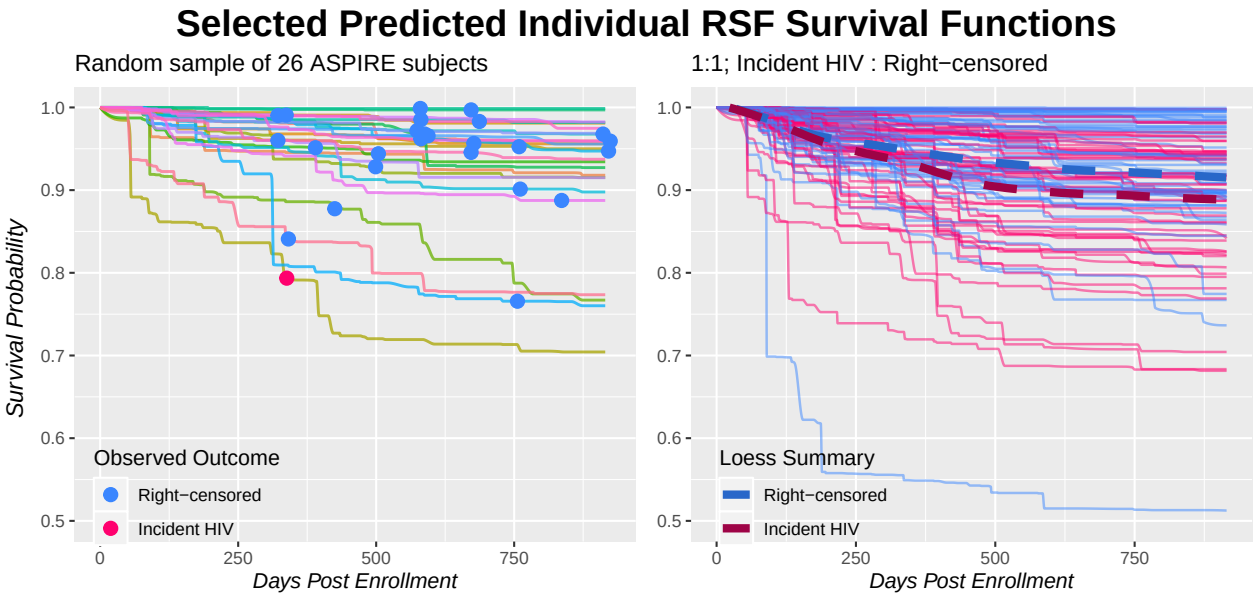

Figure 7.

Discussion

Individual risk evaluation may play an increasing role in the HIV response. With data from the ASPIRE cohort, we modeled individual-specific HIV survival distribution (ISD) using accelerated failure time, Cox regression with a specified baseline hazard, and random survival forests. Though all three exhibited modest ability to rank order survival times, they could not convincingly support statements on absolute personal risk and on predicting individual survival time, underperformed relative to simple Kaplan-Meier analysis.

The coming decade will be pivotal for the HIV response in East and South Africa. More than twenty years of significant investment in this space has delivered steady returns in terms of epidemiologic metrics and rapid progress on the biopharmaceutical front in the development of novel prevention interventions. Sophisticated new medicines to treat and prevent HIV are safe and extremely effective. However, barriers to the comprehensive delivery of these therapies to candidates, including logistical challenges and budget

shortfalls, as well as issues surrounding uptake, including treatment initiation hesitancy and poor adherence, remain formidable and underscore the need to further develop efficiencies where possible.

Typical recommendations from the HIV response community are to use coarse risk metrics to optimize the delivery of PrEP, ART, and PMTC services at the community level[44,45]. However, where HIV prevalence is relatively low (i.e., in the general population relative to KPs) and PrEP delivery is challenging or expensive (as is the case in much of East and South Africa), individual-specific risk evaluation to identify and target individuals at the highest risk for HIV may increase the cost-effectiveness of PrEP initiatives[46].

In this respect, the three models fit above appear modestly capable, as evidenced by estimates of their concordance, which range from 0.65 (RSF) to 0.69 (AFT_{weib}). Though the ability to rank subjects by risk is potentially useful for prioritizing PrEP candidates, this interpretation is not unique to ISD models. Standard methods, including multiple regression, can provide this type of intelligence and, being simpler, are preferable where this is the only goal.

However, in estimating the entire distribution of survival times, ISD models enable statements regarding absolute personal risk as well. Calculating the mean or median of this distribution gives the predicted individual survival time. The utility of a high-performing estimator of this quantity is evident. However, the estimated l1-losses of the models we fit (point estimates for which ranged from 6,136 to 8,048 days) are far too high to have any confidence in their predictions of individual survival time. In fact, these models substantially underperformed in this regard relative to simple Kaplan-Meier analysis, which returned a mean cross-validated l1-loss of 4,920 days.

These models may also be used to return the probability an individual is surviving at any time point. The performance of predictions of this type is captured by the models' Brier score. The AFT_{weib}, RSF, and Cox_{KPEN} models fit above are essentially equivalent to simple Kaplan-Meier on Brier score.

It is clear that as specified, these models do not provide accurate predictions of absolute individual risk. This is undoubtedly caused partly by the high degree of censoring (92%) and a relatively small number of individuals (2,629) in the cohort and potentially by issues with feature selection and engineering.

A potentially productive application for models of this type is in enhancing communication between providers and patients by providing intuitive, visual summaries of personal risk. An understanding of one's health risks (which in the case of HIV is typically poor[47]) and the factors driving them is a precondition to health-promoting behaviors common to most theoretical models of health psychology[48]. Furthermore, enhanced communication in clinical settings can increase treatment uptake and improve adherence, which are persistent problems with PrEP and ART. There may even be a role for these tools outside the clinic. A recent

experiment in providing HIV risk scores freely over the internet reported moderate success in inducing HIV self-testing[49].

However, as with any public health intervention, such a data instrument must be carefully evaluated beyond merely it's technical performance before exposing it to the public.

Despite the ever-expanding taxonomy of analytic methods and data sources, the criteria for a good risk prediction tool for a public health application is a bit more static. Criteria necessarily include high-performance (can the output be trusted?), practicality (can the covariates included be reasonably measured?), and usefulness of interpretation. Advocates of tools that return individual survival functions (the type fit above) tout their performance with regard to that last point, interpretation. They argue three features of any risk score define its interpretation, and in each, individual survival functions offer certain improvements over the status quo. These features of risk scores are their demographic resolution, temporal resolution, and summary output type.

Demographic resolution concerns how customized the tool is to the individual under evaluation. Does it, for example, incorporate broad demographic and behavioral attributes and thus likely to score many people equivalently? Or, on the other hand, is the set of personal features considered enough to make the output unique to the individual? Second is the temporal resolution. Certain risk scores make estimates at only discrete time points (i.e., 1-year or 5-year) while others are compatible with continuous time. Lastly, is the type of output. Some risk scores output relative measures that only take on meaning in relation to another (for example, the Hazard Ratio) or are binned into severity grades determined a priori (see PHQ9 discussed above for an example of such risk stratification) while still others return an actual probability. ISDs offer probabilities at any time point for an individual.

Finally, predictive models should be considered public health interventions subject to robust ethical standards. Digital interventions are not without the ability to cause harm. Potential abuses of such a tool are unfortunately not hard to fathom, and serious issues surrounding privacy and discrimination, especially given a profit motive, must be proactively addressed. Any high-performing model must be subject to stringent data security measures. Advanced predictive models fed by a shockingly pervasive and ever-expanding surveillance infrastructure are, and will increasingly become, powerful tools to advance public health. Now it the time to grapple with the consequences of advanced predictive models and carefully plot a human-centric path for their future.

By necessity, the sophisticated use of data is likely to play an increasing role in the HIV response. HIV can be eradicated with medicine on the market today. Victory is within sight, but barriers to care remain at the

systemic, community, pairwise, and individual levels. Initiatives in precision resource allocation could help accommodate shrinking budgets. Probabilistic case identification could accent traditional testing campaigns. Alongside tools to predict low adherence, ART failure, and disease progression, individual risk assessments could improve prescribing, reduce treatment hesitancy, and improve communication within the critically important doctor-patient relationship. Self-administered online risk assessments could guide risk-reducing behavior, collect data, and offer testing services or referral to care. Successful use of data may well put HIV eradication in our grasp. However, such tools are not without risk and should be developed and implemented cautiously and conscientiously.

References

1. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *New England Journal of Medicine*. 2011;365(6):493-505. doi:10.1056/NEJMoa1105243

2. Gray R, Kigozi G, Kong X, et al. The effectiveness of male circumcision for HIV prevention and effects on risk behaviors in a posttrial follow-up study. *AIDS*. 2012;26(5):609-615. doi:10.1097/QAD.0b013e3283504a3f

3. Grant RM, Lama JR, Anderson PL, et al. Preexposure Chemoprophylaxis for HIV Prevention in Men Who Have Sex with Men. *New England Journal of Medicine*. 2010;363(27):2587-2599. doi:10.1056/NEJMoa1011205

4. Rozhnova G, Heijne J, Bezemer D, et al. Elimination prospects of the Dutch HIV epidemic among men who have sex with men in the era of preexposure prophylaxis. *Aids*. 2018;32(17):2615-2623. doi:10.1097/QAD.0000000000002050

5. Kusejko K, Marzel A, Hampel B, et al. Quantifying the drivers of HIV transmission and prevention in men who have sex with men: a population model-based analysis in Switzerland. *HIV Medicine*. 2018;19(10):688-697. doi:10.1111/hiv.12660

6. Li J, Peng L, Gilmour S, et al. A mathematical model of biomedical interventions for HIV prevention among men who have sex with men in China. *BMC Infectious Diseases*. 2018;18(1):1-9. doi:10.1186/s12879-018-3516-8

7. Nel A, Van Niekerk N, Kapiga S, et al. Safety and efficacy of a dapivirine vaginal ring for HIV prevention in women. *New England Journal of Medicine*. 2016;375(22):2133-2143. doi:10.1056/NEJMoa1602046

8. Baeten J, Mgodil N, Palanee-Phillips T. A Phase 3B Open-Label Follow-on Trial to Assess the Continued

Safety of and Adherence to a Vaginal Ring Containing Dapivirine in Women. Published online 2014.

9. UNAIDS. UNAIDS Data 2018. Published online 2018:1-376.

10. UNAIDS. *Making the end of AIDS real: consensus building around what we mean by "epidemic control" — a meeting convened by the UNAIDS Science Panel. Glion, Switzerland, 4–6 October 2017.*; 2017.

11. Ghys PD, Williams BG, Over M, Hallett TB, Godfrey-Faussett P. Epidemiological metrics and benchmarks for a transition in the HIV epidemic. *PLOS Medicine*. 2018;15(10):e1002678. doi:10.1371/journal.pmed.1002678

12. OMB. *A Budget for America's Future.*; 2020.

13. UNAIDS. UNAIDS 2019 Data. Published online 2019:1-248. doi:978-92-9173-945-5

14. Cuadros DF, Sartorius B, Hall C, Akullian A, Bärnighausen T, Tanser F. Capturing the spatial variability of HIV epidemics in South Africa and Tanzania using routine healthcare facility data. *International Journal of Health Geographics*. 2018;17(1):1-9. doi:10.1186/s12942-018-0146-8

15. Wolf RC, Bingham T, Millett G, Wilcher R. Building the evidence base to optimize the impact of key population programming across the HIV cascade. *Journal of the International AIDS Society*. 2018;21:e25146. doi:10.1002/jia2.25146

16. WHO. Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations. *WHO Guidelines*. 2014;(July):184. http://apps.who.int/iris/bitstream/10665/128048/1/9789241507431{_}eng.pdf?ua=1

17. UNAIDS. *Report on the global AIDS epidemic*. Joint United Nations Programme on HIV/AIDS; 2010.

18. Dunkle KL, Stephenson R, Karita E, et al. New heterosexually transmitted HIV infections in married or cohabiting couples in urban Zambia and Rwanda: an analysis of survey and clinical data. *The Lancet*. 2008;371(9631):2183-2191. doi:10.1016/S0140-6736(08)60953-8

19. UNAIDS. AIDSinfo.

20. Mugo NR, Ngure K, Kiragu M, Irungu E, Kilonzo N. PrEP for Africa: What we have learnt and what is needed to move to program implementation. *Current Opinion in HIV and AIDS*. 2016;11(1):80-86. doi:10.1097/COH.0000000000000224

21. Cowan FM, Delany-Moretlwe S, Sanders EJ, et al. PrEP implementation research in Africa: What is new? *Journal of the International AIDS Society*. 2016;19(Suppl 6). doi:10.7448/IAS.19.7.21101

22. Mboup A, Béhanzin L, Guédou FA, et al. Early antiretroviral therapy and daily pre-exposure prophylaxis for HIV prevention among female sex workers in Cotonou, Benin: a prospective observational demonstration

study. *Journal of the International AIDS Society.* 2018;21(11):1-11. doi:10.1002/jia2.25208

23. Heffron R and, Ngure K, Odoyo J, et al. Pre-exposure prophylaxis for HIV-negative persons with partners living with HIV: uptake, use, and effectiveness in an open-label demonstration project in East Africa. *Gates Open Research.* 2017;1:3. doi:10.12688/gatesopenres.12752.1

24. Sinead DM, Hosek S, Celum C, Wilson CM, Kapogiannis B, Bekker LG. Preventing HIV among adolescents with oral PrEP: Observations and challenges in the United States and South Africa. *Journal of the International AIDS Society.* 2016;19(Suppl 6):1-7. doi:10.7448/IAS.19.7.21107

25. Mugwanya KK, Pintye J, Kinuthia J, et al. Integrating preexposure prophylaxis delivery in routine family planning clinics: A feasibility programmatic evaluation in Kenya. *PLoS Medicine.* 2019;16(9):22-29. doi:10.1371/journal.pmed.1002885

26. Celum CL, Delany-Moretlwe S, McConnell M, et al. Rethinking HIV prevention to prepare for oral PrEP implementation for young African women. *Journal of the International AIDS Society.* 2015;18(Suppl 3):1-10. doi:10.7448/IAS.18.4.20227

27. Folkman S, Lazarus RS, Gruen RJ, DeLongis A. Appraisal, coping, health status, and psychological symptoms. *Journal of personality and social psychology.* 1986;50(3):571-579. doi:10.1037//0022-3514.50.3.571

28. Kreuter MW, Strecher VJ. Changing inaccurate perceptions of health risk: Results from a randomized trial. *Health Psychology.* 1995;14(1):56-63. doi:10.1037/0278-6133.14.1.56

29. Goetzel RZ, Staley P, Ogden L, et al. A Framework for Patient- Centered Health Risk Assessments. *Centers For Disease Control and Prevention.* Published online 2008.

30. Krist AH, Glasgow RE, Heurtin-Roberts S, et al. The impact of behavioral and mental health risk assessments on goal setting in primary care. *Translational Behavioral Medicine.* 2016;6(2):212-219. doi:10.1007/s13142-015-0384-2

31. Baeten JM, Palanee-Phillips T, Brown ER, et al. Use of a vaginal ring containing dapivirine for HIV-1 prevention in women. *New England Journal of Medicine.* 2016;375(22):2121-2132. doi:10.1056/NEJMoa1506110

32. HOPE. *Results of open-label study of a vaginal ring for HIV prevention suggest women are interested in and willing to use it.* Vol 3315. Microbicide Trials Network; 2019:1-7.

33. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data.* Vol 31.; 1982:278.

doi:10.2307/2988004

34. Lin DY. On the Breslow estimator. *Lifetime Data Analysis*. 2007;13(4):471-480. doi:10.1007/s10985-007-9048-y

35. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning*. Vol 103. Springer New York; 2013. doi:10.1007/978-1-4614-7138-7

36. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *Journal of Statistical Software*. 2011;39(5):1-13. doi:10.18637/jss.v039.i05

37. Basha L, Hoxha F. Kernel Estimation of the Baseline Function in the Cox Model. *European Scientific Journal ESJ*. 2019;15(6):105-118. doi:10.19044/esj.2019.v15n6p105

38. Mogensen U, Ishwaran H, Gerds T. Evaluating Random Forests for Survival Analysis Using Prediction Error Curves. *Journal of Statistical Software*. 2012;50(11):301-325.

39. Raykar VC, Steck H, Krishnapuram B, Dehing-Oberije C, Lambin P. On ranking in survival analysis: Bounds on the concordance index. *Advances in Neural Information Processing Systems 20 - Proceedings of the 2007 Conference*. Published online 2009:1-8.

40. Haider H, Hoehn B, Davis S, Greiner R. Effective Ways to Build and Evaluate Individual Survival Distributions. Published online 2018:1-60. http://arxiv.org/abs/1811.11347

41. Gupta R, Bradley D. Representing the mean residual life in terms of the failure rate. *Mathematical and Computer Modelling*. 2003;37(12-13):1271-1280. doi:10.1016/S0895-7177(03)90038-0

42. Murphy AH. A New Vector Partition of the Probability Score. *Journal of Applied Meteorology (1962-1982)*. 1973;12(4):595-600.

43. Graf E, Schmoor C, Sauerbrei W, Schumacher M. Assessment and comparison of prognostic classification schemes for survival data. *Statistics in Medicine*. 1999;18(1718):2529-2545. doi:10.1002/(sici)1097-0258(19990915/30)18:17/18<2529::aid-sim274>3.3.co;2-x

44. WHO. *World Health Organization. The use of antiretroviral drugs for treating and preventing hiv infection.*; 2016:99-152, 402. https://apps.who.int/iris/bitstream/handle/10665/208825/9789241549684{_}eng.pdf?sequence=1{

45. Joseph Davey DL, Bekker L-G, Gorbach PM, Coates TJ, Myer L. Delivering preexposure prophylaxis to pregnant and breastfeeding women in Sub-Saharan Africa: the implementation science frontier. *AIDS (London, England)*. 2017;31(16):2193-2197. doi:10.1097/QAD.0000000000001604

46. Wong NS, Kwan TH, Tsang OT, et al. Pre-exposure prophylaxis (PrEP) for MSM in low HIV incidence

places: should high risk individuals be targeted? *Scientific Reports*. 2018;8(1):1-9. doi:10.1038/s41598-018-30101-9

47. Stringer EM, Sinkala M, Kumwenda R, et al. Personal Risk Perception, HIV Knowledge and Risk Avoidance Behavior, and Their Relationships to Actual HIV Serostatus in an Urban African Obstetric Population. *Journal of Acquired Immune Deficiency Syndromes*. 2004;35(1):60-66. doi:10.1097/00126334-200401010-00009

48. Macintyre K, Rutenberg N, Brown L, Karim A. Understanding perceptions of HIV risk among adolescents in KwaZulu-Natal. *AIDS and Behavior*. 2004;8(3):237-250. doi:10.1023/B:AIBE.0000044072.71361.b3

49. De Boni RB, Veloso VG, Fernandes NM, et al. An internet-based HIV self-testing program to increase HIV testing uptake among men who have sex with men in Brazil: Descriptive cross-sectional analysis. *Journal of Medical Internet Research*. 2019;21(8). doi:10.2196/14145