Introduction: Successful management of HIV is underpinned by suppression of circulating virus in infected individuals. Current treatment regimens have the ability to suppress viral load to levels undetectable by assay measurement, sometimes for many years, but frequently periods of viraemia are interspersed with intervals when viral load again reaches detectable levels. In this study of patients on long-term therapy we aim to estimate and compare distributions of aviremic periods overlapping fixed time points in their treatment history. As not all individuals are followed through to completion of their spell of aviremia we need to take account of the censoring, but note that assumptions underlying standard survival techniques may not be valid due to correlations between the censoring and observed duration times induced by the sampling scheme. We therefore adopt a convolution-based approach to minimize potential biases.

Methods

Problem Definition:

- Interest centres on the distribution associated with $T_D$, the duration of observed spells of aviremia overlapping a fixed time point $t_0$.
- Let $T_D$ denote the backward recurrence time from the start of the spell to $t_0$ for the $i$th individual, and $T_F$ the forward recurrence time from $t_0$ to the end of the spell.
- Then $T_D = T_W + T_F$.
- When data may be censored:
  - We observe the pair $(C(t_0),t_0)$ where $t_0$ is the censoring indicator and $C(t_0) = T_W + Z(t_0)$ for $Z(t_0) = \min(T_W,C(t_0))$ with $C(t_0)$ the time from $t_0$ to end of follow-up.
  - Given individuals are only included in analysis if observed to be aviremic at $t_0$, observed duration time and censoring are not independent and standard survival analysis is not appropriate.

Aviremia indicated by heavy time

$T_D$ Duration time observed

$T_F$ Duration time censored

Not included in analysis as not aviremic at $t_0$.

Inference Based on Backward and (Censored) Forward Recurrence Times

Providing the conditional distribution of $T_D$ given $T_W$ is the same irrespective of the forward censoring time, we note that from standard theory we can derive the likelihood in terms of its conditional distribution:

$$L = \prod_i L_i = \prod_i f(T_D|t_0)g(t_0)g(T_W|t_0)f(T_W),$$

and write $S(t_0) = f(T_D > t_0)$ as a convolution of $f_T(t_0|x)$ and $S(t_0|x)$. Estimation of $S(t_0)$

For analyses presented here we utilize Weibull regression modelling for inference and estimation of $f(t_0)$ and $S(t_0)$ and discrete time duration for derivation of the convolution:

Suppose $0 < d_1 < \ldots < d_j$ divides the support of $T_D$ and let $t_j$ denote the $j$th interval. Then $\hat{f}(t_0|x) = \sum_{i=1}^j \hat{f}_i(t_0|x), \hat{S}(t_0|x) = \sum_{i=1}^j \hat{S}_i(t_0|x)$.

Data characteristics:

- Data collected from patients of the Western Australian HIV Cohort study undergoing regular viral testing as part of standard clinic practice, and receiving one of 2 standard treatment classes: either a protease inhibitor (PI) or a non-nucleoside reverse transcriptase inhibitor (NNRTI).
- Patients’ viral control assessed at a fixed time $t_0$ from commencement of therapy, with the time point depending on cohort of treatment initiation: at 12 years post-initiation for those commencing therapy in the years 1996-1999; at 8 years post-initiation for those commencing 2000-2003; at 4 years post-initiation for those commencing 2004-2007.
- Patients included in analysis if aviremic at $t_0$ with aviremia defined as any viral presence in blood being below detectable limits.

Simulations

1-sample simulations

- Direct modelling of true time: $T_D$
- Convolution of backward and forward recurrence time distribution: $S_T(t_0)$
- Direct modelling of censored duration time:
  - $\hat{C}_0 = T_D + \min(T_D,C(t_0))$.

Figure 1.

- Distributions of estimated sample medians. Mean of true median values overlaid in red. The convolution-based estimates showed a small downwards bias, especially with less censoring. The direct modelling of the observed censored duration time yielded increasingly large biases with increasing censorship.

2-sample simulations

- Simulation example
  - Convolution $\hat{f}_T(t_0|x)$ and $\hat{S}_T(t_0|x)$ where $d$ denotes group membership.
  - Direct modelling of censored duration time: $\hat{C}_0 = T_D + \min(T_D,C(t_0))$.

Figure 2.

- Difference in estimated medians
  - $\hat{C}_0$ vs $\hat{C}_0$
  - Increasing biases in estimated covariate effect observed with increasing censorship for direct modelling of censored duration time.

Application

Analyses:

- Outcome of interest is estimation of duration of aviremic period about $t_0$, with assessment of covariate impact.
- Weibull regression utilising conducted separate analyses of the backward recurrence times, the time from the start of the aviremic period to $t_0$, and the times from $t_0$ to the end of the period or end of follow-up whichever has occurred first.
- The total distribution of aviremic period is then achieved by a convolution of the estimated distribution of the backward recurrence times and the conditional distribution of $T_W$.
- Comparison made with the direct method of simply applying standard survival modelling to sum of backward and censored forward recurrence times, ignoring issue of bias.

Figure 3.

- Time from start of aviremic period to $t_0$
- Time from $t_0$ to end of aviremic period or end of follow-up
- Time from start of aviremic period to end of period or loss of follow-up

Figure 4.

- Median duration of aviremic periods, demonstrating impact of significant covariates:
  1) High vs low baseline viral load (left and right interval endpoints, respectively).
  2) Treatment differences – those on NNRTI therapy having longer periods of aviremia in general cohort ; and
  3) Cohort effects, with evidence of interaction with treatment.

Direct estimation of censored duration times (dotted lines) nearly always yielded higher estimates than those obtained by the convolution method (heavy lines).

Figure 5.

- PI Rx vs NNRTI Rx
- Kaplan-Meier plots of backward recurrence times, (censored) forward recurrence times and their sum, according to treatment choice and cohort of commencement therapy. For these untreated times, between-cohort differences are more evident amongst individuals receiving a NNRTI-based treatment.

- Convoluted vs direct method
- Relative median duration estimates derived for individuals receiving PI therapy, and commencing treatment from 2003-2008. Plots ($a$) contrast differences between the 2 estimation methods, and ($b$) demonstrate relative impact of baseline viral load, with between-method differences that are only evident in one group translating to almost double the size of between-group estimates.

Summary:

- Modelling of both backward and forward recurrence times that assumed proportional hazard covariate effects found duration of aviremia to be significantly associated with viral load values at commencement of therapy, cohort of initiation, and class of treatment. When the same covariates were included in direct analysis of the total observed periods, the resulting estimates of mean duration were consistently larger than those obtained by the convolution method. Moreover, variation in real impact of the different covariates was also observed.