Analysis of censored duration intervals overlapping a fixed time point where censoring and survival times are correlated

Elizabeth McKinnon*, Ian James, David Nolan

Centre for Clinical Immunology and Biomedical Statistics, Murdoch University & Royal Perth Hospital, Perth, Western Australia

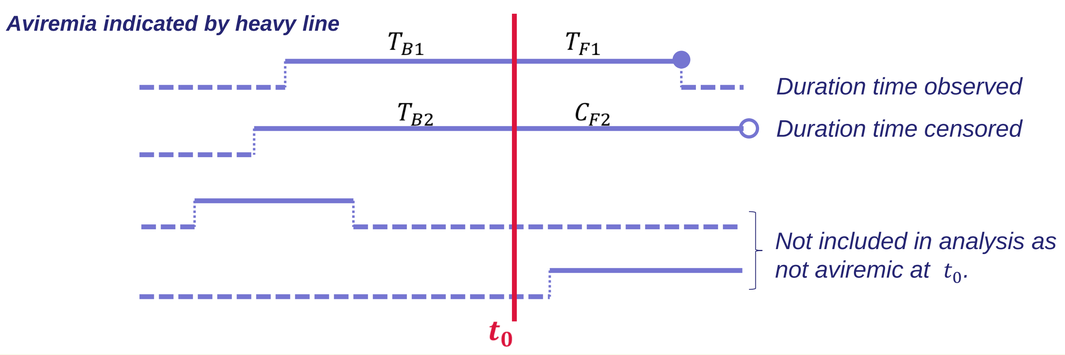
**Email: E.McKinnon@murdoch.edu.au*

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 aviremia 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_{Bi} denote the backward recurrence time from the start of the spell to t_0 for the ith individual, and T_{Fi} the forward recurrence time from t_0 to the end of the spell. Then $T_{Di} = T_{Bi} + T_{Fi}$.
- When data may be censored:
  - We observe the pair (C_{Di}, δ_i) where δ_i is the censoring indicator and $C_{Di} = T_{Bi} + Z_{Fi}$ for $Z_{Fi} = min(T_{Fi}, C_{Fi})$ with C_{Fi} 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.



Inference Based on Backward and (Censored) Forward Recurrence Times

Providing the conditional distribution of T_F given T_B 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:

$$l = \prod_{i=1}^{n} \{ f_F(z_{Fi}|\boldsymbol{x}_i, t_{Bi})^{\delta_i} S_F(z_{Fi}|\boldsymbol{x}_i, t_{Bi})^{1-\delta_i} \} f_B(t_{Bi}|\boldsymbol{x}_i) ,$$

and write $S_D = P(T_D > t|\boldsymbol{x})$ as a convolution of $f_B(t_B|\boldsymbol{x})$ and $S_F(t_F|\boldsymbol{x}, t_B)$.

Estimation of S_D

For analyses presented here we utilize Weibull regression modelling for inference and estimation of f_B and S_F, and discretize duration time for derivation of the convolution:

Suppose $0 < d_1 < d_2 \ldots d_J$ divides the support of T_D, and let I_j denote the j^{th} interval. Then

$$\hat{P}(T_D > d_k) \approx \sum_{j=1}^{k} \hat{P}_B(t_B \in I_j \mid \boldsymbol{x}) \; \hat{P}_F(T_F > d_k - d_j | t_B \in I_j; \boldsymbol{x}).$$

Simulations

T_{Di} generated from Weibull (shape=0.5, scale=5), $t_{0i} = T_{Bi}$ uniform over T_{Di}, $T_{Fi} = T_{Di} - t_{oi}$ and C_{Fi} as exponential. Weibull modelling utilized for both direct and convolution methods

1-sample simulations

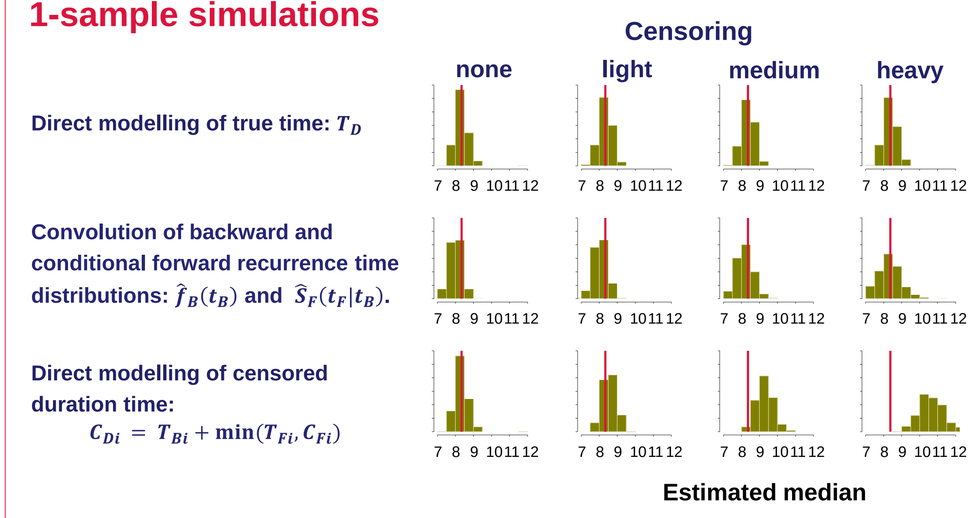
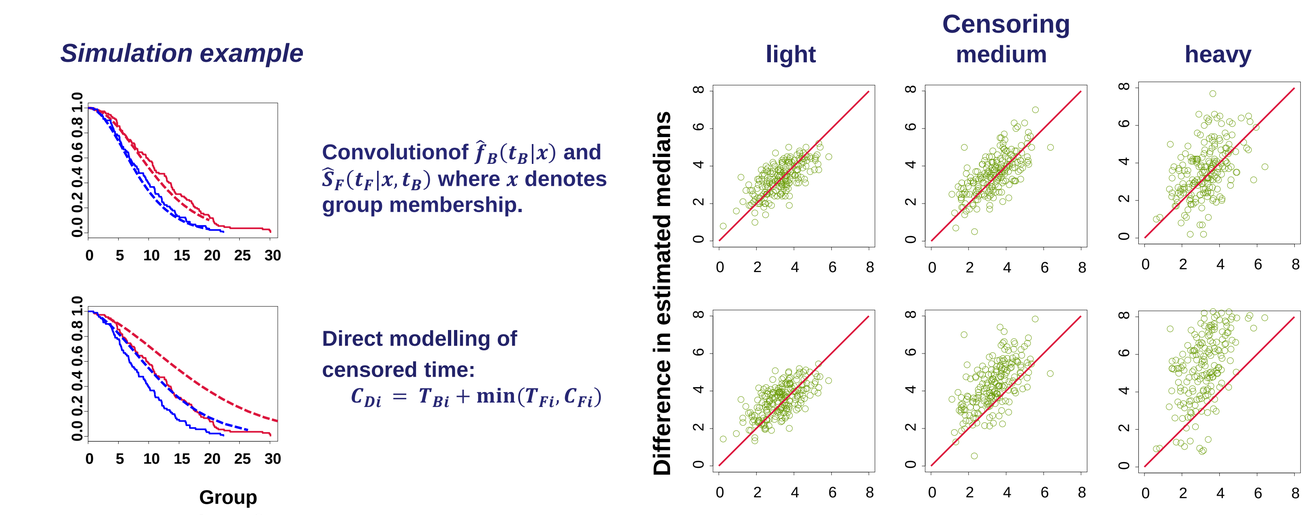


Figure 1.

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

Figure 2.



Increasing biases in estimated covariate effect observed with increasing censorship for direct modelling of censored duration time

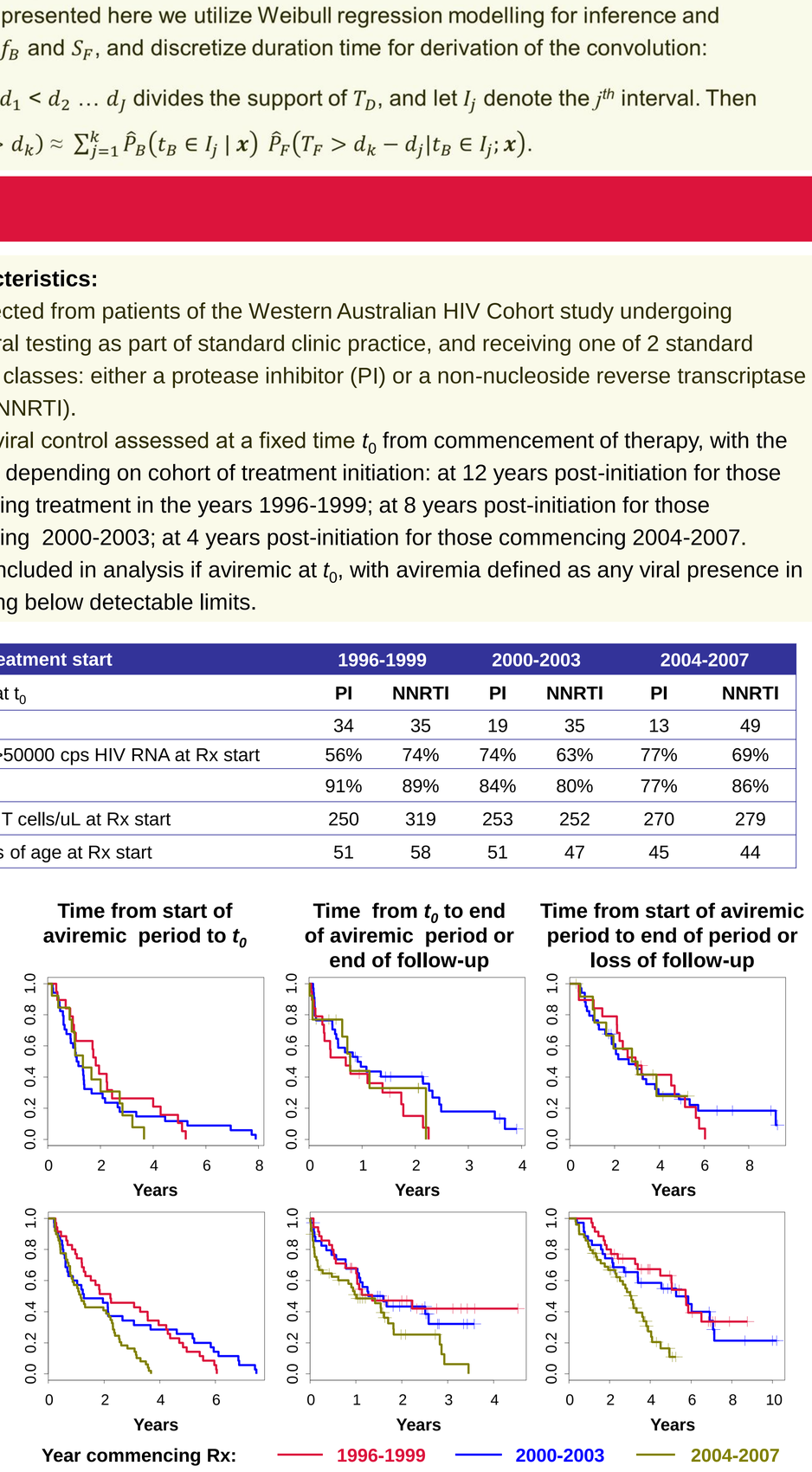
Application

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 treatment 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.

Cohort of treatment start	1996-1999		2000-2003		2004-2007	
Class of Rx at t_0	PI	NNRTI	PI	NNRTI	PI	NNRTI
N	34	35	19	35	13	49
% of N with >50000 cps HIV RNA at Rx start	56%	74%	74%	63%	77%	69%
% male	91%	89%	84%	80%	77%	86%
Median CD4 T cells/uL at Rx start	250	319	253	252	270	279
Median years of age at Rx start	51	58	51	47	45	44

Figure 3.



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

Analyses:

- Outcome of interest is estimation of duration of aviremic period about t_0, with assessment of covariate impact.
- Weibull regression modelling utilized to conduct 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.
- Distribution of the total aviremic period is then achieved by a convolution of the estimated distribution of the backward recurrence times and the conditional distribution of the forward recurrence times.
- 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.

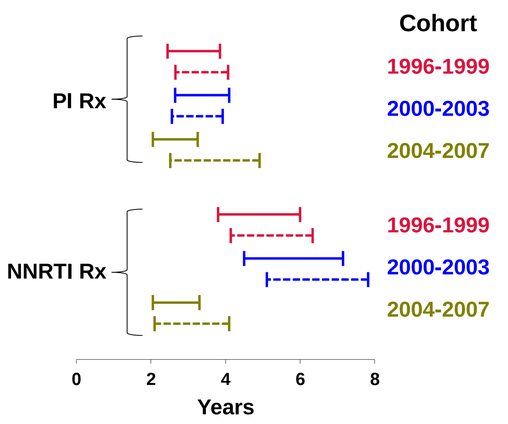


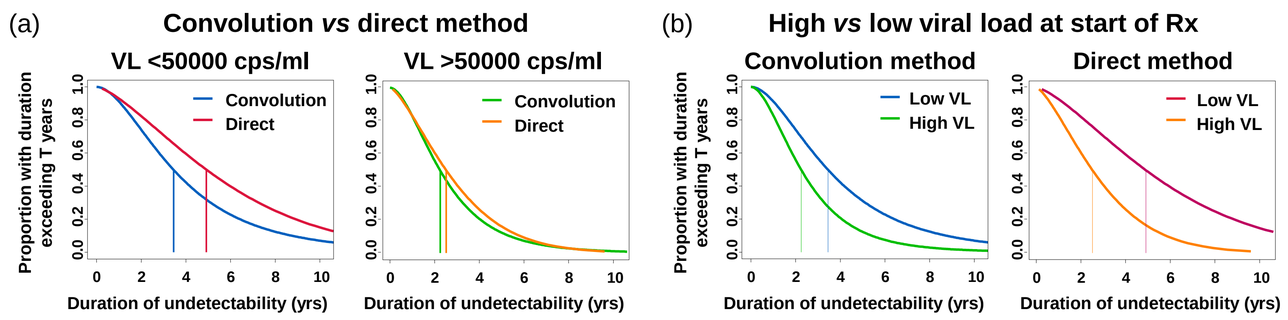
Figure 4.

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 n 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.



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 relative impact of the different covariates was also observed.