Appendix 1

The sequence of available data for Fiebig stages consists of time of last HIV-1 uninfected sample ; time and stage of first HIV-infected sample, and a series of subsequent times and stage of infection up to the first Fiebig stage 6 sample (. Each person’s “infection interval” is defined as the time between the last HIV-uninfected and first HIV-infected sample .

Time to each Fiebig stage, after (unobserved) time of infection is assumed to follow an Exponential waiting time distribution with mean . The time of infection is assumed to be Uniformly distributed in the infection interval (. The parametric survival distribution for the time to Stage k is:

For each person, for Fiebig stage *k*, the contribution to the likelihood is

1. 1 - S(t1; ) if the stage observed at t1 is greater than k
2. S(ti-1; ) – S(ti; ) if stage observed at prior time ti-1 is < k and at time ti is > k
3. 0, otherwise.