

# Time-of-Infection Estimated by a Time-Continuous HIV-1 IgG Model

Helena Skar,  
Thomas Leitner, T-6

The incidence of an infectious disease is vital for informed and targeted prevention, and knowing the time-of-infection is important for estimating the incidence in a population. For acute infections, like influenza, it is relatively straightforward to estimate time-of-infection because it occurred just shortly before the diagnosis. For chronic infections, like that of Human Immunodeficiency Virus type 1 (HIV-1), time-of-infection estimation is more complicated because most diagnosed persons have an established HIV-1 infection of unknown duration. In the last few years, there has been considerable interest in the development of biomarker assays that can determine if an HIV-1 infection is recent in order to estimate the incidence in a population. These serological assays are based on the knowledge about the development and maturation of the HIV-1 antibody response in infected persons. Until now, these assays have provided a binary result, that is, recent versus long-term infection, rather than a quantitative estimation of time since infection.

In this project, we developed a model describing the production of a specific type of IgG antibodies that can be detected by the IgG capture BED enzyme immunoassay (BED assay). The assay name “BED” signifies that it is based on a trimeric branched peptide with each branch derived from the immunodominant region of the gp41 glycoprotein of HIV-1 subtype B, circulating recombinant form (CRF) 01\_AE or subtype D.

Similar to many biological systems where the rate of reproduction is proportional to the existing population and limited resources, the growth of BED-specific IgG (BED IgG) following infection can be modeled by a logistic function. In order to account for patient variability of the BED IgG growth following HIV-1 infection, we developed a time-continuous mixed-effect logistic model describing the production of BED IgG. The model was trained on a large cohort metadata set from a previously published study [1] that included 2,927 longitudinal BED assay results from 756 patients. This model was then validated on a second large dataset, representing cross-sectional BED data from 819 Swedish patients newly diagnosed with HIV-1 infection between 2002 and 2010.

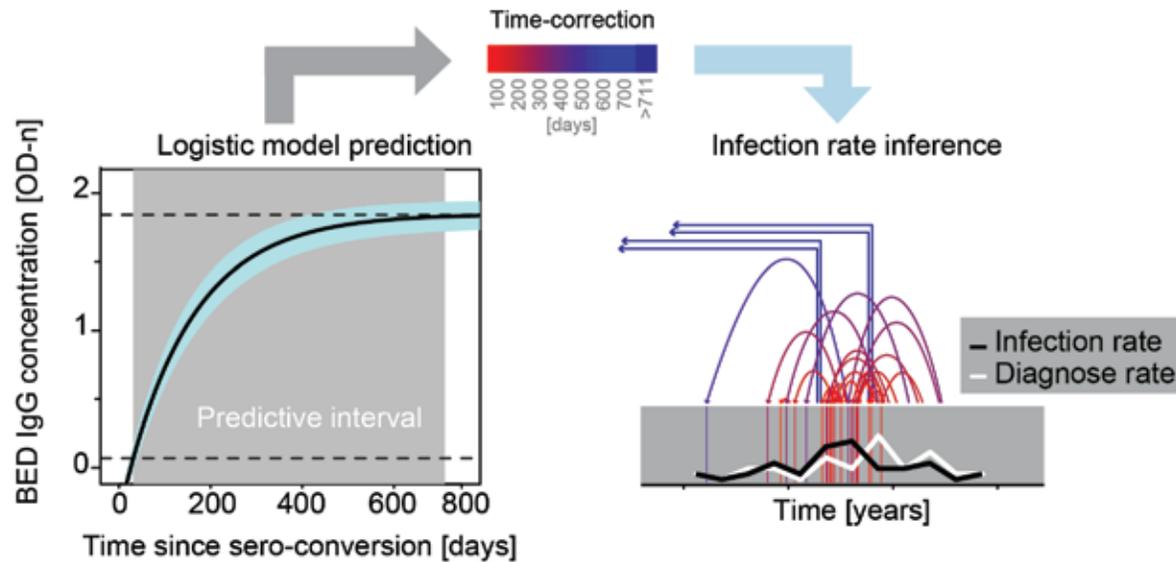
The validation data showed that the best parameterization of the logistic IgG model, which supports the biological intuition that there is no patient variation in BED IgG concentration (measured in normalized optical density, OD-n) at  $t=0$ . Thus, the typical patient is represented by a logistic growth of the BED-detected HIV-specific IgG following infection (Fig. 1). The model is informative of time since infection when the BED assay negative control is within an acceptable range, corresponding to  $\hat{t} = (0, 52)$  days, but specific to each batch of BED measurements, and

99% of the asymptote. The informative BED IgG OD-n interval translates into a continuous time interval with predictive power in  $t = (31, 711)$  days since seroconversion to when BED testing was done. Our model estimated the growth rate at  $r=0.00672$  OD-n units per day, and the asymptote at  $K=1.85$  OD-n units.

We next investigated whether the inferred time since seroconversion of our validation data could be described by the serological interval—that is, the time interval constrained by the last negative and first positive HIV-1 testing. Naturally, barring the time from infection to detectable HIV-1 by a valid method, infection must have happened sometime between these dates. The validation data was not from a cohort study, but rather data from patients detected by regular public health diagnostic testing. However, for this type of data it is not obvious that the population average time-of-infection is in the middle of the serological interval. Indeed, the non-cohort validation data shows a clear bias of infection-time shifted towards the date of diagnosis. Thus, this shows that (1) BED test results are applicable to infer the time since seroconversion in non-cohort type data, but (2) estimating date of infection as the midpoint between the last negative and the first positive HIV-1 test result is inaccurate and misleading in this type of data.

In conclusion, we have created a model that quantifies the time since seroconversion based on a simple serological assay, the BED assay. While we informed the model with BED assay results, because it is currently the most used biomarker for recency estimation, our model could easily be adjusted to other, future biomarkers. Our model is

applicable to biomarker results from patients included in cohort studies as well as patients diagnosed as a result of public health services. We expect that our method can improve incidence estimates, and thus provide valuable information for HIV-1 surveillance and prevention.



*Fig. 1. Graphical abstract. Lower left: logistic modeling of IgG-capture BED-enzyme immunoassay absorbance as a function of recency period. The resulting logistic model is predictive when BED OD-n = (0.07, 1.84), corresponding to recency periods of 31–711 days. This mixed-effects model describes the typical patient, where parameter values correspond to the whole population. Upper middle: correction-time (time between estimated time of infection and time of diagnosis) color legend. Lower right: Example from the Swedish data. The arrows show length of time-correction starting at the time of diagnosis and pointing at the inferred infection time, colored according to length of correction time. Lower part shows the inferred rate of diagnosis (grey line) and rate of infection (black line) for each quarterly year.*