Adding CD4 data to the Testing History estimates

Goal

To increase the precision of undiagnosed estimates by using CD4 counts to modify the probability of infection within the possible infection window.

Background

The testing history method assigns long infection windows to 1) cases reporting never having had a negative test and 2) cases reporting a last negative test (LNT) many years prior to diagnosis. Distributing the probability of infection across this window is a key factor in determining the time from infection to diagnosis (TID). Our current base case assumes a uniform distribution. We would like to investigate whether additional biomarker information from CD4 or viral load could be used to identify cases that may have been infected closer to the time of diagnosis. The literature suggests that viral load is complicated due to variation in set-point viral load (SPVL) across persons, and the fact that peak load can indicate either acute or late stage infection. CD4 looks more promising because on aggregate, CD4 trajectories monotonically decrease after infection, though their overall level and rate of decline is correlated with SPVL. We therefore propose to explore whether integrating CD4 data into the testing history method would improve our estimates of the undiagnosed fraction.

Approach

We will construct a new "CD4 Case" estimate, which shifts the probability of infection for long infection windows closer to the time of diagnosis when indicated by CD4 count. This will influence the undiagnosed fraction estimate by changing the population TID (time from infection to diagnosis) curve.

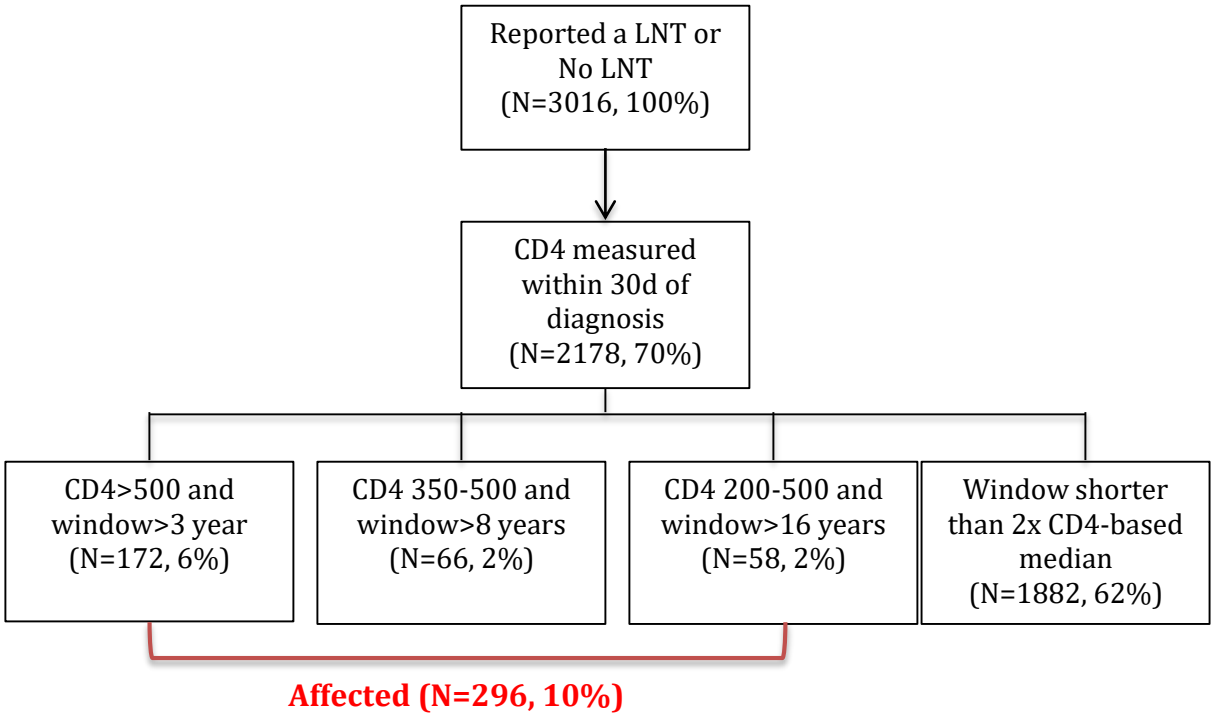
Expected Impact

We project that about 10% of the WA State cases in 2005-2014 that contribute data to the TID would have their infection probability shifted closer to the time of diagnosis under the CD4 Case. The breakdown of the cases is shown in Figure 2. More details are shown in the appendix.

Time Estimate

Developing the CD4 Case and applying to WA State data will take about 50 hrs of work with 10 hrs of supervision. Based on our prior contract, this translates to roughly $5000.

Figure 2. Sample sizes and distribution of cases impacted by the proposed CD4 Case. The full sample has N=5148, of which 3016 (59%) have non-missing testing histories. The window cutoffs for each CD4 category are defined by 2 times the median time from infection to CD4 (see point 3 in the "Proposed Procedure" section).



Appendix: Additional details

A. Proposed procedure:

The procedure will be based on the analysis of CD4 prior to treatment reported in the CASCADE cohort (Lodi et al 2011, (1)) and the ATHENA cohort (Cori et al 2015 (2)), and will account for the heterogeneity in CD4 levels and trajectories by using infection probability distributions that reflect the degree of SPVL variation.

1. Categorize cases with a CD4 count measured within 30days of diagnosis into <200, 200-350, 350-500, and 500+.

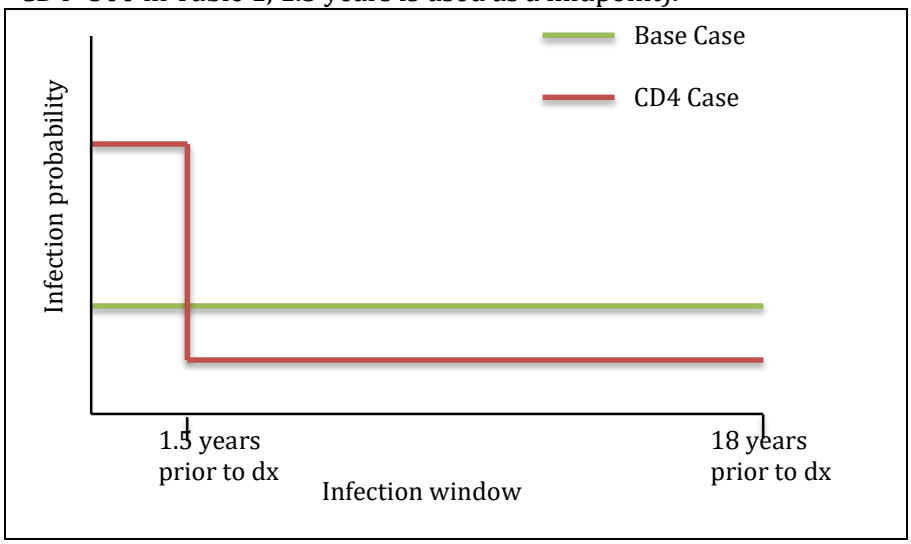
2. For those with "long" infection windows (defined in point #3), the CD4 Case will shift 50% of the probability of infection into the median time from infection to CD4 category at diagnosis (Table 1), as depicted in the example in Figure 1.

Table 1. Median times from infection to CD4 count

CD4 category	Years by which ~50% of cases reach lower bound (Ref 1)	Median years from infection to CD4 category (Ref 2)*
>500	1.19	2.3
350-500	4.19	4.17
200-350	7.93	7.98

*Based on translating the estimated transition rates between CD4 compartments into median times spent in each compartment

Figure 1. Change in infection probability distribution for a case with an 18-year window but a CD4 count above 500. 50% of the probability would be distributed across 0-1.5 years prior to diagnosis; the remaining 50% would be distributed across 1.5-18 years prior. (Note: given the discrepancy in median times for CD4>500 in Table 1, 1.5 years is used as a midpoint).



The remaining 50% of infection probability will be distributed between that median time and the end of the infection window. For the CD4>500 group, we will do a sensitivity analysis using each of the suggested medians. If the impact is minor, we will use 1.5 years as a midpoint estimate, and for simplicity we use 1.5 in this proposal.

3. The cases with "short" infection windows will receive the typical Base Case probability distribution, uniformly distributed between diagnosis and LNT. "Short" windows are those for which the uniform probability distribution places more infection probability closer to diagnosis than the CD4-based approach. The latter places 50% probability of infection within the median time to CD4 count. The former places 50% probability within the midpoint of the window. Cases with windows shorter than 2 times the median time to CD4 will thus have more probability of infection near diagnosis using a uniform distribution. For example, under the Base Case, a case with CD4>500 with an LNT 2 years prior to diagnosis would have 50% of their infection probability fall within 1 years prior to diagnosis, and the CD4-based median is 1.5 years. This case would need an infection window of 3 years or more in order to have a "long" infection window for their CD4 category.

B. Expected Impact

The CD4 Case will increase in the precision of our estimates, in terms of our confidence that individuals with long infection windows are being given reasonable infection scenarios. However, the magnitude of impact on the undiagnosed estimates depends on the fraction of the population with long windows and high CD4 counts. In this population, the impact of the CD4 Case will likely be small to moderate.

There are 3016 cases in 2005-2014 with testing history (59% of the full sample, N=5148). As noted above, about 10% of these cases would have their infection probability shifted closer to the time of diagnosis under the CD4 Case. There is greater potential for impact in populations with a greater percentage of individuals reporting never having had a last negative test. There is thus a slightly higher impact among non-MSM (17%) than among MSM (8%) (Table 3). Table 3 also shows that about 13% of cases contributing to the TID get the 18-year window assumption due to never having had a negative test. Only 30% of those cases would be impacted by the CD4 Case, given the high proportion of these cases with low CD4 counts (Table 4).

Table 3. Sample sizes and distributions of cases affected by the CD4 data inclusion, among the 3016 cases with non-missing testing history. Column percents unless indicated otherwise.

	By Exposure Group		*By Window Boundary*		
	MSM	Non-MSM	LNT	18years	age-16
Affected					
CD4>500 and window> 3 years	112 (5%)	60 (8%)	106 (4%)	34 (9%)	32 (14%)
CD4 350-500 and window>8 years	35 (2%)	31 (4%)	14 (0.6%)	38 (10%)	14 (6%)
CD4 200-500 and window>16 years	24 (1%)	34 (5%)	6 (0.3%)	47 (12%)	5 (2%)
Cumulative %	8%	17%	5%	30%	23%
Not affected					
Window < CD4 median estimate	1487 (66%)	395 (53%)	1611 (67%)	173 (44%)	98 (43%)
CD4 missing or > 30d	611 (27%)	227 (30%)	657 (27%)	104 (26%)	77 (34%)
Column Totals					
N	2269	747	2394	396	226
% of total	75%	25%	79%	13%	7%

Table 4. CD4 distributions by window boundary, including missing testing history (N=5148). Column percents shown.

	Window of possible infection based on:			
CD4 category	LNT	18 years	age-16	Missing
>500	26%	9%	15%	17%
350-500	18%	10%	11%	11%
200-350	14%	12%	19%	13%
0-200	15%	44%	21%	26%
Missing*	27%	26%	34%	34%
Column Total N	2394	396	226	2132

* includes cases not measured within 30 days of diagnosis

References

1. Lodi S, Phillips A, Touloumi G, Geskus R, Meyer L, Thiébaut R, et al. Time from human immunodeficiency virus seroconversion to reaching CD4+ cell count thresholds <200, <350, and <500 Cells/mm^3: assessment of need following changes in treatment guidelines. Clin Infect Dis Off Publ Infect Dis Soc Am. 2011 Oct;53(8):817–25.

2. Cori A, Pickles M, van Sighem A, Gras L, Bezemer D, Reiss P, et al. CD4+ cell dynamics in untreated HIV-1 infection: overall rates, and effects of age, viral load, sex and calendar time. AIDS Lond Engl. 2015 Nov 28;29(18):2435–46.