Adding CD4 data to the testing history method

# Goal

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

# Background

Currently, the testing history method has long infection windows for 1) cases reporting never having had a negative test and 2) cases reporting a long infection window, i.e. a last negative test (LNT) may years prior to diagnosis. Some of these cases may in fact have been infected closer to the time of diagnosis, and this may be reflected in their CD4 and/or viral load. However, the literature suggests that using CD4 and viral load to predict an individuals’ stage of infection is complicated. Viral load is complicated due to the variation in set-point viral load (SPVL) across persons, and the fact that it peaks twice, once during the acute stage (indicating recent infection) and again at the start of the late stage (indicating less recent infection). CD4 is somewhat less complicated, because the CD4 trajectories are monotonically decreasing after infection, even if their overall level and rate of decline is correlated with SPVL. While neither indicator is perfect, there is some suggestion in the literature that CD4, when used at the aggregate level, may provide a better foundation for back calculation. We therefore propose to explore whether integrating CD4 data into the testing history method would provide better estimates of the undiagnosed fraction.

# Approach

We will use CD4 to create a “CD4 Case” alternative to the Base Case, in which we shift the probability of infection within 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. Our procedure will also account for the heterogeneity in CD4 levels and trajectories by using infection probability distributions that reflect the degree of SPVL variation observed in cohort studies of CD4 trajectories.

The procedure outlined below is 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)).

## Proposed procedure:

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). |
| 18 years prior to dx  1.5 years prior to dx  Infection window  Base Case  CD4 Case  Infection probability |

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.

1. 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.

# Expected Impact

There are 3016 cases in 2005-2014 with testing history (59% of the full sample, N=5148). Their breakdown by CD4-count and infection window length is shown in Figure 2. About 10% of these cases that contribute data to the TID would have their infection probability shifted closer to the time of diagnosis under the CD4 Case. There is 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).

The CD4 Case represents a valuable 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 *in this population* will likely be minor. There is greater potential for impact in populations with a greater percentage of individuals reporting never having had a last negative test.

|  |
| --- |
| **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). |
| **Impacted (N=296, 10%)**  CD4>500 and window>3 year  (N=172, 6%)  CD4 350-500 and window>8 years  (N=66, 2%)  CD4 200-500 and window>16 years  (N=58, 2%)  Window shorter than 2x CD4-based median  (N=1882, 62%)  CD4 measured within 30d of diagnosis  (N=2178, 70%)  Reported a LNT or No LNT  (N=3016, 100%) |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 3**. Sample sizes and distributions of cases impacted by the CD4 Case by MSM status and LNT status, among the 3016 cases with non-missing testing history. Percents are column percents unless indicated otherwise. | | | | | |
|  | *By MSM Status* | | *By LNT Status* | | |
|  | MSM | Non-MSM | LNT | LNT (18y) | LNT (age-16) |
| **Impacted** |  |  |  |  |  |
| 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 impacted** |  |  |  |  |  |
| Window shorter than CD4-based median | 1487 (66%) | 395 (53%) | 1611 (67%) | 173 (44%) | 98 (43%) |
| CD4 missing or not within 30d | 611 (27%) | 227 (30%) | 657 (27%) | 104 (26%) | 77 (34%) |
| **Column Total** |  |  |  |  |  |
| N | 2269 | 747 | 2394 | 396 | 226 |
| As a % of the  3016 cases | 75% | 25% | 79% | 13% | 7% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 4**. CD4 distributions by LNT status, including missing testing history (N=5148). Percents are column percents. | | | | |
|  |  |  |  |  |
| CD4 category | LNT | LNT (18y) | LNT (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 or not measured within 30d | 27% | 26% | 34% | 34% |
| Column Total N | 2394 | 396 | 226 | 2132 |

# Time Estimate

|  |  |
| --- | --- |
| Program and debug the new CD4 Case | 25h |
| Analyze and present results | 25h |
| Total | 50h |

# 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/mm3: 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.