Integrating CD4 data into undiagnosed estimates

author: Martina Morris & Jeanette Birnbaum date: autosize: true transition: fade width: 1920 height: 1080

Project Goal

To use CD4 data to *increase the precision* of undiagnosed estimates from the testing history model

Why - Negative tests that are many years prior to diagnosis, i.e. "long infection windows", are not very informative regarding time of infection - CD4 at diagnosis can indicate recency of infection

How - From literature, identify typical times to infection for various CD4 counts - Use this to update the probability of infection within long infection windows

Impact in WA - Our pre-analysis of CD4 measurements in WA indicates that using CD4 will increase the precision of our undiagnosed estimates but will not significantly change the estimated number of undiagnosed cases - The main impact of integrating CD4 will be to increase our confidence in the estimates

Reminder: the testing history method

type:sub-section

Base Case distributes probability uniformly across the infection window

The hazard (instantaneous rate) of infection at any point within the window is 1/(window length), shown by the red line for two different window lengths. Time=0 refers to time of diagnosis.

The red line is lower when the window is longer, since the probability of infection is spread out over a longer time period.

Longer windows have less probability assigned to the recent period

Cumulative probability of diagnosis is the area under the curve. At 1 year prior to diagnosis, for example, the 2-year window assigns greater probability of infection than the 4-year window.

Since 1 year is 1/2 of 2 years, 50% of infection probability occurs within 1 year for the 2-year window. In contrast, only 1/4 or 25% of infection probability occurs within 1 year for the 4-year window.

Never-tester assumption gives a window of age-16 or 18 years

For never-testers, we assume a window of 18 years or age-16, whichever is smaller.

Example of windows for never-testers of different ages

The age distribution of never-testers thus influences what window lengths they contribute to the population TID (time from infection to diagnosis).

Reminder: testing histories in WA, 2005-2014

type:sub-section

59% of cases have testing history (2005-2014)

Testing histories = Tested + Never Tested

Breakdown of testing histories in WA HIV cases 2005-2014.

Total N=5148 N with testing history N=3016.

43% of non-MSM have testing histories, versus 66% of MSM

Testing history breakdown by MSM status. Percents are within MSM status group. Total N=5148.

MSM comprise about 2/3 of cases. What is driving differential response rates?

Half the windows are 2 years or shorter

Distribution of infection window lengths in years (N=3016). Labels indicate bounds, e.g. (0,1] includes windows>0 and <=1. The 18-year windows imputed because of the assumption for never-testers are labeled as 18 (NT).

50% have windows between 0 and 2 years (red plus gold). Our assumption for never-testers almost doubles the number of windows >=5 years (blue plus purple).

More than half of non-MSM windows are over 5 years

Distribution of infection window lengths for non-MSM (top) and MSM. Percents are within MSM status groups.

The 18-year assumption for never-testers is almost 4x more common among non-MSM (31%) than among MSM (7%).

Longer non-MSM windows reflect more concurrent dx

Percent of cases with a concurrent AIDS diagnosis at time of HIV diagnosis, by MSM status.

Longer non-MSM windows also reflect older never-testers

Age distribution of never-testers in WA 2005-2015, by MSM status (pink=MSM).

Mean age at diagnosis among never-testers is 42 in non-MSM, versus 36 in MSM Remember the assumption for never-testers: window=min(age-16,18). Cases older than 16+18=34 will all have 18-year windows.

In WA, non-MSM are where precision can be increased

Whether it's using BED, CD4, or concurrent AIDS dx, etc - Short testing intervals in MSM = low uncertainty regarding time of infection

Increasing precision will not necessarily decrease undiagnosed estimates - More precise estimates will deviate from the Base Case inasmuch as the additional information changes the probability of infection within the window

Total population estimates will not change greatly, but our confidence in them will increase - If people with long windows tested due to risky exposure, CD4 should help pick up on that

Methods: integrating CD4 into the testing history method

type: sub-section

Under Base Case, long windows = low probability of recent infection

Window Length	If CD4 unknown, infection probability is mostly:
Short	Recent
Long	Non-recent

Long window but high CD4 indicates a higher likelihood of recent infection

Window Length	If CD4 unknown, infection probability is mostly:	CD4=Low	CD4=High
Short	Recent	Recent (fast progressor?)	Recent
Long	Non-recent	Non-recent	**Recent**

How to define "short" and "long", "low" and "high"?

Median times to CD4 from the literature

Years it takes for 50% of cases to reach CD4 level:

CD4 Category	Lodi 2011	Cori 2015	Our approximation^*
>500	1.19	2.3	1.5
350-500	4.19	4.17	4
200-350	7.93	7.98	8
0-200		11.5	9

- Simplified median times that are consistent with both sources. For CD4 0-200, we will retain our maximum-window assumption of 18 years. This sets a maximum median time of 9 years

Interpretation: 50% of infection probability should occur within the median time. **Implication**: If a case has CD4>500 at diagnosis, for example, we expect 50% of his infection probability to occur in the 1.5 years prior to dx

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 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.

Strategy: re-allocate infection probability given CD4 at diagnosis

Base Case: 50% infection probability is in each half of the window

In the example of a window of 18 years, 50% probability of infection is spread across years 0-9 prior to dx

Strategy: re-allocate infection probability given CD4 at diagnosis

CD4 Case: 50% of infection probability is shifted into the CD4-based median window

For a window of 18 years and various CD4 bins at diagnosis, the graphs show the Base Case (red) versus CD4 Case (blue) distribution of the probability of infection within the window. Shading indicates the new placement of 50% of infection probability.

Impact will come from high CD4s in people with "long" windows

CD4 Category	CD4 Median	Impacts windows longer than
>500	1.5	3
350-500	4	8
200-350	8	16

Degree of impact will depend on how much longer windows are than 2x the CD4-based median * 18-year windows among individuals with CD4>500 will provide much greater impact than 4-year windows, for example * Even 18-year windows will have minimal impact among CD4 200-350, since the Base Case median for 18-year windows is 9 years and the CD4-based median is 8 years (not much difference)

Results: CD4 in WA cases

type: sub-section

72% of cases with testing history also have useable CD4 data

What is happening with the other 28%?

Distribution of time of CD4 measurement (x-axis, in days since diagnosis) and CD4 counts (colors) among the 838 individuals who did not have a CD4 count within 30 days. High CD4s in delayed measurement groups may reflect initiation of treatment

Low CD4 is correlated with long windows, but with high variability

Scatterplot of first CD4 count versus infection window length, among CD4-eligible cases (N=2178). Panels are testing status (columns) and MSM status (rows). * Never testers get a window length of min(age-16, 18 years). * Low CD4 and long windows both imply low probability of recent infection

Low CD4 is more common in never-testers than testers

CD4 distribution density by testing history status (colors) and MSM status (panels). Never-testers have CD4 distributions that are much more skewed towards low CD4.

However, some never-testers do have high CD4

CD4 distributions by testing history status (panels) and MSM vs non-MSM status (colors)

Never-testers with high CD4 are where the CD4 Case will have greatest impact. Never-testers with CD4 200-350 will have minimal impact.

10% of cases have CD4 indicating more recent infection probability than Base Case assumes

CD4 Category	CD4 Median	Impacts windows longer than
>500	1.5	3
350-500	4	8
200-350	8	16

CD4 impacts 7% of MSM versus 17% of non-MSM

The lowest-impact group are the CD4 250-300 cases with the longest windows * MSM: 1%, or 1/7 = 14% of impacted cases * non-MSM: 5%, or

5/17 = 29% of impacted cases

Results: CD4 impact on median infection point

type: sub-section

Average median infection point decreases by 0.1-0.8 years using CD4 Case

Median infection point = time by with 50% of infection probability has occurred

Among all cases with testing history (N=3016), average median infection points under the Base Case (orange) and CD4 Case (green), by CD4 bin. More details included in an appendix slide

Among impacted cases, MSM have larger decreases in median infection points

Among impacted cases (N=296), differences in the median infection points under the Base Case versus CD4 Case. Lines indicate the 50% percentile. The tall bins reflect the never-testers. Their 9-year median decreases to 8 years for CD4 200-450 (difference=1 year), to 4 years for CD4 350-500 (difference=5 years), and 1.5 years for CD4>500 (difference=7.5). 50th percentile decrease is larger for MSM (2.34 years) than for non-MSM (1.94 years)

Results: CD4 impact on TID estimate

type: sub-section

CD4 Case decreases % undiagnosed at 1 year by 1-2 percentage points

1 year after diagnosis

Population	% Undx at 1 year, Base Case	% Undx at 1 Year, CD4 Case	Difference	Percent Change
Total	40.8	39.4	1.4	3.4
MSM	33.1	31.9	1.2	3.6
non-MSM	64.3	62.0	2.3	3.6

5 years after diagnosis

Population	% Undx at 5 years, Base Case	% Undx at 5 Years, CD4 Case	Difference	Percent Change
Total	16.4	15.0	1.4	8.5
MSM	11.2	10.2	1.0	8.9
non-MSM	32.0	29.7	2.3	7.2

Average time undiagnosed decreases by about 6% in both groups

Population	Mean time undx, Base Case	Mean time undx, CD4 Case	Difference	Percent Change
MSM	1.829	1.717	0.1122	6.13
non-MSM	4.378	4.127	0.2511	5.74

TID actually changes by 0.39 more percentage points for MSM than non-MSM

Overall impact on TID is subtle

MSM *Time from infection to diagnosis: probability curve (top) and undiagnosed fraction curve (bottom)*

Overall impact on TID is very subtle

non-MSM *Time from infection to diagnosis: probability curve (top) and undiagnosed fraction curve (bottom)*

Results: CD4 impact on undiagnosed estimates

type: sub-section

Mean undiagnosed estimates decrease 5-6% in 2014

Population	Base Case	CD4 Case	Difference	Percent Change
Total	1319.0	1247.0	72.0	5.5
MSM	604.7	568.4	36.3	6.0
non-MSM	714.3	678.2	36.1	5.1

Subgroup decreases closely parallel the decreases in mean undiagnosed time (6.13% for MSM and 5.74% for non-MSM)

Undiagnosed fractions decrease 0.5-0.7 percentage points in 2014

Population	Base Case	CD4 Case	Absolute Difference	Percent Change
Total	9.4	8.9	0.5	5.3
MSM	6.2	5.8	0.4	6.5
non-MSM	17.1	16.4	0.7	4.0

The fractions take into account diagnosed PLWH. Since the Base Case undiagnosed fraction is much higher for non-MSM, it is less sensitive than the MSM fraction to the CD4 Case's decrease of about 36 undiagnosed cases.

Conclusions, limitations and possibilities

type: sub-section

CD4 data indicate that most WA cases with long windows were not recently infected

Window Length	If CD4 unknown, infection probability is mostly:	CD4=Low	CD4=High
Short	Recent	Recent (fast progressor?)	Recent
Long	Non-recent	Non-recent	**Recent**

In WA, only 10% of cases with testing history had infection windows that indicated less probability of recent infection than indicated by their CD4 count

CD4 Case did not impact subgroups differentially

We expected to see greater impact in non-MSM than MSM * non-MSM did have a higher fraction of cases impacted (17% for non-MSM versus 7% for MSM) * But among the impacted cases, the MSM had larger decreases in their median infection points (50th percentiles: 1.94 years for non-MSM versus 2.34 years for MSM) * This translated into slightly higher impacts on mean TID and mean undiagnosed estimates for MSM

Limitations

We prioritize testing history data * If testing history indicates more recent infection than CD4 does, we use the testing history * Little is known about the accuracy of self-reports of the last negative test before diagnosis

We use CD4 conservatively * Literature on CD4 trajectories is fairly sparse * Data indicate plenty of heterogeneity, hence the need for using a fairly conservative approach for the CD4 Case

Future work

CD4 in cases with missing testing history * Use CD4 to examine the "missing at random" assumption among cases with missing testing history

Multiple markers * Use a Bayesian approach to combine information from BED, CD4 and concurrent diagnoses

Appendix: Slide Bank

type: section

Reminder: 58% of cases have testing history

Among testing histories, 79% report have a prior negative test

Testing histories come disproportionately from MSM

MSM are 66% of all cases.

Recall that never-testers get a window that is either age-16 or 18 years.

Median infection point decreases by 0.1-0.8 years

Median infection point = time prior to dx within which 50% of infection probability occurs

CD4 Bin

CD4 Bin Size

Number Impacted

% of CD4 Bin Impacted

Base Case Median

CD4 Case Median

Difference

Percent Change

Proportion in CD4 Bin

[0,200)

582

0.0

0.0

4.8

4.8

0.0

0.0

0.2

[200,350)

423

58.0

13.7

2.6

2.5

0.1

3.8

0.1

[350,500)

482

66.0

13.7

1.8

1.3

0.5

27.8

0.2

[500,2e+03]

691

172.0

24.9

1.5

0.7

0.8

53.3

0.2

The impacted cases represent (172+66+58)/3016 = 10% of the cases contributing testing histories.

Median windows decrease more in non-MSM

Mode

CD4 Bin

CD4 Bin Size

Number Impacted

% of CD4 Bin Impacted

Base Case Median

CD4 Case Median

Difference

Percent Change

MSM

[0,200)

369

0.0

0.0

4.0

4.0

0.0

0.0

MSM

[200,350)

317

24.0

7.6

1.9

1.8

0.1

5.3

MSM

[350,500)

396

35.0

8.8

1.3

1.0

0.3

23.1

MSM

[500,2e+03]

576

112.0

19.4

1.3

0.6

0.7

53.8

non-MSM

[0,200)

213

0.0

0.0

6.1

6.1

0.0

0.0

non-MSM

[200,350)

106

34.0

32.1

4.7

4.4

0.3

6.4

non-MSM

[350,500)

86

31.0

36.0

3.9

2.3

1.6

41.0

non-MSM

[500,2e+03]

115

60.0

52.2

2.9

1.1

1.8

62.1

non-MSM median windows decrease up to an average of 1.8 years

Among all cases with testing history (N=3016), median infection points by Case and MSM status. The impacted MSM are 7% of MSM testing histories, whereas the impacted non-MSM are 17% of all testing histories.