Late Diagnosis Exploration Using WA State Data

Martina Morris & Jeanette Birnbaum

07 April, 2017

Contents

1 Summary of Results

Late diagnoses are ones defined by an AIDS diagnosis within 1 year of the HIV diagnosis. The goal of this project was to integrate late HIV diagnoses into the testing history method for estimating undiagnosed cases of HIV, and to report the impact of the new information on the estimate. We did this by stratifying the estimation by late diagnosis status, which allows each stratum to have a different, and more accurate, distributions of time from infection to diagnosis (TID).

- The adjusted undiagnosed estimates fall within 5% of the original estimates for most years, a relatively small change. While stratifying by late diagnoses paints a clearer picture of the late and non-late strata, the overall results are similar to a weighted average of the stratum-specific results, which is what the original analysis produces.
  - The original analysis appears to reasonably reflect the contribution of late diagnoses to the overall average TID and undiagnosed cases. The original analysis stratifies by MSM/non-MSM status, which is highly correlated with late diagnoses. Thus the additional stratification by late diagnoses has only minor influence.
- The direction of the change in estimates varies across years. In earlier years, incorporating late diagnoses increases undiagnosed estimates slightly, but in later years, it decreases undiagnosed estimates. We believe this is because it picks up on a declining trend in late diagnoses that was masked in the unstratified analysis.

1.1 Estimates of undiagnosed HIV (using Base Case assumptions)

Table 1: Total Population Undiagnosed

	Cases			Fraction		
Year	Original	With Late Dx Strata	% Change	Original	With Late Dx Strata	% Change
2010	1483	1528	3.0	12.0	12.3	2.5
2011	1449	1475	1.8	11.6	11.8	1.7
2012	1401	1405	0.3	11.1	11.1	0.0
2013	1355	1339	-1.2	10.5	10.4	-0.9
2014	1330	1290	-3.0	10.3	10.0	-2.9

Table 2: MSM Undiagnosed

	Cases			Fraction		
Year	Original	With Late Dx Strata	% Change	Original	With Late Dx Strata	% Change
2010	707	750	6.0	7.9	8.4	6.3
2011	684	719	5.2	7.6	7.9	4.0
2012	656	677	3.2	7.2	7.4	2.8
2013	630	637	1.0	6.7	6.8	1.5
2014	616	602	-2.1	6.5	6.4	-1.5

Table 3: non-MSM Undiagnosed

	Cases			Fraction		
Year	Original	With Late Dx Strata	% Change	Original	With Late Dx Strata	% Change
2010	776	778	0.3	22.5	22.6	0.4
2011	766	756	-1.3	22.1	21.9	-0.9
2012	745	728	-2.3	21.4	21.0	-1.9
2013	725	702	-3.1	20.6	20.1	-2.4
2014	715	688	-3.8	20.1	19.5	-3.0

2 Data description and basic trends

2.1 Analytic sample

Analytic data set has 5,149 cases for 2005-2014. We note that 79 cases who were 16 or younger and had no observed date of LNT were excluded from the original dataset, along with diagnoses prior to 2005. We limit this analysis to diagnoses occuring by 2014 so that all cases have the opportunity to present with AIDS within 1 year of diagnosis, our definition of late diagnosis.

2.2 Diagnoses by MSM/Late Dx subgroups

The figure below shows that 33% of all diagnoses in 2005-2014 were late diagnoses, but non-MSM have disproportionately more late diagnoses than MSM.

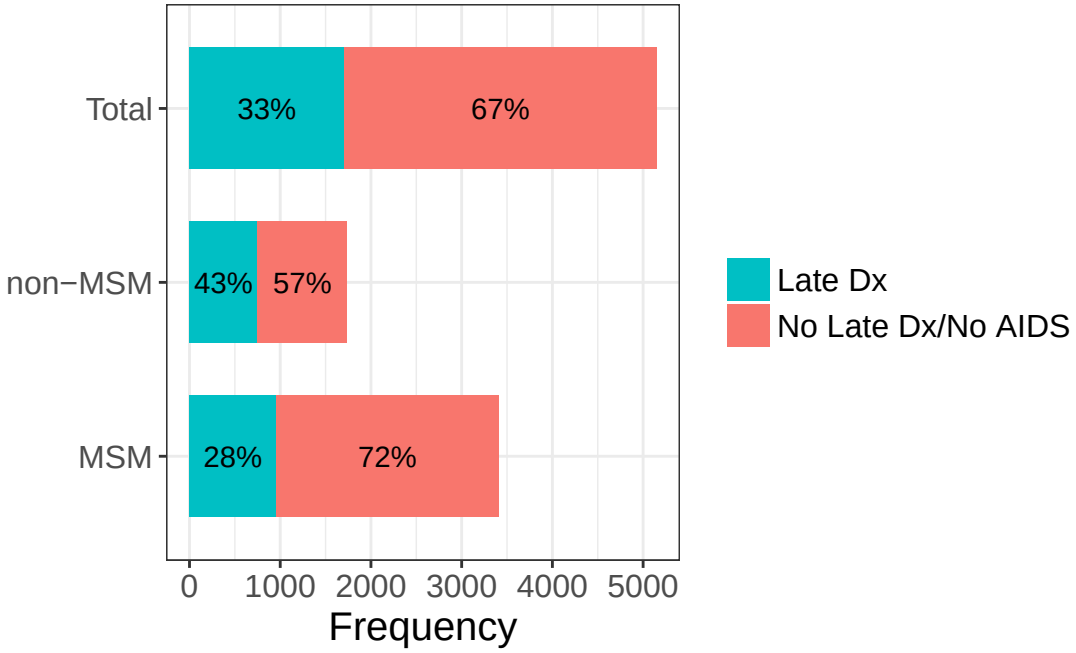


Figure 1: Number of diagnoses overall (top bar) and separately by non-MSM and MSM subgroups. Colors indicate late diagnosis and percentages sum to 100 percent within each bar.

2.3 Diagnoses over time by MSM/Late Dx subgroups

When we look over time, we see a slight downward trend in absolute numbers of late diagnoses as well as the percent of diagnoses that are late. This will impact our estimation in the late diagnoses strata - the method will model a decline in incidence to match this trend.

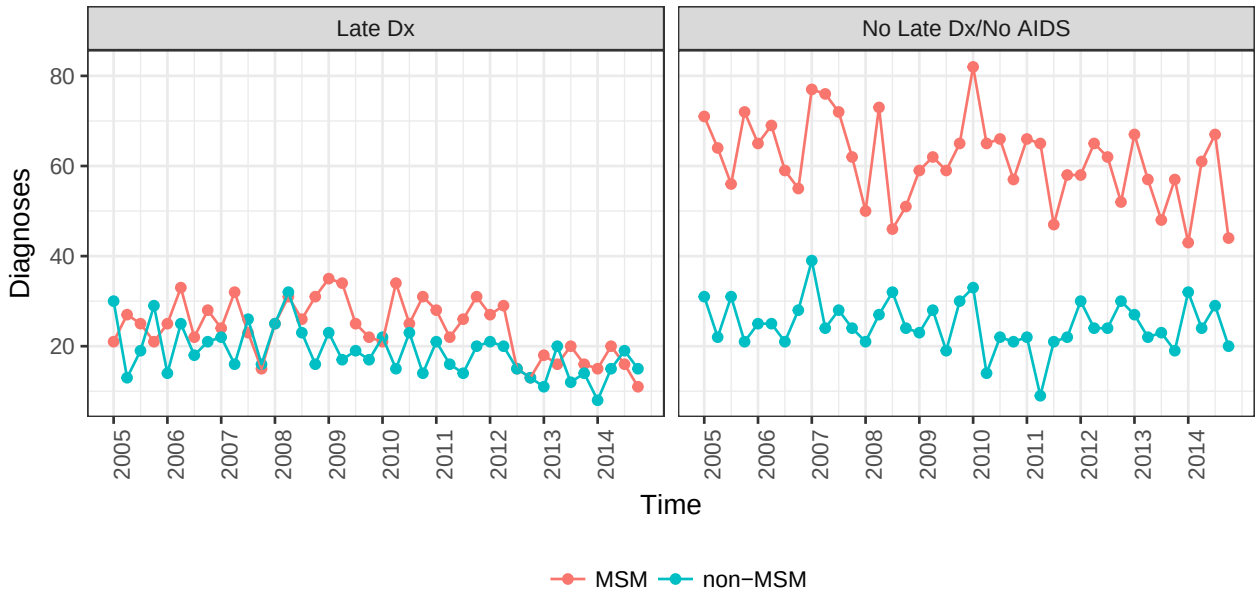


Figure 2: Diagnoses per quarter by MSM/non-MSM subgroups. Colors indicate late diagnoses.

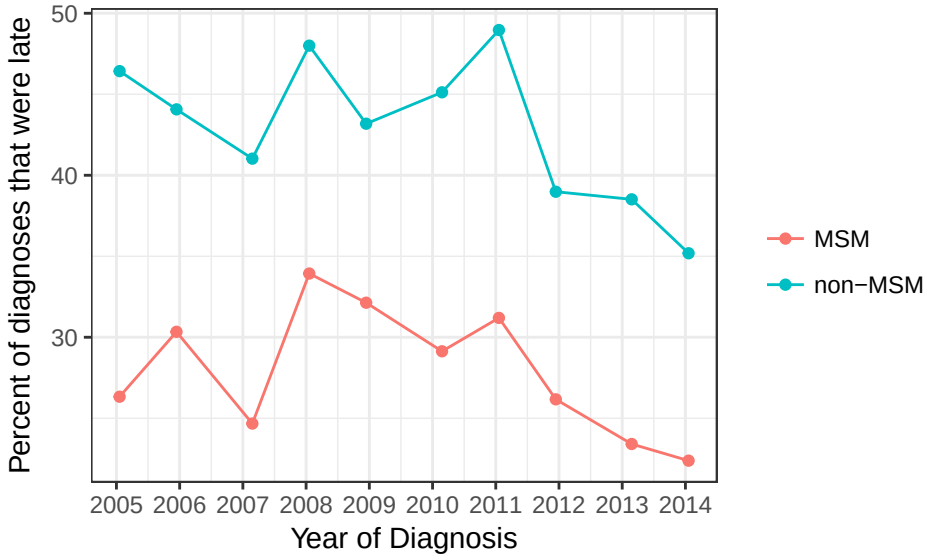


Figure 3: Within MSM/non-MSM subgroups, percent of diagnoses that were late

3 Testing Histories

3.1 Responses to "Ever had a negative test?"

The table below gives the breakdown of testing histories in the sample, overall (last row) and by MSM/late diagnosis subgroups. The subgroup data are also shown in the following bar plot, where the percentages displayed in the bars correspond to the row percentages in the table.

Late diagnoses are more likely to have missing testing history and more likely to have no testing history, for both MSM and non-MSM subgroups. However the contrast is much greater for MSM, e.g. the 6% reporting no prior negative test more than doubles to 17% among late diagnoses.

Given our missing-at-random assumption for testing history, the differential patterns of never-testing are potentially significant. When we do not stratify the estimation by late diagnosis, we assume that the late diagnoses have the same TID as the rest of the cases (within MSM/non-MSM subgroups). The greater ratio of No:Yes testing history among late-diagnosed MSM suggests otherwise. Stratifying the estimation by late diagnosis status thus allows us to be more accurate in how missing data are handled.

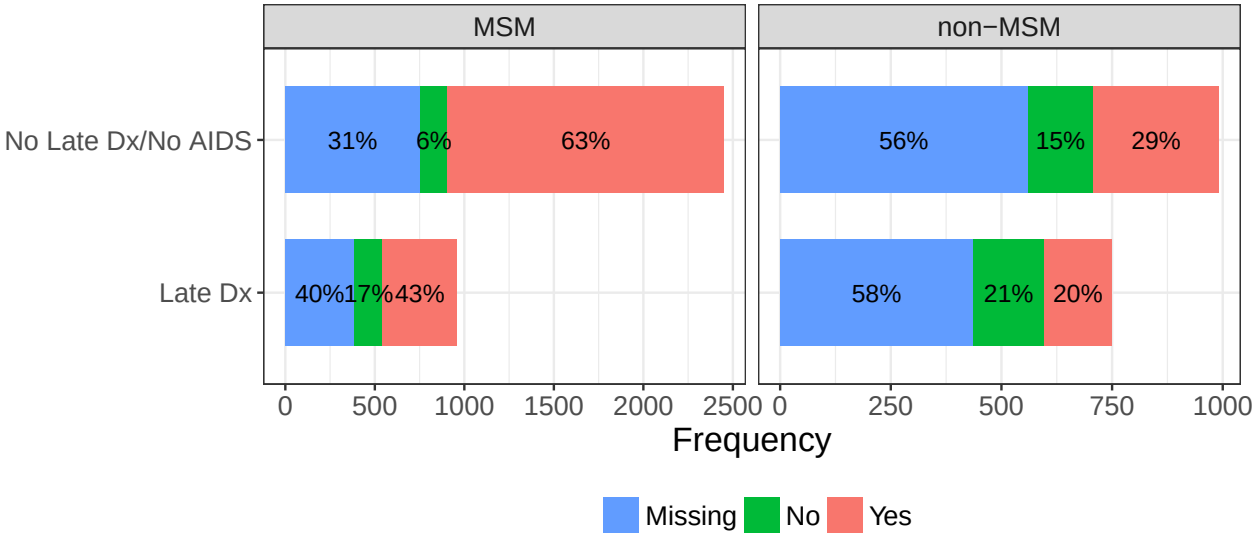


Figure 4: Testing history response frequencies by MSM and late diagnosis subgroups.

Table 4: Breakdown of testing histories in the sample. Column percent shows the composition of the total sample. The Percent Yes, Percent No and Percent Missing columns indicate the row percents of the three possible testing history statuses within each subgroup

				Ever Had a Negative Test		
Mode	Late Diagnosis	N	Column %	% Yes	% No	% Missing
MSM	No	2450	48	63	6	31
	Yes	959	19	43	17	40
non-MSM	No	991	19	29	15	56
	Yes	749	15	20	21	58
All	All	5149	100	46	12	42

3.2 Reported infection windows

The plot below shows how reported testing windows are distributed between 0-1, 1-5, and 5-18 year categories within MSM and late diagoses subgroups (percents sum to 100 within each subgroup). We see that shorter windows are less common among late diagnoses, but the contrast is much greater in non-MSM. Close to half of non-MSM late diagnoses with non-missing testing history are never-testers for whom the min(age-16, 18) assumption gives them an 18-year window. These infection window intervals suggest that late diagnoses will have very different TIDs from the non-late diagnoses.

We note that the literature suggests that only 10% of HIV cases are "rapid progressors" who progress to AIDS within 2-3 years [1][2]. It thus seems unlikely that 34% of MSM late diagnoses have progressed from HIV to AIDS within 2.5 years, but that is 34% *of late diagnoses with non-missing testing history*. Of the 959

[1]Haynes BF, Pantaleo G, Fauci AS. Toward an understanding of the correlates of protective immunity to HIV infection. Science. 1996 Jan 19;271(5247):324–8.

[2]According to the CDC's gamma(2,4) AIDS incubation distribution, 10% of cases will progress to AIDS in 2.12 years; only 2.6% within 1 year.

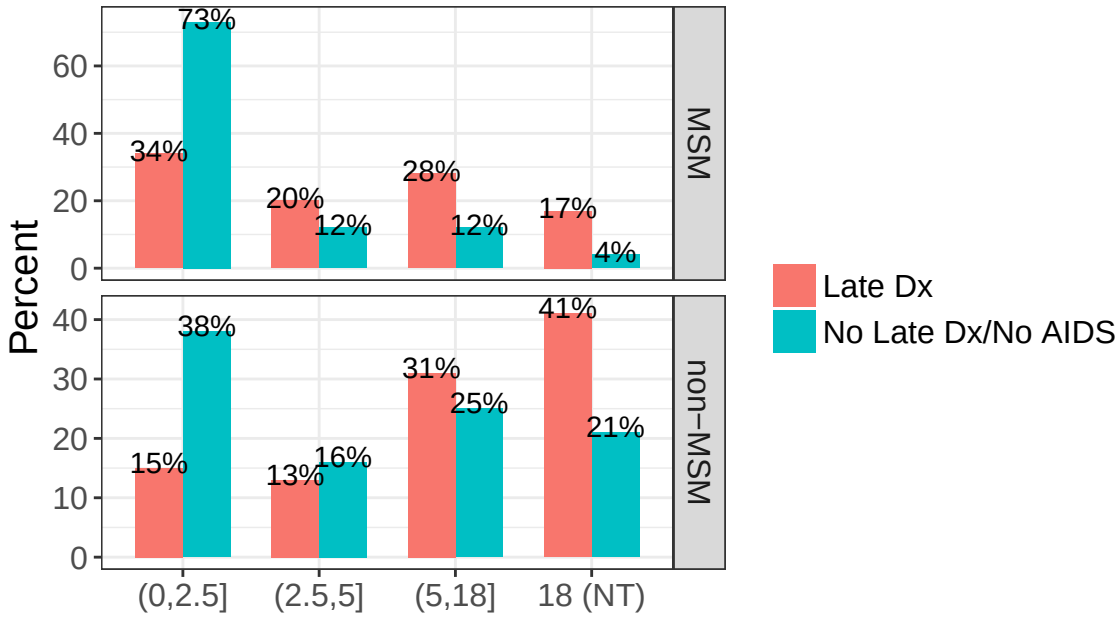
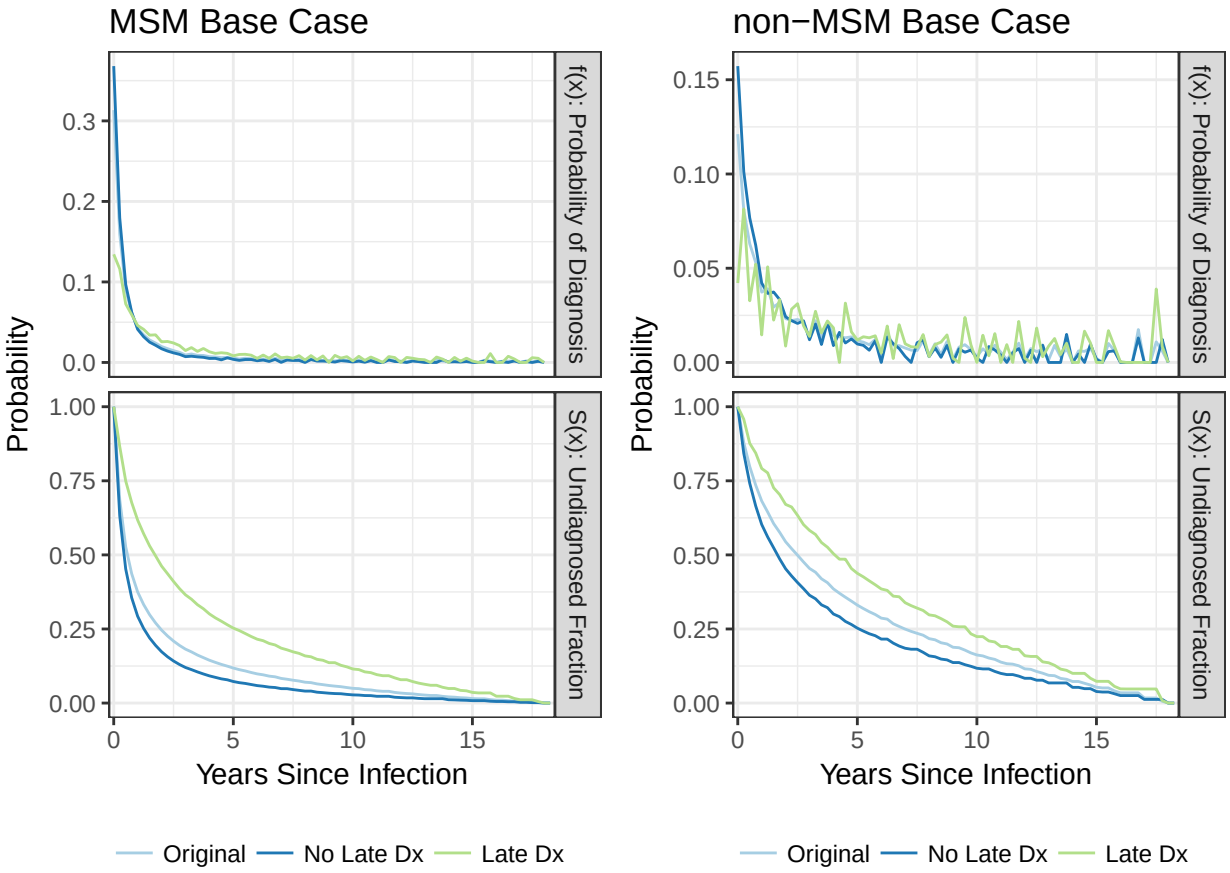


Figure 5: Inter-test interval window lengths among those with non-missing testing history. 18 (NT) refers to the 18-year windows that arise from never-testers

MSM late diagnoses, 40% have missing testing history. Given the typical times to AIDS progression, these late diagnoses with missing testing history probably have longer inter-test intervals than the late diagnoses who do report testing history.

4 Time from Infection to Diagnosis (TID)

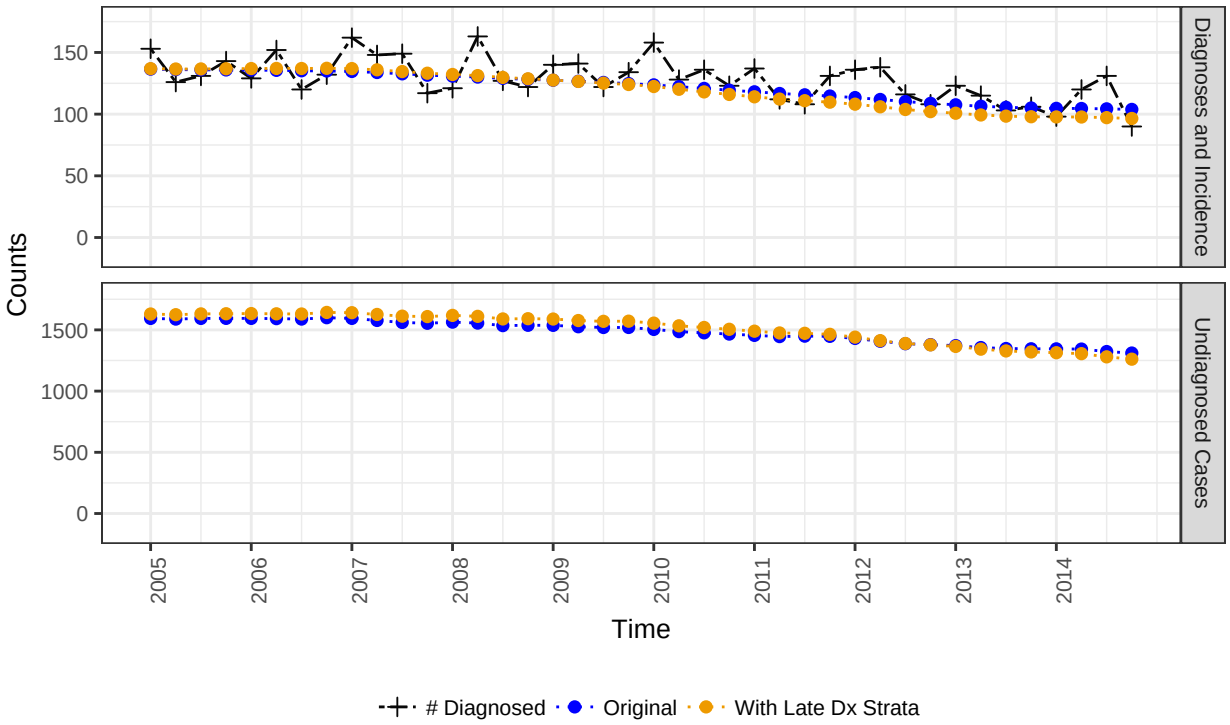
The Base Case TID plots show how the longer infection windows for late diagnoses that we saw in the previous section translate into TID distributions that are shifted towards longer times between infection and diagnosis. For MSM, the percent remaining undiagnosed at 5 years is about 25% for late diagnoses but only about 10% for others. In non-MSM at 5 years, over 40% of late diagnoses remain undiagnosed compared to only 25% of others.



Comparing the stratified TIDs to the original single TID highlights the fact that the single TID is essentially a weighted average of the stratified ones. Thus, results for the total population may be similar if patterns of missing data are not very different across the strata. This is because the stratified approach applies the stratified TIDs to their respective strata, including cases with missing testing history. If the late-diagnoses strata had longer TIDs and disproportiately more cases with missing testing history, then more diagnoses would get longer TIDs than in the unstratified approach. Table 4 suggests that this is not the case in WA State; within MSM/non-MSM strata, the percent of cases missing testing history is similar for late and non-late diagnosis groups.

5 Incidence and undiagnosed counts

The upper plot panels show diagnoses and estimated base case incidence for the original and adjusted methods, while the lower plot panels show base case undiagnosed estimated. While both methods estimate declining incidence and undiagnosed counts, incorporating late diagnoses leads to a slightly steeper decline. This is presumably because the stratified estimation is more sensitive to the decline in late diagnoses in the most recent years.



6 Conclusions and Future Work

While stratifying the testing history by late diagnosis status allows us to be more precise, particularly with our assumption that testing histories are missing at random, the overall impact on undiagnosed estimates in WA State is minor. The original approach represents undiagnosed in the total population well. A useful next step would be to more closely investigate the characteristics of cases with missing testing history and test for statistically significant differences from cases with testing history. These results could inform an application of standard missing data adjustments to correct for missing testing history bias.

7 Tables of all undiagnosed HIV estimates (including Upper Bound)

Table 5: Base Case: Total Population

	Cases			Fraction		
Year	Original	With Late Dx Strata	% Change	Original	With Late Dx Strata	% Change
2010	1483	1528	3.0	12.0	12.3	2.5
2011	1449	1475	1.8	11.6	11.8	1.7
2012	1401	1405	0.3	11.1	11.1	0.0
2013	1355	1339	-1.2	10.5	10.4	-0.9
2014	1330	1290	-3.0	10.3	10.0	-2.9

Table 6: Upper Bound: Total Population

	Cases			Fraction		
Year	Original	With Late Dx Strata	% Change	Original	With Late Dx Strata	% Change
2010	2850	2805	-1.6	20.8	20.5	-1.4
2011	2790	2717	-2.6	20.2	19.8	-2.0
2012	2709	2614	-3.5	19.5	18.9	-3.1
2013	2639	2521	-4.5	18.6	17.9	-3.8
2014	2593	2451	-5.5	18.2	17.4	-4.4

Table 7: Base Case: MSM

	Cases			Fraction		
Year	Original	With Late Dx Strata	% Change	Original	With Late Dx Strata	% Change
2010	707	750	6.0	7.9	8.4	6.3
2011	684	719	5.2	7.6	7.9	4.0
2012	656	677	3.2	7.2	7.4	2.8
2013	630	637	1.0	6.7	6.8	1.5
2014	616	602	-2.1	6.5	6.4	-1.5

Table 8: Upper Bound: MSM

	Cases			Fraction		
Year	Original	With Late Dx Strata	% Change	Original	With Late Dx Strata	% Change
2010	1381	1418	2.7	14.4	14.7	2.1
2011	1339	1354	1.1	13.8	14.0	1.4
2012	1293	1283	-0.8	13.3	13.2	-0.8
2013	1253	1219	-2.7	12.5	12.2	-2.4
2014	1230	1167	-5.1	12.3	11.7	-4.9

Table 9: Base Case: non-MSM

	Cases			Fraction		
Year	Original	With Late Dx Strata	% Change	Original	With Late Dx Strata	% Change
2010	776	778	0.3	22.5	22.6	0.4
2011	766	756	-1.3	22.1	21.9	-0.9
2012	745	728	-2.3	21.4	21.0	-1.9
2013	725	702	-3.1	20.6	20.1	-2.4
2014	715	688	-3.8	20.1	19.5	-3.0

Table 10: Upper Bound: non-MSM

	Cases			Fraction		
Year	Original	With Late Dx Strata	% Change	Original	With Late Dx Strata	% Change
2010	1468	1387	-5.5	35.5	34.2	-3.7
2011	1451	1363	-6.1	34.9	33.5	-4.0
2012	1417	1330	-6.1	34.1	32.7	-4.1
2013	1385	1302	-6.0	33.1	31.8	-3.9
2014	1364	1284	-5.9	32.5	31.2	-4.0