Impact of stratifying by MSM/non-MSM and using yearly time steps

Jeanette Birnbaum

May 17, 2016

1 Overview

This report first compares the results for WA State when stratifying by MSM/non-MSM subgroups and, secondly, compares results for quarterly versus yearly time steps. These investigations will provide insights useful for assessing how to best generate estimates for subgroups with small sample sizes.

2 Data by MSM strata

2.1 Presence of testing histories

Stratifying the estimates by MSM/non-MSM should have a noticeable impact on estimates when testing history patterns are substantially different for the two subgroups.

Table 1: Composition of analytic sample by MSM vs non-MSM

				Ever Had a Negative Test		
Characteristic	Subgroup	N	Column %	% Yes	% No	% Missing
All	All	5176	100	46	12	42
Risk Group	MSM	3403	66	57	9	34
	non-MSM	1773	34	25	18	57

3 Stratified vs Unstratified Analysis

3.1 Stratified TID

Figure 1 shows the Base Case (left) and Upper Bound (right) TIDs separately for MSM vs non-MSM. Time steps are quarters.

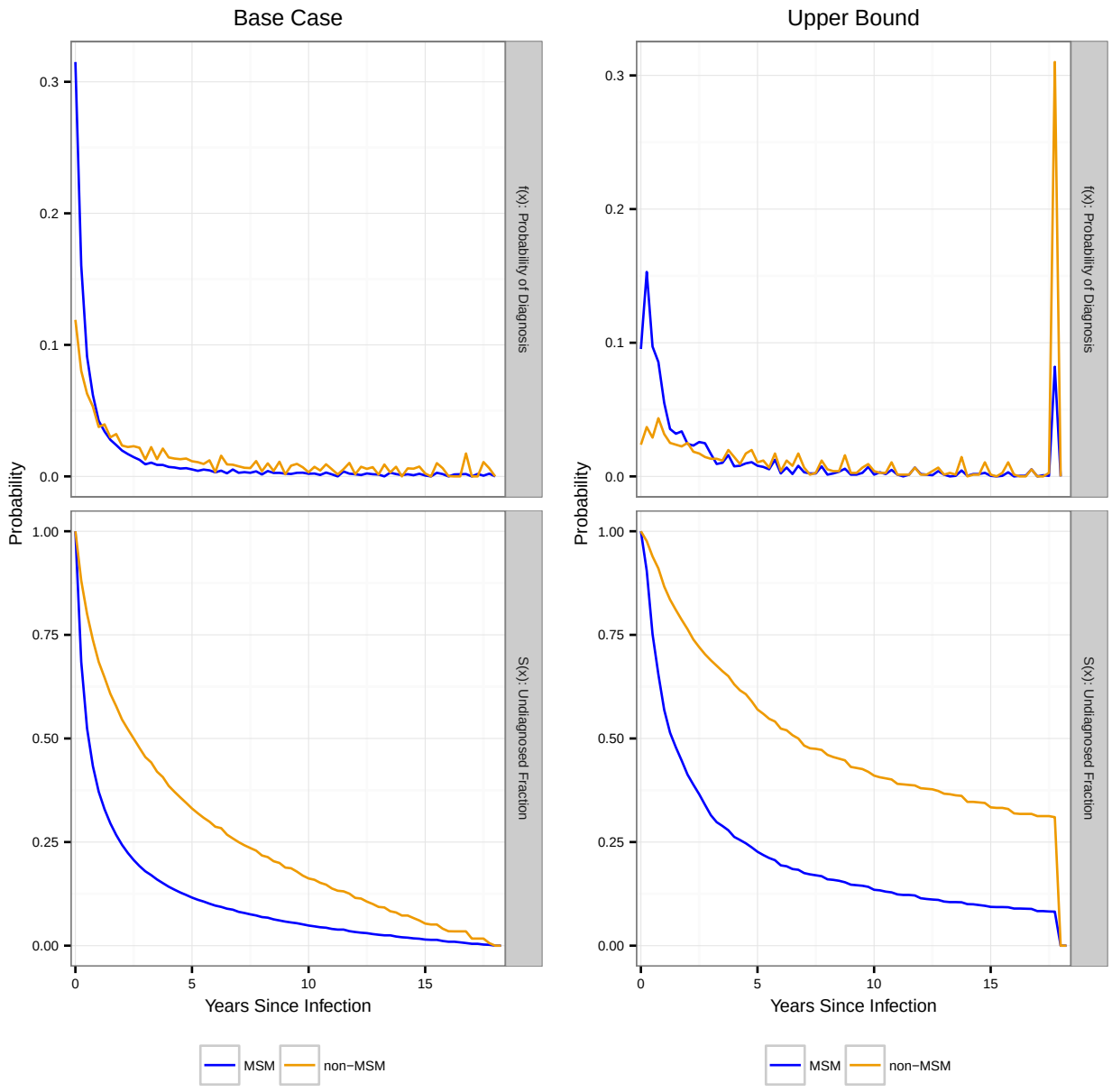

Figure 1: Time from infection to diagnosis (TID) for Base Case and Upper Bound, MSM vs non-MSM, quarterly time steps

3.2 Results

Table 2 shows the impact of stratification on undiagnosed count estimates. Percent change is relative to the unstratified results. PLWH in 2014 was 12,671, so an additional 100 cases undiagnosed represents a change in the undiagnosed fraction from 9% to 10%.

Table 2: Impact of MSM/non-MSM stratification on mean undiagnosed estimates. Percent change is relative to unstratified results

Year	Case	Stratified	Unstratified	Difference	Percent Change
2005	Base Case	1593.0	1449.0	144.0	10.0
2005-2014	Base Case	1490.0	1371.0	119.0	9.0
2006	Base Case	1595.0	1454.0	141.0	10.0
2007	Base Case	1571.0	1436.0	135.0	9.0
2008	Base Case	1546.0	1423.0	123.0	9.0
2009	Base Case	1525.0	1411.0	114.0	8.0
2010	Base Case	1484.0	1374.0	110.0	8.0
2011	Base Case	1453.0	1344.0	109.0	8.0
2012	Base Case	1410.0	1303.0	107.0	8.0
2013	Base Case	1373.0	1268.0	105.0	8.0
2014	Base Case	1351.0	1251.0	100.0	8.0
2005	Upper Bound	3086.0	2867.0	219.0	8.0
2005-2014	Upper Bound	2887.0	2704.0	183.0	7.0
2006	Upper Bound	3070.0	2862.0	208.0	7.0
2007	Upper Bound	3019.0	2824.0	195.0	7.0
2008	Upper Bound	2977.0	2795.0	182.0	7.0
2009	Upper Bound	2938.0	2762.0	176.0	6.0
2010	Upper Bound	2874.0	2696.0	178.0	7.0
2011	Upper Bound	2822.0	2643.0	179.0	7.0
2012	Upper Bound	2750.0	2576.0	174.0	7.0
2013	Upper Bound	2690.0	2523.0	167.0	7.0
2014	Upper Bound	2647.0	2492.0	155.0	6.0

4 Quarterly vs Yearly Time Step Analysis

For the stratified results, we now look at the impact of using yearly time steps instead of quarter time steps.

4.1 Stratified TID, Yearly Time Steps

Figure 2 shows the MSM vs non-MSM TIDs for yearly time steps, instead of the quarter time steps in 1.

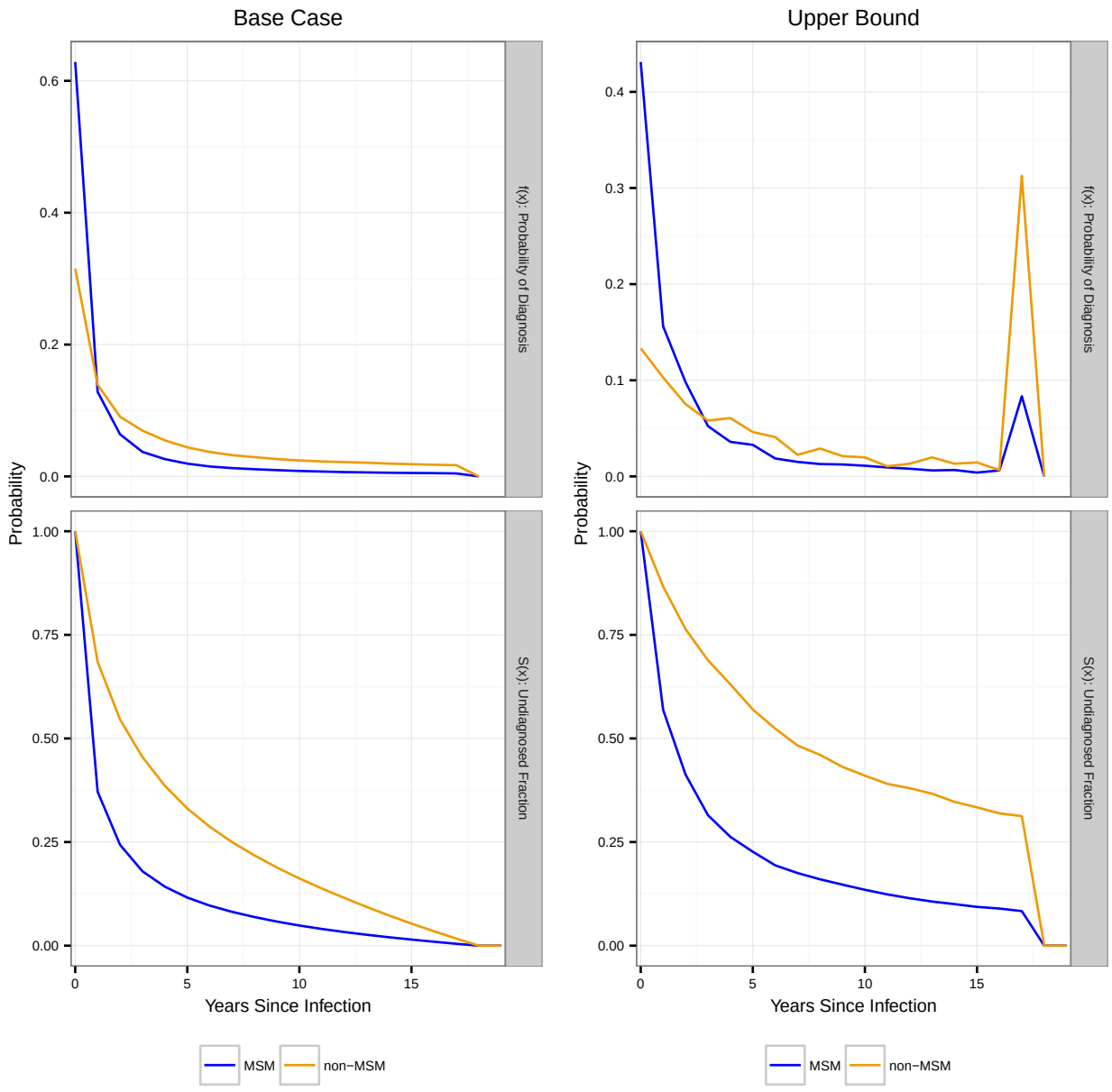

Figure 2: Time from infection to diagnosis (TID) for Base Case and Upper Bound, MSM vs non-MSM, yearly time steps

4.2 Results - Full Sample, Half-Year Aggregation

Table 3 shows the impact of aggregating to half-year time steps. Percent change is relative to quarterly estimates.

Table 3: Impact of half-year vs quarter time steps on mean undiagnosed estimates. Percent change is relative to quarter estimates

Year	Case	Half-Year	Quarter	Difference	Percent Change
2005	Base Case	1567.0	1593.0	-26.0	-2.0
2005-2014	Base Case	1476.0	1490.0	-14.0	-1.0
2006	Base Case	1559.0	1595.0	-36.0	-2.0
2007	Base Case	1542.0	1571.0	-29.0	-2.0
2008	Base Case	1508.0	1546.0	-38.0	-2.0
2009	Base Case	1491.0	1525.0	-34.0	-2.0
2010	Base Case	1464.0	1484.0	-20.0	-1.0
2011	Base Case	1440.0	1453.0	-13.0	-1.0
2012	Base Case	1409.0	1410.0	-1.0	-0.0
2013	Base Case	1392.0	1373.0	19.0	1.0
2014	Base Case	1393.0	1351.0	42.0	3.0
2005	Upper Bound	3000.0	3086.0	-86.0	-3.0
2005-2014	Upper Bound	2850.0	2887.0	-37.0	-1.0
2006	Upper Bound	2979.0	3070.0	-91.0	-3.0
2007	Upper Bound	2943.0	3019.0	-76.0	-3.0
2008	Upper Bound	2896.0	2977.0	-81.0	-3.0
2009	Upper Bound	2869.0	2938.0	-69.0	-2.0
2010	Upper Bound	2825.0	2874.0	-49.0	-2.0
2011	Upper Bound	2790.0	2822.0	-32.0	-1.0
2012	Upper Bound	2746.0	2750.0	-4.0	-0.0
2013	Upper Bound	2726.0	2690.0	36.0	1.0
2014	Upper Bound	2730.0	2647.0	83.0	3.0

4.3 Results - Full Sample, Full-Year Aggregation

Table 4 shows the impact of aggregating to yearly time steps. Percent change is relative to quarterly estimates.

The impact here is greater than the half-year aggregation, but in both cases, the aggregation results in lower undiagnosed estimates for earlier years but higher undiagnosed estimates for the last few years.

Lower undiagnosed estimates arise from the lengthened gap between infection and diagnosis that still constitutes being diagnosed at time=0, e.g., for quarterly estimates, infection in Q1 and diagnoses in Q2 means 1 quarter of undiagnosed time, but for yearly estimates, that case would contribute no undiagnosed time.

However, the aggregation treats all diagnoses made within the longer interval equally, which overestimates undiagnosis when there are many cases being diagnosed in shorter intervals. In these data, 27% of cases are actually diagnosed within one quarter under the base case, but their time to diagnosis is considered to be 1 year. This makes the most difference for the most recent years of the estimates, when the undiagnosed cases arise only from recent infections. In older years, undiagnosed cases are a mix of recent and long-standing infections.

Table 4: Impact of year vs quarter time steps on mean undiagnosed estimates. Percent change is relative to quarter estimates

Year	Case	Quarter	Year	Difference	Percent Change
2005	Base Case	1593.0	1553.0	-40.0	-3.0
2005-2014	Base Case	1490.0	1486.0	-4.0	-0.0
2006	Base Case	1595.0	1543.0	-52.0	-3.0
2007	Base Case	1571.0	1534.0	-37.0	-2.0
2008	Base Case	1546.0	1497.0	-49.0	-3.0
2009	Base Case	1525.0	1486.0	-39.0	-3.0
2010	Base Case	1484.0	1470.0	-14.0	-1.0
2011	Base Case	1453.0	1449.0	-4.0	-0.0
2012	Base Case	1410.0	1443.0	33.0	2.0
2013	Base Case	1373.0	1437.0	64.0	5.0
2014	Base Case	1351.0	1451.0	100.0	7.0
2005	Upper Bound	3086.0	2904.0	-182.0	-6.0
2005-2014	Upper Bound	2887.0	2816.0	-71.0	-2.0
2006	Upper Bound	3070.0	2890.0	-180.0	-6.0
2007	Upper Bound	3019.0	2867.0	-152.0	-5.0
2008	Upper Bound	2977.0	2826.0	-151.0	-5.0
2009	Upper Bound	2938.0	2811.0	-127.0	-4.0
2010	Upper Bound	2874.0	2787.0	-87.0	-3.0
2011	Upper Bound	2822.0	2767.0	-55.0	-2.0
2012	Upper Bound	2750.0	2759.0	9.0	0.0
2013	Upper Bound	2690.0	2761.0	71.0	3.0
2014	Upper Bound	2647.0	2789.0	142.0	5.0

4.4 Results - Racial Subgroups, Full-Year Aggregation

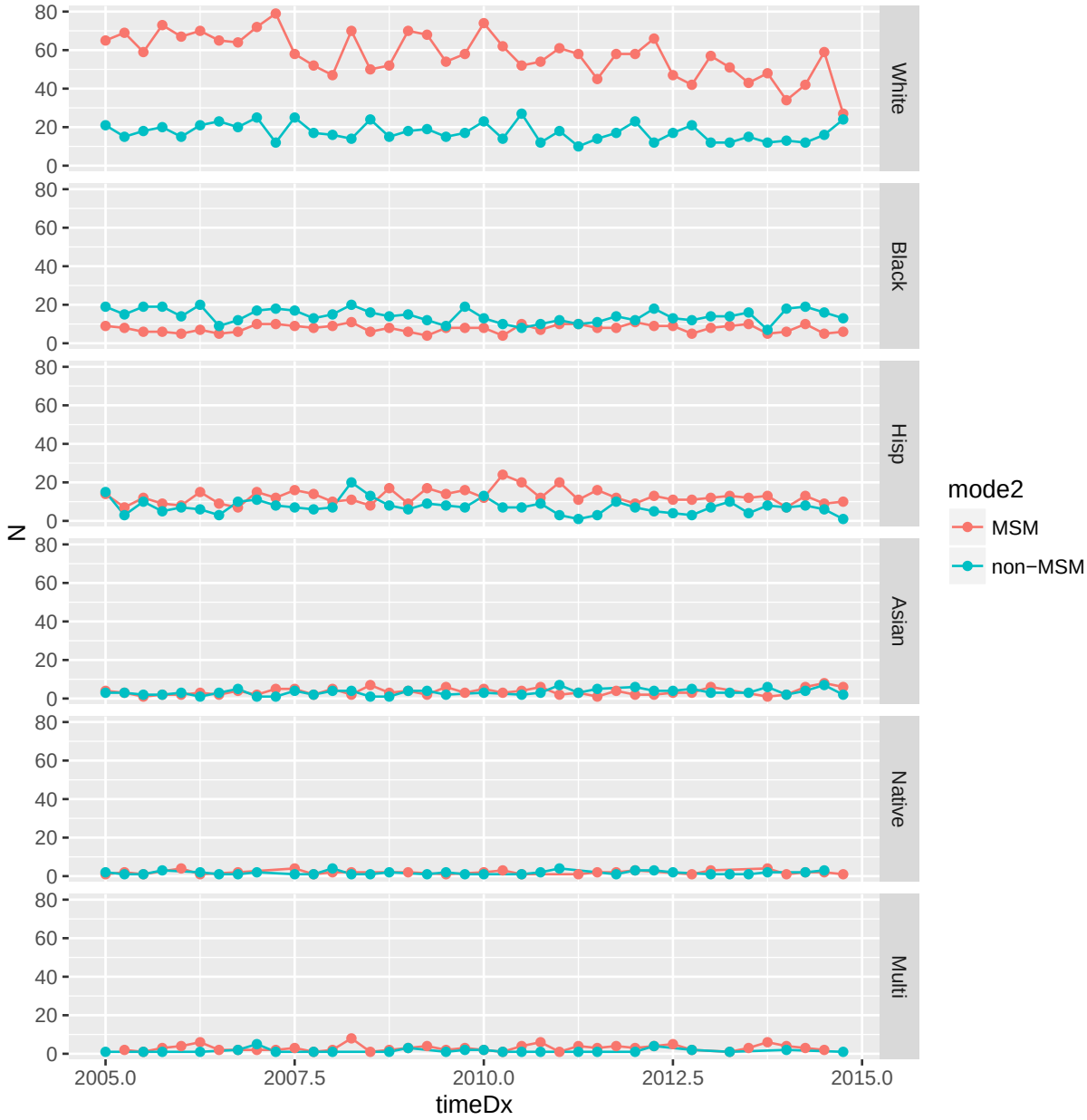

Figure 3: Quarterly diagnosis counts by race, stratified by MSM vs non-MSM

4.4.1 Whites

Table 5: WHITES: Impact of year vs quarter time steps on mean undiagnosed estimates. Percent change is relative to quarter estimates

Year	Case	Quarter	Year	Difference	Percent Change
2005	Base Case	807.3	788.2	-19.1	-2.0
2005-2014	Base Case	733.8	723.6	-10.2	-1.0
2006	Base Case	806.6	775.9	-30.7	-4.0
2007	Base Case	780.0	753.8	-26.2	-3.0
2008	Base Case	772.7	731.3	-41.4	-5.0
2009	Base Case	765.5	735.4	-30.1	-4.0
2010	Base Case	737.4	715.7	-21.7	-3.0
2011	Base Case	714.6	694.8	-19.8	-3.0
2012	Base Case	681.1	684.5	3.4	0.0
2013	Base Case	644.5	674.2	29.7	5.0
2014	Base Case	628.1	682.5	54.4	9.0
2005	Upper Bound	1586.0	1448.0	-138.0	-9.0
2005-2014	Upper Bound	1432.0	1357.0	-75.0	-5.0
2006	Upper Bound	1569.0	1425.0	-144.0	-9.0
2007	Upper Bound	1520.0	1392.0	-128.0	-8.0
2008	Upper Bound	1504.0	1370.0	-134.0	-9.0
2009	Upper Bound	1480.0	1366.0	-114.0	-8.0
2010	Upper Bound	1429.0	1338.0	-91.0	-6.0
2011	Upper Bound	1388.0	1315.0	-73.0	-5.0
2012	Upper Bound	1330.0	1302.0	-28.0	-2.0
2013	Upper Bound	1274.0	1297.0	23.0	2.0
2014	Upper Bound	1243.0	1317.0	74.0	6.0

4.4.2 Blacks

Table 6: BLACKS: Impact of year vs quarter time steps on mean undiagnosed estimates. Percent change is relative to quarter estimates

Year	Case	Quarter	Year	Difference	Percent Change
2005	Base Case	380.1	342.6	-37.5	-10.0
2005-2014	Base Case	337.5	331.1	-6.4	-2.0
2006	Base Case	368.2	336.4	-31.8	-9.0
2007	Base Case	366.4	339.5	-26.9	-7.0
2008	Base Case	348.3	327.4	-20.9	-6.0
2009	Base Case	325.7	317.0	-8.7	-3.0
2010	Base Case	310.6	320.8	10.2	3.0
2011	Base Case	314.3	330.8	16.5	5.0
2012	Base Case	316.1	332.4	16.3	5.0
2013	Base Case	320.6	332.5	11.9	4.0
2014	Base Case	324.4	331.3	6.9	2.0
2005	Upper Bound	693.7	643.0	-50.7	-7.0
2005-2014	Upper Bound	645.0	630.3	-14.7	-2.0
2006	Upper Bound	679.1	637.8	-41.3	-6.0
2007	Upper Bound	671.5	637.3	-34.2	-5.0
2008	Upper Bound	645.3	622.5	-22.8	-4.0
2009	Upper Bound	624.8	613.5	-11.3	-2.0
2010	Upper Bound	619.4	621.0	1.6	0.0
2011	Upper Bound	629.8	631.8	2.0	0.0
2012	Upper Bound	629.9	633.0	3.1	0.0
2013	Upper Bound	630.0	633.1	3.1	0.0
2014	Upper Bound	626.1	630.4	4.3	1.0

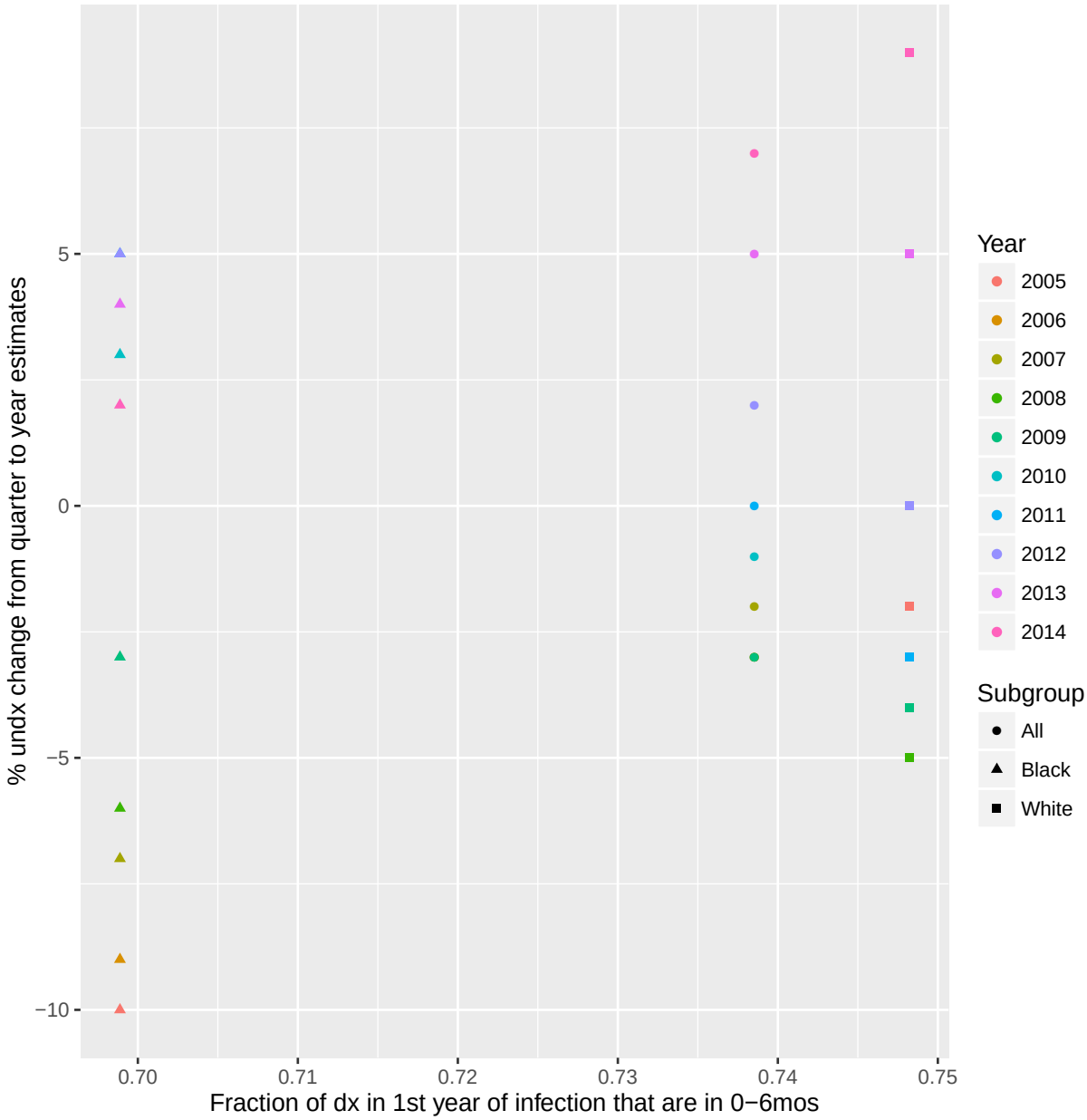

Figure 4: Percent change between year vs quarterly Base Case estimates as a function of the fraction of diagnosis probability within the first year that falls within the first 6 mos