

## **Supplementary Material: Part 2**

### **Assessing transmission attribution risk from simulated sequencing data in HIV molecular epidemiology**

Fabírcia F. Nascimento<sup>1\*</sup>, Sanjay R. Mehta<sup>2</sup>, Susan J. Little<sup>2</sup>, and Erik M. Volz<sup>1\*</sup>

<sup>1</sup>MRC Centre for Global Infectious Disease Analysis and the Department of Infectious Disease Epidemiology, Imperial College London, London, UK

<sup>2</sup>Division of Infectious Diseases, University of California San Diego, San Diego, CA, USA

**\*Corresponding authors**

**Table S2.1: Natural mortality rates:** Natural mortality rate per person per day from 1980 to 2016 used in our network simulations for MSM individuals in *region* (or San Diego).

Age Group	Year				
	1980	1981	1982	1983	1984
18–19	$2.02 \times 10^{-7}$	$1.77 \times 10^{-7}$	$1.78 \times 10^{-7}$	$1.49 \times 10^{-7}$	$1.89 \times 10^{-7}$
20–24	$3.00 \times 10^{-7}$	$2.23 \times 10^{-7}$	$2.34 \times 10^{-7}$	$2.28 \times 10^{-7}$	$2.08 \times 10^{-7}$
25–34	$3.24 \times 10^{-7}$	$2.62 \times 10^{-7}$	$2.32 \times 10^{-7}$	$2.39 \times 10^{-7}$	$2.68 \times 10^{-7}$
35–44	$4.61 \times 10^{-7}$	$4.13 \times 10^{-7}$	$3.94 \times 10^{-7}$	$3.52 \times 10^{-7}$	$3.81 \times 10^{-7}$
45–54	$1.16 \times 10^{-6}$	$1.18 \times 10^{-6}$	$1.19 \times 10^{-6}$	$1.09 \times 10^{-6}$	$1.08 \times 10^{-6}$
55–64	$2.69 \times 10^{-6}$	$2.80 \times 10^{-6}$	$2.79 \times 10^{-6}$	$2.67 \times 10^{-6}$	$2.74 \times 10^{-6}$
65–74	$6.26 \times 10^{-6}$	$6.22 \times 10^{-6}$	$6.11 \times 10^{-6}$	$5.72 \times 10^{-6}$	$5.91 \times 10^{-6}$
75–80	$1.40 \times 10^{-5}$	$1.40 \times 10^{-5}$	$1.37 \times 10^{-5}$	$1.39 \times 10^{-5}$	$1.42 \times 10^{-5}$

Age Group	Year				
	1985	1986	1987	1988	1989
18–19	$1.35 \times 10^{-7}$	$1.64 \times 10^{-7}$	$1.47 \times 10^{-7}$	$1.66 \times 10^{-7}$	$1.88 \times 10^{-7}$
20–24	$2.39 \times 10^{-7}$	$2.34 \times 10^{-7}$	$2.28 \times 10^{-7}$	$2.52 \times 10^{-7}$	$2.09 \times 10^{-7}$
25–34	$2.72 \times 10^{-7}$	$3.20 \times 10^{-7}$	$3.04 \times 10^{-7}$	$3.21 \times 10^{-7}$	$3.53 \times 10^{-7}$
35–44	$4.15 \times 10^{-7}$	$4.95 \times 10^{-7}$	$4.82 \times 10^{-7}$	$5.06 \times 10^{-7}$	$5.26 \times 10^{-7}$
45–54	$1.02 \times 10^{-6}$	$9.64 \times 10^{-7}$	$9.12 \times 10^{-7}$	$9.67 \times 10^{-7}$	$9.02 \times 10^{-7}$
55–64	$2.49 \times 10^{-6}$	$2.51 \times 10^{-6}$	$2.51 \times 10^{-6}$	$2.44 \times 10^{-6}$	$2.46 \times 10^{-6}$
65–74	$5.89 \times 10^{-6}$	$5.67 \times 10^{-6}$	$5.44 \times 10^{-6}$	$5.49 \times 10^{-6}$	$5.19 \times 10^{-6}$
75–80	$1.38 \times 10^{-5}$	$1.33 \times 10^{-5}$	$1.34 \times 10^{-5}$	$1.31 \times 10^{-5}$	$1.24 \times 10^{-5}$

Age Group	Year				
	1990	1991	1992	1993	1994
18–19	$1.56 \times 10^{-7}$	$2.23 \times 10^{-7}$	$1.86 \times 10^{-7}$	$1.93 \times 10^{-7}$	$1.66 \times 10^{-7}$
20–24	$2.09 \times 10^{-7}$	$1.85 \times 10^{-7}$	$1.67 \times 10^{-7}$	$2.35 \times 10^{-7}$	$2.24 \times 10^{-7}$
25–34	$3.64 \times 10^{-7}$	$3.49 \times 10^{-7}$	$3.33 \times 10^{-7}$	$3.81 \times 10^{-7}$	$3.52 \times 10^{-7}$
35–44	$5.62 \times 10^{-7}$	$6.23 \times 10^{-7}$	$6.04 \times 10^{-7}$	$6.50 \times 10^{-7}$	$6.48 \times 10^{-7}$
45–54	$9.69 \times 10^{-7}$	$9.40 \times 10^{-7}$	$9.50 \times 10^{-7}$	$9.25 \times 10^{-7}$	$9.79 \times 10^{-7}$
55–64	$2.41 \times 10^{-6}$	$2.26 \times 10^{-6}$	$2.36 \times 10^{-6}$	$2.28 \times 10^{-6}$	$2.17 \times 10^{-6}$
65–74	$5.27 \times 10^{-6}$	$5.18 \times 10^{-6}$	$5.25 \times 10^{-6}$	$5.05 \times 10^{-6}$	$5.29 \times 10^{-6}$
75–80	$1.25 \times 10^{-5}$	$1.19 \times 10^{-5}$	$1.23 \times 10^{-5}$	$1.20 \times 10^{-5}$	$1.25 \times 10^{-5}$

Age Group	Year				
	1995	1996	1997	1998	1999
18–19	$1.70 \times 10^{-7}$	$1.29 \times 10^{-7}$	$1.23 \times 10^{-7}$	$1.27 \times 10^{-7}$	$1.19 \times 10^{-7}$
20–24	$2.10 \times 10^{-7}$	$2.07 \times 10^{-7}$	$2.07 \times 10^{-7}$	$1.81 \times 10^{-7}$	$1.69 \times 10^{-7}$
25–34	$3.29 \times 10^{-7}$	$2.57 \times 10^{-7}$	$2.25 \times 10^{-7}$	$1.80 \times 10^{-7}$	$1.73 \times 10^{-7}$
35–44	$6.19 \times 10^{-7}$	$5.42 \times 10^{-7}$	$4.27 \times 10^{-7}$	$3.97 \times 10^{-7}$	$4.28 \times 10^{-7}$
45–54	$9.17 \times 10^{-7}$	$8.69 \times 10^{-7}$	$8.06 \times 10^{-7}$	$8.67 \times 10^{-7}$	$8.54 \times 10^{-7}$
55–64	$2.18 \times 10^{-6}$	$2.12 \times 10^{-6}$	$2.15 \times 10^{-6}$	$2.00 \times 10^{-6}$	$1.82 \times 10^{-6}$
65–74	$5.03 \times 10^{-6}$	$4.99 \times 10^{-6}$	$4.68 \times 10^{-6}$	$4.79 \times 10^{-6}$	$4.84 \times 10^{-6}$
75–80	$1.18 \times 10^{-5}$	$1.18 \times 10^{-5}$	$1.16 \times 10^{-5}$	$1.17 \times 10^{-5}$	$1.16 \times 10^{-5}$

Age Group	Year				
	2000	2001	2002	2003	2004
18–19	$1.27 \times 10^{-7}$	$1.34 \times 10^{-7}$	$1.12 \times 10^{-7}$	$1.42 \times 10^{-7}$	$1.36 \times 10^{-7}$
20–24	$1.72 \times 10^{-7}$	$1.68 \times 10^{-7}$	$1.84 \times 10^{-7}$	$1.61 \times 10^{-7}$	$1.76 \times 10^{-7}$
25–34	$1.86 \times 10^{-7}$	$1.89 \times 10^{-7}$	$1.86 \times 10^{-7}$	$1.79 \times 10^{-7}$	$1.76 \times 10^{-7}$
35–44	$3.78 \times 10^{-7}$	$3.59 \times 10^{-7}$	$3.74 \times 10^{-7}$	$3.69 \times 10^{-7}$	$3.55 \times 10^{-7}$
45–54	$8.37 \times 10^{-7}$	$8.65 \times 10^{-7}$	$8.22 \times 10^{-7}$	$8.34 \times 10^{-7}$	$8.47 \times 10^{-7}$
55–64	$1.83 \times 10^{-6}$	$1.87 \times 10^{-6}$	$1.74 \times 10^{-6}$	$1.73 \times 10^{-6}$	$1.71 \times 10^{-6}$
65–74	$4.80 \times 10^{-6}$	$4.56 \times 10^{-6}$	$4.43 \times 10^{-6}$	$4.31 \times 10^{-6}$	$4.11 \times 10^{-6}$
75–80	$1.11 \times 10^{-5}$	$1.12 \times 10^{-5}$	$1.06 \times 10^{-5}$	$1.08 \times 10^{-5}$	$1.06 \times 10^{-5}$

Age Group	Year				
	2005	2006	2007	2008	2009
18–19	$8.92 \times 10^{-8}$	$1.34 \times 10^{-7}$	$1.02 \times 10^{-7}$	$9.14 \times 10^{-8}$	$9.25 \times 10^{-8}$
20–24	$2.02 \times 10^{-7}$	$1.87 \times 10^{-7}$	$1.62 \times 10^{-7}$	$1.63 \times 10^{-7}$	$1.41 \times 10^{-7}$
25–34	$1.85 \times 10^{-7}$	$1.69 \times 10^{-7}$	$1.86 \times 10^{-7}$	$1.67 \times 10^{-7}$	$1.66 \times 10^{-7}$
35–44	$3.48 \times 10^{-7}$	$3.33 \times 10^{-7}$	$3.49 \times 10^{-7}$	$2.91 \times 10^{-7}$	$2.79 \times 10^{-7}$
45–54	$8.59 \times 10^{-7}$	$7.41 \times 10^{-7}$	$8.53 \times 10^{-7}$	$7.80 \times 10^{-7}$	$7.57 \times 10^{-7}$
55–64	$1.75 \times 10^{-6}$	$1.68 \times 10^{-6}$	$1.72 \times 10^{-6}$	$1.79 \times 10^{-6}$	$1.72 \times 10^{-6}$
65–74	$3.96 \times 10^{-6}$	$3.85 \times 10^{-6}$	$3.75 \times 10^{-6}$	$3.65 \times 10^{-6}$	$3.58 \times 10^{-6}$
75–80	$1.02 \times 10^{-5}$	$1.02 \times 10^{-5}$	$9.73 \times 10^{-6}$	$9.29 \times 10^{-6}$	$8.82 \times 10^{-6}$

Age Group	Year				
	2010	2011	2012	2013	2014
18–19	$8.85 \times 10^{-8}$	$7.12 \times 10^{-8}$	$8.96 \times 10^{-8}$	$8.00 \times 10^{-8}$	$5.45 \times 10^{-8}$
20–24	$1.39 \times 10^{-7}$	$1.44 \times 10^{-7}$	$1.42 \times 10^{-7}$	$1.41 \times 10^{-7}$	$1.42 \times 10^{-7}$
25–34	$1.29 \times 10^{-7}$	$1.69 \times 10^{-7}$	$1.82 \times 10^{-7}$	$1.69 \times 10^{-7}$	$1.61 \times 10^{-7}$
35–44	$2.69 \times 10^{-7}$	$2.67 \times 10^{-7}$	$2.69 \times 10^{-7}$	$2.64 \times 10^{-7}$	$2.62 \times 10^{-7}$
45–54	$6.52 \times 10^{-7}$	$7.46 \times 10^{-7}$	$7.05 \times 10^{-7}$	$6.97 \times 10^{-7}$	$6.90 \times 10^{-7}$
55–64	$1.61 \times 10^{-6}$	$1.57 \times 10^{-6}$	$1.66 \times 10^{-6}$	$1.65 \times 10^{-6}$	$1.52 \times 10^{-6}$
65–74	$3.39 \times 10^{-6}$	$3.21 \times 10^{-6}$	$3.17 \times 10^{-6}$	$3.26 \times 10^{-6}$	$3.10 \times 10^{-6}$
75–80	$8.89 \times 10^{-6}$	$9.07 \times 10^{-6}$	$8.86 \times 10^{-6}$	$8.81 \times 10^{-6}$	$8.39 \times 10^{-6}$

Age Group	Year	
	2015	2016
18–19	$7.49 \times 10^{-8}$	$6.39 \times 10^{-8}$
20–24	$1.62 \times 10^{-7}$	$1.61 \times 10^{-7}$
25–34	$1.57 \times 10^{-7}$	$1.65 \times 10^{-7}$
35–44	$2.46 \times 10^{-7}$	$2.95 \times 10^{-7}$
45–54	$6.68 \times 10^{-7}$	$5.99 \times 10^{-7}$
55–64	$1.55 \times 10^{-6}$	$1.62 \times 10^{-6}$
65–74	$3.16 \times 10^{-6}$	$3.16 \times 10^{-6}$
75–80	$8.29 \times 10^{-6}$	$8.25 \times 10^{-6}$

**Table S2.2: Testing probabilities:** Probability between 0 and 1 an MSM would be tested per day.

Year	Prob. per day
1989	0.00067
1990	0.00101
1991	0.00083
1992	0.00061
1993	0.00064
1994	0.00064
1995	0.00064
1996	0.00050
1997	0.00054
1998	0.00063
1999	0.00056
2000	0.00062
2001	0.00088

**Table S2.3: Total number of pairs analyzed for consensus sequences for sampling strategy 1 for an average of 250 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals not on ART for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 1.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 1 / True trees / 5	3 / 0	5 / 2	4,749 / 4,752	239 / 242	3 / 0
Group 1 / 1,000bp / 5	3 / 0	6 / 2	4,556 / 4,560	207 / 210	2 / 0
Group 1 / 10,000bp / 5	3 / 0	5 / 2	4,727 / 4,730	233 / 236	3 / 0
Group 1 / True trees / 10	14 / 4	5 / 1	8,925 / 8,929	896 / 906	12 / 4
Group 1 / 1,000bp / 10	11 / 4	4 / 1	8,726 / 8,729	765 / 772	10 / 4
Group 1 / 10,000bp / 10	14 / 4	3 / 1	8,779 / 8,781	888 / 898	12 / 4
Group 1 / True trees / 20	53 / 13	26 / 4	16,580 / 16,602	2,695 / 2,735	48 / 13
Group 1 / 1,000bp / 20	67 / 24	48 / 8	16,506 / 16,546	2,335 / 2,378	60 / 23
Group 1 / 10,000bp / 20	54 / 15	28 / 3	16,330 / 16,355	2,614 / 2,653	48 / 14
Group 1 / True trees / 30	94 / 26	51 / 15	23,101 / 23,137	5,692 / 5,760	88 / 25
Group 1 / 1,000bp / 30	115 / 33	75 / 13	23,483 / 23,545	4,835 / 4,917	103 / 32
Group 1 / 10,000bp / 30	97 / 29	52 / 14	22,970 / 23,008	5,501 / 5,569	89 / 28
Group 1 / True trees / 40	201 / 56	85 / 15	28,337 / 28,407	9,565 / 9,710	165 / 50
Group 1 / 1,000bp / 40	246 / 74	117 / 26	29,089 / 29,180	8,116 / 8,288	202 / 65
Group 1 / 10,000bp / 40	208 / 57	83 / 16	28,075 / 28,142	9,324 / 9,475	174 / 51
Group 1 / True trees / 50	340 / 97	115 / 23	32,515 / 32,607	13,778 / 14,021	284 / 90
Group 1 / 1,000bp / 50	406 / 131	169 / 37	33,881 / 34,013	11,750 / 12,025	337 / 116
Group 1 / 10,000bp / 50	347 / 106	126 / 23	32,476 / 32,579	13,265 / 13,506	292 / 97
Group 1 / True trees / 60	446 / 111	133 / 26	36,933 / 37,040	18,512 / 18,847	380 / 101
Group 1 / 1,000bp / 60	523 / 146	178 / 43	39,260 / 39,395	15,819 / 16,196	445 / 128
Group 1 / 10,000bp / 60	453 / 112	140 / 31	36,908 / 37,017	17,857 / 18,198	391 / 103
Group 1 / True trees / 70	643 / 173	181 / 35	40,261 / 40,407	23,939 / 24,409	539 / 161
Group 1 / 1,000bp / 70	732 / 233	240 / 63	43,524 / 43,701	20,256 / 20,755	605 / 205
Group 1 / 10,000bp / 70	636 / 173	186 / 41	40,460 / 40,605	23,000 / 23,463	537 / 161
Group 1 / True trees / 80	811 / 200	228 / 40	43,776 / 43,964	29,231 / 29,842	698 / 176
Group 1 / 1,000bp / 80	973 / 303	301 / 76	47,453 / 47,678	24,837 / 25,507	835 / 270
Group 1 / 10,000bp / 80	818 / 196	235 / 45	43,735 / 43,925	28,298 / 28,920	695 / 174
Group 1 / True trees / 90	1,021 / 269	238 / 45	45,836 / 46,029	35,577 / 36,329	875 / 250
Group 1 / 1,000bp / 90	1,176 / 381	347 / 76	50,401 / 50,672	30,322 / 31,117	992 / 349
Group 1 / 10,000bp / 90	1,041 / 282	244 / 41	46,320 / 46,523	34,157 / 34,916	895 / 258

**Table S2.4: Total number of pairs analyzed for consensus sequences for sampling strategy 1 for an average of 250 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals not on ART for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 2.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 2 / True trees / 5	9 / 2	3 / 0	8,365 / 8,368	523 / 530	8 / 2
Group 2 / 1,000bp / 5	8 / 3	2 / 0	8,564 / 8,566	458 / 463	7 / 3
Group 2 / 10,000bp / 5	9 / 4	2 / 1	8,538 / 8,539	537 / 542	8 / 4
Group 2 / True trees / 10	15 / 2	9 / 2	15,887 / 15,894	1,859 / 1,872	14 / 1
Group 2 / 1,000bp / 10	15 / 6	14 / 0	16,728 / 16,742	1,503 / 1,512	13 / 5
Group 2 / 10,000bp / 10	20 / 6	11 / 2	16,237 / 16,246	1,770 / 1,784	19 / 5
Group 2 / True trees / 20	80 / 23	30 / 8	28,538 / 28,560	6,646 / 6,703	68 / 21
Group 2 / 1,000bp / 20	92 / 26	25 / 5	30,897 / 30,917	5,386 / 5,452	78 / 24
Group 2 / 10,000bp / 20	85 / 23	30 / 4	28,828 / 28,854	6,397 / 6,459	71 / 21
Group 2 / True trees / 30	136 / 29	57 / 5	39,063 / 39,115	13,014 / 13,121	124 / 28
Group 2 / 1,000bp / 30	162 / 51	83 / 8	42,821 / 42,896	10,524 / 10,635	147 / 47
Group 2 / 10,000bp / 30	140 / 35	66 / 8	39,756 / 39,814	12,302 / 12,407	130 / 34
Group 2 / True trees / 40	299 / 85	88 / 14	46,966 / 47,040	21,889 / 22,103	266 / 81
Group 2 / 1,000bp / 40	355 / 114	106 / 32	53,024 / 53,098	17,407 / 17,648	304 / 100
Group 2 / 10,000bp / 40	307 / 88	94 / 12	47,920 / 48,002	20,845 / 21,064	271 / 83
Group 2 / True trees / 50	482 / 144	114 / 22	53,812 / 53,904	30,914 / 31,252	429 / 133
Group 2 / 1,000bp / 50	558 / 181	158 / 36	62,352 / 62,474	24,914 / 25,291	493 / 164
Group 2 / 10,000bp / 50	487 / 143	118 / 15	55,214 / 55,317	29,479 / 29,823	433 / 136
Group 2 / True trees / 60	660 / 189	163 / 32	58,965 / 59,096	41,890 / 42,361	603 / 179
Group 2 / 1,000bp / 60	796 / 254	221 / 62	70,413 / 70,572	33,336 / 33,878	717 / 239
Group 2 / 10,000bp / 60	678 / 190	167 / 35	60,923 / 61,055	39,878 / 40,366	620 / 181
Group 2 / True trees / 70	890 / 250	202 / 29	63,654 / 63,827	53,472 / 54,112	800 / 232
Group 2 / 1,000bp / 70	1,114 / 349	291 / 60	77,275 / 77,506	43,074 / 43,839	974 / 314
Group 2 / 10,000bp / 70	932 / 253	214 / 33	65,686 / 65,867	50,978 / 51,657	836 / 231
Group 2 / True trees / 80	1,206 / 337	253 / 38	66,747 / 66,962	65,792 / 66,661	1,063 / 316
Group 2 / 1,000bp / 80	1,414 / 457	366 / 94	84,586 / 84,858	52,610 / 53,567	1,239 / 413
Group 2 / 10,000bp / 80	1,236 / 347	224 / 56	69,608 / 69,776	63,166 / 64,055	1,095 / 323
Group 2 / True trees / 90	1,656 / 450	282 / 61	69,096 / 69,317	79,050 / 80,256	1,462 / 425
Group 2 / 1,000bp / 90	1,871 / 627	393 / 86	89,507 / 89,814	63,505 / 64,749	1,644 / 568
Group 2 / 10,000bp / 90	1,705 / 500	296 / 56	72,568 / 72,808	75,607 / 76,812	1,510 / 465

**Table S2.5: Total number of pairs analyzed for consensus sequences for sampling strategy 1 for an average of 500 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals not on ART for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 1.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 1 / True trees / 5	4 / 0	1 / 0	4,655 / 4,656	182 / 186	4 / 0
Group 1 / 1,000bp / 5	7 / 1	2 / 0	4,470 / 4,472	163 / 169	7 / 1
Group 1 / 10,000bp / 5	5 / 0	1 / 0	4,561 / 4,562	173 / 178	5 / 0
Group 1 / True trees / 10	12 / 5	5 / 2	8,723 / 8,726	760 / 767	10 / 4
Group 1 / 1,000bp / 10	24 / 8	8 / 1	8,292 / 8,299	660 / 676	18 / 6
Group 1 / 10,000bp / 10	13 / 5	4 / 2	8,552 / 8,554	749 / 757	11 / 4
Group 1 / True trees / 20	46 / 12	16 / 2	16,256 / 16,270	2,404 / 2,438	40 / 11
Group 1 / 1,000bp / 20	55 / 19	22 / 6	15,938 / 15,954	2,141 / 2,177	45 / 15
Group 1 / 10,000bp / 20	45 / 10	19 / 2	16,035 / 16,052	2,343 / 2,378	39 / 9
Group 1 / True trees / 30	100 / 21	37 / 5	22,757 / 22,789	4,736 / 4,815	76 / 16
Group 1 / 1,000bp / 30	114 / 31	54 / 12	23,066 / 23,108	4,014 / 4,097	86 / 25
Group 1 / 10,000bp / 30	97 / 21	36 / 6	22,556 / 22,586	4,569 / 4,645	73 / 15
Group 1 / True trees / 40	154 / 48	67 / 12	28,563 / 28,618	8,130 / 8,236	122 / 41
Group 1 / 1,000bp / 40	184 / 55	99 / 20	29,269 / 29,348	6,966 / 7,095	146 / 43
Group 1 / 10,000bp / 40	171 / 47	72 / 14	28,358 / 28,416	7,905 / 8,029	133 / 40
Group 1 / True trees / 50	252 / 64	97 / 14	33,551 / 33,634	11,900 / 12,088	220 / 57
Group 1 / 1,000bp / 50	290 / 83	138 / 26	34,574 / 34,686	10,320 / 10,527	249 / 72
Group 1 / 10,000bp / 50	247 / 65	92 / 18	33,438 / 33,512	11,449 / 11,631	214 / 59
Group 1 / True trees / 60	388 / 94	108 / 22	37,746 / 37,832	16,084 / 16,378	331 / 84
Group 1 / 1,000bp / 60	482 / 149	157 / 36	39,589 / 39,710	13,680 / 14,013	398 / 124
Group 1 / 10,000bp / 60	406 / 101	113 / 19	37,665 / 37,759	15,478 / 15,783	348 / 86
Group 1 / True trees / 70	503 / 135	131 / 22	42,207 / 42,316	20,697 / 21,065	429 / 121
Group 1 / 1,000bp / 70	580 / 182	196 / 47	44,560 / 44,709	17,672 / 18,070	471 / 161
Group 1 / 10,000bp / 70	518 / 138	134 / 25	42,252 / 42,361	19,854 / 20,234	438 / 123
Group 1 / True trees / 80	676 / 180	170 / 32	45,864 / 46,002	25,698 / 26,194	569 / 159
Group 1 / 1,000bp / 80	807 / 263	258 / 65	48,866 / 49,059	21,973 / 22,517	657 / 226
Group 1 / 10,000bp / 80	690 / 182	178 / 32	46,060 / 46,206	24,812 / 25,320	577 / 161
Group 1 / True trees / 90	835 / 237	177 / 29	48,527 / 48,675	31,127 / 31,725	700 / 207
Group 1 / 1,000bp / 90	1,005 / 309	301 / 77	52,147 / 52,371	26,763 / 27,459	823 / 266
Group 1 / 10,000bp / 90	863 / 226	192 / 34	48,628 / 48,786	30,149 / 30,786	721 / 206

**Table S2.6: Total number of pairs analyzed for consensus sequences for sampling strategy 1 for an average of 500 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals not on ART for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 2.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 2 / True trees / 5	2 / 2	1 / 0	7,641 / 7,642	392 / 392	1 / 1
Group 2 / 1,000bp / 5	2 / 2	1 / 0	7,827 / 7,828	316 / 316	1 / 1
Group 2 / 10,000bp / 5	3 / 2	1 / 0	7,975 / 7,976	381 / 382	2 / 1
Group 2 / True trees / 10	15 / 3	7 / 3	14,871 / 14,875	1,473 / 1,485	11 / 3
Group 2 / 1,000bp / 10	19 / 5	14 / 4	15,548 / 15,558	1,227 / 1,241	13 / 3
Group 2 / 10,000bp / 10	15 / 3	6 / 2	15,146 / 15,150	1,417 / 1,429	11 / 3
Group 2 / True trees / 20	55 / 7	17 / 3	26,851 / 26,865	5,355 / 5,403	50 / 6
Group 2 / 1,000bp / 20	76 / 13	27 / 8	28,673 / 28,692	4,480 / 4,543	67 / 13
Group 2 / 10,000bp / 20	55 / 7	18 / 3	27,416 / 27,431	5,179 / 5,227	49 / 6
Group 2 / True trees / 30	128 / 37	36 / 3	36,824 / 36,857	10,858 / 10,949	116 / 34
Group 2 / 1,000bp / 30	160 / 45	47 / 8	40,419 / 40,458	8,896 / 9,011	145 / 42
Group 2 / 10,000bp / 30	139 / 38	40 / 7	37,384 / 37,417	10,521 / 10,622	126 / 35
Group 2 / True trees / 40	223 / 54	67 / 13	45,273 / 45,327	17,847 / 18,016	189 / 50
Group 2 / 1,000bp / 40	276 / 88	111 / 24	50,623 / 50,710	14,494 / 14,682	232 / 81
Group 2 / 10,000bp / 40	233 / 60	85 / 14	46,263 / 46,334	17,049 / 17,222	196 / 52
Group 2 / True trees / 50	382 / 89	81 / 14	52,459 / 52,526	25,402 / 25,695	338 / 84
Group 2 / 1,000bp / 50	437 / 128	116 / 23	60,224 / 60,317	20,693 / 21,002	378 / 118
Group 2 / 10,000bp / 50	377 / 92	81 / 15	53,659 / 53,725	24,427 / 24,712	340 / 87
Group 2 / True trees / 60	507 / 136	140 / 31	58,382 / 58,491	34,645 / 35,016	451 / 128
Group 2 / 1,000bp / 60	625 / 200	175 / 39	68,097 / 68,233	28,037 / 28,462	539 / 188
Group 2 / 10,000bp / 60	542 / 148	129 / 28	60,085 / 60,186	33,260 / 33,654	474 / 140
Group 2 / True trees / 70	706 / 205	115 / 19	63,223 / 63,319	44,294 / 44,795	630 / 191
Group 2 / 1,000bp / 70	814 / 225	216 / 41	75,642 / 75,817	35,970 / 36,559	708 / 210
Group 2 / 10,000bp / 70	723 / 204	150 / 33	65,360 / 65,477	42,659 / 43,178	631 / 191
Group 2 / True trees / 80	907 / 235	184 / 41	67,404 / 67,547	55,117 / 55,789	785 / 219
Group 2 / 1,000bp / 80	1,105 / 338	294 / 50	82,170 / 82,414	44,795 / 45,562	946 / 297
Group 2 / 10,000bp / 80	953 / 241	195 / 39	69,917 / 70,073	52,769 / 53,481	829 / 228
Group 2 / True trees / 90	1,207 / 318	202 / 35	71,090 / 71,257	66,239 / 67,128	1,064 / 299
Group 2 / 1,000bp / 90	1,450 / 473	303 / 77	87,963 / 88,189	53,554 / 54,531	1,251 / 425
Group 2 / 10,000bp / 90	1,209 / 335	214 / 43	73,688 / 73,859	63,753 / 64,627	1,061 / 316

**Table S2.7: Total number of pairs analyzed for consensus sequences for sampling strategy 1 for an average of 750 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals not on ART for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 1.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 1 / True trees / 5	3 / 1	1 / 0	4,653 / 4,654	111 / 113	2 / 1
Group 1 / 1,000bp / 5	3 / 1	1 / 0	4,275 / 4,276	103 / 105	2 / 1
Group 1 / 10,000bp / 5	3 / 1	0 / 0	4,596 / 4,596	101 / 103	2 / 1
Group 1 / True trees / 10	5 / 2	2 / 0	8,614 / 8,616	583 / 586	4 / 2
Group 1 / 1,000bp / 10	9 / 2	6 / 1	8,398 / 8,403	525 / 532	8 / 2
Group 1 / 10,000bp / 10	6 / 1	2 / 1	8,508 / 8,509	560 / 565	5 / 1
Group 1 / True trees / 20	46 / 12	13 / 5	15,973 / 15,981	2,254 / 2,288	40 / 11
Group 1 / 1,000bp / 20	44 / 18	24 / 6	15,920 / 15,938	1,962 / 1,988	37 / 15
Group 1 / 10,000bp / 20	45 / 11	14 / 5	15,766 / 15,775	2,165 / 2,199	39 / 10
Group 1 / True trees / 30	72 / 19	25 / 3	22,695 / 22,717	4,368 / 4,421	56 / 18
Group 1 / 1,000bp / 30	79 / 34	43 / 6	22,605 / 22,642	3,767 / 3,812	60 / 26
Group 1 / 10,000bp / 30	84 / 19	33 / 3	22,451 / 22,481	4,250 / 4,315	65 / 18
Group 1 / True trees / 40	137 / 36	46 / 14	28,692 / 28,724	7,087 / 7,188	115 / 32
Group 1 / 1,000bp / 40	172 / 53	74 / 23	29,248 / 29,299	6,012 / 6,131	146 / 43
Group 1 / 10,000bp / 40	140 / 39	55 / 13	28,311 / 28,353	6,826 / 6,927	119 / 32
Group 1 / True trees / 50	249 / 69	79 / 14	33,643 / 33,708	10,417 / 10,597	212 / 65
Group 1 / 1,000bp / 50	305 / 100	120 / 19	34,764 / 34,865	8,913 / 9,118	255 / 89
Group 1 / 10,000bp / 50	249 / 69	80 / 15	33,506 / 33,571	10,049 / 10,229	213 / 65
Group 1 / True trees / 60	335 / 77	80 / 10	38,392 / 38,462	14,599 / 14,857	279 / 70
Group 1 / 1,000bp / 60	408 / 128	147 / 30	40,073 / 40,190	12,316 / 12,596	326 / 107
Group 1 / 10,000bp / 60	339 / 91	94 / 14	38,346 / 38,426	14,063 / 14,311	286 / 78
Group 1 / True trees / 70	424 / 108	123 / 17	42,721 / 42,827	18,696 / 19,012	366 / 98
Group 1 / 1,000bp / 70	493 / 153	191 / 36	44,623 / 44,778	16,009 / 16,349	410 / 134
Group 1 / 10,000bp / 70	440 / 119	125 / 17	42,561 / 42,669	17,980 / 18,301	379 / 110
Group 1 / True trees / 80	582 / 146	149 / 30	46,501 / 46,620	23,396 / 23,832	481 / 129
Group 1 / 1,000bp / 80	710 / 220	228 / 53	49,220 / 49,395	19,956 / 20,446	581 / 189
Group 1 / 10,000bp / 80	588 / 152	161 / 28	46,547 / 46,680	22,662 / 23,098	491 / 136
Group 1 / True trees / 90	772 / 209	150 / 23	49,954 / 50,081	28,302 / 28,865	620 / 175
Group 1 / 1,000bp / 90	921 / 302	229 / 60	53,579 / 53,748	24,127 / 24,746	749 / 246
Group 1 / 10,000bp / 90	770 / 222	158 / 24	50,142 / 50,276	27,398 / 27,946	627 / 186



**Table S2.8: Total number of pairs analyzed for consensus sequences for sampling strategy 1 for an average of 750 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals not on ART for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 2.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 2 / True trees / 5	3 / 2	2 / 0	7,072 / 7,074	381 / 382	3 / 2
Group 2 / 1,000bp / 5	3 / 2	3 / 0	7,241 / 7,244	309 / 310	3 / 2
Group 2 / 10,000bp / 5	3 / 2	1 / 0	7,383 / 7,384	369 / 370	3 / 2
Group 2 / True trees / 10	10 / 4	3 / 2	13,833 / 13,834	1,218 / 1,224	10 / 4
Group 2 / 1,000bp / 10	16 / 3	7 / 3	14,565 / 14,569	982 / 995	15 / 3
Group 2 / 10,000bp / 10	11 / 4	3 / 2	14,261 / 14,262	1,175 / 1,182	11 / 4
Group 2 / True trees / 20	49 / 14	16 / 2	25,536 / 25,550	4,545 / 4,580	44 / 12
Group 2 / 1,000bp / 20	57 / 19	23 / 5	27,397 / 27,415	3,677 / 3,715	49 / 18
Group 2 / 10,000bp / 20	51 / 13	16 / 3	26,074 / 26,087	4,373 / 4,411	46 / 12
Group 2 / True trees / 30	103 / 32	45 / 8	35,661 / 35,698	9,283 / 9,354	92 / 30
Group 2 / 1,000bp / 30	119 / 42	60 / 13	39,246 / 39,293	7,547 / 7,624	106 / 39
Group 2 / 10,000bp / 30	103 / 33	42 / 10	36,504 / 36,536	8,937 / 9,007	92 / 31
Group 2 / True trees / 40	208 / 46	64 / 11	44,058 / 44,111	15,264 / 15,426	183 / 45
Group 2 / 1,000bp / 40	237 / 74	86 / 18	49,702 / 49,770	12,417 / 12,580	211 / 70
Group 2 / 10,000bp / 40	210 / 55	68 / 13	45,044 / 45,099	14,674 / 14,829	188 / 53
Group 2 / True trees / 50	301 / 79	96 / 15	51,300 / 51,381	22,377 / 22,599	255 / 74
Group 2 / 1,000bp / 50	387 / 134	122 / 30	58,646 / 58,738	18,003 / 18,256	317 / 115
Group 2 / 10,000bp / 50	311 / 84	103 / 18	52,619 / 52,704	21,377 / 21,604	265 / 79
Group 2 / True trees / 60	397 / 106	97 / 23	57,623 / 57,697	30,173 / 30,464	346 / 101
Group 2 / 1,000bp / 60	466 / 137	156 / 37	67,230 / 67,349	24,296 / 24,625	396 / 119
Group 2 / 10,000bp / 60	413 / 113	112 / 21	59,372 / 59,463	28,907 / 29,207	359 / 107
Group 2 / True trees / 70	555 / 146	125 / 19	63,383 / 63,489	39,133 / 39,542	477 / 131
Group 2 / 1,000bp / 70	683 / 203	192 / 38	75,332 / 75,486	31,657 / 32,137	587 / 178
Group 2 / 10,000bp / 70	567 / 148	130 / 23	65,206 / 65,313	37,463 / 37,882	488 / 132
Group 2 / True trees / 80	783 / 186	171 / 40	68,131 / 68,262	47,995 / 48,592	670 / 175
Group 2 / 1,000bp / 80	935 / 293	269 / 56	82,195 / 82,408	38,801 / 39,443	785 / 258
Group 2 / 10,000bp / 80	799 / 194	175 / 39	70,725 / 70,861	45,781 / 46,386	686 / 183
Group 2 / True trees / 90	986 / 249	178 / 26	72,406 / 72,558	57,812 / 58,549	853 / 230
Group 2 / 1,000bp / 90	1,253 / 398	273 / 53	87,999 / 88,219	47,211 / 48,066	1,062 / 358
Group 2 / 10,000bp / 90	1,059 / 257	192 / 35	74,942 / 75,099	55,385 / 56,187	928 / 241

**Table S2.9: Total number of pairs analyzed for consensus sequences for sampling strategy 2 for an average of 250 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals independent of ART-status for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 1.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 1 / True trees / 5	7 / 4	2 / 0	10,950 / 10,952	1,057 / 1,060	7 / 4
Group 1 / 1,000bp / 5	11 / 3	3 / 0	10,479 / 10,482	935 / 943	11 / 3
Group 1 / 10,000bp / 5	6 / 3	2 / 0	10,740 / 10,742	1,042 / 1,045	6 / 3
Group 1 / True trees / 10	21 / 5	11 / 3	19,737 / 19,745	3,889 / 3,905	17 / 4
Group 1 / 1,000bp / 10	23 / 6	18 / 7	19,608 / 19,619	3,385 / 3,402	17 / 5
Group 1 / 10,000bp / 10	22 / 5	12 / 3	19,526 / 19,535	3,776 / 3,793	19 / 4
Group 1 / True trees / 20	71 / 10	27 / 4	32,795 / 32,818	13,221 / 13,282	63 / 9
Group 1 / 1,000bp / 20	85 / 18	37 / 11	34,081 / 34,107	11,347 / 11,414	72 / 17
Group 1 / 10,000bp / 20	74 / 12	28 / 4	32,614 / 32,638	12,754 / 12,816	66 / 11
Group 1 / True trees / 30	181 / 38	62 / 3	41,892 / 41,951	25,893 / 26,036	156 / 33
Group 1 / 1,000bp / 30	229 / 65	95 / 12	45,269 / 45,352	21,845 / 22,009	189 / 56
Group 1 / 10,000bp / 30	188 / 40	66 / 6	42,000 / 42,060	24,962 / 25,110	162 / 36
Group 1 / True trees / 40	334 / 73	82 / 12	48,008 / 48,078	41,420 / 41,681	296 / 68
Group 1 / 1,000bp / 40	401 / 103	128 / 24	54,146 / 54,250	34,799 / 35,097	347 / 91
Group 1 / 10,000bp / 40	343 / 74	85 / 17	48,603 / 48,671	39,795 / 40,064	303 / 68
Group 1 / True trees / 50	531 / 110	105 / 27	52,025 / 52,103	58,267 / 58,688	454 / 96
Group 1 / 1,000bp / 50	630 / 166	148 / 35	61,476 / 61,589	48,622 / 49,086	519 / 146
Group 1 / 10,000bp / 50	548 / 117	112 / 25	53,054 / 53,141	55,846 / 56,277	472 / 104
Group 1 / True trees / 60	691 / 187	153 / 32	53,549 / 53,670	77,999 / 78,503	611 / 175
Group 1 / 1,000bp / 60	885 / 260	206 / 45	66,562 / 66,723	65,351 / 65,976	753 / 236
Group 1 / 10,000bp / 60	761 / 192	160 / 31	55,478 / 55,607	74,569 / 75,138	645 / 176
Group 1 / True trees / 70	962 / 236	169 / 37	54,065 / 54,197	99,038 / 99,764	823 / 217
Group 1 / 1,000bp / 70	1177 / 331	256 / 64	70,924 / 71,116	82,323 / 83,169	1,012 / 299
Group 1 / 10,000bp / 70	967 / 242	181 / 41	56,967 / 57,107	95,069 / 95,794	835 / 218
Group 1 / True trees / 80	1,285 / 278	190 / 41	52,866 / 53,015	120,813 / 121,820	1,100 / 259
Group 1 / 1,000bp / 80	1,536 / 435	294 / 66	74,124 / 74,352	100,150 / 101,251	1,303 / 383
Group 1 / 10,000bp / 80	1,290 / 314	207 / 34	56,485 / 56,658	115,938 / 116,914	1,108 / 286
Group 1 / True trees / 90	1,652 / 389	223 / 47	50,265 / 50,441	144,672 / 145,935	1,420 / 356
Group 1 / 1,000bp / 90	1,971 / 556	339 / 85	76,621 / 76,875	120,083 / 121,498	1,647 / 492
Group 1 / 10,000bp / 90	1,646 / 412	225 / 48	55,369 / 55,546	138,894 / 140,128	1,394 / 370

**Table S2.10: Total number of pairs analyzed for consensus sequences for sampling strategy 2 for an average of 250 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals independent of ART-status for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 2.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 2 / True trees / 5	12 / 4	1 / 0	14,923 / 14,924	1,564 / 1,572	9 / 3
Group 2 / 1,000bp / 5	12 / 3	7 / 0	15,215 / 15,222	1,358 / 1,367	10 / 2
Group 2 / 10,000bp / 5	12 / 4	3 / 0	15,001 / 15,004	1,532 / 1,540	9 / 3
Group 2 / True trees / 10	21 / 4	13 / 2	26,871 / 26,882	5,617 / 5,634	20 / 4
Group 2 / 1,000bp / 10	29 / 6	18 / 5	27,914 / 27,927	4,763 / 4,786	26 / 6
Group 2 / 10,000bp / 10	29 / 5	10 / 2	27,014 / 27,022	5,397 / 5,421	27 / 5
Group 2 / True trees / 20	114 / 35	48 / 6	44,976 / 45,018	18,778 / 18,857	104 / 33
Group 2 / 1,000bp / 20	156 / 44	56 / 16	49,166 / 49,206	15,722 / 15,834	138 / 41
Group 2 / 10,000bp / 20	120 / 39	49 / 7	45,481 / 45,523	18,140 / 18,221	109 / 36
Group 2 / True trees / 30	223 / 48	84 / 18	56,990 / 57,056	36,613 / 36,788	198 / 47
Group 2 / 1,000bp / 30	281 / 73	110 / 31	65,752 / 65,831	30,659 / 30,867	246 / 68
Group 2 / 10,000bp / 30	222 / 54	90 / 18	58,358 / 58,430	34,994 / 35,162	196 / 52
Group 2 / True trees / 40	434 / 108	101 / 18	65,285 / 65,368	57,840 / 58,166	384 / 99
Group 2 / 1,000bp / 40	519 / 145	144 / 26	78,328 / 78,446	47,917 / 48,291	442 / 135
Group 2 / 10,000bp / 40	462 / 116	99 / 21	67,597 / 67,675	55,650 / 55,996	426 / 107
Group 2 / True trees / 50	702 / 181	119 / 22	71,043 / 71,140	81,316 / 81,837	640 / 172
Group 2 / 1,000bp / 50	813 / 246	185 / 49	88,953 / 89,089	67,845 / 68,412	723 / 227
Group 2 / 10,000bp / 50	697 / 173	126 / 30	73,392 / 73,488	79,015 / 79,539	625 / 163
Group 2 / True trees / 60	924 / 229	174 / 42	72,636 / 72,768	108,890 / 109,585	825 / 216
Group 2 / 1,000bp / 60	1,130 / 298	250 / 67	96,042 / 96,225	90,196 / 91,028	977 / 272
Group 2 / 10,000bp / 60	948 / 233	188 / 35	76,062 / 76,215	105,170 / 105,885	838 / 216
Group 2 / True trees / 70	1,259 / 322	185 / 46	71,327 / 71,466	139,529 / 140,466	1,125 / 30
Group 2 / 1,000bp / 70	1,529 / 436	292 / 61	102,638 / 102,869	114,311 / 115,404	1,342 / 40
Group 2 / 10,000bp / 70	1,265 / 310	207 / 44	77,189 / 77,352	133,415 / 134,370	1,128 / 29
Group 2 / True trees / 80	1,672 / 430	228 / 52	69,682 / 69,858	170,076 / 171,318	1,491 / 40
Group 2 / 1,000bp / 80	1,949 / 594	325 / 72	102,839 / 103,092	135,143 / 136,498	1,708 / 55
Group 2 / 10,000bp / 80	1,690 / 459	211 / 41	76,145 / 76,315	163,556 / 164,787	1,516 / 43
Group 2 / True trees / 90	2,180 / 500	234 / 39	66,016 / 66,211	202,744 / 204,424	1,914 / 46
Group 2 / 1,000bp / 90	1,759 / 504	269 / 82	73,159 / 73,346	110,487 / 111,742	1,547 / 46
Group 2 / 10,000bp / 90	2,169 / 555	234 / 45	73,944 / 74,133	194,565 / 196,179	1,925 / 52

**Table S2.11: Total number of pairs analyzed for consensus sequences for sampling strategy 2 for an average of 500 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals independent of ART-status for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 1.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 1 / True trees / 5	7 / 3	2 / 0	10,136 / 10,138	879 / 883	6 / 3
Group 1 / 1,000bp / 5	8 / 5	2 / 0	9,720 / 9,722	810 / 813	8 / 5
Group 1 / 10,000bp / 5	7 / 3	3 / 1	9,971 / 9,973	861 / 865	7 / 3
Group 1 / True trees / 10	16 / 2	10 / 0	18,458 / 18,468	3,200 / 3,214	13 / 2
Group 1 / 1,000bp / 10	23 / 5	13 / 3	18,231 / 18,241	2,869 / 2,887	17 / 4
Group 1 / 10,000bp / 10	20 / 2	12 / 1	18,146 / 18,157	3,104 / 3,122	15 / 2
Group 1 / True trees / 20	64 / 13	17 / 1	32,033 / 32,049	10,508 / 10,559	59 / 13
Group 1 / 1,000bp / 20	84 / 19	35 / 5	33,005 / 33,035	9,112 / 9,177	73 / 17
Group 1 / 10,000bp / 20	69 / 14	17 / 2	32,015 / 32,030	10,103 / 10,158	61 / 14
Group 1 / True trees / 30	149 / 37	43 / 12	41,887 / 41,918	21,309 / 21,421	124 / 31
Group 1 / 1,000bp / 30	213 / 64	67 / 15	44,451 / 44,503	18,181 / 18,330	170 / 55
Group 1 / 10,000bp / 30	153 / 39	50 / 12	41,816 / 41,854	20,551 / 20,665	126 / 32
Group 1 / True trees / 40	270 / 69	69 / 10	48,905 / 48,964	34,112 / 34,313	232 / 63
Group 1 / 1,000bp / 40	339 / 97	119 / 23	53,443 / 53,539	28,899 / 29,141	278 / 89
Group 1 / 10,000bp / 40	273 / 69	76 / 11	49,336 / 49,401	32,691 / 32,895	231 / 63
Group 1 / True trees / 50	434 / 108	92 / 18	54,512 / 54,586	48,408 / 48,734	364 / 97
Group 1 / 1,000bp / 50	521 / 152	134 / 25	61,520 / 61,629	40,921 / 41,290	423 / 134
Group 1 / 10,000bp / 50	451 / 112	81 / 21	55,189 / 55,249	46,531 / 46,870	379 / 101
Group 1 / True trees / 60	553 / 130	113 / 25	57,879 / 57,967	64,839 / 65,262	474 / 120
Group 1 / 1,000bp / 60	711 / 185	188 / 34	67,814 / 67,956	54,443 / 54,975	590 / 163
Group 1 / 10,000bp / 60	598 / 145	112 / 17	59,254 / 59,349	62,320 / 62,773	502 / 125
Group 1 / True trees / 70	822 / 190	153 / 25	59,913 / 60,041	81,836 / 82,468	692 / 172
Group 1 / 1,000bp / 70	997 / 288	227 / 49	73,485 / 73,663	68,945 / 69,654	832 / 247
Group 1 / 10,000bp / 70	861 / 221	154 / 22	61,784 / 61,916	79,197 / 79,837	728 / 200
Group 1 / True trees / 80	1,055 / 232	172 / 38	60,778 / 60,912	101,439 / 102,262	915 / 211
Group 1 / 1,000bp / 80	1,331 / 375	275 / 66	77,285 / 77,516	85,271 / 86,237	1101 / 324
Group 1 / 10,000bp / 80	1,124 / 279	181 / 42	62,887 / 63,026	97,804 / 98,649	946 / 247
Group 1 / True trees / 90	1,355 / 341	199 / 28	60,163 / 60,334	121,581 / 122,595	1,159 / 309
Group 1 / 1,000bp / 90	1,619 / 479	311 / 73	81,153 / 81,391	101,383 / 102,523	1,342 / 421
Group 1 / 10,000bp / 90	1,400 / 339	217 / 40	64,395 / 64,572	116,420 / 117,481	1,169 / 309

**Table S2.12: Total number of pairs analyzed for consensus sequences for sampling strategy 2 for an average of 500 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals independent of ART-status for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 2.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 2 / True trees / 5	5 / 0	1 / 0	13,243 / 13,244	1,269 / 1,274	5 / 0
Group 2 / 1,000bp / 5	8 / 0	4 / 1	13,358 / 13,361	1,126 / 1,134	6 / 0
Group 2 / 10,000bp / 5	6 / 0	1 / 0	13,431 / 13,432	1,214 / 1,220	5 / 0
Group 2 / True trees / 10	14 / 3	5 / 1	24,495 / 24,499	4,318 / 4,329	14 / 3
Group 2 / 1,000bp / 10	22 / 3	13 / 4	25,081 / 25,090	3,696 / 3,715	19 / 3
Group 2 / 10,000bp / 10	13 / 2	6 / 1	24,796 / 24,801	4,229 / 4,240	13 / 2
Group 2 / True trees / 20	88 / 23	18 / 3	42,136 / 42,151	14,434 / 14,499	78 / 22
Group 2 / 1,000bp / 20	117 / 37	24 / 8	45,566 / 45,582	12,069 / 12,149	104 / 35
Group 2 / 10,000bp / 20	90 / 25	25 / 3	42,641 / 42,663	13,832 / 13,897	81 / 24
Group 2 / True trees / 30	186 / 49	51 / 10	54,849 / 54,890	28,538 / 28,675	158 / 47
Group 2 / 1,000bp / 30	244 / 76	81 / 19	61,315 / 61,377	24,146 / 24,314	197 / 67
Group 2 / 10,000bp / 30	192 / 52	55 / 10	55,687 / 55,732	27,514 / 27,654	161 / 48
Group 2 / True trees / 40	313 / 71	88 / 12	63,748 / 63,824	45,679 / 45,921	277 / 69
Group 2 / 1,000bp / 40	432 / 100	131 / 29	74,495 / 74,581	38,226 / 38,564	384 / 96
Group 2 / 10,000bp / 40	335 / 84	74 / 13	65,858 / 65,919	44,201 / 44,452	300 / 83
Group 2 / True trees / 50	539 / 121	126 / 23	70,438 / 70,541	64,935 / 65,353	482 / 112
Group 2 / 1,000bp / 50	564 / 146	170 / 40	85,248 / 85,374	54,414 / 54,836	490 / 138
Group 2 / 10,000bp / 50	496 / 126	116 / 18	72,612 / 72,710	62,906 / 63,276	437 / 116
Group 2 / True trees / 60	711 / 169	143 / 28	74,381 / 74,496	87,317 / 87,859	635 / 166
Group 2 / 1,000bp / 60	900 / 239	180 / 51	94,240 / 94,343	72,560 / 73,235	801 / 226
Group 2 / 10,000bp / 60	766 / 170	150 / 33	77,846 / 77,963	84,068 / 84,664	693 / 163
Group 2 / True trees / 70	955 / 247	128 / 21	76,794 / 76,901	111,023 / 111,731	838 / 231
Group 2 / 1,000bp / 70	1,182 / 353	209 / 54	101,861 / 101,994	91,880 / 92,705	1017 / 31
Group 2 / 10,000bp / 70	1,004 / 262	149 / 27	81,233 / 81,355	106,792 / 107,534	885 / 242
Group 2 / True trees / 80	1,165 / 285	169 / 31	76,857 / 76,995	136,545 / 137,425	1,025 / 27
Group 2 / 1,000bp / 80	1,468 / 422	255 / 51	103,975 / 104,209	109,406 / 110,412	1,270 / 38
Group 2 / 10,000bp / 80	1,242 / 303	178 / 35	82,628 / 82,771	131,272 / 132,211	1,097 / 29
Group 2 / True trees / 90	1,622 / 390	171 / 22	76,351 / 76,500	163,524 / 164,756	1,451 / 37
Group 2 / 1,000bp / 90	1,698 / 530	251 / 75	96,971 / 109,019	116,518 / 132,334	1,460 / 48
Group 2 / 10,000bp / 90	1,626 / 395	206 / 30	82,394 / 82,570	157,352 / 158,583	1,421 / 36

**Table S2.13: Total number of pairs analyzed for consensus sequences for sampling strategy 2 for an average of 750 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals independent of ART-status for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 1.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 1 / True trees / 5	5 / 1	2 / 0	9,498 / 9,500	693 / 697	5 / 1
Group 1 / 1,000bp / 5	4 / 0	3 / 2	9,145 / 9,146	624 / 628	3 / 0
Group 1 / 10,000bp / 5	5 / 1	2 / 0	9,444 / 9,446	665 / 669	5 / 1
Group 1 / True trees / 10	10 / 5	8 / 0	17,804 / 17,812	2,600 / 2,605	6 / 5
Group 1 / 1,000bp / 10	15 / 8	15 / 1	17,215 / 17,229	2,315 / 2,322	9 / 6
Group 1 / 10,000bp / 10	12 / 6	9 / 0	17,709 / 17,718	2,518 / 2,524	7 / 5
Group 1 / True trees / 20	69 / 23	17 / 3	31,549 / 31,563	8,577 / 8,623	59 / 20
Group 1 / 1,000bp / 20	91 / 32	29 / 4	32,059 / 32,084	7,581 / 7,640	74 / 27
Group 1 / 10,000bp / 20	73 / 21	21 / 3	31,199 / 31,217	8,319 / 8,371	63 / 18
Group 1 / True trees / 30	135 / 25	26 / 8	41,212 / 41,230	18,099 / 18,209	115 / 21
Group 1 / 1,000bp / 30	177 / 36	55 / 13	42,639 / 42,681	15,671 / 15,812	147 / 29
Group 1 / 10,000bp / 30	143 / 28	33 / 9	40,979 / 41,003	17,491 / 17,606	122 / 23
Group 1 / True trees / 40	252 / 61	54 / 14	49,708 / 49,748	28,596 / 28,787	209 / 55
Group 1 / 1,000bp / 40	291 / 79	88 / 24	52,698 / 52,762	24,519 / 24,731	240 / 71
Group 1 / 10,000bp / 40	254 / 63	62 / 14	49,698 / 49,746	27,642 / 27,833	213 / 56
Group 1 / True trees / 50	326 / 73	77 / 12	55,663 / 55,728	41,160 / 41,413	278 / 67
Group 1 / 1,000bp / 50	439 / 99	129 / 34	61,151 / 61,246	35,027 / 35,367	369 / 89
Group 1 / 10,000bp / 50	353 / 76	78 / 15	56,092 / 56,155	39,595 / 39,872	295 / 69
Group 1 / True trees / 60	500 / 124	98 / 30	59,882 / 59,950	55,346 / 55,722	432 / 116
Group 1 / 1,000bp / 60	624 / 184	182 / 41	67,952 / 68,093	47,040 / 47,480	526 / 167
Group 1 / 10,000bp / 60	513 / 126	133 / 38	61,433 / 61,528	52,779 / 53,166	418 / 117
Group 1 / True trees / 70	695 / 172	145 / 22	63,261 / 63,384	71,077 / 71,600	592 / 157
Group 1 / 1,000bp / 70	863 / 234	213 / 44	73,313 / 73,482	60,103 / 60,732	725 / 206
Group 1 / 10,000bp / 70	703 / 186	154 / 22	65,124 / 65,256	68,375 / 68,892	598 / 170
Group 1 / True trees / 80	880 / 217	167 / 34	66,051 / 66,184	86,342 / 87,005	713 / 191
Group 1 / 1,000bp / 80	1061 / 297	228 / 49	78,714 / 78,893	74,091 / 74,855	889 / 272
Group 1 / 10,000bp / 80	937 / 248	166 / 21	67,550 / 67,695	83,861 / 84,550	791 / 226
Group 1 / True trees / 90	1,230 / 308	191 / 42	66,987 / 67,136	104,344 / 105,266	1,038 / 277
Group 1 / 1,000bp / 90	1,481 / 412	305 / 63	83,267 / 83,509	87,895 / 88,964	1,204 / 363
Group 1 / 10,000bp / 90	1,259 / 304	191 / 42	69,693 / 69,842	100,447 / 101,402	1,044 / 269

**Table S2.14: Total number of pairs analyzed for consensus sequences for sampling strategy 2 for an average of 750 migrants per year to and from *region*.** Absolute number for true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) estimated for infector probabilities using a threshold of 80% and 90%. Here we quantified whether a pair represented a transmission pair independent of who infected whom. From those pairs classified as TP, we also report the total number of pairs in which who infected whom (wiw) was correctly identified. Results are reported for different sampling depth of 5 to 90% of the diagnosed individuals independent of ART-status for true trees, ML estimated with 1,000bp and 10,000bp for parameter group 2.

Parameters / Tree / Perc.	Threshold 80% / 90%				
	TP	FP	TN	FN	WIW
Group 2 / True trees / 5	6 / 0	2 / 0	12,062 / 12,064	894 / 900	5 / 0
Group 2 / 1,000bp / 5	8 / 0	3 / 0	12,235 / 12,238	768 / 776	7 / 0
Group 2 / 10,000bp / 5	3 / 0	2 / 0	12,294 / 12,296	895 / 898	3 / 0
Group 2 / True trees / 10	24 / 5	5 / 1	22,777 / 22,781	3,374 / 3,393	21 / 5
Group 2 / 1,000bp / 10	32 / 10	8 / 1	23,664 / 23,671	2,864 / 2,886	26 / 9
Group 2 / 10,000bp / 10	23 / 7	4 / 1	23,074 / 23,077	3,281 / 3,297	20 / 6
Group 2 / True trees / 20	72 / 15	27 / 8	39,949 / 39,968	11,644 / 11,701	60 / 13
Group 2 / 1,000bp / 20	87 / 22	35 / 6	43,303 / 43,332	9,937 / 10,002	72 / 17
Group 2 / 10,000bp / 20	73 / 15	24 / 8	40,348 / 40,364	11,291 / 11,349	61 / 14
Group 2 / True trees / 30	144 / 32	45 / 6	53,415 / 53,454	22,950 / 23,062	121 / 31
Group 2 / 1,000bp / 30	192 / 53	74 / 19	59,638 / 59,693	19,214 / 19,353	163 / 44
Group 2 / 10,000bp / 30	149 / 37	43 / 10	54,383 / 54,416	22,091 / 22,203	129 / 36
Group 2 / True trees / 40	265 / 55	72 / 13	63,412 / 63,471	37,171 / 37,381	226 / 51
Group 2 / 1,000bp / 40	344 / 88	106 / 23	73,122 / 73,205	31,112 / 31,368	288 / 80
Group 2 / 10,000bp / 40	267 / 53	67 / 11	64,857 / 64,913	35,995 / 36,209	236 / 49
Group 2 / True trees / 50	377 / 97	96 / 21	70,662 / 70,737	53,905 / 54,185	330 / 91
Group 2 / 1,000bp / 50	517 / 139	159 / 42	84,743 / 84,860	45,357 / 45,735	442 / 129
Group 2 / 10,000bp / 50	411 / 106	85 / 13	72,587 / 72,659	52,065 / 52,370	362 / 98
Group 2 / True trees / 60	592 / 148	113 / 27	76,307 / 76,393	71,974 / 72,418	525 / 136
Group 2 / 1,000bp / 60	748 / 190	213 / 44	93,997 / 94,166	60,062 / 60,620	651 / 173
Group 2 / 10,000bp / 60	609 / 165	139 / 34	78,579 / 78,684	69,445 / 69,889	545 / 153
Group 2 / True trees / 70	798 / 205	131 / 32	79,905 / 80,004	92,262 / 92,855	701 / 196
Group 2 / 1,000bp / 70	971 / 270	195 / 49	102,215 / 102,361	77,165 / 77,866	820 / 241
Group 2 / 10,000bp / 70	840 / 189	138 / 30	83,580 / 83,688	89,208 / 89,859	735 / 174
Group 2 / True trees / 80	1,060 / 250	163 / 23	81,879 / 82,019	114,070 / 114,880	955 / 242
Group 2 / 1,000bp / 80	1,356 / 359	245 / 60	109,403 / 109,588	95,024 / 96,021	1,183 / 33
Group 2 / 10,000bp / 80	1,118 / 279	165 / 32	86,313 / 86,446	110,102 / 110,941	989 / 262
Group 2 / True trees / 90	1,312 / 357	186 / 36	83,676 / 83,826	137,006 / 137,961	1,146 / 32
Group 2 / 1,000bp / 90	1,721 / 529	292 / 70	114,986 / 115,208	113,525 / 114,717	1,469 / 48
Group 2 / 10,000bp / 90	1,428 / 368	208 / 46	88,640 / 88,802	131,578 / 132,638	1,247 / 34