ONLINE SUPPLEMENT TO: Mixed method estimation of population level HIV viral suppression rate in the Western Cape, South Africa

Authors:

Elton Mukonda MPhil[1], Nei-Yuan Hsiao[2,3], Lara Vojnov PhD[4], Landon Myer PhD[1,5], Maia Lesosky PhD[1]

Affiliations:

[1]Division of Epidemiology & Biostatistics, School of Public Health & Family Medicine, University of Cape Town, Cape Town, South Africa;

[2]National Health Laboratory Services, Cape Town, South Africa;

[3]Division of Medical Virology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa;

[4]World Health Organization, Geneva, Switzerland;

[5]Centre for Infectious Disease Epidemiology & Research, School of Public Health & Family Medicine, University of Cape Town, Cape Town, South Africa;

Correspondence address:

Maia Lesosky, Division of Epidemiology & Biostatistics, School of Public Health & Family Medicine, University of Cape Town Faculty of Health Sciences, Anzio Road, Observatory 7925, South Africa. Email: maia.lesosky@uct.ac.za; Phone: +27612504532

SUPPLEMENT METHODS

Data processing and record linkage

Data cleaning and preparation stage

The cleaning process includes trimming of blank spaces, removal of titles in name fields, removal of special characters, splitting of multiple given names in one variable into multiple variables, joining of surnames with prefixes into one name, converting all string variables to lower case and standardization of date formats.

Exact linkage stage

As movement between facilities is likely, the viral load tests for each individual were identified using the patient name, year of birth, gender and in some cases, existing patient identifiers. The main steps included creating four separate datasets based on the aggregate number of tests corresponding to each grouping of the reported patient identifier, and the combination of the last name and the first name initials. These datasets were dataset 1: Matching totals after grouping for the patient identifier and for the combination of the last name, and first name initial (≥ 2 tests); dataset 2: Matching totals after grouping for the patient identifier and for the combination of the last name, and first name initial (1 test); dataset 3: Non-matching totals after grouping for the patient identifier and for the combination of the last name, and first name initial; dataset 4: No patient identifier.

Using dataset 1 as the reference, pairwise evaluations between the reference set and the comparative data sets were made and similarity scores using the Jaro-Winkler (JW) algorithm [1] and hierarchical clustering using the Jaro-Winkler distance metric were calculated for first names, second names independently. Scores were combined, and any score above a specified threshold was used to indicate a match. Comparing the names from dataset 1 to matched names in datasets 2-4, records were then classified as exact matches, probable matches, plausible matches, and non-matches. For example, records were exact matches if they had a summary similarity score of 1, the same year of birth, gender, and were taken from the same healthcare facility.

Records were determined to be from the same individual if the total similarity score was above a threshold value, the difference between the years of birth was not more than 3-years and gender was the same for the two entities. Matches were integrated into the reference set, and

the algorithm run iteratively until no further matches could be identified. Unique identifiers are assigned to all matched records and all remaining records for which insufficient data exists for any type of meaningful linking to be possible. The same procedure was performed for comparing names from dataset 2 vs 3-4 (dataset 2 as the reference), and then for dataset 3 vs 4 (dataset 3 as the reference). Hierarchical clustering was used for classifying the remaining records in dataset 4.

3

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Global Health

Supplement Table 1: Summary and comparison of test-level and person-level viral suppression (VL<400 copies/mL) between 2008 and 2018 for all tests and for individuals in care, disaggregated by sex.

	Male		Females		Overall	
Year	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)
2008	14167(81.3)	10479(82.1)	27814(82.5)	20716(83.8)	41981(82.1)	31195(83.2)
2009	26156(82.1)	16142(81.8)	51986(83.6)	31720(83.9)	78142(83.1)	47862(83.2)
2010	25581(80.7)	19252(82.9)	50053(81.4)	37706(83.8)	75634(81.2)	56958(83.5)
2011	32761(76.6)	26330(81.9)	65585(77.3)	52705(83.1)	98346(77)	79035(82.7)
2012	40673(78.5)	33426(83.2)	83376(80.4)	68338(85.5)	124049(79.8)	101764(84.7)
2013	49262(78.5)	40214(83.7)	101471(80.1)	83787(85.6)	150733(79.5)	124001(85)
2014	50515(78.6)	43153(83.9)	111995(81.3)	94042(86.2)	162510(80.4)	137195(85.5)
2015	59449(79.6)	51397(84.5)	138215(82.9)	114684(87.2)	197664(81.9)	166081(86.4)
2016	70554(80.3)	61199(85.8)	161286(83.2)	132835(88.1)	231839(82.3)	194034(87.4)
2017	76561(80.9)	67864(86.3)	177907(84.2)	148875(88.8)	254468(83.2)	216739(88)
2018	62346(82.6)	59317(86.5)	143828(85.5)	130980(88.8)	206174(84.6)	190297(88.1)

Supplement Table 2: Summary and comparison of test-level and person level viral suppression (VL<1000 copies/mL) between 2008 and 2018 disaggregated for individuals <25 years.

Age	<12 months		12-23 months		24-59 months		5-14 years		15-24 years	
Year	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)
2008	279(38.3)	201(36.8)	403(63.5)	288(64.1)	1395(73.5)	1033(76.6)	2496(80.1)	1854(82.5)	1947(77.4)	1497(78.4)
2009	404(41.4)	270(41.6)	630(65.5)	384(65.3)	2037(78)	1205(78.8)	4424(82.6)	2594(82.7)	3754(79.8)	2403(79.8)
2010	287(36.9)	204(36.5)	526(67.3)	372(69.8)	1590(76.2)	1187(78.5)	3988(81.1)	2927(84.3)	3658(77.4)	2742(78.7)
2011	225(31.9)	147(32.7)	521(58)	382(63.2)	1696(70.8)	1325(76.7)	4703(77.2)	3718(83.4)	5166(72.4)	3958(76.5)
2012	233(32.2)	159(35.1)	406(51.9)	301(58.6)	1652(69.1)	1336(76.1)	5099(79)	4139(84.4)	6961(76.7)	5561(80.7)
2013	223(30.8)	149(32.5)	374(50.3)	261(56.4)	1647(69)	1269(75.8)	5632(78)	4560(84.6)	8678(74.8)	7003(79.5)
2014	189(29.8)	133(31.7)	366(51.3)	271(56)	1446(68.1)	1107(73.5)	5390(78)	4477(84)	10966(76.8)	8703(80.6)
2015	317(49.7)	218(57.8)	381(50.1)	269(55.3)	1387(69.1)	1048(74.3)	5804(78.4)	4915(85)	14204(80)	10856(83.7)
2016	241(61)	202(66.9)	220(47.4)	156(52.5)	1296(66.4)	1003(74.5)	5977(78.9)	5167(86.2)	15623(80.6)	11924(84.7)
2017	198(65.6)	156(69.6)	251(51.2)	180(57)	1103(63.2)	873(70.2)	5912(78.3)	5196(86.6)	17365(81.3)	13449(85)
2018	158(63.7)	133(68.2)	182(53.4)	148(59)	760(64.5)	671(70.6)	4372(80.6)	4092(85.3)	13542(81.7)	11764(85.3)
Overall	2754(40.2)	1972(42.6)	4260(56.2)	3012(60.4)	16009(70.3)	12057(75.3)	53797(79.1)	43639(84.7)	101864(78.9)	79860(82.6)

5

Supplement Table 3: Summary and comparison of test-level and person level viral suppression (VL<1000 copies/mL) between 2008 and 2018 disaggregated for individuals > 25 years.

	25-34 years		35-44 years		45-54 years		55-64 years		65+ years	
Year	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)	Tests <1000 copies/mL N (%)	Individuals with VL<1000 copies/mL N (%)
2008	16030(85.3)	11860(86.1)	14163(87.6)	10510(88.5)	5113(89.2)	3738(89.7)	1225(91.6)	896(91.5)	202(95.3)	158(94)
2009	29603(85.7)	18097(85.7)	27119(88.2)	16612(88.2)	10082(90.4)	6190(90.2)	2384(91.9)	1443(92.4)	386(93.9)	254(92.7)
2010	28344(83.5)	21131(85.1)	26875(86.7)	20151(88.2)	10456(88.8)	7769(89.9)	2478(90.5)	1837(91.6)	393(91.8)	290(92.1)
2011	35839(78.1)	28404(83.5)	35398(81.5)	28576(86.6)	14081(84.8)	11221(88.9)	3354(87)	2720(91)	562(88.4)	448(90.1)
2012	45163(81.6)	36496(86.2)	45338(84)	37189(88.4)	17882(85.9)	14698(89.5)	4453(88.2)	3657(91.3)	769(87.8)	630(89.6)
2013	54147(81)	43922(85.9)	55659(84)	45550(88.7)	22315(85.8)	18426(89.9)	5830(88)	4819(91.7)	1046(89.3)	869(92.3)
2014	59080(82.5)	48327(86.7)	58641(84.8)	49988(89.1)	23970(85.7)	20646(89.8)	6480(88)	5583(91.2)	1123(89.2)	973(91.5)
2015	70128(84.2)	56490(87.8)	71672(85.9)	61138(89.7)	29943(85.8)	26296(89.7)	8423(88.4)	7359(91.3)	1452(87.6)	1280(89.6)
2016	79631(84.5)	62895(88.5)	85587(86)	72677(90.5)	37596(86.2)	33004(90.7)	10948(88.2)	9578(92.2)	1978(88.8)	1732(92.7)
2017	84405(85.4)	67790(89.2)	94034(86.7)	81374(91)	43571(86.9)	38855(91.2)	12880(89.1)	11506(92.4)	2413(90.4)	2155(93.5)
2018	65137(86.3)	57477(89.2)	77355(88)	72144(91.1)	37033(88.5)	35388(91.5)	11048(89.2)	10596(92)	2184(91.5)	2081(94)
Overall	567507(83.6)	452889(87.3)	591841(85.8)	495909(89.7)	252042(86.7)	216231(90.4)	69503(88.7)	59994(91.8)	12508(89.8)	10870(92.3)

6

Supplement Table 4: Summary and comparison of test-level and person level viral suppression (VL<400 copies/mL) between 2008 and 2018 disaggregated

for individuals <25 years.

Year	<12 months		12-23 months		24-59 months		5-14 years		15-24 years	
	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)
2008	267(36.6)	192(35.2)	381(60)	273(60.8)	1335(70.4)	994(73.7)	2402(77.1)	1785(79.4)	1869(74.3)	1438(75.3)
2009	365(37.4)	250(38.5)	593(61.6)	358(60.9)	1960(75.1)	1159(75.8)	4239(79.2)	2478(79)	3587(76.2)	2299(76.4)
2010	255(32.8)	186(33.3)	476(60.9)	350(65.7)	1493(71.5)	1141(75.4)	3793(77.1)	2832(81.5)	3492(73.9)	2633(75.6)
2011	193(27.3)	133(29.6)	457(50.8)	346(57.3)	1596(66.6)	1269(73.5)	4513(74.1)	3622(81.2)	4983(69.8)	3841(74.3)
2012	209(28.9)	146(32.2)	362(46.2)	282(54.9)	1535(64.2)	1270(72.3)	4851(75.2)	3998(81.6)	6737(74.2)	5413(78.5)
2013	187(25.8)	125(27.3)	328(44.1)	240(51.8)	1539(64.5)	1215(72.6)	5371(74.3)	4416(82)	8333(71.8)	6786(77.1)
2014	163(25.7)	117(27.9)	317(44.4)	243(50.2)	1348(63.5)	1067(70.8)	5070(73.4)	4310(80.9)	10560(73.9)	8442(78.2)
2015	281(44)	206(54.6)	344(45.2)	247(50.8)	1268(63.1)	988(70)	5517(74.5)	4759(82.3)	13658(76.9)	10517(81.1)
2016	229(58)	196(64.9)	201(43.3)	148(49.8)	1191(61)	947(70.3)	5671(74.9)	4999(83.4)	15034(77.6)	11558(82.1)
2017	184(60.9)	152(67.9)	224(45.7)	166(52.5)	1015(58.2)	820(66)	5624(74.5)	5036(83.9)	16687(78.1)	13003(82.2)
2018	151(60.9)	127(65.1)	163(47.8)	138(55)	695(58.9)	616(64.8)	4139(76.3)	3924(81.8)	13011(78.5)	11371(82.4)
Overall	2484(36.2)	1830(39.6)	3846(50.8)	2791(56)	14975(65.8)	11486(71.7)	51190(75.3)	42159(81.9)	97951(75.8)	77301(79.9)

7

Supplement Table 5: Summary and comparison of test-level and person level viral suppression (VL<400 copies/mL) between 2008 and 2018 disaggregated

for individuals > 25 years.

	25-34 years		35-44 years		45-54 years		55-64 years		65+ years	
Year	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)	Tests <400 copies/mL N (%)	Individuals with VL<400 copies/mL N (%)
2008	15481(82.4)	11508(83.5)	13826(85.5)	10291(86.6)	5019(87.6)	3676(88.2)	1201(89.8)	881(90)	200(94.3)	157(93.5)
2009	28526(82.6)	17460(82.7)	26330(85.6)	16161(85.8)	9832(88.2)	6038(88)	2336(90)	1412(90.4)	374(91)	247(90.1)
2010	27263(80.3)	20478(82.4)	25981(83.8)	19668(86.1)	10097(85.7)	7586(87.8)	2400(87.7)	1799(89.7)	384(89.7)	285(90.5)
2011	34700(75.6)	27707(81.5)	34396(79.2)	28008(84.9)	13690(82.4)	10991(87)	3274(84.9)	2677(89.6)	544(85.5)	441(88.7)
2012	43796(79.1)	35630(84.1)	44079(81.6)	36431(86.6)	17378(83.5)	14375(87.5)	4344(86)	3592(89.6)	758(86.5)	627(89.2)
2013	52487(78.6)	42929(84)	54011(81.6)	44580(86.8)	21734(83.5)	18108(88.3)	5713(86.2)	4743(90.3)	1030(88)	859(91.3)
2014	57319(80.1)	47282(84.8)	56971(82.4)	48991(87.3)	23319(83.3)	20279(88.2)	6350(86.2)	5508(90)	1093(86.8)	956(89.9)
2015	68117(81.8)	55257(85.9)	69751(83.6)	59926(87.9)	29112(83.4)	25722(87.8)	8193(86)	7205(89.4)	1423(85.8)	1254(87.8)
2016	77273(82)	61548(86.6)	83128(83.6)	71222(88.7)	36529(83.8)	32327(88.8)	10657(85.8)	9387(90.3)	1925(86.4)	1701(91)
2017	82018(83)	66293(87.3)	91470(84.4)	79769(89.2)	42337(84.4)	38080(89.4)	12542(86.8)	11298(90.7)	2367(88.7)	2122(92.1)
2018	63508(84.2)	56275(87.3)	75472(85.9)	70723(89.3)	36078(86.2)	34663(89.6)	10810(87.2)	10408(90.3)	2147(89.9)	2052(92.7)
Overall	550488(81.1)	442367(85.3)	575415(83.4)	485770(87.8)	245125(84.4)	211845(88.6)	67820(86.5)	58910(90.1)	12245(87.9)	10701(90.8)

8

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Global Health

Supplement Table 6: Summary of sensitivity analysis due to choice of linkage methods and

resulting cohort characteristics

| Linkage method	Number of unique individuals	Median (IQR) tests
Deterministic: Full name	478,560	7 (4, 10)
Deterministic: Existing ID	508,955	6 (3, 9)
Deterministic + Probabilistic (JW=1)	476,080	7 (4, 10
Deterministic + Probabilistic (JW=0.9)	474,595	7 (4, 10)
Deterministic + Probabilistic (JW=0.8)	469,565	7 (4, 10)
Deterministic + Probabilistic (JW=0.7)	465,457	7 (4, 10)
Deterministic + Probabilistic (JW=0.6)	463,476	7 (4, 10)
Deterministic + Probabilistic (JW=0.5)	462,859	7 (4, 10)

Supplement Table 7: Estimated population viral suppression rate for all people living with HIV

in the Western Cape using VL<400 as threshold

Year	End of year populations^	Total HIV Prevalence#	HIV infected population	VL<400 Copies/ mL	Proportion virally suppressed^^
2008	5308512	4,9%	261332	31095	11.9%
2009	5288519	5,2%	275266	47761	17.4%
2010	5254963	5,5%	287057	58904	20.5%
2011	5455552	5,7%	310433	79031	25.5%
2012	5819473	5,9%	342790	101803	29.7%
2013	6063550	6,1%	368182	124120	33.7%
2014	6157898	6,2%	383989	137261	35.7%
2015	6222410	6,4%	397189	166253	41.9%
2016	6396236	6,5%	416515	194069	46.6%
2017	6564961	6,6%	434634	216991	49.9%
2018*	6655523	6,7%	446303	190485	42.7%
^ Interpolated from SA mid-year population estimates [2-12]					
*Estimates for the end of September 2018					
# Obtained from the Thembisa model [13]					
^^ Assumes those not in care have VL>400 cps/mL					

9

Supplement Table 8: Estimated population viral suppression rate (VL<1000) for all people
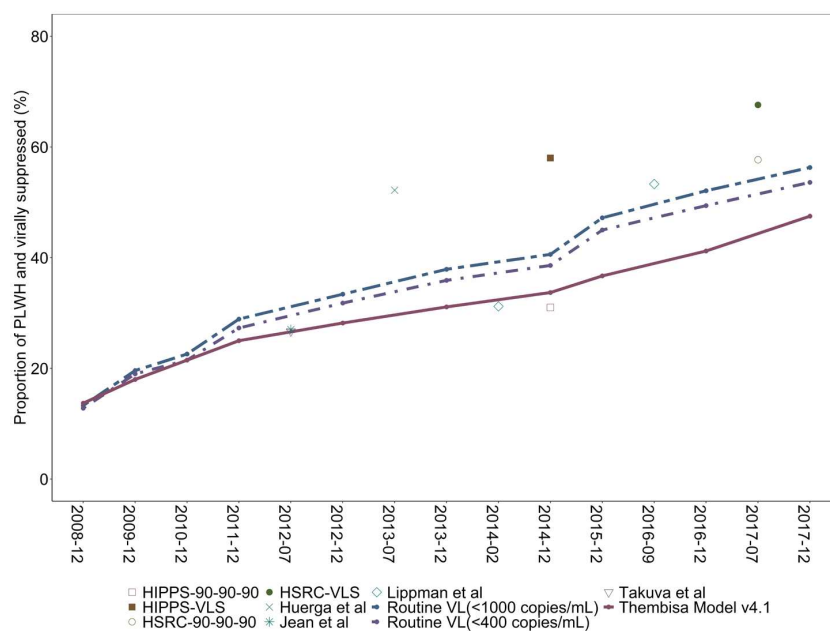
living with HIV in the Western Cape by linkage method

| Year | Deterministic | | Deterministic + Probabilistic | | | | | | No Linkage |
	Full name	Existing ID	1	0,9	0,8	0,7	0,6	0,5	
2008	12.1%	12.4%	12.3%	12.3%	12.2%	12.2%	12.2%	12.1%	14.3%
2009	17.7%	19.0%	18.0%	18.0%	17.9%	17.9%	17.7%	17.6%	20.9%
2010	20.3%	20.0%	20.4%	20.4%	20.4%	20.3%	20.1%	20.0%	23.8%
2011	25.8%	25.6%	26.1%	26.1%	26.0%	25.9%	25.7%	25.6%	30.8%
2012	30.0%	30.2%	30.4%	30.4%	30.4%	30.3%	30.2%	30.1%	35.0%
2013	33.9%	34.3%	34.5%	34.4%	34.4%	34.4%	34.3%	34.2%	39.6%
2014	35.9%	36.5%	36.5%	36.5%	36.5%	36.4%	36.3%	36.3%	41.8%
2015	42.3%	42.6%	42.8%	42.8%	42.7%	42.7%	42.6%	42.5%	48.4%
2016	46.7%	47.3%	47.6%	47.6%	47.6%	47.5%	47.4%	47.3%	53.3%
2017	49.9%	50.8%	51.0%	51.0%	50.9%	50.9%	50.8%	50.7%	56.7%
2018	42.8%	43.5%	43.6%	43.6%	43.6%	43.5%	43.5%	43.4%	48.4%
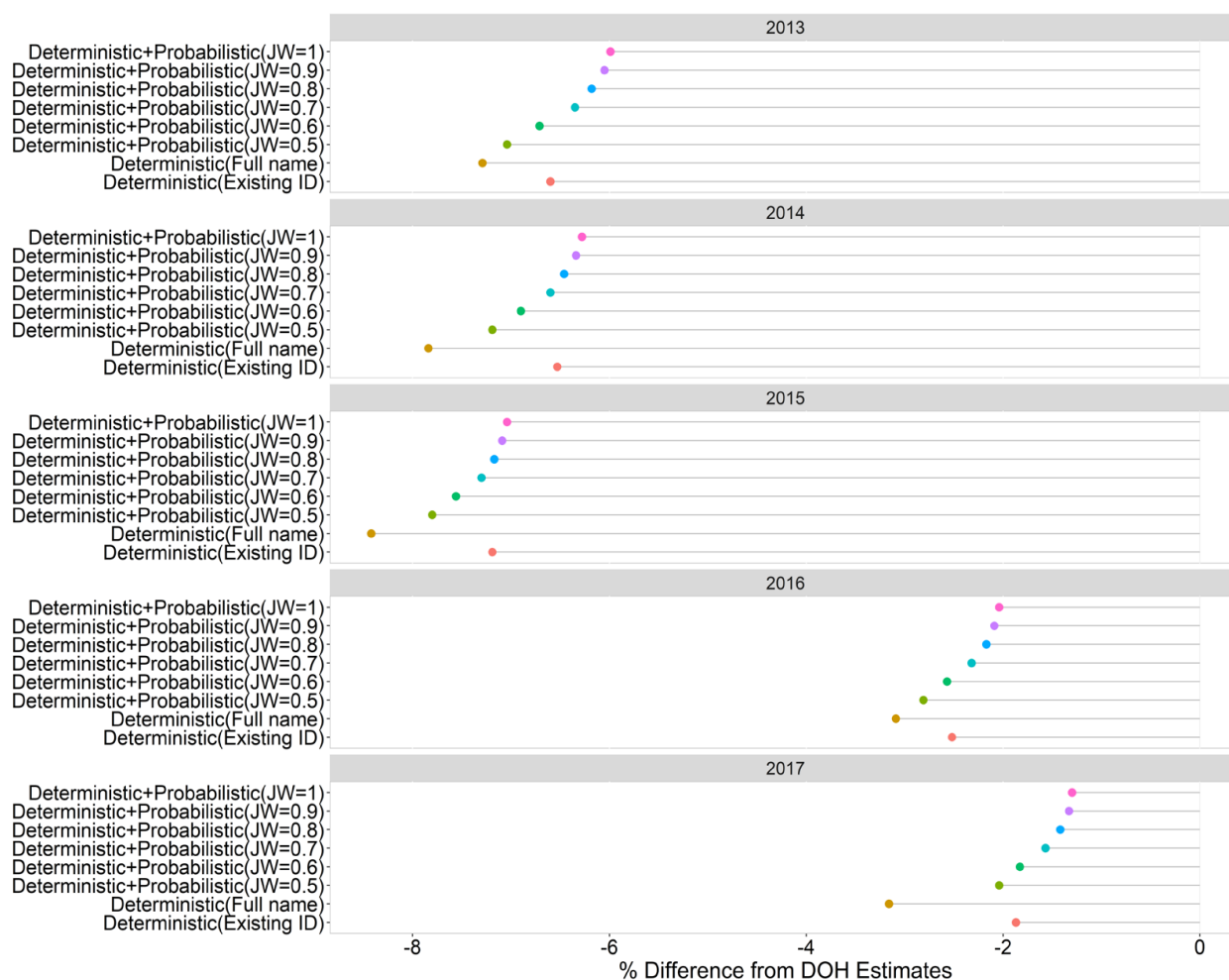
Supplement Figure 1: Percentage difference of the number of individuals estimated to be living with HIV and in care in the public sector from the linkage algorithms and those reported by the Western Cape DOH [14] and linkage algorithms. Estimates obtained from a combination of Deterministic and Probabilistic linkage were closest to those given in the DOH reports for each of the years available. Deterministic linkage only yields estimates that are furthest from the DOH estimates.

11

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Global Health
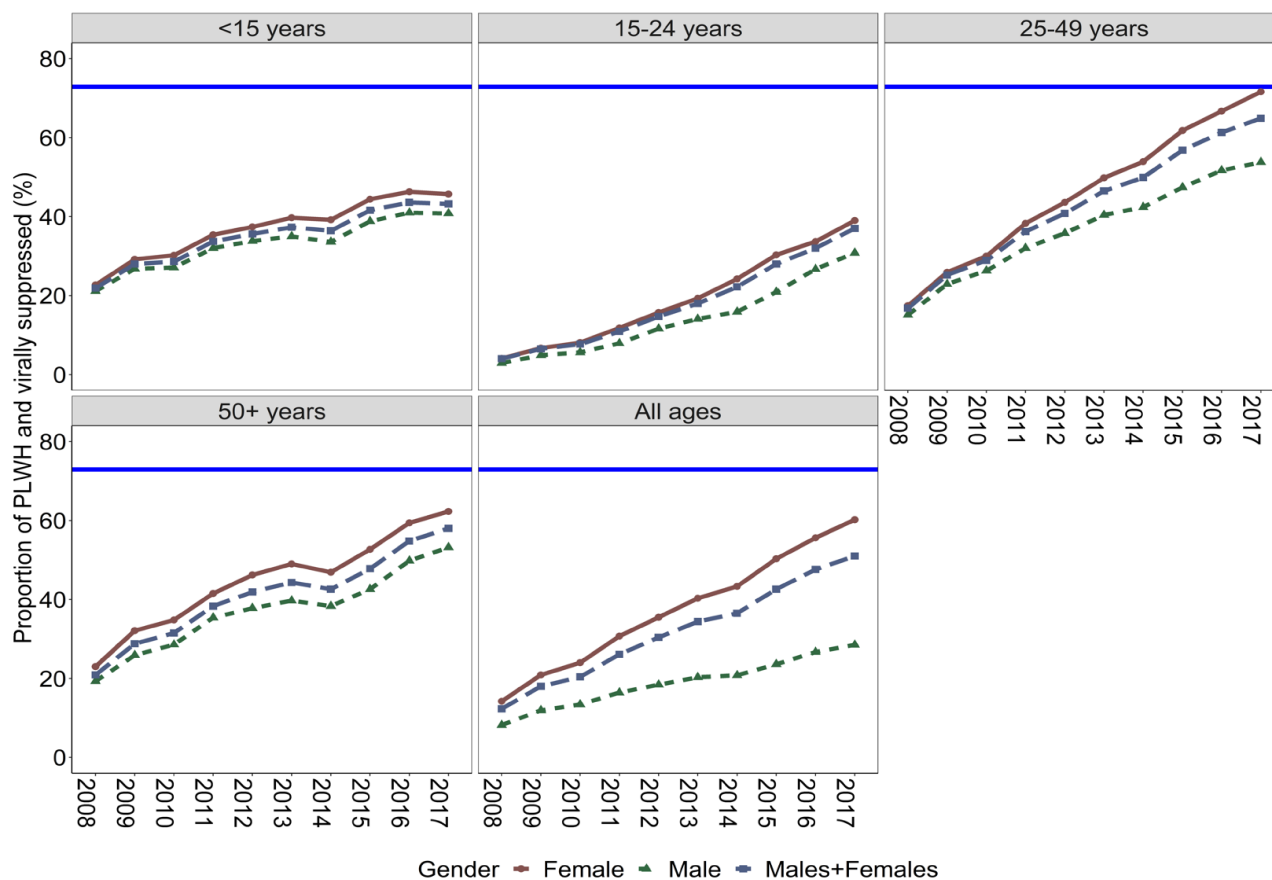
Supplement Figure 2 : Estimated population viral suppression in the Western Cape, South Africa using routinely collected viral load data (2008 - 2017). This figure gives the estimated proportion of individuals living with HIV (PLWH) who were virally suppressed in the Western Cape, South Africa. Estimates were obtained from a combination of routinely collected viral load data and model-based estimates. Population viral suppression estimates were disaggregated by age (< 15 years, 15-24 years, 25-49 years, >=50 years, All ages),and sex (Male, Female, Males + Females). The blue horizontal line represents the overall 90-90-90 target (72.9%).

Supplement Figure 3: Method based estimation of population viral suppression among all females living with HIV (PLWH) in the Western Cape (dashed lines), compared to mathematical model based (purple solid line) and cross-sectional survey estimates (points) from 2008 - 2017.
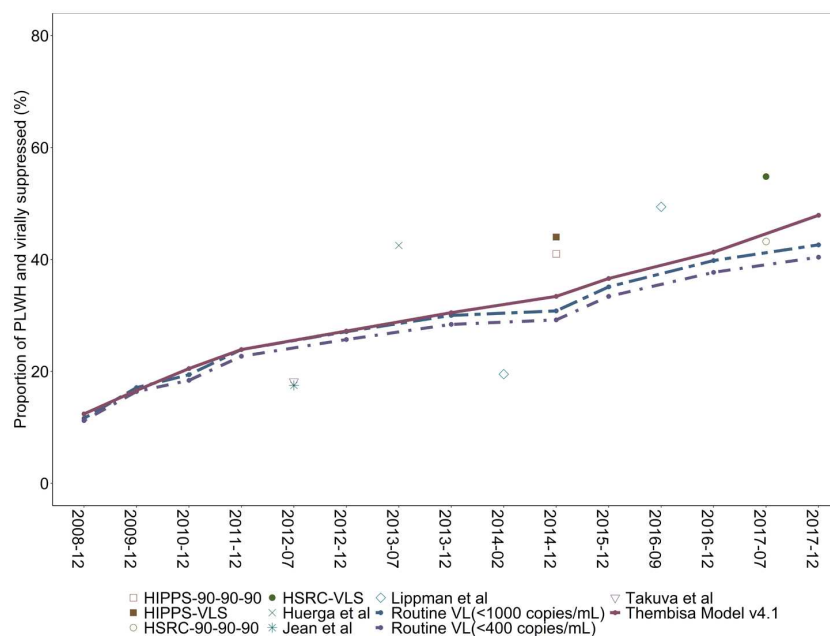
Supplement Figure 4: Method based estimation of population viral suppression among all males living with HIV (PLWH) in the Western Cape (dashed lines), compared to mathematical model based (purple solid line) and cross-sectional survey estimates (points) from 2008 - 2017.

13

References for online supplement

1.   Winkler WE. Overview of Record Linkage and Current Research Directions. Washington DC: Statistical Research Division, U.S. Census Bureau;2006.

2.   Statistics South Africa. Mid-year population estimates, 2019. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2019.

3.   Statistics South Africa. Mid-year population estimates, 2018. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2018.

4.   Statistics South Africa. Mid-year population estimates, 2017. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2017.

5.   Statistics South Africa. Mid-year population estimates, 2016. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2016.

6.   Statistics South Africa. Mid-year population estimates, 2015. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2015.

7.   Statistics South Africa. Mid-year population estimates, 2014. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2014.

8.   Statistics South Africa. Mid-year population estimates, 2013. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2013.

9.   Statistics South Africa. Mid-year population estimates, 2011. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2011.

10.   Statistics South Africa. Mid-year population estimates, 2010. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2010.

11.   Statistics South Africa. Mid-year population estimates, 2009. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2009.

12.   Statistics South Africa. Mid-year population estimates, 2008. In: Statistics South Africa, ed. Pretoria: Statistics South Africa; 2008.

13.   Johnson LF, Chiu C, Myer L, et al. Prospects for HIV control in South Africa: a model-based analysis. *Glob Health Action.* 2016;9:30314.

14.   Western Cape Government Health. Western Cape Government Health, Annual Report 2017/2018. Cape Town.2018