

Data Codebook 4 Codebook_HIVafter40_Twin_Epidemics_HIV_Hyp

Brian Houle and Jane Menken

8 July 2022

This codebook is for the Stata replication dataset

Data Dataset 6 HIVafter40_Twin_Epidemics_HIV_Hyp

analyzed in the paper: *Twin epidemics: The effects of HIV and systolic blood pressure on mortality risk in rural South Africa, 2010-2019.*

The replication dataset was created from the non-public

AHDSS_HaNakekela_HIV_Hyp dataset (Data Codebook 2)

using

Data Dofile 3 DatasetBuilding_Public.

The non-public dataset contains variables extracted from the full HaNakekela survey dataset linked to the AHDSS Census 2009 and mortality-migration followup from AHDSS Censuses through 2019. Final status of individuals (resident, migrated, dead) is adjusted to July 31, 2019.

Variable			Label
Prepared 26 Apr 2022 using STATA codebook command			
Number of observations: 32,491			
Outcome			Survey
outcome			
Freq.	Value	Label	
23,035	1	participant	
7,553	2	not found	
1,903	4	refused	
test_outcomeN			0-4 Neg/Pos/not found/refusal test/
Freq.	Value	Label	
6,270	0	negative	
5,278	1	positive	
1,487	2	refused test	
1,903	3	refused survey	
7,553	4	not found	
dead			Death recorded in AHDSS
Freq.	Value	Label	
29,122	0	person alive at last observation	
3,369	1	person's death recorded in AHDSS	
male			Sex
Freq.	Value	Label	
17,666	0	female	
14,825	1	male	
edu			Edu Low/Primary/Secondary+
Freq.	Value	Label	
13,166	1	None/very low (<= 3 years)	
8,386	2	Primary(4-8 years)	
10,939	3	Secondary school or higher (>8 years)	
south_african			South African nativity
Freq.	Value	Label	
9,448	0	other	
23,043	1	south african	

working						Currently working
	Freq.	Value	Label			
	19,121	0	not working			
	11,592	1	working			
	1,778	.	missing			
mar3						Single/Married_Cohabiting/Divorced_Widowed
	Freq.	Value	Label			
	4,560	1	Single			
	17,985	2	Married/cohabiting			
	9,934	3	Widowed/divorced			
	12	.	missing			
qses3						Household wealth quintile 1-2/3/4-5
	Freq.	Value	Label			
	10,690	1	quint 1/2			
	6,587	2	quint 3			
	15,098	3	quint 4/5			
	116	.	missing			
systolic2						Mean systolic mmHg
	range:	[79,255]	missing .: 9,775/32,491			
	mean:	134.279				
	std. dev:	22.1416				
	percentiles:	10%	25%	50%	75%	90%
		109	118.5	132	146.5	163.5
histbpmed						History of blood pressure medication
	Freq.	Value	Label			
	18,693	0	no			
	4,342	1	yes			
	9,456	.	missing			
diastolic2						Diastolic mmHg - mean of measures 2 and 3
	range:	[44,170.5]	missing .: 9,732/32,491			
	mean:	88.6243				
	std. dev:	15.5617				
	percentiles:	10%	25%	50%	75%	90%
		70	78.5	87.5	98	108.5
hyp_simple						Normotensive/Hypertensive
	Freq.	Value	Label			
	12,091	0	Normotensive			
	5,659	1	Hypertensive: reported or treated			
	14,741	.	missing			
sbp_cat2						Systolic cutoffs 120/140/160
	Freq.	Value	Label			
	6,164	1	<120			
	8,506	2	120-139			
	5,066	3	140-159			
	2,755	4	160+			
	10,000	.	missing			
hiv_positive						HIV status
	Freq.	Value	Label			
	16,270	0	Negative			
	5,278	1	Positive			
	10,943	.	missing			

vl_simple_400			Viral load: Neg/Suppressed<=400/Unsuppressed>400
Freq.	Value	Label	
16,270	0	HIV Negative	
2,200	1	Suppressed<=400	
2,960	2	Unsuppressed>400	
11,061	.	missing	

vl_simple_2013			2013 Viral load: Neg/Suppressed<=1000/Unsuppressed>1000
Freq.	Value	Label	
16,270	0	HIV Negative	
2,642	1	Suppressed<=1000	
2,518	2	Unsuppressed>1000	
11,061	.	missing	

_t0							Person's age at 20100801 (start of first observation year)
range:			[31.238877,103.13758]		missing .: 0/32,491		
mean:			53.4855				
std. dev:			14.0666				
percentiles:			10%	25%	50%	75%	90%
			37.692	41.8891	49.6865	63.4634	73.9165

m			Person's Observation year
Freq.	Value	Label	
3,104	1	observation year 1	
3,225	2	observation year 2	
3,319	3	observation year 3	
3,362	4	observation year 4	
3,425	5	observation year 5	
3,479	6	observation year 6	
3,485	7	observation year 7	
3,507	8	observation year 8	
3,491	9	observation year 9	
2,094	10	observation year 10	

M			Person's total observation years
Freq.	Value	Label	
20	1	1 year of observation	
174	2	2 years of observation	
308	3	3 years of observation	
410	4	4 years of observation	
535	5	5 years of observation	
593	6	6 years of observation	
735	7	7 years of observation	
1,103	8	8 years of observation	
11,243	9	9 years of observation	
17,370	10	10 years of observation	

EndYear			Status at end of person-year
Freq.	Value	Label	
31,859	0	Survived	
632	1	Died	

AgeYear							Age at start of person-year
range:			[40,110]		missing .: 0/32,491		
mean:			57.3664				
std. dev:			13.8092				
percentiles:			10%	25%	50%	75%	90%
			42	46	54	67	78

agecat			Age category in person-year
Freq.	Value	Label	
19,864	2	40-59	

8,149	3	60-74
4,478	4	75+

wt_new					Person-year weight
range:	[.00152364,110.59737]	missing .:	1,335/32,491		
mean:	5.05991				refer to definition
std. dev:	5.79362				
percentiles:	10%	25%	50%	75%	90%
	1.73071	2.77001	4.08326	5.38247	7.3641

The finalwt was calculated for each person and depends on the probability the person was selected into the sample and the predicted probability the person participated. See Data Codebook 2 and Data 1 for further description of finalwt.

For wt_new, finalwt is multiplied by the proportion of this person-year the person was observed. Details of the calculation are in Data Dofile 3.
