

Data 1

README: File and Project Descriptions HIVafter40 and Ha Nakekela (We Care)



INDEX

File Descriptions	Page
Data Codebook 2	1
Codebook_AHDSS_HaNakekea_HIV_Hyp	
Data Dofile 3	2
DatasetBuilding_Public	
Data Codebook 4	3
Codebook_HIVafter40_Twin_Epidemics_HIV_Hyp	
Data Dofile 5	3
Analysis_HIVafter40_Twin_Epidemics_Public	
Data Dataset 6	3
HIVafter40_Twin_Epidemics_HIV_Hyp	
Project Descriptions	3

Data Codebook 2

Codebook_AHDSS_HaNakekea_HIV_Hyp.docx

The AHDSS_Ha_Nakekela_HIV_Hyp.dta is a non-public STATA dataset that includes information from the

Agincourt Health and socioDemographic Surveillance System (AHDSS)

- on the age/sex stratified probability sample of 7657 individuals from the 2009 AHDSS Census selected for the Ha Nakekela Survey
- demographic and socioeconomic characteristics of individuals in the sample taken from the 2009 and earlier AHDSS Censuses
- dates of departure (death, migration, censoring) of individuals through July 31, 2019 taken from 2009-2020 AHDSS Censuses

2010-2011 Ha Nakekela (We Care) HIV and NCD Survey

- survey outcome (participant, not found, dead before survey date, refused, ineligible)
- self-reported and measured HIV status and blood pressure
- the disease risk (STEPS Health Risk Questionnaire) and sexual behavior (BERIS Sexual Behaviors Questionnaire) data collection instruments are available at <https://hivafter40.princeton.edu/study-instruments>
- variables in the codebook beginning with Q refer to the same-named questions in the data collection instruments.

For each of the 7,662 people in the Ha Nakekela original sample, we used the full AHDSS 2009 census to calculate the survey weight as the inverse of the probability the person was selected for the sample. To adjust for survey nonresponse, we used a logistic regression model of factors from the 2009 census to predict the probability of participation and its inverse, the inverse

probability sampling weight (IPSW). Final weights (finalwt) for participants are the product of the IPSW and the survey weight.

Data Dofile3

DatasetBuilding_Public.do

This do-file creates a person-year dataset for the HIVafter40 study of the impact of the twin epidemics of HIV and hypertension on mortality. We have run it on both STATA 15 and STATA 16. It

- drops sample individuals with outcomes not found, dead before survey date, ineligible
- recodes/creates variables
- prepares and retains data on person-years in which a participant in the Ha Nakekela Survey was age 40+(including those who refused HIV test)
- creates the weight for the person-year (wt_new) as the product of the weight for the person (finalwt) and the proportion of the person-year that the person was observed
- and creates the publicly available STATA person-year data set:
HIVafter40_Twin_Epidemics_HIV_NCD.dta

The dataset is the basis for the paper:

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

BMC Public Health (2022) 22:387

<https://doi.org/10.1186/s12889-022-12791-z>

Brian Houle^{1,2*}
Chodziwadziwa W Kabudula²
Andrea M Tilstra^{3,4,5}
Sanyu A Mojola^{2,6}
Enid Schatz^{2,7}
Samuel J Clark^{2,8}
Nicole Angotti^{2,9}
F Xavier Gómez-Olivé²
Jane Menken⁴

1 School of Demography, The Australian National University, Canberra, Australia.

2 MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), Faculty of Health Sciences, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa.

3 Nuffield College, University of Oxford, Oxford, UK.

4 Institute of Behavioral Science and Department of Sociology, University of Colorado Boulder, Boulder, USA.

5 Leverhulme Centre for Demographic Science, Department of Sociology, University of Oxford, Oxford, UK.

6 Department of Sociology, School of Public and International Affairs, and Office of Population Research, Princeton University, Princeton, USA.

7 Department of Public Health, University of Missouri, Columbia, USA.

8 Department of Sociology, The Ohio State University, Columbus, USA

9 Department of Sociology, American University, Washington, USA.

Data Codebook 4

Codebook_HIVafter40_Twin_Epidemics_HIV_Hyp

Data Codebook 4 is the codebook for Data Dataset 6. Both were created with Data Dofile 3.

Data Dofile 5

Analysis_HIVafter40_Twin_Epidemics_Public

This dofile was used to prepare all analyses for the Twin Epidemics paper. It employs Data Dataset 6.

Data Dataset 6

HIVafter40_Twin_Epidemics_HIV_Hyp

Project Descriptions - Ha Nakekela and HIVafter40

The HIVafter40_Twin_Epidemics_HIV_NCD Dataset was created as part of the HIVafter40 project, the successor to the Ha Nakekela Project, both nested within the Agincourt Health and socio-Demographic System (AHDSS) located in northeastern South Africa.

The projects use mixed methods to measure prevalence and incidence of HIV and NCDs, investigate health and sexual behaviors, and study how people have coped with the abrupt introduction of a new and deadly health threat and efforts to prevent and ameliorate it.

The website <https://hivafter40.princeton.edu/> is regularly updated and contains the most complete information on the projects.

Team

The Ha Nakekela and HIVafter40 projects are carried out by a team that is described in full at <https://hivafter40.princeton.edu/research-team>

Here we draw attention to the Investigative Team (as of June 2022):

Ha Nakekela and HIVafter40

Nicole Angotti	American University
Samuel Clark	The Ohio State University
F. Xavier Gómez-Olivé	University of the Witwatersrand
Brian Houle	The Australian National University
Chodziwadziwa Kabudula Whiteson	University of the Witwatersrand
Jane Menken	University of Colorado Boulder
Sanyu Mojola	Princeton University
Enid Schatz	University of Missouri
Andrea Tilstra	Oxford University

Ha Nakekela

Jill Williams	University of Colorado Boulder
---------------	--------------------------------

HIVafter40

Vusi Dlamini	University of Maryland, College Park
Erin Ice	University of Michigan

Data

The AHDSS has, since 1992, collected an annual census from on households in a defined geographic area of Mpumalanga Province. The 2009 census served as the sampling frame for the 2011 Ha Nakekela Project. The AHDSS provides rapid measures of vital statistics and is the site of a wide range of more detailed studies. As of June 2018, the study site consisted of 31 research villages with a population of 116,549 people, living in 22,721 households. Further information is available at <https://www.agincourt.co.za/> and, specifically for census data collection, at https://www.agincourt.co.za/?page_id=1805

The Ha Nakekela project collected two datasets:

- The Ha Nakekela Survey is based on an age/sex stratified sample of respondents in the 2009 AHDSS Census. Carried out in 2010-11, the cross-sectional survey included 5,080 respondents, of whom 2,080 are aged 40 – 93. It has 3 parts:
 - Behavioral Risk Assessment (based on **BERIS Sexual Behavior Questionnaire**)
 - Chronic Disease Risk Factor Surveillance (based on **STEPS Health Risk Questionnaire**)
 - HIV and Non-Communicable Disease (NCD) biometric data
- The Izindaba za Badala (Matters that Concern Older People) Study (IZB) is nested within the Ha Nakekela Survey. It consisted of three related interview studies carried out in 2013:
 - IZB Life History Interviews: 60 life history interviews with people with and without HIV aged 40-84 who participated in the Ha Nakekela Survey
 - IZB Community Focus Group Interviews: 9 community focus group interviews with a total of 77 respondents
 - IZB Key informant Interviews: 9 health workers interviews in 3 local health clinics

The HIVafter40 project added two follow-up datasets:

- AHDSS_Ha_Nakekela_HIV_Hyp_Sample.dta contains information on the Ha Nakekela sample in which individuals were linked to their location/outcome reports in successive AHDSS censuses through 2020. This dataset permits examination of mortality in the nine years subsequent to participation in the Ha Nakekela Survey. HIVafter40_Twin_Epidemics_HIV_NCD.dta is a subset of this dataset.
- Izindaba za Badala Round 2
 - IZB Life History Interviews: 25 in-depth follow-up interviews in 2018
 - IZB Community Focus Group Interviews: 10 community focus group interviews with a total of 84 participants in 2018
 - IZB Key informant Interviews (cancelled due to the COVID-19 epidemic): 10 community focus group interviews with a total of 84 participants planned for 2020.

All data collection instruments are available at <https://hivafter40.princeton.edu/study-instruments>

Funding

We gratefully acknowledge funding from the following sources (as of June 2022):

National Institute on Aging - R01, AG049634 - PI: Sanyu A. Mojola - HIV after 40 in Rural South Africa: Aging in the Context of an HIV Epidemic

National Institute on Aging - R24, AG032112-05 - PI: Jane Menken - Partnership for Social Science AIDS Research in South Africa's Era of ART Rollout

Wellcome Trust - 058893/Z/99/A, 069683/Z/02/Z, and 085477/Z/08/Z (PI: Stephen Tollman) - AHDSS

University of Colorado Boulder Innovative Seed Grant - PI: Sanyu Mojola – HIV after 40 in Rural South Africa

William and Flora Hewlett Foundation (2008-1840) – PI: Jane Menken - African Population Studies Research and Training Program

SPARC Fellowship, University of the Witwatersrand (Nicole Angotti)

A fully up-to-date listing of funding is available at
<https://hivafter40.princeton.edu/project-funding> for

Publications

The Data Note references publications through July 2022. The fully up-to-date list of publications from the HIVafter40 and the Ha Nakekela projects is available at
<https://hivafter40.princeton.edu/publications-0>

Ethical approvals

HIVafter40/Ha Nakekela received IRB approvals from the University of Colorado Boulder, the University of Michigan, Princeton University, the University of the Witwatersrand Human Research Ethics Committee, and the Department of Health, Mpumalanga Provincial Government, South Africa.