

South Africa

Dr Abraham Jacobus Herbst, Africa Health Research Institute

Prof. Willem Hanekom, Africa Health Research Institute

Prof. Janet Seeley, Africa Health Research Institute

Prof. Maryam Shahmanesh, Africa Health Research Institute

Dr. Guy Harling, Africa Health Research Institute

Dr. Mark Siedner, Africa Health Research Institute

AHRI HDSS:Individual Health Surveillance HIV 2005

Study Documentation

August 24, 2023

Metadata Production

Metadata Producer(s)	Africa Health Research Institute (AHRI)
Identification	DDI.AHRI.RD05-02.HDSS.HIV.2005

Table of Contents

Overview.....	4
Scope & Coverage.....	5
Producers & Sponsors.....	5
Sampling.....	6
Data Collection.....	6
Data Processing & Appraisal.....	6
Accessibility.....	6
Files Description.....	7
RD05-02 ACDIS HIV 2005.....	7
Variables List.....	8
RD05-02 ACDIS HIV 2005.....	8
Variables Description.....	9
RD05-02 ACDIS HIV 2005.....	10

AHRI HDSS:Individual Health Surveillance HIV 2005

Overview	
Identification	AHRI.RD05-02.HDSS.HIV.2005
Version	v1.0.0
<p>Abstract</p> <p>The Africa Centre Demographic Information System (ACDIS) was established in 2000 by the Africa Health Research Institute (AHRI; formerly the Africa Centre for Health and Population Studies) covering a population of approximately 85,0000 residents and non-residents from an area 438km² in size in uMkhanyakude district, KwaZulu-Natal province, South Africa.</p> <p>In 2005, the “Eligibility List” for Round 2 of the HIV Surveillance used 10 January 2005 as its reference date. The criteria for inclusion were different for Residents and non-Residents. For Residents all males aged 15-54 and all females aged 15-49 on the same date were selected. For Non-Residents, a stratified sampling procedure was adopted with the overall objective of obtaining a representative sample of 12.5% of all non-residents. The 10 strata were defined by sex and by ?? and five 'return patterns', i.e., self-stated frequencies of returning home to the household in the DSA ('annual', 'occasional', 'month end', 'frequent', and 'other'). Other criteria were also used, and full details are available separately.</p> <p>The total number eligible was 31,271, and this is the number of rows in the dataset.</p> <p>The field visits took place from January 2005 to February 2006. At the time of visit some people included in the Eligibility List were found to in fact be ineligible - perhaps because they were dead or had out migrated - events either not known about at the time the Eligibility List was compiled or occurring between the drawing up of the Eligibility List and the actual visit. These 'retrospective ineligibilities' are identifiable in the dataset as 'Premature Completions with reasons such as “Death or “Outmigration”. There were also some individuals who could not be contacted, even after repeated visits and 'tracking' attempts for those who were reported to have moved. These are identifiable as <code>PrematureCompletionReason = N</code> (i.e., Non-Contact) Additionally a small number were contacted but refused to participate in the survey. They are identifiable by <code>VisitType = Refused</code>. Finally there are some who participated in the survey, but refused the HIV test offered. They are identifiable as <code>HIVRefused = 'Y'</code></p> <p>This dataset includes only those in the Eligibility List. It does not include the small number (x) of 'Voluntary walk-in Testers' i.e. individuals who were not part of the Eligibility List, but who visited Counselling Centres and requested a test.</p> <p>Individuals were free to refuse to participate in and to withdraw from the individual surveys, without any impact on routine health care or other services to which they were entitled. If a household refused to participate in the household component, its members were not invited to participate in the individual component during that same household visit.</p> <p>Note: Users of these datasets are strongly encouraged to refer to the questionnaires included in the documentation for the dataset to check carefully for any changes in question wording, ordering, options offered etc. This document does not attempt to describe every single change and variation.</p> <p>For more refer:</p> <p>1. Tanser F, Hosegood V, Barnighausen T, Herbst K, Nyirenda M, Muhwava W, Newell C, Viljoen J, Mutevedzi T, Newell ML. Cohort Profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. <i>Int J Epidemiol.</i> 2008 Oct;37(5):956-62. doi: 10.1093/ije/dym211. Epub 2007 Nov 12. PMID: 17998242; PMCID: PMC2557060.</p>	
Kind of Data	Surveillance data
Unit of Analysis	AHRI HDSS Individual surveillance eligible individuals

Scope & Coverage	
Keywords	South Africa, Surveillance, Epidemiology, HIV, Population Cohort
Topics	HIV, Population, Surveillance, HIV Incidence, HIV prevalence, Epidemiology, Africa, Demography, information systems, South Africa, KwaZulu-Natal, Medical, Population Intervention
Time Period(s)	2005-2006
Countries	South Africa
Geographic Coverage AHRI's Demographic Surveillance Area (DSA) is situated in Mtubatuba local authority, uMkhanyakude district, KwaZulu-Natal province which is approximately 200km north of Durban. For more refer: 1. Dickman Gareta, Kathy Baisley, Thobeka Mngomezulu, Theresa Smit, Thandeka Khoza, Siyabonga Nxumalo, Jaco Dreyer, Sweetness Dube, Nomathamsanqa Majozi, Gregory Ordington-Jespersion, Eugene Ehlers, Guy Harling, Maryam Shahmanesh, Mark Siedner, Willem Hanekom, Kobus Herbst, Cohort Profile Update: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey, International Journal of Epidemiology, Volume 50, Issue 1, February 2021, Pages 33-34, < https://doi.org/10.1093/ije/dyaa264 > 2. Tanser F, Hosegood V, Barnighausen T, Herbst K, Nyirenda M, Muhwava W, Newell C, Viljoen J, Mutevedzi T, Newell ML. Cohort Profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. Int J Epidemiol. 2008 Oct;37(5):956-62. doi: 10.1093/ije/dym211. Epub 2007 Nov 12. PMID: 17998242; PMCID: PMC2557060. 3. Iwuji CC, Orne-Gliemann J, Larmarange J, Balestre E, Thiebaut R, Tanser F, Okesola N, Makowa T, Dreyer J, Herbst K, McGrath N, Barnighausen T, Boyer S, De Oliveira T, Rekacewicz C, Bazin B, Newell ML, Pillay D, Dabis F; ANRS 12249 TasP Study Group. Universal test and treat and the HIV epidemic in rural South Africa: a phase 4, open-label, community cluster randomised trial. Lancet HIV. 2018 Mar;5(3):e116-e125. doi: 10.1016/S2352-3018(17)30205-9. Epub 2017 Nov 30. PMID: 29199100.	
Universe All resident household members from AHRI's DSA who met eligibility criteria and gave informed consent for their surveillance participation.	

Producers & Sponsors	
Primary Investigator(s)	Dr Abraham Jacobus Herbst, Africa Health Research Institute Prof. Willem Hanekom, Africa Health Research Institute Prof. Janet Seeley, Africa Health Research Institute Prof. Maryam Shahmanesh, Africa Health Research Institute Dr. Guy Harling, Africa Health Research Institute Dr. Mark Siedner, Africa Health Research Institute
Other Producer(s)	Africa Health Research Institute (AHRI)
Funding Agency/ies	Welcome Trust (WT) , Funder DSI-MRC South African Population Research Infrastructure Network (SAPRIN) , Funder
Other Acknowledgment(s)	Thobeka Mngomezulu , Data collection , AHRI Phumzile Dlamini , Data collection , AHRI Lindiwe Sithole , Data collection , AHRI Bonginkosi Ntimane , Data quality , AHRI Njabulo Myeni , Data quality , AHRI Eugene Ehlers , Software Development , AHRI Eugene Prenzler , Data Analytics , AHRI Brendan Gilbert , IT infrastructure , AHRI

Sweetness Dube , Data documentation , AHRI
 Siyabonga Nxumalo , Research Data management , AHRI
 Dickman Gareta , Research Data management , AHRI
 Kathy Baisley , Statistics , AHRI

Sampling

Sampling Procedure

All individuals meeting the eligibility criteria described in the abstract

Data Collection

Data Collection Dates	start 2005-01-18 end 2006-02-08
------------------------------	------------------------------------

Data Processing & Appraisal

Data Editing

Surveillance data is processed and stored on servers under the physical control of AHRI until datasets are made available on the data repository. The data is de-identified and can then be downloaded for processing on the data user's computer.

Data is stored on industry-standard relational databases with data integrity and user authentication for access control. Data is replicated on at least a daily basis to the Durban site of the Institution to provide secure offsite storage of data. Transactional logs are backed up every 30 minutes to enable recovery of data in the event of equipment failure.

All users of the system are authenticated through individual passwords with minimum complexity and regular change rules (passwords must be at least eight digits, with a mix of small and capital letters, at least one numeric or non-alphabetic digit and changed at least every 45 days). AHRI uses industry standard malware and intrusion detection with at least annual penetration tests by a reputable outside security audit company.

Both at the Institute and for the clinic-based data collection, a client-server architecture is implemented where data is not stored on laptops or local workstation, but only on a central server with restricted physical access. Specifically, at the clinics the local server is enclosed in a tamper-proof enclosure kept under lock and key. The server hard disk is encrypted.

Accessibility

Access Conditions

Access to the data requires accurate completion of the online data access application form accessible on the AHRI Data repository(<<https://data.ahri.org/>>). Data users are required to abide by the data use conditions stipulated on the application for access to the data. Failure to do so may result in their data access privileges being revoked by the Data Custodian. In order to recognise the effort and intellectual contributions of AHRI investigators in producing and curating the data, users of AHRI data must acknowledge the source of the data and abide by the terms and conditions under which the data is accessed and must cite the dataset in publication using the citation provided as part of this documentation. All analytical datasets published on the AHRI Data Repository are assigned digital object identifier (DOIs) and the DOIs can be found on the Data Repository under Study Description tab - Access policy. AHRI data users are required to always cite the dataset using the relevant DOI.

Citation Requirements

Herbst, K., Hanekom, W., Seeley, J., Shahmanesh, M., Harling, G., & Siedner, M. (2023). AHRI HDSS:Individual Health Surveillance HIV 2005 [Data set]. Africa Health Research Institute. DOI:<https://doi.org/10.23664/AHRI.RD05-02.HDSS.HIV.2005>

Files Description

Dataset contains 1 file(s)

RD05-02 ACDIS HIV 2005	
# Cases	31271
# Variable(s)	19

Variables List

Dataset contains 19 variable(s)

File RD05-02 ACDIS HIV 2005							
#	Name	Label	Type	Format	Valid	Invalid	Question
1	Question ..	Questionnaire Id	continuous	numeric-12.0	31271	0	-
2	BMFStatu ..	BMFStatusObservation Id	continuous	numeric-12.0	31224	47	-
3	Prematur ..	Premature Completion Reason	discrete	numeric-12.0	31224	47	-
4	IIntId	Unique Internal Id of Individual	continuous	numeric-12.0	31271	0	-
5	Sex	Sex	discrete	numeric-12.0	31271	0	-
6	DateOfBi ..	Date of birth	discrete	character-11	31271	-	-
7	Resident	Residency status	discrete	numeric-12.0	31271	0	-
8	BSIntId	Unique Internal Id of BoundedStructures	continuous	numeric-12.0	31271	0	-
9	VisitDate	Visit Date	discrete	character-11	31224	-	-
10	VisitType	Visit Type	discrete	numeric-12.0	31224	47	-
11	Attempt1 ..	First Attempt Date	discrete	character-11	12221	-	-
12	Attempt2 ..	Second Attempt Date	discrete	character-11	6395	-	-
13	Attempt3 ..	Third Attempt Date	discrete	character-11	3607	-	-
14	Attempt4 ..	Fourth Attempt Date	discrete	character-11	1346	-	-
15	LatestAt ..	Latest Attempt Date	discrete	character-11	12221	-	-
16	HIVRefused	HIV Refused	discrete	numeric-12.0	31224	47	-
17	HIVRefus ..	HIV Refused By	discrete	numeric-12.0	31224	47	-
18	HIVSampl ..	HIV Sample Available	discrete	numeric-12.0	31224	47	-
19	HIVResult	HIV Result	discrete	numeric-12.0	31107	164	-

Variables Description

Dataset contains 19 variable(s)

File : RD05-02 ACDIS HIV 2005

QuestionnaireId: Questionnaire Id

Information	[Type= continuous] [Format=numeric] [Range= 468704-46263639] [Missing=*]
Statistics [NW/ W]	[Valid=31271 /-] [Invalid=0 /-] [Mean=529100.387 /-] [StdDev=818043.979 /-]

BMFStatusObservation: BMFStatusObservation Id

Information	[Type= continuous] [Format=numeric] [Range= 26654-106790] [Missing=*]
Statistics [NW/ W]	[Valid=31224 /-] [Invalid=47 /-] [Mean=43471.014 /-] [StdDev=9473.286 /-]

PrematureCompletionReason: Premature Completion Reason

Information	[Type= discrete] [Format=numeric] [Range= 1-99] [Missing=*]
Statistics [NW/ W]	[Valid=31224 /-] [Invalid=47 /-]

Value	Label	Cases	Percentage
1	Broken Down BS	55	0.2%
2	Mentally Disabled	187	0.6%
3	Outmigration	4259	13.6%
4	Non-contact	581	1.9%
5	Other Reasons	2243	7.2%
6	Very Sick	74	0.2%
7	Not Tracked (too far)	37	0.1%
93	Refused	421	1.3%
97	Not Applicable	23367	74.8%
99	Missing	0	
Sysmiss		47	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

IntId: Unique Internal Id of Individual

Information	[Type= continuous] [Format=numeric] [Range= 17-264757] [Missing=*]
Statistics [NW/ W]	[Valid=31271 /-] [Invalid=0 /-] [Mean=55505.13 /-] [StdDev=34378.234 /-]

Sex: Sex

Information	[Type= discrete] [Format=numeric] [Range= 1-99] [Missing=*]
Statistics [NW/ W]	[Valid=31271 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Male	14218	45.5%
2	Female	17053	54.5%
99	Unknown	0	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

DateOfBirth: Date of birth

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=31271 /-]

Resident: Residency status

Information	[Type= discrete] [Format=numeric] [Range= 1-99] [Missing=*]
Statistics [NW/ W]	[Valid=31271 /-] [Invalid=0 /-]

Value	Label	Cases	Percentage
1	Yes	28385	90.8%

File : RD05-02 ACDIS HIV 2005

Resident: Residency status

Value	Label	Cases	Percentage
2	No	2885	9.2%
90	SpecialQuery	0	
92	Deprecated	0	
93	Refused	0	
94	Query	0	
95	Default	0	
96	Other	0	
97	Not Applicable	0	
98	Don't know	0	
99	Missing	1	0.0%

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

BSIntId: Unique Internal Id of BoundedStructures

Information	[Type= continuous] [Format=numeric] [Range= 12-13061] [Missing=*]
Statistics [NW/ W]	[Valid=31271 /-] [Invalid=0 /-] [Mean=6173.627 /-] [StdDev=3660.941 /-]

VisitDate: Visit Date

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=31224 /-]

VisitType: Visit Type

Information	[Type= discrete] [Format=numeric] [Range= 1-999] [Missing=*]
Statistics [NW/ W]	[Valid=31224 /-] [Invalid=47 /-]

Value	Label	Cases	Percentage
1	Avoided Visit	4	0.0%
2	Unsuccessful Visit	3	0.0%
3	Regular Visit	25673	82.2%
4	Special Task Visit	4	0.0%
5	Problem Solving	0	
6	Quality Control	0	
7	Spot Check	0	
8	Tracking	4719	15.1%
9	Verbal Autopsy	1	0.0%
10	Mapping Visit	1	0.0%
11	COUNSELLING VISIT	0	
12	Illness And Death	0	
13	Bounded Structure Verification	0	
30	Community Liaison Office Visit	0	
92	Deprecated	1	0.0%
93	Refused Visit	818	2.6%
94	Query	0	
95	Default	0	
96	Other	0	
97	Not Applicable	0	

File : RD05-02 ACDIS HIV 2005

VisitType: Visit Type

Value	Label	Cases	Percentage
99	Missing	0	
999	Unknown	0	
Sysmiss		47	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

Attempt1Date: First Attempt Date

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=12221 /-]

Attempt2Date: Second Attempt Date

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=6395 /-]

Attempt3Date: Third Attempt Date

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=3607 /-]

Attempt4Date: Fourth Attempt Date

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=1346 /-]

LatestAttemptDate: Latest Attempt Date

Information	[Type= discrete] [Format=character] [Missing=*]
Statistics [NW/ W]	[Valid=12221 /-]

HIVRefused: HIV Refused

Information	[Type= discrete] [Format=numeric] [Range= 1-99] [Missing=*]
Statistics [NW/ W]	[Valid=31224 /-] [Invalid=47 /-]

Value	Label	Cases	Percentage
1	Yes	13829	44.3%
2	No	9538	30.5%
90	SpecialQuery	0	
92	Deprecated	0	
93	Refused	0	
94	Query	0	
95	Default	0	
96	Other	0	
97	Not Applicable	7857	25.2%
98	Don't know	0	
99	Missing	0	
Sysmiss		47	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

HIVRefusedBy: HIV Refused By

Information	[Type= discrete] [Format=numeric] [Range= 1-99] [Missing=*]
Statistics [NW/ W]	[Valid=31224 /-] [Invalid=47 /-]

File : RD05-02 ACDIS HIV 2005

HIVRefusedBy: HIV Refused By

Value	Label	Cases	Percentage
1	Self	12883	41.3%
2	Partner	184	0.6%
3	HH head	299	1.0%
4	BS owner	244	0.8%
5	Other	217	0.7%
97	Not applicable	17397	55.7%
99	Missing	0	
Sysmiss		47	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

HIVSampleAvailable: HIV Sample Available

Information	[Type= discrete] [Format=numeric] [Range= 1-2] [Missing=*]
Statistics [NW/ W]	[Valid=31224 /-] [Invalid=47 /-]

Value	Label	Cases	Percentage
1	Yes	9537	30.5%
2	No,but there is a BMFStatusObs	21687	69.5%
Sysmiss		47	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.

HIVResult: HIV Result

Information	[Type= discrete] [Format=numeric] [Range= 0-99] [Missing=*]
Statistics [NW/ W]	[Valid=31107 /-] [Invalid=164 /-]

Value	Label	Cases	Percentage
0	Negative	7560	24.3%
1	Positive	1926	6.2%
2	Indeterminate	32	0.1%
3	Insufficient sample volume	0	
4	Spurious (probably not this person)	11	0.0%
5	Spurious negative(HIV reverse seroconversion)	0	
6	Spurious positive(HIV reverse seroconversion)	1	0.0%
97	Not applicable	21577	69.4%
99	Missing (Sample lost)	0	
Sysmiss		164	

Warning: these figures indicate the number of cases found in the data file. They cannot be interpreted as summary statistics of the population of interest.