THIS 2022-2023
Data Use Manual Supplement

THIS 2022 2023

TANZANIA HIV IMPACT SURVEY

THIS 2022-2023 Collaborating Institutions
National Public Health Laboratory (NPHL), Tanzania
National AIDS Control Programme (NACP), Tanzania
Zanzibar Integrated HIV, Hepatitis, Tuberculosis, and Leprosy Program (ZIHHTLP)
National Bureau of Statistics (NBS), Tanzania
Office of the Chief Government Statistician (OCGS), Tanzania
Tanzania Commission for AIDS (TACAIDS)
Zanzibar AIDS Commission (ZAC)
National Institute for Medical Research (NIMR), Tanzania
Zanzibar Health Research Institute (ZAHRI)
President's Office Regional Administration and Local Government (PO-RALG), Tanzania
Centers for Disease Control and Prevention (CDC), Tanzania
Centers for Disease Control and Prevention (CDC), Atlanta
ICAP at Columbia University, New York
ICAP at Columbia University, Tanzania
Westat
Bugando Medical Centre, Tanzania

Donor Support

Suggested Citation

Manual
Tanzania HIV Impact Survey 2022-2023 (THIS 2022-2023) Data Use Manual Supplement. New York, NY. [Month and year of release].

Single dataset
Tanzania Commission for AIDS (TACAIDS), Zanzibar AIDS Commission (ZAC). Tanzania HIV Impact Survey 2022-2023 (THIS 2022-2023) [Dataset]. <Insert dataset name used, e.g. this2022adultbio>. Dar es Salaam, Tanzania: National Public Health Laboratory, National AIDS Control Programme, Zanzibar Integrated HIV, Hepatitis, Tuberculosis, and Leprosy Program, National Bureau of Statistics, Office of the Chief Government Statistician, Tanzania Commission for AIDS, Zanzibar AIDS Commission, National Institute for Medical Research, Zanzibar Health Research Institute, President's Office Regional Administration and Local Government, Bugando Medical Centre, PEPFAR, CDC, Westat, ICAP at Columbia University [Producers]. ICAP at Columbia University [Distributor], [year of release].

Multiple datasets (more than 3)
Tanzania Commission for AIDS (TACAIDS), Zanzibar AIDS Commission (ZAC). Tanzania HIV Impact Survey 2022-2023 (THIS 2022-2023) (various) [Datasets]. Dar es Salaam, Tanzania: National Public Health Laboratory, National AIDS Control Programme, Zanzibar Integrated HIV, Hepatitis, Tuberculosis, and Leprosy Program, National Bureau of Statistics, Office of the Chief Government Statistician, Tanzania Commission for AIDS, Zanzibar AIDS Commission, National Institute for Medical Research, Zanzibar Health Research Institute, President's Office Regional Administration and Local Government, Bugando Medical Centre, PEPFAR, CDC, Westat, ICAP at Columbia University [Producers]. ICAP at Columbia University [Distributor], [year of release].

Access this Manual Online
<insert THIS 2022-2023 link, once available>

Contact Information
ICAP at Columbia University
722 West 168th Street
New York, NY 10032
Website: icap.columbia.edu
Email: icap-communications@columbia.edu

Table of Contents

Abbreviations

ART	Antiretroviral Therapy
Anti-HBc	Hepatitis B core antigen antibody
ARV	Antiretroviral
CAP/CTM	COBAS AmpliPrep/COBAS Taqman HIV-1 Qualitative Test
CD4	CD4+ T-Cell
CI	Confidence Interval
DHS	Demographic and Health Surveys
DNA	Deoxyribonucleic Acid
EA	Enumeration Area
HIV	Human Immunodeficiency Virus
HCV	Hepatitis C Virus
ID	Identification
LAg-EIA	Limiting-Antigen Avidity Enzyme Immunoassay
THIS	Tanzania HIV Impact Survey
OVC	Orphans and Vulnerable Children
PCA	Principal Components Analysis
PCR	Polymerase Chain Reaction
PHIA	Population-based HIV Impact Assessment

1 Background

1.1 What is THIS 2022-2023?

The Tanzania HIV Impact Survey 2022-2023 was a cross-sectional household-based survey conducted in Tanzania. THIS 2022-2023 is part of the PHIA Project, a series of population-based surveys, which are designed to assess the burden of HIV disease and impact of the health sector response to national HIV epidemics.

Please note that THIS 2022-2023 was the second PHIA survey in Tanzania. The previous PHIA in Tanzania is referred to as THIS 2016-2017.

1.2 Purpose of the THIS 2022-2023 Data Use Manual Supplement

The ***THIS 2022-2023 Data Manual Supplement*** (hereafter, ***Supplement***) contains THIS 2022-2023 survey specifications, including survey-specific eligibility criteria, sampling approaches and measures, and survey-specific technical documentation such as codebooks and questionnaires. It is intended to accompany the ***PHIA Data Manual*** (hereafter, "***Manual***"), which contains information on PHIA data generally applicable to all PHIA surveys, including general information on the data and documentation packages and their contents, a guide to getting started with the PHIA data, and details on the files and variables available for all PHIAs. Users should refer to both the ***Manual*** and this ***Supplement*** for a complete discussion of the THIS 2022-2023 data.

1.3 Other documentation and resources

In addition to the ***Manual*** and this ***Supplement***, users should refer to THIS 2022-2023 publications such as the ***THIS 2022-2023 Summary Sheet*** and ***THIS 2022-2023 Final Report***. The ***THIS 2022-2023 Summary Sheet*** contains highlights and summary results from the survey. The ***THIS 2022-2023 Final Report*** contains detailed results from THIS 2022-2023 along with information on survey data collection procedures, establishing participation by the household head, procedures for individual consent, maintaining confidentiality during data collection and testing procedures, procedures for returning/obtaining test results, and referral for or direct linkage to services.

Important THIS 2022-specific documentation attached to this ***Supplement*** includes:

- ***THIS 2022-2023 Survey Questionnaires***: The THIS 2022-2023 household, roster, and adult questionnaires. These questionnaires illustrate the questionnaire's structure, including the order that the questions were asked, each question's wording, variable names and labels, value coding and labels, and skip patterns. The question number on the questionnaire is referenced in the variable label on the datasets and in the "variable label" of the codebook, where applicable.
- ***THIS 2022-2023 Codebook with Frequencies***: Codebooks are provided for each dataset and document each variable's name, category (i.e., the questionnaire module or source data of the variable), full question text or variable description, variable label (i.e., a condensed label used on the datasets), type and width (e.g., numeric, text), coding values and labels, and the frequency and percent of records containing each value. Summary statistics are provided in the coding values and labels for selected numeric variables, such as counts.

- ***THIS 2022-2023 Analytic Variable Flow Diagrams***: Flow diagrams illustrate the logic used to create key analytic variables.
- ***THIS 2022-2023 Testing Methodology Diagram***: Flow diagram illustrating household-based HIV testing algorithm.
- ***THIS 2022-2023 Sampling and Weighting Technical Report***: Details of THIS 2022-2023 sampling and weighting procedures.
- ***THIS 2022-2023 Survey-Specific Table Specifications***: Table shells and technical specifications for THIS 2022-2023 specific final report tabulations. (Shells and specifications for all standard final report tabulations are included in the ***Manual***.)

Utility statistical programs are provided for Stata, SAS, or R depending on the file format requested.

- ***THIS 2022-2023 Stata Intro Code.do***: Stata do-file
- ***THIS 2022-2023 SAS Intro Code.sas***: SAS program
- ***THIS 2022-2023 R Intro Code.R***: R script

For Stata, values labels are included in the datasets. For SAS, there is a second statistical program containing code to label all values for variables on each of the data sets.

- ***THIS 2022-2023 SAS Formats.sas***: SAS program

2   Survey design and data collection

THIS 2022-2023 was a nationally representative, cross-sectional, two-stage, population-based survey of households across Tanzania. Its target population corresponded to adults, defined in this survey as those aged 15 years and older.

Table 1. THIS 2022-2023 survey design characteristics

Survey design characteristics	Description
Survey design	
Data source for survey weighting[1]	Pre-cartographic update for 2022 Tanzania census.
Sampling stratum	Region[1]
Primary sampling unit	Census Enumeration Areas (EA)
Urban/rural categorization	Urban/rural
Survey administration	
Data collection dates	November 2022-March 2023
Languages	English and Kiswahili
Sample size[2]	
Number of selected EAs	567
Number of selected households	19,845
Number of rostered individuals	81,788
Survey participation	
Number of completed household interviews	17,301
Number of completed individual interviews	35,960
Number of completed biomarker tests	33,663

[1] Zanzibar regions were collapsed and treated as a single sampling stratum.

[2] See **Attachment 9.5: *Sampling and Weighting Technical Report*** for more details on the survey weighting approach and for response rates.

Exceptions to the general PHIA design

Sampling

THIS 2022-2023 used a more complex sample design specification compared to most other PHIA surveys. Regions which were classified into three groups based on HIV prevalence and/or high rates of unsuppressed HIV viral load, with each group having different precision or minimum sample size requirements. For the five regions with high HIV prevalence (6% or more), the number of EAs per region was set to achieve a VLS precision of +/- 10% in each region, with the total resulting number of EAs then re-allocated equally across the regions. For the 13 regions with medium HIV prevalence of 3-6%, a similar process was carried out with a target precision of +/- 20% for VLS. For the remaining 12 low HIV prevalence regions, 12 EAs were selected per region with no specific VLS precision target. For full details of the sample design, see the sampling and weighting technical report.

One of the 567 selected EAs was a wildlife management area and contained no households. The THIS 2022-2023 datasets contain data from the remaining 566 EAs.

An error was made during the household interviews where the wrong household was visited. This error was not corrected before the weighting process, but it was excluded from the final datasets, which means that the survey weights do not precisely add up to the specified control totals.

Hepatitis Testing

In THIS 2022-2023 hepatitis B and hepatitis C testing was carried out in addition to the standard HIV tests. This testing was carried out in the laboratory using the blood samples collected during the household visits.

Hepatitis B testing was carried out for all participants, except for those where only 1 mL or 4 mL whole blood was collected. Hepatitis B core antigen (Anti-HBc) antibody testing was carried out using an immunoassay, followed by confirmation with the SD Bioline hepatitis B surface antigen (HBsAg) rapid test for those positive for Anti-HBc antibodies.

Hepatitis C testing was carried out for all participants, except for those where only 1 mL or 4 mL whole blood was collected. Antibody testing was performed using the SD Bioline Hepatitis C Virus (HCV) antibody test. Those testing positive had diagnostic HCV RNA viral load quantitative PCR testing performed using the Roche platform. Participants who were HIV positive with CD4 counts less than 200 cells/mL, who could be antibody negative, were also tested for HCV RNA viral load using PCR.

Questionnaire Changes

There were several country-specific additions to the questionnaire in THIS 2022-2023. Most new questions related to the following topics:

1. Pediatric elicitation ratio
2. HIV-related knowledge and attitudes
3. Hepatitis diagnosis, treatment, and risk
4. COVID-19 vaccination

These new questions, and those with differences that could cause misinterpretation or incomparability with the corresponding questions in other PHIA countries have had their dataset variables renamed to use a "_tz" suffix. The full list of questions and variables having country-specific changes is as follows:

New questions and variables for THIS 2022-2023

Adult Interview Dataset	
Question	**Variable Name**
Pediatric elicitation	
Do you have any children to whom you have given birth who are currently living with you?	childlivew_tz
How many children to whom you have given birth live with you currently?	childlivewnum_tz
Do you have any children to whom you have given birth who are alive but do not currently live with you?	childnotlivew_tz

How many children to whom you have given birth are alive but do not currently live with you?	childnotlivewnum_tz
Just to make sure that I have this right: you have TOTAL [childtotal] children who are alive today to whom you have given birth?	childtotalconf_tz
How many of your children who are alive today are 19 years of age or younger?	child19yr_tz
How many of your children who are alive today are 14 years of age or younger?	child14yr_tz

HIV-related knowledge and attitudes

Would you buy fresh vegetables from a shopkeeper or vendor if you knew that this person had HIV?	buyfood_tz
Do you think that children living with HIV should be able to attend school with children who are HIV negative?	kidsschool_tz
Do you agree or disagree with the following statement: "I would be ashamed if someone in my family had HIV."	famshame_tz
Do you agree or disagree with the following statement: "When taken as prescribed by a health worker, HIV medications decrease the amount of HIV in the blood of people living with HIV. Therefore, the amount of virus in their blood becomes too low to detect in a laboratory test."	attviralload_tz
Do you agree or disagree with the following statement: "A person living with HIV who is taking HIV medications cannot pass HIV to a sexual partner once a laboratory test can no longer detect the HIV virus in their blood."	atttransmit_tz
Do you agree or disagree with the following statement: "In the last 12 months, I have felt ashamed because of my HIV status."	shame12mo_tz
In the last 12 months, have people talked badly about you because of your HIV status?	talkbad12mo_tz
In the last 12 months, did someone else disclose your HIV status without your permission?	discloseoth12m_tz o
In the last 12 months, have you been verbally insulted, harassed, or threatened because of your HIV status?	insult12mo_tz
In the last 12 months, have you lost your job or another source of income because of your HIV status?	lostjob12mo_tz
In the last 12 months, have you been forced to change your place of residence or been unable to rent accommodation because of your HIV status?	changres12mo_tz
In the last 12 months, have you been denied health services because of your HIV status?	deniedserv12mo_tz

Hepatitis diagnosis, treatment, and risk

Have you previously been diagnosed with a chronic liver disease?	chronliver_tz
Have you previously been diagnosed with hepatitis B?	hepb_tz
Have you previously been diagnosed with hepatitis C?	hepc_tz
Have you ever lived with someone who has been diagnosed with hepatitis B or hepatitis C?	hepbclivew_tz

Have you ever been prescribed treatment for hepatitis B virus infection?	hepbtreat_tz
Have you ever been prescribed treatment for hepatitis C virus infection?	hepctreat_tz
Have you been vaccinated against hepatitis B?	hepbvac_tz
Have you ever heard about hepatitis from media or other community campaigns?	hepmedia_tz
Have you ever received a blood or blood product transfusion or organ transplant?	bloodtransf_tz
In the last 12 months, how many injections did you receive, for any reason?	inj12monum_tz
Have you received an injection from any of these persons in the last 12 months?	inj12mopers_[a-f, x]_tz
Have you ever received dialysis for kidney disease?	dialysis_tz
Do you have an ear or body piercing?	piercing_tz
Do you have a tattoo, cosmetic marking, or cuts/scarification?	tattoocuts_tz
Have you ever had a surgery or dental procedure?	surgeryproc_tz
COVID-19 vaccination	
Have you received a COVID-19 vaccine?	covvac_tz
How many doses of COVID-19 vaccine have you received?	covvacdose_tz
Which COVID-19 vaccine did you FIRST receive?	covvactype_tz
During what month/year did you receive the first dose of COVID-19 vaccine?	covvacfirstm_tz, covvacfirsty_tz
During what month/year did you receive the second dose of COVID-19 vaccine?	covvacsecm_tz, covvacsecy_tz
Do you think it will be easy or difficult to get a COVID-19 vaccine for yourself?	covvacdiff_tz
What makes it difficult for you to get a COVID-19 vaccine?	covvacdiffrsn_[a-l, x]_tz
What would motivate you to get vaccinated for COVID-19?	covvacmotiv_[a-j, w, x]_tz
Do you think most of your friends and family will get a COVID-19 vaccine, if it is recommended for them?	covvacfam_tz
Do you know where to get accurate, timely information about COVID-19 vaccines?	covvacinfo_tz
What are your top 3 most trusted sources of information about COVID-19 vaccines?	covvactrust_[a-k, w, x]_tz
Other THIS2-specific interview questions	
In the last 30 days, have you had any drinks containing alcohol?	alc30days_tz
In the last 12 months, have you had any drinks containing alcohol?	alc12mo_tz
How often do you have five or more drinks on one occasion?	alcfivemore_tz
Where do you normally consume alcohol?	alcloc_[a-g, x]_tz
Do you currently smoke any form of tobacco on a daily basis, less than daily, or not at all?	smoketob_tz

Adult Biomarker Dataset	
Question	**Variable Name**
Hepatitis biomarker variables	
Final hepatitis B test result	hepbstatusfinal_tz
Hepatitis B antibody test result	hepb_ab_tz
Hepatitis B confirmatory antigen test result	hepb_ag_tz
Final hepatitis C test result	hepcstatusfinal_tz
Hepatitis C antibody test result	hepc_ab_tz
Hepatitis C confirmatory antigen test result	hepc_ag_tz
Hepatitis C viral load test result	hepc_vl_tz

Questions with changes in wording or response options in THIS 2022-2023	
Household Dataset	
Question	**Variable Name**
What kind of toilet facility do members of your household usually use?	toilettype_tz
Main material of roof	matroof_tz
Does any member of your household own: [list of items]	hhqown_[a-f]_tz
Altogether, how many other animals (camels, horses, donkeys) do members of your household own?	ownhorsenum_tz
Roster Dataset	
Question	**Variable Name**
What region is [person name] in currently?	liveregion_tz
What country is [person name] in currently?	livecountry_tz
Adult Interview Dataset	
Question	**Variable Name**
Language of questionnaire	lngvqx_lng_tz
Language of interview	lngvint_lng_tz
Where did you deliver your child/children?	chdeliver_loc_[a-d,x]_tz
What is the highest level of school you attended?	schlhi_tz
What is the highest grade you have completed?	schcom_tz
You mentioned that you have a wife or wives who live elsewhere. Where are they?	wifewhere_tz

Is your husband or partner living with you now or is he staying elsewhere?	huslivew_tz
Over the past two weeks, how often have you been bothered by having little interest in doing things? Would you say: Not at all, several days, more than half the days, or nearly every day?	littleinterest_tz
Over the past two weeks, how often have you felt down, depressed or hopeless? Would you say: Not at all, several days, more than half the days, or nearly every day?	depressed_tz
Over the past two weeks, how often have you felt nervous, anxious or on edge? Would you say: Not at all, several days, more than half the days, or nearly every day?	anxiety_tz
Over the past two weeks, how often have you not been able to stop or control worrying? Would you say: Not at all, several days, more than half the days, or nearly every day?	worry_tz
Have you ever been told by a doctor or health worker that you have any of the following chronic health conditions?	chroniccond_[a-i, x]_tz
Are you currently taking medication for any of the following chronic health conditions?	chronicmed_[a-i, x]_tz
Have you taken part in any of the following prevention or treatment programs?	adhivprev_[a-d, w,x]_tz
Why is it difficult for you to get condoms?	condomnoteasyrsn_[a-j, x]_tz

3 Overview of survey questionnaires

In participating households, a household questionnaire was administered to the designated household head. Household head is defined as the person who is recognized within the HH as being the head and is 18 years and older or is considered an emancipated minor (an emancipated minor is a person aged 15-17 years old who is married or no longer depends on the parents and therefore does not require parental permission to participate in the survey). The household head provided consent for the household to participate in the survey, after which individual members were rostered during the household interview.

Then, adult individual questionnaires were administered to eligible and consenting individuals aged 15 years and older in the household. Consent criteria are determined in each country. It should be noted that non-emancipated minors provided consent via a different process than adults although they are grouped as adults for sampling and reporting. The consent criteria included:

- Women and men aged 18 years and older, living in or visiting the selected households, who slept in the household the night before the survey and who were willing and able to provide verbal informed consent.
- Minors aged 15-17 years, living in or visiting the selected households who slept in the household the night before the survey, were willing and able to provide verbal assent, and whose parents or guardians were willing and able to provide verbal informed consent/permission
  - For emancipated minors 15-17 years, parental permission was not required, but they had to be willing and able to provide verbal informed consent.

Modules included in each questionnaire and their associated eligibility criteria are listed in the table below. The content and order of each module may differ between THIS 2022-2023 and other PHIA surveys. Users can refer to each PHIA survey's *Survey Questionnaires* and *Codebooks* provided as attachments to this document.

Table 2. THIS 2022-2023 questionnaire

Questionnaire module	Eligibility criteria
Household questionnaire	Sample of households within selected EAs
Household roster	
Household roster for minors	
Deaths in the household	
Household characteristics	
Economic support	
Individual questionnaire – adults (15 years and older)	All eligible[1] and consenting individuals
Respondent background	
Marriage	
Reproductive history	All women
Male circumcision	All men
Sexual activity	
HIV testing history	
HIV status, care and treatment	All self-reporting HIV-positive adults
Tuberculosis and other health issues	
Hepatitis diagnosis, treatment, and risk	
Alcohol and tobacco use	
Exposure to prevention intervention	All individuals age 15-24

COVID-19 vaccination

[1] Household members are eligible if they were confirmed to have slept in the household the night before the interview.

4 Biomarker testing

In THIS 2022-2023, biomarker testing was offered to all rostered and consenting adults (15+ years). Eligibility criteria for receiving tests for specific biomarkers are provided in the table below.

Table 3. THIS 2022-2023 biomarker testing

Biomarker	Eligibility criteria
HIV serostatus[1]	All participants
Limiting Antigen Enzyme (LAg-Avidity)[2]	All HIV+ individuals
CD4+ cell count	All HIV+ individuals
HIV RNA viral load	All HIV+ individuals
Antiretroviral (ARV) drug presence	All HIV+ individuals
ARV drug resistance	All HIV+ individuals with viral load ≥ 200 copies/mL
Hepatitis B	All participants
Hepatitis C	All participants

[1] See HIV testing algorithm below.
[2] Recency of HIV infection is determined via a combination of Limiting Antigen Enzyme (LAg-Avidity) Immunoassay, viral load and ARV results. See "New HIV infections and annual HIV incidence" in the **PHIA Data Use Manual**.

THIS 2022-2023 HIV testing algorithm

For participants 15 years of age and older, initial household-based HIV testing was performed with the national HIV testing algorithm using two HIV rapid tests, see **Attachment 9.4: HIV testing methodology diagram**. Individuals with a nonreactive result on an initial HIV rapid screening test (SD Bioline HIV-1/2) were classified as HIV-negative. Then, individuals with a reactive screening test underwent confirmatory testing using the UniGold rapid HIV test. Individuals with a reactive screening test result followed by a non-reactive confirmatory test results were retested in series in the field. Individuals with both tests non-reactive were classified as HIV-negative. Those with both tests reactive were classified as HIV-positive. Individuals with repeated positive SD Bioline and negative UniGold tests were considered to have inconclusive results and referred to a health facility for retesting after 14 days as per the national guidelines.

All HIV positives identified in the field received confirmatory testing in the satellite laboratory using the BioRad Geenius™ HIV 1/2 Supplemental Assay. Individuals who self-reported being HIV-positive but tested HIV-negative received additional laboratory-based HIV testing via DNA qualitative polymerase chain reaction (PCR) test (Roche COBAS AmpliPrep/COBAS Taqman (CAP/CTM) HIV-1 Qualitative Test).

5 Data confidentiality

As noted in the ***Manual***, various risk mitigation actions were used to protect the privacy and confidentiality of respondents in the public use data. Some of these actions apply to all PHIA surveys, while other actions are data-driven decisions motivated by various risk disclosure concerns. These concerns include small counts as a result of certain combinations of variables and values which may introduce individual disclosure risk concerns. This section outlines the variables that have been identified for disclosure risk remediation and the specific data action taken to address the risk concern.

The following date variables were redacted for all PHIA surveys prior to public release:

Table 4. Date variables redacted for all PHIA surveys

Dataset(s)	Filename	Variable
Household	this2022hh	dieddated_01-05
Household	this2022hh	dieddatem_01-05
Adult individual	this2022adultind	surveystday
Adult individual	this2022adultind	birthday

Top-coding is the process of re-coding values above an upper bound to the value of the upper bound. Age for all respondents was top coded at 80. There was also top-coding to collapse small counts with nearby values, in which the data were re-coded so that the highest category contains at least 25 cases or 1 percent of households or individuals reporting the category. Variables that underwent top-coding are listed below:

Table 5. Variables that underwent top-coding

Dataset(s)	Filename	Variable	Top-coding upper bound
Household	this2022hh	ownchiknnum	30
Household	this2022hh	owncownum	30
Household	this2022hh	owndognum	4
Household	this2022hh	owngoatnum	20
Household	this2022hh	ownhorsenum_tz	3
Household	this2022hh	roomsleep	8
Household	this2022hh	diedagey_01-03	80
Roster	this2022roster	age	80
Adult individual	this2022adultind	age	80
Adult individual	this2022adultind	agemar	40
Adult individual	this2022adultind	arvsmissdays	4
Adult individual	this2022adultind	firstsxage	30
Adult individual	this2022adultind	husnwif	5
Adult individual	this2022adultind	lifetimesex	40
Adult individual	this2022adultind	liveb	11
Adult individual	this2022adultind	mcage	40
Adult individual	this2022adultind	monthtimes	7
Adult individual	this2022adultind	numwif	3
Adult individual	this2022adultind	part12monum	6
Adult individual	this2022adultind	partage1-3	80

Dataset(s)	Filename	Variable	Top-coding upper bound
Adult individual	this2022adultind	wifliveew	3
Adult individual	this2022adultind	arvamtm	6
Adult individual	this2022adultind	chtsthivagelastm1	24
Adult individual	this2022adultind	chtsthivagem1	24
Adult individual	this2022adultind	inj12monum_tz	10
Adult individual	this2022adultind	livetimey	61
Adult biomarker	this2022adultbio	age	80

Bottom-coding is the process of re-coding values below a lower bound to the value of the lower bound. Bottom-coding was used collapse small counts with nearby values, in which the data were re-coded so that the lowest category contains at least 25 cases or 1 percent of households or individuals reporting the category. Variables that underwent bottom-coding are listed below:

Table 6. Variables that underwent bottom-coding

Dataset(s)	Filename	Variable	Bottom-coding lower bound
Household	this2022hh	diedagey_01-03	5
Roster	this2022roster	liveawayy	2018
Adult individual	this2022adultind	agemar	14
Adult individual	this2022adultind	arvfty	2004
Adult individual	this2022adultind	cervcntsy	2010
Adult individual	this2022adultind	firstsxage	14
Adult individual	this2022adultind	hivcly	2021
Adult individual	this2022adultind	hivlastnegy	2000
Adult individual	this2022adultind	hivtesty	2000
Adult individual	this2022adultind	hivtfposy	2000
Adult individual	this2022adultind	monthwheny	2000
Adult individual	this2022adultind	partage1-3	14

The following variables and values were combined with the code for "other" due to small counts or percentages:

Table 7. Variables and values collapsed in to the "other" classification

Dataset(s)	Filename	Variable	Value(s)
Household	this2022hh	cookingfuel	5, 8
Household	this2022hh	matexwalls	23, 24
Household	this2022hh	matfloor	12, 22, 31
Household	this2022hh	matroof_tz	22, 32, 35
Household	this2022hh	toilettype_tz	41, 51, 62
Roster	this2022roster	livecountry_tz	3, 5, 8, 9, 12, 14, 15
Roster	this2022roster	liveregion_tz	8, 15, 21, 24, 26, 51, 52, 54, 55
Roster	this2022roster	relattohh	9
Adult individual	this2022adultind	hivtstlocation	6, 7, 11
Adult individual	this2022adultind	chronicmed_tz	H
Adult individual	this2022adultind	cmethod	B, H, K
Adult individual	this2022adultind	inj12mopers_tz	E
Adult individual	this2022adultind	workind	9
Adult individual	this2022adultind	partrelation1-3	7
Adult individual	this2022adultind	arvloc	3, 4, 5

The following variables were redacted entirely due to small counts or percentages:

Table 8. Variables that were redacted

Dataset(s)	Filename	Variable
Household	this2022hh	dieddatem_01-05
Adult individual	this2022adultind	childalive3-4
Adult individual	this2022adultind	childbrstfd2-4
Adult individual	this2022adultind	childbrstfdnow3-4
Adult individual	this2022adultind	chtsthivbirth2-4
Adult individual	this2022adultind	chtsthivbrstfd2
Adult individual	this2022adultind	chtsthivresultlast2
Adult individual	this2022adultind	deathagemo2
Adult individual	this2022adultind	deathageyr1-2
Adult individual	this2022adultind	firstsxagec
Adult individual	this2022adultind	hivposprovm
Adult individual	this2022adultind	chtsthivagelastm2
Adult individual	this2022adultind	chtsthivagem2
Adult individual	this2022adultind	lngvint_lng_tz
Adult individual	this2022adultind	lngvqx_lng_tz
Adult individual	this2022adultind	outregionwhrc_tz
Adult individual	this2022adultind	whereoutc_tz

Table 9. Variables with new categories

Dataset(s)	Filename	Variable	Values	New category

Roster	this2022roster	relattohh	9	10 – Other Relative
Adult individual	this2022adultind	arvloc	3, 4, 5	96 – Other
Adult individual	this2022adultind	arvamtm	0, 1	1 – 0-1 month
			2	2 – 2 months
			3	3 – 3 months
			4,5	4 – 4-5 months
			6 or greater	5 – 6+ months
Adult individual	this2022adultind	cerncnrslt	2, 3	Abnormal/Positive/Suspect Cancer
			4, 5	Inconclusive/Not Received

6 Dataset specifications

Table 10. THIS 2022-2023 dataset specifications

Dataset	Filename	Number of observations	Number of variables
Household	this2022hh	19,819	353
Roster	this2022roster	81,777	65
Adult individual	this2022adultind	35,957	686
Adult biomarker	this2022adultbio	33,663	332
Drug resistance [not yet available]	this2022dr	TBD	TBD
Household intermediary weights	this2022hhintermediarywts	19,819	283
Individual intermediary weights	this2022indintermediarywts	81,777	1,120

Dataset specification	Description
Two-letter country code prefix for ID variables	TZ
Survey weights provided (variable prefix)	Household (hhwt)
	Individual interview (intwt)
	Blood test (btwt)
	Drug resistance (drwt) [not yet available]
Number of jackknife replicate weights	277
Selected variable parameters	
Household characteristics used for wealth index construction	*See next section*
Mean duration recent infection (MDRI) used for HIV incidence estimation	130 days (95% CI 118-142 days, standard error 37.48575911)

7   Wealth index

As described in the **Manual**, a wealth index is constructed using principal component analysis (PCA) on household characteristics and asset ownership variables. The details of these variables vary by country. The table below lists the variables used to construct the wealth index for THIS 2022-2023.

Table 11. Household characteristics used for wealth index construction in THIS 2022-2023

Indicator variable	Type	Description
memsleep	Continuous	Number of household members per sleeping room[1]
matroof_tz	Categorical	Dwelling roofing material
matexwalls	Categorical	Dwelling wall material
matfloor	Categorical	Dwelling floor material
toilettype_tz	Categorical	Type of toilet used by the household
watersource	Categorical	Source of water used by the household
cookingfuel	Categorical	Type of cooking fuel used by the household
econsup12	Binary	Any external economic support
For the remainder of the variables:		*Does this household have/own…?*
hhqitems (option A)	Binary	Electricity
hhqitems (option B)	Binary	A working radio
hhqitems (option C)	Binary	A working television
hhqitems (option D)	Binary	A working telephone/mobile telephone
hhqitems (option E)	Binary	A working refrigerator
hhqown_tz (option A)	Binary	A bicycle
hhqown_tz (option B)	Binary	A working motorcycle or motor scooter
hhqown_tz (option C)	Binary	A working car or truck
hhqown_tz (option D)	Binary	A working boat with a motor
hhqown_tz (option E)	Binary	An animal-drawn cart
		How many of the following does this household have/own?[2]
ownchiknnum	Continuous	Chicken
owncownum	Continuous	Cows
owndognum	Continuous	Dogs
owngoatnum	Continuous	Goats
ownhorsenum	Continuous	Other animals (camels, horses, donkeys)

[1]Rounded to the nearest integer.
[2]For wealth index calculation, continuous variables for animal ownership have been changed into binary (yes/no). For example, the households that had any chickens were assigned "yes" and the households that had no chickens were assigned "no". This was done to be consistent with the DHS computation of wealth index (MoHCDGEC, 2016).

Wealth scores and model performance

The first component of the PCA model is interpreted as an index of household wealth. However, it does not explain a large proportion of the total variance: it accounts for only around 6.08% of the total variance in the common model, 4.85% for the urban model, and 4.71% for the rural

model. Howe et al. note that this figure is "often less than 20%" (Howe et. Al., 2008). The results from THIS 2022-2023 are consistent with those of other DHS studies in similar settings (Vyas and Kumaranayake, 2006; Filmer D and Pritchett, 2001; MoHCDGEC, 2016).

The PCA method does not guarantee the extraction of an index that is actually well-correlated with wealth but results from the PCA can be used to check whether the interpretation of the model makes sense. The component loading for each asset variable describes the association between that asset and the wealth index. The following table shows the most influential variables as measured by absolute value of their loading in each model.

Table 12. PCA results for THIS 2022-2023 wealth quintile

		Component loading		
Variable	Category	Common model	Urban model	Rural model
Floor material	Earth/Sand	-0.1384	-0.1659	-0.2064
Cooking fuel	Charcoal from wood	0.1332	0.1443	0.1560
Electricity	Yes	0.1245	0.2495	0.1730
Television	Yes	0.1205	0.1440	0.1335
Floor material	Cement/Terazo	0.1058	0.1140	0.1218
Toilet type	Flush to pit latrine	0.0962	0.0763	0.1114
Cooking fuel	Firewood/Straw	0.0865	0.1881	0.0779
Cooking fuel	LPG/Natural gas	-0.0836	-0.0860	-0.0692
Radio	Yes	0.0831	0.1165	0.0727

In THIS 2022-2023, floor material and cooking fuel were particularly important for the determination of wealth score. Note that variables with negative component loadings are associated with lower wealth, while those with positive loadings indicate a wealthier household.

The distribution of wealth index values from the model is shown in the figures below, first the wealth index for households in rural and urban areas separately, and then an overlayed distribution of the final wealth score with urban and rural households in contrasting colors. The distribution for the composite wealth index shows that urban households generally have higher wealth than rural households, though there is substantial overlap.

Distribution of continuous wealth scores (separate)

Distribution of continuous wealth scores (composite)
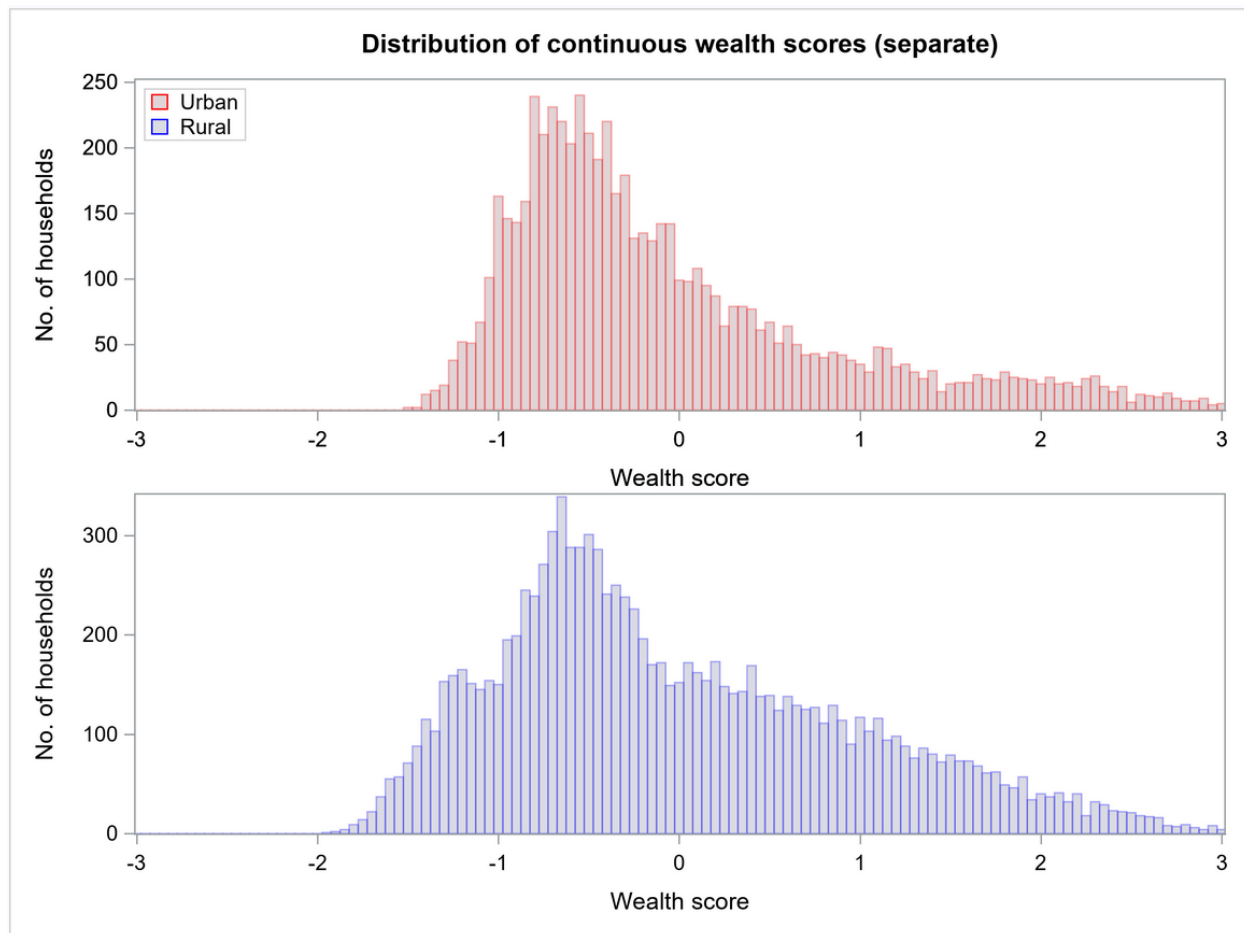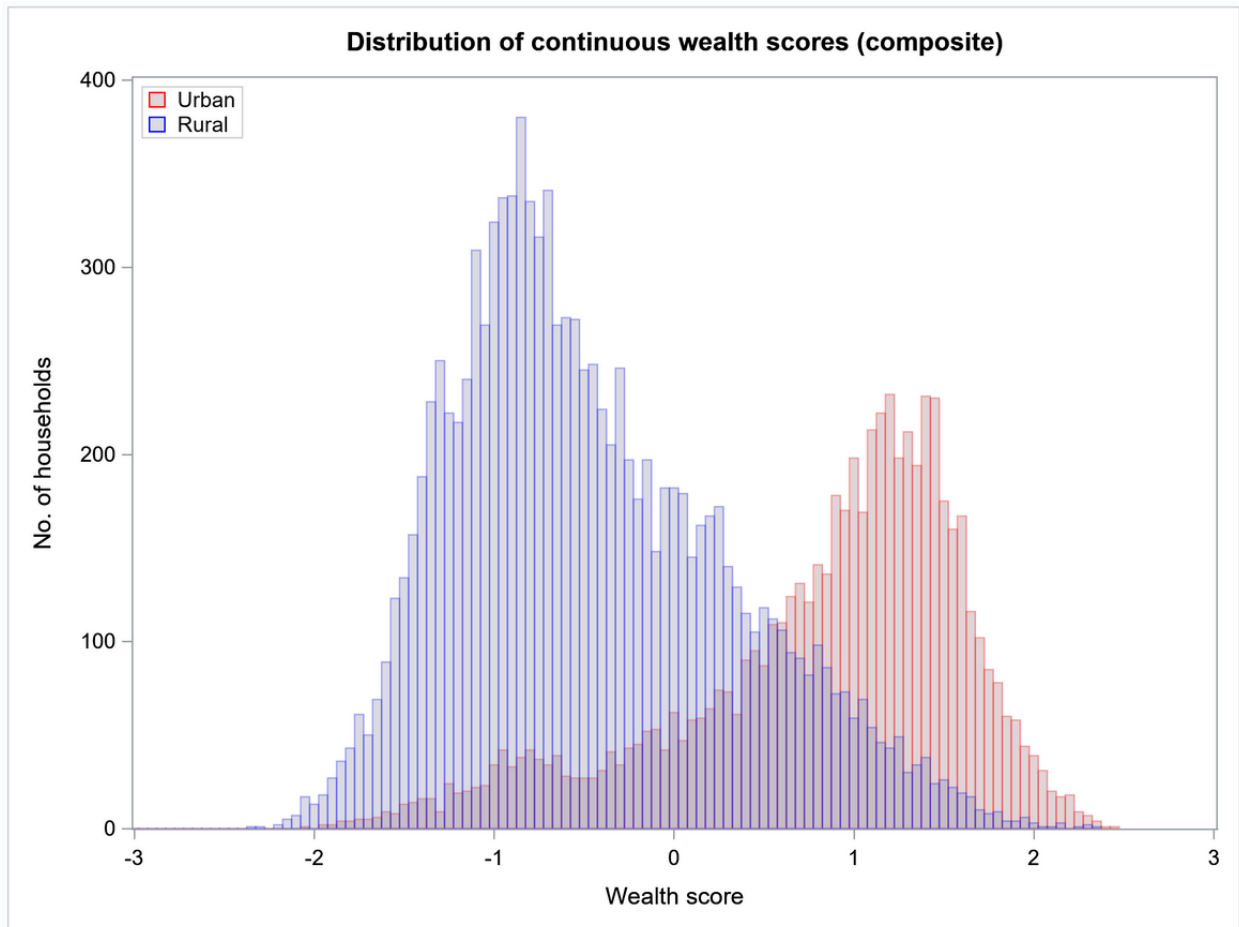
8 References

Filmer D, Pritchett LH. Estimating wealth effects without expenditure data - or tears: An application to educational enrollments in states of India. Demography. 2001;38(1):115-132.

Howe LD, Hargreaves JR, Huttly SR. Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. Emerg Themes Epidemiol. 2008; 5:3.(doi):10.1186/1742-7622-1185-1183.

Tanzania Commission for AIDS (TACAIDS), Zanzibar AIDS Commission (ZAC). Tanzania HIV Impact Survey (THIS) 2016-2017: Final Report. Dar es Salaam, Tanzania. December 2018.

Tanzania Commission for AIDS (TACAIDS), Zanzibar AIDS Commission (ZAC), Ministry of Health [Tanzania Mainland], Ministry of Health [Zanziabar]. Tanzania HIV Impact Survey (THIS) 2022-2023: Summary Sheet. Dar es Salaam, Tanzania. December 2023.

Tanzania Commission for AIDS (TACAIDS), Zanzibar AIDS Commission (ZAC), Ministry of Health [Tanzania Mainland], Ministry of Health [Zanziabar]. Tanzania HIV Impact Survey (THIS) 2022-2023: Final Report. Dar es Salaam, Tanzania. [TBD] 2024.

Vyas S, Kumaranayake L. Constructing socio-economic status indices: how to use principal components analysis. Health Policy Plan. 2006;21(6):459-468. Epub 2006 Oct 2009.

Ministry of Health (MoH) [Tanzania Mainland], Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS), and ICF. 2023 Tanzania Demographic and Health Survey and Malaria Indicator Survey 2022 Key Indicators Report. Dodoma, Tanzania, and Rockville, Maryland, USA: MoH, NBS, OCGS, and ICF.

Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC) [Tanzania Mainland], Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS), and ICF. 2016. Tanzania Demographic and Health Survey and Malaria Indicator Survey (TDHS-MIS) 2015-16. Dar es Salaam, Tanzania, and Rockville, Maryland, USA:MoHCDGEC, MoH, NBS, OCGS, and ICF.

9 Attachments

9.1 Questionnaires

Supplement Attachment 1 - THIS 2022-2023 Questionnaires

9.2 Codebook with frequencies

Supplement Attachment 2 - THIS 2022-2023 Codebook

9.3 Flow Diagrams for selected analytic variables

Supplement Attachment 3 - THIS 2022-2023 Flow Diagrams for Analytic Variables.pdf

9.4 HIV Testing Methodology Diagram

Supplement Attachment 4 - THIS 2022-2023 Testing Methodology Diagram.pdf

9.5 Sampling and Weighting Technical Report

Supplement Attachment 5 - THIS 2022-2023 Sampling and Weighting Technical Report.docx

9.6 THIS 2022-2023 Survey-specific table specifications

Supplement Attachment 6 - THIS 2022-2023 Additional Table Specifications.xlsx

9.7 Requesting data

THIS 2022-2023 data can be requested for use in research and analysis under the following conditions:

- Recipient will use this data only for the purpose of the research and analysis described in this data request. The recipient will submit a new request if they intend to use the data for another purpose.
- Recipient will not share this data with other researchers, with the exception of those listed in this data request as co-researchers for the project.
- Recipient will ensure that co-researchers are aware of and follow the terms of this data use agreement.
- Recipient will treat all data as confidential. Recipient will not use the data to deliberately compromise or otherwise infringe on the anonymity of participants' information and their right to privacy and will not attempt to identify any individual, household, or community in the survey based upon these data.
- Recipient will not publish any result in which participants, EAs or communities can be identified.
- Recipient will keep data in a secure location where it cannot be accessed by unauthorized users.
- Recipient will not use this data for any commercial venture.
- Recipient agrees that this agreement terminates immediately upon any breach by the recipient of the data or any co-researchers.

To see a demonstration of the data request process, watch the video here. The process is described in detail below.

To make a data request, first create an account at https://phia-data.icap.columbia.edu/ using the "Register" button and login using the button at the top right of the page. Once logged in, click "Data Sets" in the top menu to see the list of countries available. For THIS 2022-2023, select "Tanzania" from the list.

The top part of the page shows the PHIA survey years and datasets available for request, and the lower part shows the available documentation. Documentation may be downloaded without submitting a request. To obtain access to datasets, select the datasets you require for your project and click "Request Access". Fill out the project title and project description, including the general aims of your research and a brief description of your planned analysis. Fill out any co-researcher details, then click "Next". Read the conditions of use and enter your name to agree to the conditions and submit your request. Requests will generally be reviewed and approved within 1-2 business days. You will receive an email confirmation of approval. Once access has been approved, the check marks beside the requested datasets will be replaced with clickable buttons which will begin downloads of the data.

Requests for PHIA geospatial data have a more rigorous approval process because of the additional privacy and confidentiality risks associated with location data. Requests for geospatial data must explain why geomasked cluster centroid data are essential to the proposed analysis and describe the specific spatial analytical methods that will be used. Refer to the PHIA Geospatial Data Use Manual, available freely on each country's data request page, for full information on the content of the geospatial datasets.

For assistance or for any questions about the data, you can use the help request section at the bottom of https://phia-data.icap.columbia.edu/help to submit a question.

9.8 Data explorer

The ICAP PHIA data site also includes data visualization tools which allow you to look up survey estimates for specific countries and to compare across countries. To access these, visit https://phia-data.icap.columbia.edu/visualization. To see a video demonstration of the data visualization tools, watch the video here. The main steps to create a data visualization are described below.

1. Choose Country

Select the country or countries you are interested in by clicking them on the map, then click "Next".

2. Choose Indicator

Use the "Indicator" drop down to choose the indicator of interest. Typing in the indicator box after clicking the drop down allows you to filter the indicators available. Many indicators include subindicators, which are selected using the subindicator drop down. For example, after selecting the "90-90-90 (self-reported ARV, Overall Percentages)" indicator, you can choose some or all of "Diagnosed", "On Treatment", and "Viral Load Suppression" as subindicators.

3. Specify Age and Gender

The age and gender drop downs allow you to subset the data visualization to include the age group and gender you are interested in.

4. Choose Stratification

Stratification categories allow you to obtain estimates broken down by a range of variables, such as age groups, education, marital status, and others. The available stratification options depend on the indicators selected.

5. Choose Visualization Type

Visualizations can be selected using the "Chart", "Table", and "Map" buttons in the top right of the display. The default is Chart, which typically displays a horizontal bar chart showing percentages with a 95% confidence interval, or for some indicators a count or median. The Table option shows the estimates in a tabular format, including columns for each selected option. The Map displays the estimates as a heat map for the selected countries.

6. Download

Chart and Table visuals can be saved by clicking the download button next to the question mark on the top right of the page. For a Chart, the download is a static image of the visual. For a table, a CSV file is generated for download.

For help with the data visualization tools, click the help button question mark in the top right of the page, or visit https://phia-data.icap.columbia.edu/help.