REPUBLIC OF ZAMBIA

ZAMPHIA

ZAMBIA POPULATION-BASED HIV IMPACT ASSESSMENT 2021

DATA MANUAL SUPPLEMENT
SAMPLING & WEIGHTING
TECHNICAL REPORT

PARTNERS

PEPFAR

CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

ZAMBIA STATISTICS AGENCY

ICF

UNIVERSITY of MARYLAND BALTIMORE

Table of Contents

List of Abbreviations iii
Chapter 1: Introduction 1
1.1 Overview of Sample Design 1
1.2 Overview of Weighting Process 1
Chapter 2: Sample Design 3
2.1 Population of Inference 3
2.2 Precision Specifications and Assumptions 3
2.2.1 Specifications 3
2.2.2 Assumptions 3
2.3 Selection of PSUs 5
2.3.1 Definition of PSUs 5
2.3.2 Selection of the PSU sample 5
2.4 Selection of Households 6
2.4.1 Listing 6
2.4.2 Household Selection 6
2.5 Selection of Individuals 8
Chapter 3: Weighting and Estimation 10
3.1 Overview of Survey Weights 10
3.2 Preparation for Weighting 11
3.3 Creation of Variables for Variance Estimation 13
3.3.1 Jackknife Replication 14
3.3.2 Taylor's Series 15
3.4 Eligible and Completed Households and Individuals 15
3.5 Development of Weights 16
3.5.1 Design Weight 16
3.5.2 Household Survey Weights 19
3.5.3 Individual Weight (indwt0) 21
3.5.4 Blood Draw Weight (bldwt) 25
3.6 Survey Weight Variables in ZAMPHIA Datasets 31
References 34
Appendix A: Reference Tables 35

List of Figures

Figure 1. Flowchart of eligibility and completion of different questionnaires and blood testing 16

List of Tables

Table 1. Sample allocation and expected levels of precision for VLS estimates by province (adults aged 15-49 years) 4
Table 2. Expected number of cases based on response rates for adults aged 15-49 years 5
Table 3. Provincial response rates and inputs based on ZAMPHIA 2016 5
Table 4. Distribution of sampled and completed PSUs and households by province 8
Table 5. Distribution of selected households by province and response status 8
Table 6. Distribution of individuals aged 15+ years by province and eligibility for data collection 9
Table 7. Distribution of individuals aged 15+ years who completed the individual questionnaire by province and completion of the blood test 9
Table 8. Variables used to determine eligibility and completion of different questionnaires and modules 12
Table 9. Number of PSUs and variance-estimation strata constructed for variance estimation 14
Table 10. Number of selected and dropped PSUs and weighted response rates by province and urban/rural 18
Table 11. Number of selected households by response groups 19
Table 12. Number of eligible adults by response groups 21
Table 13. Variables considered for the LASSO models and selected by LASSO and CHAID 23
Table 14. Number of variables used in LASSO/CHAID and number of nonresponse adjustment cells 24
Table 15. Projected and weighted totals of individual aged 15+ years and calibration factors by gender and age 24
Table 16. Number of eligible adults for blood testing by response groups 25
Table 17. Variables considered for the LASSO models and selected by LASSO and CHAID by gender 26
Table 18. Number of variables used in LASSO/CHAID and number of nonresponse adjustment cells 30
Table 19. Projected and weighted totals of individuals aged 15+ years and calibration factors by gender and age 31
Table 20. Survey weights and replicate weights: ZAMPHIA 32
Table 21. Survey Intermediary weights and replicate weights: ZAMPHIA 33
Table A.1. Interview weights: Nonresponse adjustment cells for adults aged 15+ years 36
Table A.2. Blood weights: Nonresponse adjustment cells for male adults aged 15+ years 41
Table A.3. Blood weights: Nonresponse adjustment cells for female adults aged 15+ years 44

List of Abbreviations

CHAID	Chi-square Automatic Interaction Detector
EA	Enumeration Area
HH	Household
ICC	Intraclass Correlation Coefficient
LASSO	Least Absolute Shrinkage and Selection Operator
PSU	Primary Sampling Unit
VLS	Viral Load Suppression
ZAMPHIA	Zambia Population-based HIV Impact Assessment

Chapter 1: Introduction

The Zambia Population-based HIV Impact Assessment (ZAMPHIA) is a two-stage stratified cluster sample survey designed to assess the prevalence of key HIV-related health indicators. ZAMPHIA 2021 had a sample size of 12,245 selected households (HHs) from 403 completed primary sampling units (PSUs) across 10 provinces. A total of 10,627 households completed a household questionnaire. Within the interviewed household, all household members aged 15+ years who slept in the house the night before were eligible to participate in the survey. A total of 22,262 eligible adults aged 15+ years completed a questionnaire for individuals. Blood samples yielding valid lab results were drawn from a total of 18,804 individuals aged 15+ years. The purpose of this report is to document the procedures used to select PSUs, households, and individuals for the study and the subsequent weighting of the respondent sample.

1.1 Overview of Sample Design

The ZAMPHIA sample was stratified and selected in two stages. The sample was stratified by province and place of residence (urban/rural), with a total of 20 sampling strata. Samples of enumeration areas (EAs), also referred to as PSUs, were selected independently in each stratum in two stages. In the first stage, EAs were selected with probability proportional to the EA size, number of households according to the census frame, and with independent selection in each sampling stratum. A household listing operation was carried out in all the selected EAs, and the resulting lists of households served as sampling frame for the selection of households in the next stage. In the second stage of selection, the number of households was selected from each EA with an equal probability systematic selection from the newly created household listing. Chapter 2 provides more details about the sampling design of ZAMPHIA.

1.2 Overview of Weighting Process

The main purpose of the survey weights calculated for ZAMPHIA was to (1) account for unequal selection probabilities at different stages of sampling, (2) adjust for nonresponse at different stages of data collection, (3) reduce the variability of the weights using a weight trimming procedure, and (4) calibrate the weights to the 2021 population projections. Taking into consideration those objectives, the process of calculating the weights started by calculating the design weights that account for the selection probabilities of the different sampling units in different sampling stages. The design weights were adjusted to account for nonresponse that occurred at the PSU and household levels. When

weights were calculated for individuals, the weights were adjusted for individual-level nonresponse to the survey questionnaire. When weights were calculated for blood testing, the weights were adjusted for nonresponse to the test. All weights were trimmed, where outliers were capped at a maximum value. Finally, all weights were calibrated based on the total distributions of the projected population. Chapter 3 provides technical details of the weighting procedures employed in ZAMPHIA.

Chapter 2: Sample Design

2.1 Population of Inference

The population of inference for ZAMPHIA was comprised of the de facto household population of Zambia. The de facto population comprised individuals who were present in households (i.e., slept in the household) on the night prior to the household interview. In contrast, the de jure population comprises individuals who are usual residents of the household, irrespective of whether they slept in the household on the night prior to the household interview.

2.2 Precision Specifications and Assumptions

The following specifications and assumptions were used to develop the sample design for ZAMPHIA.

2.2.1 Specifications

- Province-level viral load suppression (VLS) among HIV-positive persons aged 15-49 years with 95% confidence interval ±10%
- National-level HIV incidence for persons aged 15-49 years with relative standard error (RSE) ≤ 0.2.

2.2.2 Assumptions

- Provincial-level VLS for HIV-positive persons aged 15-49 years = 0.5
- National-level and provincial-level HIV prevalence estimates for persons aged 15-49 years as specified in Table 1, which are based on the ZAMPHIA 2016 survey (Ministry of Health, Zambia 2019)
- National-level annual HIV incidence for persons aged 15-49 years = 0.64%, which is based on the ZAMPHIA 2016 survey (Ministry of Health, Zambia 2019)
- Intraclass correlation coefficient (ICC) for VLS = 0.061, which is based on the ZAMPHIA 2016 survey and is the ICC associated with VLS, discounting the effect of differential weighting (Ministry of Health, Zambia 2019)
- ICC for HIV prevalence = 0.017, which is based on the ZAMPHIA 2016 survey and is the ICC associated with HIV prevalence, discounting the effect of differential weighting (Ministry of Health, Zambia 2019)
- ICC for HIV incidence = 0.0, which is based on the ZAMPHIA 2016 survey and is the ICC associated with HIV incidence, discounting the effect of differential weighting (Ministry of Health, Zambia 2019)

- Mean duration of recent infection = 130 days
- Probability false recent = 0.000001%
- An adjustment factor = 1.018 to account for mean duration of recent infection and probability false recent is included for national HIV incidence estimation and associated variance calculations
- Average number of selected dwelling units per cluster = 30, which should yield an average of approximately 24 responding HHs per cluster
- Actual number of selected dwelling units per cluster will reflect changes in the measure of size between the sampling frame and HH listing
- Average number of de facto HH household members aged 15-49 years = 2.07, based on the ZAMPHIA 2016 survey (Ministry of Health, Zambia 2019)
- Provincial-level average number of de facto HH members aged 15-49 years as specified in Table 3 based on the ZAMPHIA 2016 survey (Ministry of Health, Zambia 2019)
- Provincial-level HH occupancy rates as specified in Table 3, based on the ZAMPHIA 2016 survey (Ministry of Health, Zambia 2019)
- Provincial-level HH interview, individual interview, and HIV testing response rates as specified in Table 3, based on the ZAMPHIA 2016 survey (Ministry of Health, Zambia 2019)

Table 1. Sample allocation and expected levels of precision for VLS estimates by province (adults aged 15-49 years)

Province	HHs (%)	Selected EAs	Blood samples	HIV prevalence (%)	HIV-positive cases (%)	Expected number of HIV-positive cases	Cluster design effect[1]	VLS 95% confidence bounds
Central	9.7%	30	1,188	12.4%	9.4%	147	1.24	±0.09
Copperbelt	14.9%	42	1,957	13.2%	16.5%	258	1.31	±0.07
Eastern	11.5%	40	1,605	7.4%	7.6%	119	1.12	±0.10
Luapula	7.1%	42	1,391	8.9%	7.9%	124	1.12	±0.10
Lusaka	18.8%	40	1,627	14.4%	15.0%	234	1.30	±0.08
Muchinga	6.1%	57	2,085	5.3%	7.1%	111	1.06	±0.10
Northern	8.5%	39	1,341	9.4%	8.1%	126	1.14	±0.10
North Western	5.3%	48	1,800	6.4%	7.4%	115	1.09	±0.10
Southern	11.9%	29	1,208	12.5%	9.7%	151	1.26	±0.09
Western	6.0%	37	1,134	15.4%	11.2%	175	1.23	±0.08
TOTAL	100.0%	404	15,336	11.4%	100.0%	1,560	1.17	±0.03

Parameters used: Average of 24.3 responding HHs per EA. Average HH size of 4.62 with 2.07 people aged 15-49 years (ZAMPHIA 2016).

[1]Total design effect = 1.17 is the product of the unequal weighting effect and the clustering effect.

Table 2. Expected number of cases based on response rates for adults aged 15-49 years

	Response rate	Expected cases[2]
Sample (dwelling units at listing)		12,120
HH occupancy	92.0%	11,093
HH response	88.6%	9,811
Total HH response	81.5%	9,811
Eligible adults (aged 15-49 years)	(2.07/HH)	19,776
Adult interview response	85.1%	17,106
Adult biomarker response	90.6%	15,336
Total adult response[1]	77.0%	15,336

[1] Total adult response rate is the product of adult interview response and biomarker rates.

[2] Based on provincial response rates and inputs in Table 3

Table 3. Provincial response rates and inputs based on ZAMPHIA 2016

Province	HH occupancy rate	HH response rate	Eligible person interview response rate	Adult biomarker response rate	Eligible persons (aged 15-49 years)/ HH
Central	0.93	0.90	0.88	0.89	2.01
Copperbelt	0.95	0.91	0.86	0.90	2.32
Eastern	0.94	0.92	0.88	0.93	1.91
Luapula	0.89	0.89	0.90	0.90	1.72
Lusaka	0.93	0.88	0.78	0.89	2.37
Muchinga	0.91	0.89	0.89	0.90	1.88
Northern	0.89	0.88	0.89	0.86	1.92
North Western	0.90	0.85	0.84	0.90	2.19
Southern	0.92	0.88	0.91	0.91	2.08
Western	0.91	0.84	0.87	0.89	1.75
ZAMBIA	0.92	0.89	0.86	0.90	2.07

2.3 Selection of PSUs

2.3.1 Definition of PSUs

The sampling frame used for ZAMPHIA was based on the Zambia Population and Housing Census conducted in 2010, provided by the Zambia Statistics Agency. Zambia is divided into 10 provinces. Each province is subdivided into districts, each district into constituencies, each constituency into wards, and each ward into mutually exclusive EAs. The census frame contained a total of 25,631 EAs containing 13,092,666 individuals living in 2,513,768 households, with an average number of 98 households per EA and 511 persons per EA.

2.3.2 Selection of the PSU sample

A stratified sample of 404 EAs was selected with probability proportional to the EA measure of size and with independent selection in each sampling stratum with the sample allocation given in Table 4. The measure of size is the number of residential households residing in the EA based on the 2010 Zambia Population and Housing Census. Implicit stratification and proportional allocation are achieved at each of the lower administrative levels by sorting the sampling frame within each sampling stratum before

sample selection according to administrative levels, and by using a probability proportional to size selection at the first stage of sampling.

2.4 Selection of Households

The selection of households for ZAMPHIA involved the following steps: (1) listing the dwelling units/households within the sampled EAs, (2) assigning eligibility codes to the listed dwelling unit/household records, and (3) selecting the samples of dwelling units/households.

For both sampling and analysis purposes, a household was defined as a group of persons who normally live and eat together. These people may or may not be related by blood but make common provision for food or other essentials for living and have only one person whom they all regard as the head of the household. Households were eligible for participation in this survey if they were within the predefined EA and were randomly selected for inclusion in the survey.

2.4.1 Listing

A household listing operation was carried out in all the selected sample EAs, and the resulting lists of households served as the sampling frame for the selection of households in the next stage. The following activities were conducted by the listers after training:

- Obtained the census EA Map of the sampled EAs from the Zambia Statistics Agency.
- Approached the local community leaders of the sampled EAs for permission to map and list households after explaining their mission.
- Beginning with the “starting point” on the EA map, identified current structures within the boundaries of the sampled EAs.
- Numbered all the buildings/structures within the EA as ZAMPHIA/Cluster Number/ Building number. Example: ZAMPHIA/300/19 on the walls.
- Explained their mission to occupants of residential buildings.
- Explained the concept of “household” to the occupant of all the residential buildings.
- Identified the households occupying the respective buildings and listed each household, indicating the address and the name of the head of household. This information was captured using electronic data capture on a tablet and streamed in real time to a central server.

2.4.2 Household Selection

Within each EA, a random sample of households was selected from households listed during the listing process. Using a systematic sampling procedure, an average of 30 households, ranging between 10 and

60 households, were sampled from each cluster in all provinces as a secondary sampling unit. Table 4 provides the allocation of the clusters and households selected by province.

To achieve equal probability samples of households, within-PSU sampling rates were required to select households that depended on the actual number of households found at the time of listing and the number based on the census frame. By using the within-PSU sampling rates, we selected more than the desired 30 households per PSU from PSUs that have experienced growth in population since the 2010 census, and fewer than the desired 30 households per PSU from PSUs that have declined in population.

The within-PSU sampling rate for a PSU *i* was calculated as follows:

$$f_{0hi} = \frac{L_{hi}}{N_{hi}}$$

where L_{hi} is the number of households listed during the household listing operation in PSU *i* in stratum *h,* and N_{hi} is the number of households according to the sampling frame in the i^{th} PSU. The initial within-PSU sampling rate was then used to calculate a sample size for PSU *i* as follows:

$$g_{0hi} = \begin{cases} 10 & 30 * f_{0hi} \leq 10 \\ 30 * f_{0hi} & 10 < 30 * f_{0hi} < 60 \\ 60 & 30 * f_{0hi} \geq 60 \end{cases}$$

The upper and lower limits (i.e., 10 and 60) were imposed on the sample size to avoid high variation in workload across the sampled PSUs. The difference between the number of households that would have been selected using the rates, f_{0hi}, and the target number of selected households was reduced by adjusting the within-PSU sampling rate as follows:

$$f_{hi} = f_{0hi} * \frac{30 * a}{\sum_{i=1}^{a} g_{0hi}}$$

where *a* is the total number of PSUs selected for the stratum. The adjusted version of the within-PSU sampling rate f_{hi} was then used for calculating the final number of households selected as follows:

$$g_{hi} = \begin{cases} 10 & 30 * f_{hi} \leq 10 \\ 30 * f_{hi} & 10 < 30 * f_{hi} < 60 \\ 60 & 30 * f_{hi} \geq 60 \end{cases}$$

Table 4. Distribution of sampled and completed PSUs and households by province

Province	Selected PSUs	Completed PSUs	Selected households	Interviewed households
Central	30	30	1,117	1,007
Copperbelt	42	42	1,215	1,133
Eastern	40	40	1,304	1,134
Luapula	42	41*	1,105	982
Lusaka	40	40	1,159	987
Muchinga	57	57	1,753	1,476
Northern	39	39	1,082	996
North Western	48	48	1,320	1,089
Southern	29	29	827	669
Western	37	37	1,363	1,154
Zambia	404	403	12,245	10,627

* One EA was dropped because it was found to be in the Democratic Republic of Congo following border re-demarcation between the Democratic Republic of Congo and Zambia.

Table 5 summarizes the number of households selected for the ZAMPHIA by final household response status. Of the 12,245 sampled households, 10,627 (86.8%) completed the household questionnaire, 926 (7.6%) were determined to be eligible for the study but did not complete the household questionnaire, and 666 (5.4%) were determined to be ineligible. For 26 households (0.2%), eligibility for the survey could not be established.

Table 5. Distribution of selected households by province and response status

Province	Eligible responding households	Eligible nonresponding households	Ineligible	Unknown eligibility status	Total
Central	1,007	64	42	4	1,117
Copperbelt	1,133	45	32	5	1,215
Eastern	1,134	118	50	2	1,304
Luapula	982	44	79	0	1,105
Lusaka	987	111	57	4	1,159
Muchinga	1,476	156	120	1	1,753
Northern	996	43	43	0	1,082
North Western	1,089	110	121	0	1,320
Southern	669	113	38	7	827
Western	1,154	122	84	3	1,363
Zambia	10,627	926	666	26	12,245

2.5 Selection of Individuals

The selection of individuals for ZAMPHIA involved the following steps: (1) compiling a list of all individuals known to reside in the household or who slept in the household during the night prior to data collection, (2) identifying rostered individuals who were eligible for data collection, and (3) selecting those individuals meeting the age and residency requirements of the study. Only those individuals who slept in the household the night before the household interview (i.e., the de facto population) were eligible for interviews.

In all households, all de facto adults aged 15+ years were eligible to complete the individual questionnaire. Table 6 summarizes the number of rostered adults aged 15+ years by province and eligibility status for the data collection. Adults who completed the individual questionnaire were eligible for blood testing for HIV. Table 7 summarizes the number of adults aged 15+ years who were eligible for blood testing by province and completion of the blood testing according to the final lab results.

Table 6. Distribution of individuals aged 15+ years by province and eligibility for data collection

Province	Eligible respondent	Eligible nonrespondent	Total
Central	2,168	246	2,414
Copperbelt	2,576	347	2,923
Eastern	2,191	487	2,678
Luapula	2,087	153	2,240
Lusaka	1,887	593	2,480
Muchinga	2,803	356	3,159
Northern	2,055	125	2,180
North Western	2,391	341	2,732
Southern	1,380	291	1,671
Western	2,724	282	3,006
Zambia	22,262	3,221	25,483

Table 7. Distribution of individuals aged 15+ years who completed the individual questionnaire by province and completion of the blood test

Province	Number of valid blood tests	Number of invalid blood tests and non-tested	Number completing interview
Central	1,824	344	2,168
Copperbelt	2,252	324	2,576
Eastern	1,739	452	2,191
Luapula	1,977	110	2,087
Lusaka	1,486	401	1,887
Muchinga	2,019	784	2,803
Northern	1,918	137	2,055
North Western	2,044	347	2,391
Southern	1,063	317	1,380
Western	2,482	242	2,724
Zambia	18,804	3,458	22,262

Chapter 3: Weighting and Estimation

In general, the purpose of weighting survey data from a complex sample design is to (1) compensate for variable probabilities of selection, (2) account for differential nonresponse rates within relevant subsets of the sample, and (3) adjust for possible under-coverage of certain population groups. Weighting is accomplished by assigning an appropriate sampling weight to each responding sampled unit (e.g., a household or person) and using that weight to calculate weighted estimates from the sample. The critical component of the sampling weight is the base weight, which is defined to be the inverse of the probability of selecting a household or person in the survey sample. The base weights are used to inflate the responses of the sampled units to population levels and are generally unbiased if there is no nonresponse or noncoverage in the sample (e.g., see Kish, 1965, p. 67). When nonresponse or noncoverage occurs in the survey, weighting adjustments are applied to the base weights to compensate for both types of sample omissions.

Nonresponse is unavoidable in virtually all surveys of human populations. For ZAMPHIA, nonresponse can occur at different stages of data collection, for example, (1) before the enumeration of individuals in the household, (2) after household enumeration and selection of persons but before completion of the individual interview, and (3) after completion of the interview but before collection of a usable blood sample. The procedures used to compensate for nonresponse at each of the relevant stages of data collection are described in Section 3.4.

Noncoverage arises when some members of the survey population have no chance of being selected for the sample. For example, noncoverage can occur if the field operations fail to enumerate dwelling units during the listing process, or if certain household members are omitted from the household rosters. To compensate for such omissions, the adjustment procedures described in Sections 3.5.3 and 3.5.4 are used to calibrate the weighted sample counts to available population projections.

3.1 Overview of Survey Weights

The following three survey weights were calculated for the ZAMPHIA data:

- **Household survey weight (hhwt0):** Weight for the household interviews
- **Interview weight (indwt0):** Individual weight for adults aged 15+ years in all households. Eligibility for the individual questionnaire is described in detail in Section 3.4.
- **Biomarker weight (bldwt0):** Weight for blood tests for adults aged 15+ years for whom a valid blood test exists

The overall weighting process included several steps, as follows:

1. **Initial checks:** Checks of the data files are carried out as part of the survey and data quality control, and the probabilities of selection for PSUs and households are calculated and checked.
2. **Creation of jackknife replicates:** The variables needed to create the jackknife replicates for variance estimation are established at this point. This step can be implemented immediately after the PSU sample has been selected.
3. **Calculation of base weight:** The weighting process begins with the calculation and checking of the sample PSU (EA) probabilities of selection and the within-EA household selection probabilities. The reciprocal of the product of the two probabilities is the base weight for all the next steps. At this step, the base weight can also be adjusted for any PSUs that could not be accessed or interviewed.
4. **Calculation of household weight (hhwt0):** The next step is to calculate household weights by adjusting the base weight for selected households for which it could not be determined whether the dwelling unit contained an eligible household and for nonresponding eligible households. This adjustment is made based on the EA in which the households are located, and the resulting weight is the final household weight.
5. **Calculation of interview weight (indwt0):** After the household weights are calculated, they are used to calculate the individual weights. The household weights are adjusted for nonresponse among the eligible individuals, trimmed, and then calibrated to compensate for under-coverage in the sampling process by weighting up to 2021 population projections.
6. **Calculation of biomarker weight (bldwt0):** For individuals who were interviewed, the uncalibrated interview weights are adjusted for nonresponse for blood testing, trimmed, and then calibrated to compensate for under-coverage.
7. **Application of weighting adjustments to jackknife replicates:** All the adjustment processes are applied to the full sample and the replicate samples so that the final set of full sample and replicate weights can be used for variance estimation that considers the complex sample design.

3.2 Preparation for Weighting

The following data files were used during the weight calculations:

- **ZMhh_household:** A SAS dataset that includes records for all selected households and data collected in the household questionnaire for completed households

- **ZMhh_individual:** A SAS dataset that contains records of all individuals rostered in the completed households collected in the household questionnaire
- **ZMin_individual:** A SAS dataset that contains interview records for individuals aged 15+ years collected in the individual questionnaire
- **ZMlb_individual:** A SAS dataset that contains biomarker records for all eligible participants who agreed to blood testing for HIV
- **selprob:** A CSV dataset that includes the selection probabilities for all the selected sampling clusters for the ZAMPHIA. The file includes the selection probabilities for the different stages (p_1, p_2, and p), and strata (VarStrat) and clusters (VarUnit) for variance estimation, and *pairs* and *pairs_fact* for the calculations of Jackknife replicates.

From these datasets, several variables were used to determine the eligibility for different questionnaires and for blood testing, as explained in Section 3.4. Table 8 presents a list of key variables used for that purpose.

Table 8. Variables used to determine eligibility and completion of different questionnaires and modules

Variable name	Value codes and label names
HRESULT *Result of household interview* Coded in ZMhh_household	1 = "Completed" 2 = "No household member at home" 3 = "Entire household absent for extended period of time" 4 = "Postponed" 5 = "Refused" 6 = "Dwelling vacant" 7 = "Address not a dwelling" 8 = "Dwelling destroyed" 9 = "Dwelling not found" 10 = "Inaccessible due to flooding or security" 11 = "Withdrawn" 12 = "Stop survey" 13 = "No competent head of household" 14 = "At home, reschedule visit" 15 = "Team in isolation/quarantine for COVID-19" 95 = "Unsuccessful HH due to COVID19" 96 = "Other"
IRESULT *Result of individual interview* Coded in ZMin_individual	1 = "Completed" 2 = "Not at home" 3 = "Postponed" 4 = "Refused" 5 = "Incapacitated" 6 = "Withdrawn" 7 = "Cognitively impaired (Ineligible)" 8 = "Stop survey" 9 = "At home, reschedule visit" 10 = "Team in isolation/quarantine for COVID-19" 95 = "Unsuccessful individual interview due to COVID19"

Variable name	Value codes and label names
	96 = "Other" 97 = "Not Revisited during data collection"
FinalResult *Final lab result of HIV testing* Coded in ZMlb_individual	1 = "HIV Positive" 2 = "HIV Negative" 3 = "No result"
Gender *Final gender of respondent* An analytic variable coded based on variables from ZMhh_individual and ZMin_individual	1 = "Male" 2 = "Female"
AGE *Final age in years* An analytic variable coded based on variables from ZMhh_individual, ZMin_individual, and ZMlb_individual	A continuous variable
hhstatus *Response Rates _ Households* An analytic variable coded based on variables from ZMhh_household	1 = "Eligible Responding Household" 2 = "Eligible Nonresponding household" 3 = "Ineligible (Vacant Household, not a Dwelling, Dwelling Destroyed)" 4 = " Unknown Eligibility Status"
indstatus *Respondent disposition code_adults* An analytic variable coded based on variables from ZMhh_individual and ZMin_individual	0 = "Incomplete household questionnaire" 1 = "Respondent" 2 = "Eligible Non-Respondent" 3 = "Ineligible (confirmed age < 15)" 4 = "Roster ineligible (roster age < 15)" 5 = "Rostered eligible from household with no questionnaire data" 6 = "Non-de-facto ineligible (Roster age 0:95)"
btstatus *Biomarker disposition code* An analytic variable coded based on variables from ZMin_individual and ZMlb_individual	1 = "Lab blood test has a definite result" 2 = "Lab blood test does not a definite result" 3 = "Eligible for biomarker but blood not drawn" 4 = "Nonrespondents to individual questionnaire" 5 = "Ineligible for biomarker"

3.3 Creation of Variables for Variance Estimation

Two general methods can be used for estimating the sampling errors of survey-based estimates derived from ZAMPHIA: jackknife replication and Taylor's Series methods. The jackknife replication variance estimation method is a widely used method for producing variance estimates using data from a complex survey. This method can correctly account for the stratification, clustering, and sample weighting, including nonresponse, trimming, and poststratification weighting adjustments from the ZAMPHIA complex sample design. The Taylor's Series method is also widely used, and it employs linear approximations to calculate the variance of a sample-derived estimate. To implement either method, certain variables required for variance estimation must be included in the weighted data files. In the case of jackknife replication, the required variables are a series of weights that correspond to each of the jackknife replicates. In the case of the Taylor's Series method, the required variables are variables that indicate the "variance stratum" and the "variance unit" to which each sampled respondent belongs.

3.3.1 Jackknife Replication

To permit the calculation of variance estimates from the survey data, a series of weights, referred to as jackknife replicate weights, are attached to each record in the data file, along with the corresponding final full sample weight. Calculation of the replicate weights first requires the construction of a set of subsamples of the full sample, referred to as "jackknife replicates." These replicates depend only on the selected PSUs, so they can be created immediately after the selection of PSUs. As described in Section 2.3, the PSUs were selected systematically from a list of PSUs that had been ordered by province, urban-rural status within province, lower administrative levels within urban-rural status, and ultimately by EA within the lowest administrative level. To account for the precision benefits of implicit stratification as fully as possible, the sampled PSUs within each province were paired off in the systematic order in which they were selected, treating each pair as a variance-estimation stratum. When there was an odd number of sampled PSUs in a province, one of the variance-estimation strata was defined to contain three sampled PSUs. To fully reflect the sample design, the formation of the variance-estimation strata was applied to all 404 of the sampled PSUs, including the PSU that was not completed during data collection (see Table 4).

For ZAMPHIA, 197 variance-estimation strata were created. A jackknife replicate was then formed by randomly deleting a PSU from a particular variance-estimation stratum *k*, for example, and retaining all the PSUs in the remaining variance-estimation strata. For a variance-estimation stratum consisting of a pair of PSUs, the weight of the retained PSU within the variance-estimation stratum *k* was doubled. For a variance-estimation stratum consisting of three PSUs, the weight of the two retained PSUs within the variance-estimation stratum was multiplied by 1.5. This process was repeated for all $r = 1, 2, ..., 197$ variance-estimation strata, resulting in a total of 197 jackknife replicates. Table 9 summarizes the number of jackknife replicates that were created for variance estimation.

Table 9. Number of PSUs and variance-estimation strata constructed for variance estimation

Province	Place of residence	No. sampling strata	Selected PSUs	No. variance strata consisting of pairs	No. variance strata consisting of triplets	No. jackknife replicates
Central	Rural	1	21	9	1	10
Central	Urban	1	9	3	1	4
Copperbelt	Rural	1	5	1	1	2
Copperbelt	Urban	1	37	17	1	18
Eastern	Rural	1	26	13	0	13
Eastern	Urban	1	14	7	0	7
Luapula	Rural	1	34	15	1	16
Luapula	Urban	1	8	4	0	4
Lusaka	Rural	1	4	2	0	2
Lusaka	Urban	1	36	18	0	18

Province	Place of residence	No. sampling strata	Selected PSUs	No. variance strata consisting of pairs	No. variance strata consisting of triplets	No. jackknife replicates
Muchinga	Rural	1	48	24	0	24
Muchinga	Urban	1	9	3	1	4
North Western	Rural	1	34	17	0	17
North Western	Urban	1	14	7	0	7
Northern	Rural	1	31	14	1	15
Northern	Urban	1	8	4	0	4
Southern	Rural	1	12	6	0	6
Southern	Urban	1	17	7	1	8
Western	Rural	1	32	16	0	16
Western	Urban	1	5	1	1	2
Zambia		20	404	188	9	197

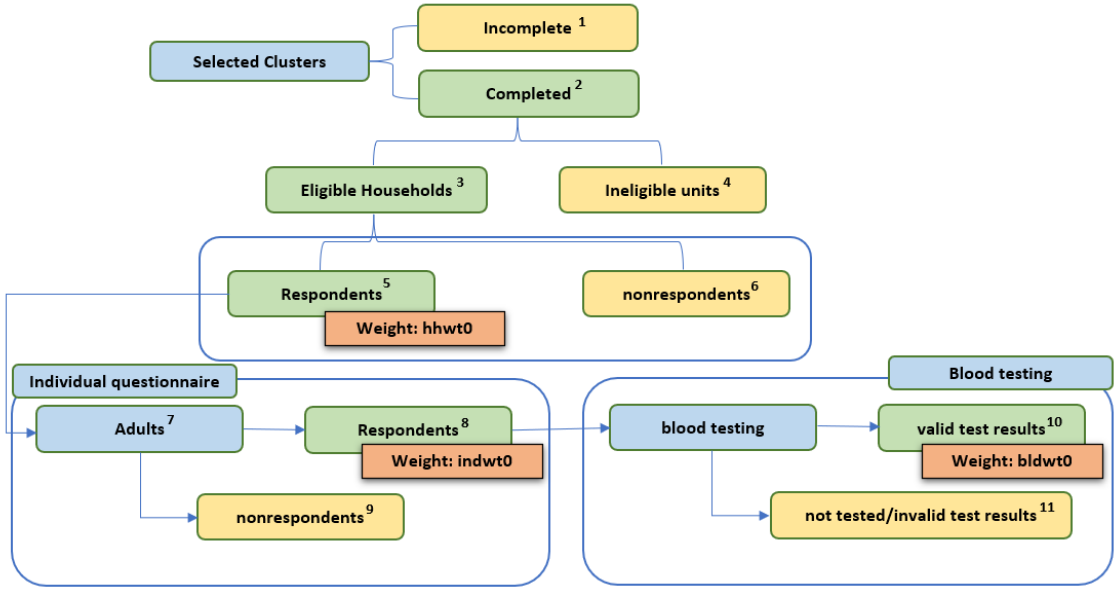
3.3.2 Taylor's Series

Even though jackknife replication is the recommended method for variance estimation, not all software packages have a replication option to produce variance estimates. For example, SPSS has built-in options for estimating variance using the Taylor's Series method, but the end user has to write a program in SPSS to produce replicate estimates of variance. Therefore, information for producing Taylor's Series estimates of variance is included in the ZAMPHIA data files. The full sample weight (see Section 3.5) is used as the weight to compute Taylor's Series variance estimates. The variable VarStrat indicates the 197 variance-estimation strata, and the variable VarUnit indicates the PSU or cluster within the variance-estimation stratum. This pair of variables allows analysts to produce variance estimates if their software does not easily accommodate replication methods but does have Taylor's Series capability. Note that the variance-estimation strata and the sampling strata are not equivalent; as shown in Table 9, the sampling strata are defined by provinces and place of residence, and the variance-estimation strata are based on groupings of PSUs within each sampling stratum.

3.4 Eligible and Completed Households and Individuals

For each weight, the calculation process starts by identifying the eligible units, households or individuals, and the completed cases, including completed questionnaires or valid blood measurements. Figure 1 presents a flowchart of the entire process, starting from the selected clusters and ending with the lab results. The flowchart highlights the eligible units or individuals for different questionnaires and blood testing, with details about survey variables used in identifying those units outlined in the endnotes that follow the flowchart. In addition, in each stage, respondents and nonrespondents are defined, such as completed and incomplete clusters, respondent and nonrespondent households, and respondent and nonrespondent individuals, for questionnaires and blood testing.

Figure 1. Flowchart of eligibility and completion of different questionnaires and blood testing



[1] Incomplete clusters: Clusters in which data collection was not executed for any reason. One cluster from Luapula-rural was dropped from the survey.

[2] Completed clusters: Clusters in which operations of household listing and data collection were successfully executed. A total of 403 clusters.

[3] Eligible households: All selected units found to be residential households that were not vacant, destroyed, or inhabited by a household that was absent for an extended period. They can be identified using the survey variables as follows: HRESULT in (1, 2,4,5,9:15,95, 96).

[4] Ineligible units: All selected units found to be vacant, destroyed, not a dwelling, or inhabited by a household that was absent for an extended period. They can be identified as HRESULT in (3, 6, 7, 8).

[5] Respondents: Households for which a household questionnaire was completed. They can be identified as HRESULT = 1.

[6] Nonrespondents: Households for which a household questionnaire was not completed. They can be identified as HRESULT in (2,4,5,9:15,95, 96).

[7] Adults: All de facto adults aged 15+ years. They can be identified as SLEEPHERE = 1 and AGE >= 15.

[8] Respondents: Eligible adults who completed the individual questionnaire. They can be identified as IRESULT = 1.

[9] Nonrespondents: Eligible adults who did not complete the individual questionnaire. They can be identified as IRESULT ≠ 1.

[10] Valid test results: Eligible adults who provided a blood draw for HIV testing that yielded a positive or negative result. They can be identified as IRESULT = 1 and FinalResult in (1,2).

[11] Nonrespondents: Eligible adults who did not provide a blood draw for HIV testing or who provided a blood draw that yielded an invalid result. They can be identified as IRESULT = 1 and (FinalResult ≠ 1 & FinalResult ≠ 2).

3.5 Development of Weights

3.5.1 Design Weight

The ZAMPHIA sample is a two-stage stratified cluster sample stratified by provinces and urban/rural, so the process of calculating the survey weights started by accounting for the different sampling probabilities, which were calculated separately for each sampling stage and each cluster. Let P_{1hi} be the

first stage's sampling probability of the i^{th} cluster in stratum *h*, P_{2hi} is the second stage's sampling probability of households within the i^{th} cluster, and P_{hi} is the overall sampling probability of any households of the i^{th} cluster in stratum *h*. The probability of selection of the PSU *i* in stratum *h* in the sample, using the probability proportional to size method, is calculated as follows:

$$P_{1hi} = \frac{a_h N_{hi}}{N_h}$$

where a_h denotes the number of clusters selected in stratum *h*, N_{hi} the number of households according to the sampling frame in the i^{th} cluster, and N_h the total number of households in stratum *h*. Now, let L_{hi} be the number of households listed during the household listing operation in cluster *i* in stratum *h,* and let g_{hi} be the number of households selected in the same cluster. The second stage's selection probability for each household in the cluster is calculated as follows:

$$P_{2hi} = \frac{g_{hi}}{L_{hi}}$$

The overall selection probability of each household in cluster *i* of stratum *h* is, therefore, the product of the two stages of selection probabilities:

$$P_{hi} = P_{1hi} \times P_{2hi}$$

Therefore, the design weight for each household in cluster *i* of stratum *h* is the inverse of its overall selection probability as follows:

$$d_{hi} = \frac{1}{P_{hi}}$$

d_{hi} was adjusted for cluster-level nonresponse to account for selected PSUs that were not completed. Let R_i identify completed PSU *i*, where $R_i = 1$ if PSU *i* was completed, and $R_i = 0$ if PSU *i* was not. Provinces by urban/rural were used to form the nonresponse adjustment cells, in which a separate adjustment factor was calculated for each adjustment cell *c* as follows:

$$A_c^{PSU} = \sum_{i=1}^{a_c} d_{hi} \Big/ \sum_{i=1}^{a_c} R_i d_{hi}$$

where a_c is the number of sampled PSUs in adjustment cell *c*. Weighted response rates are presented in Table 10. For the completed PSUs, the nonresponse adjusted design weight for PSU *i* in stratum *h* was then computed as follows:

$$D_{hi} = A_c^{PSU} \, d_{hi}$$

The design weight D_{hi} is the base for all the survey weights explained in the following sections.

As indicated in Table 9, 197 jackknife replicates were formed from the 404 sampled PSUs. For variance estimation, replicate-specific PSU weights, $D_{(r)hi}, r = 1,2, \ldots, 197$ were created to provide the basis for calculating the required replicate weights in subsequent stages of the weighting process. Let *h* denote one of the variance-estimation strata created for jackknife replication (Section 3.3.1), and let *i* denote the PSU within variance-estimation stratum *h*. For a given jackknife replicate, *r* = 1, 2, ..., 197, the corresponding replicate-specific PSU design weight was computed as follows:

$D_{(r)hi}$ = aD_{hi} if *h* = *r* and PSU *i* in variance-estimation stratum *h* is included in replicate *r*

= 0 if *h* = *r* and PSU *i* in variance-estimation stratum *h* is not included in replicate *r*

= D_{hi} if *h* ≠ *r*

where the coefficient *a* = 2 or 1.5, depending on whether the variance-estimation stratum consisted of two or three PSUs, respectively.

Table 10. Number of selected and dropped PSUs and weighted response rates by province and urban/rural

Province	Place of residence	Selected PSUs	Dropped PSUs	Weighted response rate ($1/A_c^{PSU}$)
Central	Rural	21	0	1
Central	Urban	9	0	1
Copperbelt	Rural	5	0	1
Copperbelt	Urban	37	0	1
Eastern	Rural	26	0	1
Eastern	Urban	14	0	1
Luapula	Rural	34	1	0.970
Luapula	Urban	8	0	1
Lusaka	Rural	4	0	1
Lusaka	Urban	36	0	1
Muchinga	Rural	48	0	1
Muchinga	Urban	9	0	1
North Western	Rural	34	0	1
North Western	Urban	14	0	1
Northern	Rural	31	0	1
Northern	Urban	8	0	1
Southern	Rural	12	0	1
Southern	Urban	17	0	1
Western	Rural	32	0	1
Western	Urban	5	0	1
Zambia		404	1	0.997

3.5.2 Household Survey Weights

The first step of calculating the household survey weight *hhwt0* was to adjust the design weight D_{hi} for household nonresponse to account for eligible nonrespondents and units with unknown eligibility. Each sampled household was assigned to one of the following household interview result codes:

1. Completed
2. No household member at home
3. Entire household absent for extended period of time
4. Postponed
5. Refused
6. Dwelling vacant
7. Address not a dwelling
8. Dwelling destroyed
9. Dwelling not found
10. Inaccessible due to flooding or security
11. Withdrawn
12. Stop survey
13. No competent head of household
14. At home, reschedule visit
15. Team in isolation/quarantine for COVID-19
95. Unsuccessful HH due to COVID19
96. Other

Table 11 shows the distribution of households by response status and maps the result codes to the following response status groups: eligible respondents, eligible nonrespondents, ineligible/out-of-scope, and unknown eligibility.

Table 11. Number of selected households by response groups

Household interview result	Number of households	Response status group
1. Completed	10,627	Eligible respondents
2. No household member at home	433	Eligible nonrespondents
3. Entire household absent for extended period of time	384	Ineligible/out-of-scope
4. Postponed	0	Eligible nonrespondents
5. Refused	370	Eligible nonrespondents
6. Dwelling vacant	234	Ineligible/out-of-scope
7. Address not a dwelling	15	Ineligible/out-of-scope
8. Dwelling destroyed	33	Ineligible/out-of-scope
9. Dwelling not found	12	Unknown eligibility

Household interview result	Number of households	Response status group
10. Inaccessible due to flooding or security	0	Eligible nonrespondents
11. Withdrawn	0	Eligible nonrespondents
12. Stop survey	58	Eligible nonrespondents
13. No competent head of household	65	Eligible nonrespondents
14. At home, reschedule visit	0	Eligible nonrespondents
15. Team in isolation/quarantine for COVID-19	3	Unknown eligibility
95. Unsuccessful HH due to COVID19	8	Unknown eligibility
96. Other	3	Unknown eligibility
Total	12,245	

The household nonresponse adjustment for eligible nonrespondents and units with unknown eligibility was implemented in one step. Each household *j* was assigned to one of the four eligible response status groups. R_{hij} identifies respondent household *j* in PSU *i* in stratum *h*, where $R_{hij} = 1$ if the household *j* was assigned to the first response group (1. Eligible respondents), and $R_{hij} = 0$ if the household *j* was assigned to the second or forth response status groups (2. Eligible nonrespondents or 4. Unknown eligibility). Households in the third response group are ineligible for the survey and therefore are not part of this adjustment or the weight calculations in general. PSUs were used to form the nonresponse adjustment cells, in which \ a separate adjustment factor was calculated for each adjustment cell *c* as follows:

$$A_c^{HH} = \sum_{j=1}^{n_c} D_{hij} \Big/ \sum_{j=1}^{n_c} R_{hij} D_{hij}$$

where n_c is the number of sampled households in adjustment cell *c,* not including the ineligible households. Adjustment factors of PSUs with weighted a response rate that was less than 50% were trimmed at 2. For the respondent households in response-status group 1, the nonresponse adjusted weight (*hhwt0*) for household *j* in PSU *i* in stratum *h* was then computed as follows:

$$W_{hij}^1 = A_c^{HH} D_{hij}$$

The corresponding replicate weights for replicate *r* = 1, 2, …, 197 (*hhwt001-hhwt197*) were computed in a similar fashion as follows:

$$W_{(r)hij}^1 = A_{(r)c}^{HH} D_{(r)hij}$$

where

$$A_{(r)c}^{HH} = \sum_{j=1}^{n_{(r)c}} D_{(r)hij} \Big/ \sum_{j=1}^{n_{(r)c}} R_{hij} D_{(r)hij}$$

3.5.3 Individual Weight (indwt0)

In completed households, all de facto adults aged 15+ years were eligible to complete the individual questionnaire. Table 12 presents the number of eligible adults by the response groups.

Table 12. Number of eligible adults by response groups

Interview result	Adults	
	n	%
Completed	22,262	87.36
Not at home	1,850	7.26
Refused	879	3.45
Incapacitated	190	0.75
Cognitively impaired (ineligible)	166	0.65
Stop survey	91	0.36
Team in isolation/quarantine for COVID-19	3	0.01
Unsuccessful individual due to COVID-19	5	0.02
Other	36	0.14
Not revisited during data collection	1	0.00
Total	25,483	100.00

The first step in the calculation of the individual weight was to use the weight (*hhwt0*) as the base weight (*ind_dsnwt0*) for the next steps—that is:

$$W^2_{hijk} = W^1_{hijk}$$

We then adjusted the base weight for individual nonresponse, where adjustment factors inflated the base weight to account for the nonrespondents. An adjustment factor A^I_k was calculated within each adjustment cell *c* as follows:

$$A^I_c = \sum_{k=1}^{m_c} W^2_{hijk} \Big/ \sum_{k=1}^{m_c} R^I_{hijk} W^2_{hijk}$$

where $R^I_{hijk} = 1$ if the individual *k* in household *j* was respondent, and $R^I_{hijk} = 0$ if the individual was a nonrespondent, and m_c is the number of eligible individuals in adjustment cell *c*. For the respondent individual, the nonresponse adjusted weight (*ind_nrwt0*) for individual *k* in household *j* in PSU *i* in stratum *h* was then computed as follows:

$$W^3_{hijk} = A^I_c \, W^2_{hijk}$$

The corresponding replicate weights for replicate *r* = 1, 2, …, 197 (*ind_nrwt001- ind_nrwt197*) were computed in a similar fashion as follows:

$$W^3_{(r)hijk} = A^I_{(r)c} \, W^2_{(r)hijk}$$

where

$$A^{I}_{(r)c} = \sum_{k=1}^{m_{(r)c}} W^{2}_{(r)hijk} \Big/ \sum_{k=1}^{m_{(r)c}} R^{I}_{hijk} W^{2}_{(r)hijk}$$

The nonresponse adjustment cells *c* were determined through a two-stage process in which 49 variables from the household questionnaire were used as covariates in a Least Absolute Shrinkage and Selection Operator (LASSO) model that was used to model the response to the individual questionnaire. One model was developed using the *HPGENSELECT* procedure (Johnston and Rodriguez, 2015) with selection *method=lasso* in SAS 9.4, with the weight set to the person base weight. Only the significant covariates of the response variable were identified and promoted for the next step. Out of 49 variables used in the initial model, the LASSO identified 43 variables to be significant predictors of response, as shown in Tables 13. Variables such as gender, age, province, and urban/rural were promoted for the next step, regardless of the results of the LASSO model.

In the next step, the promoted variables were inserted as inputs for a Chi-square Automatic Interaction Detector (CHAID) algorithm (Magidson, 2005) that identified the final nonresponse adjustment weighting cells. CHAID classifies the sampled individuals (i.e., the respondents and nonrespondents) into "cells" based on information available (promoted variables) for all sample persons, using a weighted log-linear modeling algorithm for the computation of chi-square statistics associated with each predictor, where the weight is the person base weight. The CHAID procedure produces a tree diagram that specifies the optimal number of final weighting cells and their definitions based on the input predictor variables.[1] A total of 157 adjustment cells were identified based on 33 variables. Adjustment cells with fewer than 30 respondents, or a weighted response rate of 50% or less (i.e., adjustment factor > 2), were combined with neighboring cells. This resulted in a reduction in the number of the adjustment cells to 155 adjustment cells. Table 13 presents the variables that were included in the LASSO and CHAID models. Tables 14 presents the number of variables used in the LASSO model and CHAID algorithm and the number of nonresponse adjustment cells. See Table A.1 in Appendix A for more details about the adjustment cells and factors.

[1] The depth limit of the tree was set to 5, and the minimum subgroup size required to allow splitting and minimum terminal node size were set to 50 observations (both respondents and nonrespondents).

Table 13. Variables considered for the LASSO models and selected by LASSO and CHAID

ID	Variable name	Description	LASSO	CHAID
enforced*	HPROVINCE	Province	1	1
	HRESIDENCE	Urban/rural	2	2
	HSEX	Gender of individual	3	3
	AGEC	Age of individual	4	4
1	HVISITS	Number of household visits	5	5
2	HLANGQ	Language of questionnaire	6	6
3	HLANGI	Language of interview	7	7
4	HLANGR	Language of respondent	8	
5	HMEMBER	Total household members	9	8
6	HWOMEN	Total eligible women in household aged 15 years and older	10	9
7	HMEN	Total eligible men in household aged 15 years and older	11	10
8	WATERSOURCEOT	Household main source of drinking water	12	11
9	TOILETTYPE	Household usual type of toilet facility	13	12
10	TOILETSHARE	Toilet shared with other households	14	13
11	HHQITEMSA	Electricity	15	14
12	HHQITEMSB	Radio		
13	HHQITEMSC	Television		
14	HHQITEMSD	Fixed phone	16	
15	HHQITEMSE	Refrigerator	17	
16	HHQITEMSF	Internet		
17	HHQITEMSG	Computer	18	15
18	HHQITEMSH	Bed		
19	HHQITEMSI	Table	19	16
20	HHQITEMSJ	Sofa	20	17
21	HHQITEMSK	Hammer mill	21	
22	HHQITEMSL	Microwave	22	18
23	COOKINGFUEL	Main household cooking fuel	23	19
24	MATFLOOROT	Main flooring material of dwelling	24	20
25	MATROOF	Main roofing material of dwelling	25	
26	MATEXWALLS	Main wall material of dwelling	26	21
27	ROOMSLEEP	Number of rooms used for sleeping	27	22
28	HHQOWNA	Bicycle	28	23
29	HHQOWNB	Motorcycle or motor scooter	29	
30	HHQOWNC	Car or truck	30	
31	HHQOWND	Boat with a motor	31	
32	HHQOWNE	Canoe	32	24
33	HHQOWNF	Donkey cart	33	
34	HHQOWNG	Tractor		
35	HHQOWNH	Plough	34	25
36	HHQOWNI	Grain-grinder	35	26
37	HHQOWNJ	Feature (mobile) phone	36	27
38	HHQOWNK	Smart phone	37	28
39	OWNCOWNUM	Cows	38	29
40	OWNGOATNUM	Goats or sheep	39	30
41	OWNCHIKNNUM	Chickens	40	31
42	OWNDOGNUM	Dogs	41	
43	OWNHORSNUM	Horses/donkeys		
44	H0120A	Any agricultural land owned	42	32
45	ECONSUP1_A	No external economic support	43	33

* Variables that are enforced in the CHAID procedures regardless of the LASSO model results

Table 14. Number of variables used in LASSO/CHAID and number of nonresponse adjustment cells

Age group	Number of				
	Variables inserted in LASSO	Variables inserted in CHAID	Variables used by CHAID	Adjustment cells	Adjustment cells after collapsing*
Adults 15+ years	49	43**	33	157	155

* Adjustment cells with fewer than 30 respondents, or a weighted response rate of 50% or less (i.e., adjustment factor > 2), were combined with neighboring cells, and cells with 100% response rates were collapsed together.
** Including variables that were enforced in the CHAID algorithm, such as province, urban/rural, gender, and age

To reduce the variability of the weights, which can lead to inflated sampling variances, the nonresponse adjusted weights W_{hijk}^{3} were trimmed where outliers (identified as greater than 3.5 times the median of the W_{hijk}^{3} within the corresponding province and age group) were capped at 3.5 times the median weight, yielding the trimmed nonresponse adjusted weight W_{hijk}^{4} (*ind_trmnrwt0*) (Valliant, Dever, and Kreuter, 2013). The corresponding replicate weights for replicate *r* = 1, 2, …, 197 (*ind_trmnrwt001- ind_trmnrwt197*) were computed in a similar fashion. Finally, the trimmed weights were calibrated to the 2021 population projections of individuals (aged 15+ years) by gender and 14 age groups, yielding the final individual interview weight (*indwt0*) as follows:

$$W_{hijk}^{5} = W_{hijk}^{4} \frac{M_c}{\sum_{k=1}^{m_c} R_{hijk}^{I} W_{hijk}^{4}}$$

where M_c is the 2021 projected population total in calibration cell *c*. Table 15 presents population projections, weighted totals and calibration factors by age and gender. The corresponding replicate weights for replicate *r* = 1, 2, …, 197 (*indwt001- indwt197*) were computed in a similar fashion as follows:

$$W_{(r)hijk}^{5} = W_{(r)hijk}^{4} \frac{M_c}{\sum_{k=1}^{m_{(r)c}} R_{hijk}^{I} W_{(r)hijk}^{4}}$$

Table 15. Projected and weighted totals of individual aged 15+ years and calibration factors by gender and age

Age	Gender	2021 population projections	Weighted totals	Calibration factors
15–19 years	Male	988,683	872,131.30	1.13
15–19 years	Female	983,185	948,266.00	1.04
20–24 years	Male	851,902	672,988.20	1.27
20–24 years	Female	866,159	868,170.60	1.00
25–29 years	Male	721,111	571,405.40	1.26
25–29 years	Female	749,357	746,130.30	1.00
30–34 years	Male	521,639	452,863.80	1.15
30–34 years	Female	599,104	565,560.00	1.06
35–39 years	Male	437,151	420,868.50	1.04
35–39 years	Female	507,649	526,015.60	0.97

Age	Gender	2021 population projections	Weighted totals	Calibration factors
40–44 years	Male	369,913	349,265.40	1.06
40–44 years	Female	381,493	408,529.40	0.93
45–49 years	Male	308,198	306,133.30	1.01
45–49 years	Female	295,741	342,711.40	0.86
50–54 years	Male	218,440	243,100.10	0.90
50–54 years	Female	205,053	247,637.60	0.83
55–59 years	Male	158,674	165,138.20	0.96
55–59 years	Female	163,770	211,854.40	0.77
60–64 years	Male	114,736	122,757.20	0.93
60–64 years	Female	129,462	148,035.50	0.87
65–69 years	Male	75,960	87,434.37	0.87
65–69 years	Female	83,432	92,302.27	0.90
70–74 years	Male	55,064	75,699.90	0.73
70–74 years	Female	67,824	74,282.03	0.91
75–79 years	Male	36,251	41,161.43	0.88
75–79 years	Female	46,739	56,890.80	0.82
80+ years	Male	51,308	47,622.33	1.08
80+ years	Female	61,193	32,179.63	1.90

3.5.4 Blood Draw Weight (bldwt)

All adults who completed the individual questionnaire were eligible for blood testing. Table 16 presents the number of adults aged 15+ years eligible for blood draws distributed over the response groups. Adults with a blood draw that resulted in a valid HIV test (positive or negative HIV status) were coded as completed. If eligible for the blood draw, adults who did not consent for the blood draw or cases with invalid HIV tests were coded as not completed.

Table 16. Number of eligible adults for blood testing by response groups

Blood test result	Adults[1]	
	n	%
Completed with valid results	18,804	84.47%
Not completed	3,458	15.53%
Total	22,262	100.00%

[1] Adults who completed the individual questionnaire

The first step in the process of calculating the blood weights was to adjust the individual survey weight for blood draw nonresponse, yielding *bld_nrwt0*, as follows:

$$W^6_{hijk} = A^B_c W^5_{hijk}$$

where A^B_k was the adjustment factor for the blood nonresponse, calculated within each adjustment cell *c* as follows:

$$A^B_c = \sum_{k=1}^{m_c} W^5_{hijk} \Big/ \sum_{k=1}^{m_c} R^B_{hijk} W^5_{hijk}$$

where $R^B_{hijk} = 1$ if the individual *k* in household *j* provided a valid blood sample, and $R^B_{hijk} = 0$ otherwise, and m_c is the number of eligible individuals in adjustment cell *c*. The corresponding replicate weights for replicate *r* = 1, 2, …, 197 (*bld_nrwt001-bld_nrwt197*) were computed in a similar fashion as follows:

$$W^6_{(r)hijk} = A^B_{(r)c} W^5_{(r)hijk}$$

where

$$A^B_{(r)c} = \sum_{k=1}^{m_{(r)c}} W^5_{(r)hijk} \Big/ \sum_{k=1}^{m_{(r)c}} R^B_{hijk} W^5_{(r)hijk}$$

Determining the adjustment cells through the LASSO/CHAID approach was done separately for males and females, with variables from the individual questionnaire added to variables from the household questionnaire and used in the LASSO model. Separate models were developed for males and females because different questions were asked. Table 17 presents the variables inserted and selected by LASSO and CHAID models. Table 18 presents the number of variables in each stage and the number of adjustment cells. See Tables A.2 and A.3 in Appendix A for more details about the adjustment cells and factors.

Table 17. Variables considered for the LASSO models and selected by LASSO and CHAID by gender

ID	Variable name	Description	Male adults		Female adults	
			LASSO	CHAID	LASSO	CHAID
enforced*	HPROVINCE	Province	1	1	1	1
	HRESIDENCE	Urban/rural	2	2	2	2
	AGEC	Age of individual	3	3	3	
1	HVISITS	Number of household visits	4	4	4	3
2	HLANGQ	Language of questionnaire	5			
3	HLANGI	Language of interview			5	4
4	HLANGR	Language of respondent	6		6	5
5	HMEMBER	Total household members	7	5	7	6
6	HWOMEN	Total eligible women in household aged 15 and older	8	6	8	7
7	HMEN	Total eligible men in household aged 15 and older	9	7	9	8
8	WATERSOURCEOT	Household main source of drinking water	10	8		
9	TOILETTYPE	Household usual type of toilet facility	11	9	10	
10	TOILETSHARE	Toilet shared with other households	12		11	9
11	HHQITEMSA	Electricity	13	10	12	10
12	HHQITEMSB	Radio	14			
13	HHQITEMSC	Television	15	11	13	
14	HHQITEMSD	Fixed phone	16			
15	HHQITEMSE	Refrigerator				
16	HHQITEMSF	Internet	17		14	11

ID	Variable name	Description	Male adults		Female adults	
			LASSO	CHAID	LASSO	CHAID
17	HHQITEMSG	Computer				
18	HHQITEMSH	Bed			15	12
19	HHQITEMSI	Table	18	12	16	13
20	HHQITEMSJ	Sofa	19		17	
21	HHQITEMSK	Hammer mill			18	
22	HHQITEMSL	Microwave				
23	COOKINGFUEL	Main household cooking fuel	20		19	
24	MATFLOOROT	Main flooring material of dwelling	21	13	20	14
25	MATROOF	Main roofing material of dwelling				
26	MATEXWALLS	Main wall material of dwelling				
27	ROOMSLEEP	Number of rooms used for sleeping	22	14		
28	HHQOWNA	Bicycle	23	15	21	15
29	HHQOWNB	Motorcycle or motor scooter	24			
30	HHQOWNC	Car or truck	25			
31	HHQOWND	Boat with a motor				
32	HHQOWNE	Canoe	26		22	
33	HHQOWNF	Donkey cart				
34	HHQOWNG	Tractor				
35	HHQOWNH	Plough				
36	HHQOWNI	Grain-grinder	27	16		
37	HHQOWNJ	Feature (mobile) phone	28		23	16
38	HHQOWNK	Smart phone	29	17	24	17
39	OWNCOWNUM	Cows	30	18		
40	OWNGOATNUM	Goats or sheep			25	
41	OWNCHIKNNUM	Chickens	31	19	26	18
42	OWNDOGNUM	Dogs			27	19
43	OWNHORSNUM	Horses/donkeys				
44	H0120A	Any agricultural land owned	32		28	
45	ECONSUP1_A	No external economic support	33	20		
46	IVISITS	Number of individual visits	34	21	29	20
47	LNGNAT_LNG	Language of respondent				
48	LNGVINT_LNG	Language of interview			30	21
49	LNGVQX_LNG	Language of questionnaire	35	22	31	
50	SCHLAT	Attended school			32	
51	SCHCOM_A	Highest level of school attended	36			
52	WORK12MO	Have you worked in past 12 months			33	22
53	WORK7DAYS	Have you worked in last 7 days				
54	NORMWORK	Location of work place	37	23		
55	EVERMAR	Ever been married or lived with someone as if married				
56	AGEMAR	Age at the time of first marriage or started living with someone as if married				
57	CURMAR	Marital status				
58	NUMWIF[M]	Number of wives or live-in partners in the household				
59	WIFELIVEEW[M]	Number of wives or partners who live elsewhere				
60	HUSLIVEW[F]	Is the husband or partner living with you or elsewhere			34	23
61	HUSOTWIF[F]	Husband or partner have other wives or other women living as if married				
62	LIVEB[F]	Number of pregnancies resulting in live births			35	24
63	CHILDA2017[F]	Number of pregnancies past 3 years				
64	PRGCARE[F]	Visited facility for ANC at last pregnancy				

ID	Variable name	Description	Male adults		Female adults	
			LASSO	CHAID	LASSO	CHAID
65	HIVTSBP[F]	Tested for HIV before pregnancy			36	25
66	HIVPSBP[F]	Test result positive				
67	HIVTRPRG[F]	Tested for HIV during pregnancy				
68	HIVRTPG[F]	Test result during last pregnancy				
69	HIVTSAD[F]	Tested for HIV after delivery of last child			37	
70	PREGNANT[F]	Are you pregnant now				
71	AVOIDPREG	Are you using family planning method	38			
72	CMETHOD_A	Method being used: Female sterilization				
73	CMETHOD_C	Method being used: Pill	39	24		
74	CMETHOD_E	Method being used: Injections	40		38	
75	CMETHOD_F	Method being used: Implant			39	26
76	CMETHOD_G	Method being used: Condom	41			
77	MCSTATUS[M]	Are you circumcised				
78	MCPLANS[M]	Planning on circumcision in next 6 months	42			
79	MCWHOTRAD[M]	Circumcised by traditional practitioner				
80	MCAGE[M]	Age at circumcision				
81	MCWHOMED[M]	Circumcised by medical provider	43	25		
82	FIRSTSXAGE	Age at which had sex first time	44	26	40	27
83	LIFETIMESEX	Number of sexual partners in lifetime	45	27	41	28
84	PAST12MONUM	Number of sexual partners in last 12 months	46	28	42	
85	HFLAST12MO	Seen a doctor, clinical officer or nurse in a health facility in the last 12 months				
86	HFHIVTSTOFFER	Offered HIV test by doctor, clinical officer or nurse during the last 12 months of visit to health facility	47	29	43	29
87	HIVTSTEVER	Ever tested for HIV				
88	HIVTSTRSLT	Result for last HIV test	48	30	44	30
89	HIVPOSPROV	Health care provider ever told you that you have HIV			45	31
90	HIVSELFTST	Ever self-tested for HIV using a self-test kit				
91	DISCLOSEA	Who have you told that you are HIV positive (Spouse or partner)				
92	DISCLOSEB	Who have you told that you are HIV positive (Health care provider)				
93	DISCLOSEC	Who have you told that you are HIV positive (Friend)				
94	DISCLOSED	Who have you told that you are HIV positive (Family member)				
95	DISCLOSEX	Who have you told that you are HIV positive (Other)				
96	PRPEVRHDR	Ever heard of PrEP before now				
97	PREPEVER	Ever taken PrEP				
98	PREPCURNT	Currently taking PrEP				
99	PREPWDTK	Would take PrEP to help prevent HIV	49	31	46	32
100	HIVCARE	Received HIV care from doctor				
101	HIVCRLTC	Currently receiving HIV care from facility				
102	CLINCHANGE	Changed clinic of care in past year				
103	TRAVELDIFF	Does travel time cause difficulty to access care				
104	ARVSTAKENEV	Ever taken ARVs				
105	ARVSCURRENT	Are you currently taking ARVs				
106	I0714A	Any period on obtaining ARV with difficult in COVID time	50	32		
107	ARVSWITCH	Ever changed or modified ARVs				
108	VLTEST	Ever taken viral load test				

ID	Variable name	Description	Male adults		Female adults	
			LASSO	CHAID	LASSO	CHAID
109	VLTESTRESULT	Received viral load test result				
110	VLTESTRESULTA	Result for viral load				
111	TBSYMPASSESSA	TB symptoms: cough				
112	TBSYMPASSESSB	TB symptoms: Fever				
113	TBSYMPASSESSC	TB symptoms: Night sweat			47	
114	TBSYMPASSESSD	TB symptoms: Weight loss				
115	TBSYMP12MO	Ever taken medicine to prevent TB			48	33
116	MEDINHCURR	Currently taking TPT				
117	TBCLINVISIT	Visited clinic for TB treatment				
118	TBCLINHIVTST	Tested for HIV during TB visit			49	
119	TBDIAGN	Told by doctor about TB in last 12 months	51			
120	CERVCNTST[F]	Tested for cervical cancer				
121	HPVVACC[F]	Ever been vaccinated for cervical cancer			50	34
122	LITTLEINTEREST	Frequency of having no interest	52	33	51	35
123	DEPRESSED	Frequency of depression	53			
124	ANXIETY	Frequency of anxiety				
125	WORRY	Frequency of worrying				
126	CHRONICCONDA	Chronic condition: Diabetes	54			
127	CHRONICMEDA	Diabetes medication				
128	CHRONICCONDB	Chronic condition: Hypertension	55			
129	CHRONICMEDB	Hypertension medication			52	
130	CHRONICCONDC	Chronic condition: Heart disease	56			
131	CHRONICMEDC	Heart disease medication				
132	ALCFREQ	How often do you have a drink containing alcohol	57	34		
133	ALCNUMDAY	Number of drinks containing alcohol consumed on a typical day			53	
134	ALCSIXMORE	How often do you have six or more drinks on one occasion	58			
135	CONDOMWHERE_A	Where can you get condoms: Clinic/hospital				
136	CONDOMWHERE_B	Where can you get condoms: Kiosk/shop	59			
137	CONDOMWHERE_C	Where can you get condoms: Pharmacy	60			
138	CONDOMWHERE_D	Where can you get condoms: Local free dispenser				
139	CONDOMWHERE_E	Where can you get condoms: Friends/peer	61	35		
140	CONDOMWHERE_Y	Where can you get condoms: Don't know				
141	CONDOMGET	Easy to get condoms	62			
142	CONDOMNOTEASYRSN_A	Reason why it's difficult to get a condom: Condoms not available/too far				
143	CONDOMNOTEASYRSN_B	Reason why it's difficult to get a condom: Not convenient				
144	CONDOMNOTEASYRSN_C	Reason why it's difficult to get a condom: Costs too much				
145	CONDOMNOTEASYRSN_D	Reason why it's difficult to get a condom: Embarrassed to get condoms				
146	CONDOMNOTEASYRSN_E	Reason why it's difficult to get a condom: Do not want others to know				
147	CONDOMNOTEASYRSN_F	Reason why it's difficult to get a condom: Do not know where to get condoms				
148	CONDOMNOTEASYRSN_X	Reason why it's difficult to get a condom: Other				
149	CONDOMNOTEASYRSN_Y	Reason why it's difficult to get a condom: Don't know				
150	ADTPSX	Ever talked about sex with a parent or guardian			54	36
151	ADDISHIV	Ever discussed HIV with a parent or guardian				

ID	Variable name	Description	Male adults		Female adults	
			LASSO	CHAID	LASSO	CHAID
152	ADHIVPREVA	Ever taken part in any prevention or treatment of Safe Spaces (Stepping Stones Curriculum)	63	36		
153	ADHIVPREVB	Ever taken part in any prevention or treatment Condom				
154	ADHIVPREVC	Ever taken part in any prevention or treatment HIV Testing	64			
155	ADHIVPREVD	Ever taken part in any prevention or treatment Family planning				
156	ADHIVPREVE	Ever taken part in any prevention or treatment Families Matter Program (parent/caregiver program)				
157	ADHIVPREVF	Ever taken part in any prevention or treatment PrEP				
158	ADHIVPREVG	Ever taken part in any prevention or treatment Educational subsidies				
159	ADHIVPREVH	Ever taken part in any prevention or treatment Enrollment in savings group				
160	ADHIVPREVX	Ever taken part in any prevention or treatment (Other)				
161	ADHIVSCHMTG	Number of times participated in a school meeting or class period about HIV/AIDS in the past 12 months	65		55	37

* Variables that are enforced in the CHAID procedures regardless of the LASSO model results
[F] Variables relevant for women only
[M] Variables relevant for men only

Table 18. Number of variables used in LASSO/CHAID and number of nonresponse adjustment cells

Age group	Gender	Number of				
		Variables inserted in LASSO	Variables inserted in CHAID	Variables used by CHAID	Adjustment cells	Adjustment cells after collapsing*
Adults aged 15+ years	Male	148	65**	36	78	74
Adults aged 15+ years	Female	154	55**	37	92	87

* Adjustment cells with fewer than 30 respondents, or a weighted response rate of 50% or less (i.e., adjustment factor > 2), were combined with neighboring cells, and cells with 100% response rates were collapsed together.
** Including variables that were enforced in the CHAID algorithm, such as province, urban/rural, and age

The nonresponse adjusted weights W_{hijk}^{6} were then trimmed where outliers (identified as greater than 3.5 times the median of the W_{hijk}^{6} within the corresponding province) were capped at 3.5 times the median weight, yielding the trimmed nonresponse adjusted weight W_{hijk}^{7}, *bld_trmnrwt0*. The corresponding replicate weights for replicate *r* = 1, 2, …, 197 (*bld_trmnrwt001- bld_trmnrwt197*) were computed in a similar fashion.

The trimmed weights were then calibrated to the 2021 population projections of individuals (aged 15+ years) by gender and 14 age groups, yielding the final biomarker weight (*bldwt0*) as follows:

$$W_{hijk}^{9} = W_{hijk}^{8} \frac{M_c}{\sum_{k=1}^{m_c} R_{hijk}^{B} W_{hijk}^{7}}$$

where M_c is the 2021 projected population total in calibration cell *c*. Table 19 presents population projections, weighted totals, and calibration factors by age and gender. The corresponding replicate weights for replicate *r* = 1, 2, …, 197 (*bldwt001- bldwt197*) were computed in a similar fashion as follows:

$$W^9_{(r)hijk} = W^8_{(r)hijk} \frac{M_c}{\sum_{k=1}^{m_{(r)c}} R^B_{hijk} W^7_{(r)hijk}}$$

Table 19. Projected and weighted totals of individuals aged 15+ years and calibration factors by gender and age

Age	Gender	2021 population projections	Weighted totals	Calibration factors
15–19 years	Male	988,683	868,682.10	1.14
15–19 years	Female	983,185	950,839.00	1.03
20–24 years	Male	851,902	674,892.90	1.26
20–24 years	Female	866,159	871,542.00	0.99
25–29 years	Male	721,111	571,465.00	1.26
25–29 years	Female	749,357	746,314.00	1.00
30–34 years	Male	521,639	434,788.20	1.20
30–34 years	Female	599,104	565,031.10	1.06
35–39 years	Male	437,151	414,556.00	1.05
35–39 years	Female	507,649	530,909.20	0.96
40–44 years	Male	369,913	353,324.60	1.05
40–44 years	Female	381,493	403,111.20	0.95
45–49 years	Male	308,198	316,819.90	0.97
45–49 years	Female	295,741	341,552.10	0.87
50–54 years	Male	218,440	248,644.10	0.88
50–54 years	Female	205,053	251,161.10	0.82
55–59 years	Male	158,674	179,099.90	0.89
55–59 years	Female	163,770	218,811.70	0.75
60–64 years	Male	114,736	121,873.50	0.94
60–64 years	Female	129,462	152,997.90	0.85
65–69 years	Male	75,960	87,103.76	0.87
65–69 years	Female	83,432	91,047.47	0.92
70–74 years	Male	55,064	70,829.48	0.78
70–74 years	Female	67,824	66,039.57	1.03
75–79 years	Male	36,251	41,882.76	0.87
75–79 years	Female	46,739	50,962.32	0.92
80 + years	Male	51,308	44,607.37	1.15
80 + years	Female	61,193	28,246.75	2.17

3.6 Survey Weight Variables in ZAMPHIA Datasets

The ZAMPHIA datasets include a set of survey weight variables to enable weighted analyses for the survey data. The final survey weights are provided in each dataset and labeled accordingly (Table 20 provides the variable names of survey weights for each level of analysis). Availability of survey

estimation procedures varies by statistical software. Users will need to use the appropriate weights for the specific analysis of interest, which is generally determined by the target population of inference.

- Household weights (*hhwt0*) can be used for analyses conducted at the household level, for example, distribution of households by urban/rural residence. Household weights can be interpreted as the number of households that the participating household represents in the population, accounting for sampling and nonresponse at the EA and household levels.
- Interview weights (*intwt0*) can be used for analyses conducted at the individual level for data collected for all potentially eligible interview participants. For example, self-reported HIV testing (i.e., ever received an HIV test prior to the survey) should be estimated using interview weights because all interview respondents received HIV testing questions. In this scenario, interview weights can be interpreted as the number of individuals that the respondent represents in the population who could have participated in the interview, accounting for sampling and nonresponse at the EA, household, and individual levels.
- Blood weights (*btwt0*) can be used for analyses conducted among blood test participants only. For example, HIV prevalence should be estimated using blood test weights even if the analysis includes predictors at the household or individual level, because not all interview respondents participated in blood tests. In this scenario, each participant's blood weight can be interpreted as the number of individuals that the participant represents in the population who could have participated in blood testing, accounting for selection and nonresponse at the EA, household, individual, and blood testing levels. In addition, if the outcome of interest comes from the interview (e.g., HIV testing history), but the analysis is restricted to those who have blood test results, blood test weights should be used.

Lastly, users interested in accessing the intermediary weights used for sample selection at each stage and for nonresponse and post-stratification adjustment will find these variables in the ZAMPHIA survey's intermediary weights datasets listed in Table 21.

Table 20. Survey weights and replicate weights: ZAMPHIA

Level	Variable name	
	Survey weight	**JK replicate weights**
Household	hhwt0	hhwt001-hhwt197
Individual interview	intwt0	intwt001-intwt197
Blood test	btwt0	btwt001-btwt197

Table 21. Survey Intermediary weights and replicate weights: ZAMPHIA

Level	Variable name	
	Survey intermediary weight	**JK replicate weights**
Household	hhbwt0	hhbwt001-hhbwt197
Individual interview	ind_dsnwt0	ind_dsnwt001-ind_dsnwt197
	ind_nrwt0	ind_nrwt001-ind_nrwt197
	ind_trmnrwt0	ind_trmnrwt001-ind_trmnrwt197
Blood test	bld_nrwt0	bld_nrwt001-bld_nrwt197
	bld_trmnrwt0	bld_trmnrwt001-bld_trmnrwt197

References

Johnston, G., & Rodriguez, R. (2015). Introducing the HPGENSELECT Procedure: Model Selection for Generalized Linear Models and More. Paper SAS1742-2015. Cary, NC: SAS Institute Inc. https://support.sas.com/resources/papers/proceedings15/SAS1742-2015.pdf

Kish, L. (1965). Survey Sampling. New York, NY: John Wiley & Sons.

Magidson, J. (2005) SI-CHAID Users Guide. Belmont, MA: Statistical Innovations Inc. https://www.statisticalinnovations.com/wp-content/uploads/SICHAIDusersguide.pdf

Ministry of Health, Zambia. (2019). Zambia Population-based HIV Impact Assessment (ZAMPHIA) 2016: Final Report. Lusaka, Zambia: Ministry of Health.

Valliant, R., Dever, J., & Kreuter, F. (2013). Practical Tools for Designing and Weighting Survey Samples. New York, NY: Springer.

Appendix A: Reference Tables

Table A.1. Interview weights: Nonresponse adjustment cells for adults aged 15+ years

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjustment factor	Final adjustment cell
HPROVINCE	FUEL				1	96	45,795	62,362	0.73	1.36	1
HPROVINCE	FUEL	HSEX	HHMEM	AGEC	2	48	30,499	34,069	0.90	1.12	2
HPROVINCE	FUEL	HSEX	HHMEM	AGEC	3	67	39,006	39,006	1.00	1.00	3
HPROVINCE	FUEL	HSEX	HHMEM	AGEC	4	56	30,330	34,178	0.89	1.13	4
HPROVINCE	FUEL	HSEX	HHMEM	AGEC	5	67	29,167	31,029	0.94	1.06	5
HPROVINCE	FUEL	HSEX	HHMEM	HLANGI	6	136	79,488	104,972	0.76	1.32	6
HPROVINCE	FUEL	HSEX	HHMEM	HLANGI	7	120	53,397	58,500	0.91	1.10	7
HPROVINCE	FUEL	HSEX	HHMEM	MNELIG	8	275	126,063	150,382	0.84	1.19	8
HPROVINCE	FUEL	HSEX	HHMEM	MNELIG	9	197	91,823	100,269	0.92	1.09	9
HPROVINCE	FUEL	HSEX	AGEC	SHARETOI	10	131	62,676	75,694	0.83	1.21	10
HPROVINCE	FUEL	HSEX	AGEC	SHARETOI	11	53	28,879	29,632	0.97	1.03	11
HPROVINCE	FUEL	HSEX	AGEC	GOAT	12	250	112,429	121,644	0.92	1.08	12
HPROVINCE	FUEL	HSEX	AGEC	GOAT	13	672	315,031	328,599	0.96	1.04	13
HPROVINCE	AGEC	SHARETOI	MICROWV	WMELIG	14	80	47,021	62,140	0.76	1.32	14
HPROVINCE	AGEC	SHARETOI	MICROWV	WMELIG	15	79	45,135	76,719	0.59	1.70	15
HPROVINCE	AGEC	SHARETOI	MICROWV	WMELIG	16	101	51,561	63,721	0.81	1.24	16
HPROVINCE	AGEC	SHARETOI	MICROWV		17	47	20,271	22,948	0.88	1.13	17
HPROVINCE	AGEC	SHARETOI	SMARTPH		18	44	31,583	39,568	0.80	1.25	18
HPROVINCE	AGEC	SHARETOI	SMARTPH		19	63	46,564	47,293	0.98	1.02	19
HPROVINCE	AGEC	HSEX	HHMEM		20	61	39,178	39,178	1.00	1.00	20
HPROVINCE	AGEC	HSEX	HHMEM		21	84	50,896	56,207	0.91	1.10	21
HPROVINCE	AGEC	HSEX	HHMEM	CHICKEN	22	171	124,408	135,648	0.92	1.09	22
HPROVINCE	AGEC	HSEX	HHMEM	CHICKEN	23	506	280,578	356,446	0.79	1.27	23
HPROVINCE	AGEC	HSEX	FLOOR	HRESIDENCE	24	72	58,741	58,741	1.00	1.00	24
HPROVINCE	AGEC	HSEX	FLOOR	HRESIDENCE	25	54	37,394	38,898	0.96	1.04	25
HPROVINCE	AGEC	HSEX	FLOOR	ECONSUP1_A	26	865	500,042	533,133	0.94	1.07	26
HPROVINCE	AGEC	HSEX	FLOOR	ECONSUP1_A	27	206	122,032	142,674	0.86	1.17	27
HPROVINCE	AGEC				28	61	36,039	36,381	0.99	1.01	28
HPROVINCE	AGEC				29	82	61,488	69,222	0.89	1.13	29
HPROVINCE	AGEC	GOAT			30	65	32,457	62,263	0.52	1.92	30
HPROVINCE	AGEC	GOAT	ECONSUP1_A	MOBILE	31	36	17,079	35,257	0.48	2.06	31
HPROVINCE	AGEC	GOAT	ECONSUP1_A	MOBILE	32	148	60,541	89,918	0.67	1.49	31
HPROVINCE	AGEC	GOAT	ECONSUP1_A		33	70	23,522	30,842	0.76	1.31	32
HPROVINCE	AGEC	HSEX	FLOOR		34	152	73,999	90,027	0.82	1.22	33
HPROVINCE	AGEC	HSEX	FLOOR	BICYC	35	56	13,532	22,486	0.60	1.66	34

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjustment factor	Final adjustment cell
HPROVINCE	AGEC	HSEX	FLOOR	BICYC	36	72	27,562	34,215	0.81	1.24	35
HPROVINCE	AGEC	HSEX	WMELIG		37	241	105,429	114,027	0.92	1.08	36
HPROVINCE	AGEC	HSEX	WMELIG	CHICKEN	38	75	30,192	34,723	0.87	1.15	37
HPROVINCE	AGEC	HSEX	WMELIG	CHICKEN	39	54	15746.35947	21709.11205	0.73	1.38	38
HPROVINCE	AGEC	WMELIG			40	74	29415.84984	30154.32267	0.98	1.03	39
HPROVINCE	AGEC	WMELIG	HSEX	HRESIDENCE	41	335	173740.7185	206146.8188	0.84	1.19	40
HPROVINCE	AGEC	WMELIG	HSEX	HRESIDENCE	42	87	17177.92447	24268.57998	0.71	1.41	41
HPROVINCE	AGEC	WMELIG	HSEX	SHARETOI	43	351	146351.2006	165142.1546	0.89	1.13	42
HPROVINCE	AGEC	WMELIG	HSEX	SHARETOI	44	197	89,179	94,114	0.95	1.06	43
HPROVINCE	AGEC	WMELIG	COWS		45	88	37,408	38,859	0.96	1.04	44
HPROVINCE	AGEC	WMELIG	COWS		46	90	28,466	32,428	0.88	1.14	45
HPROVINCE	WMELIG	AGEC			47	133	44,399	55,118	0.81	1.24	46
HPROVINCE	WMELIG	AGEC	ROOMS	MOBILE	48	116	40,526	41,765	0.97	1.03	47
HPROVINCE	WMELIG	AGEC	ROOMS	MOBILE	49	72	24,739	24,739	1.00	1.00	48
HPROVINCE	WMELIG	AGEC	ROOMS	HSEX	50	176	60,831	68,763	0.88	1.13	49
HPROVINCE	WMELIG	AGEC	ROOMS	HSEX	51	209	71,529	74,329	0.96	1.04	50
HPROVINCE	WMELIG	AGEC	LAND		52	66	21,614	22,395	0.97	1.04	51
HPROVINCE	WMELIG	AGEC	LAND	ECONSUP1_A	53	165	56,352	65,381	0.86	1.16	52
HPROVINCE	WMELIG	AGEC	LAND	ECONSUP1_A	54	61	22,903	24,004	0.95	1.05	53
HPROVINCE	WMELIG	AGEC			55	205	65,500	68,196	0.96	1.04	54
HPROVINCE	WMELIG	ELECTRI	BICYC	HSEX	56	54	16,845	20,016	0.84	1.19	55
HPROVINCE	WMELIG	ELECTRI	BICYC	HSEX	57	192	63,531	66,913	0.95	1.05	56
HPROVINCE	WMELIG	ELECTRI	BICYC	HVISITS	58	382	153,045	159,672	0.96	1.04	57
HPROVINCE	WMELIG	ELECTRI	BICYC	HVISITS	59	51	22,712	22,712	1.00	1.00	58
HPROVINCE	WMELIG	ELECTRI	LAND		60	81	28,304	29,026	0.98	1.03	59
HPROVINCE	WMELIG	ELECTRI	LAND		61	124	53,210	53,210	1.00	1.00	60
HPROVINCE	HSEX	SOFA	HLANGI		62	134	86,549	122,407	0.71	1.41	61
HPROVINCE	HSEX	SOFA	HLANGI		63	135	112,483	132,093	0.85	1.17	62
HPROVINCE	HSEX	SOFA	WMELIG		64	69	41,924	53,716	0.78	1.28	63
HPROVINCE	HSEX	SOFA	WMELIG	MICROWV	65	356	226,967	349,470	0.65	1.54	64
HPROVINCE	HSEX	SOFA	WMELIG	MICROWV	66	54	39,227	80,255	0.49	2.05	64
HPROVINCE	HSEX	WMELIG	TOILET	HLANGI	67	85	55,791	68,640	0.81	1.23	65
HPROVINCE	HSEX	WMELIG	TOILET	HLANGI	68	68	40,940	44,474	0.92	1.09	66
HPROVINCE	HSEX	WMELIG	TOILET	SMARTPH	69	190	135,298	146,316	0.92	1.08	67
HPROVINCE	HSEX	WMELIG	TOILET	SMARTPH	70	139	86,125	87,648	0.98	1.02	68
HPROVINCE	HSEX	WMELIG	AGEC		71	76	49,274	70,041	0.70	1.42	69
HPROVINCE	HSEX	WMELIG	AGEC		72	59	38,458	41,068	0.94	1.07	70

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjustment factor	Final adjustment cell
HPROVINCE	HSEX	WMELIG	AGEC	HLANGQ	73	76	50,625	66,198	0.76	1.31	71
HPROVINCE	HSEX	WMELIG	AGEC	HLANGQ	74	116	72,668	81,710	0.89	1.12	72
HPROVINCE	HSEX	WMELIG	MICROWV	AGEC	75	223	138,510	183,417	0.76	1.32	73
HPROVINCE	HSEX	WMELIG	MICROWV	AGEC	76	69	44,443	48,183	0.92	1.08	74
HPROVINCE	HSEX	WMELIG	MICROWV		77	38	26,927	48,535	0.55	1.80	75
HPROVINCE	AGEC	HSEX	HRESIDENCE	BICYC	78	61	9,184	16,454	0.56	1.79	76
HPROVINCE	AGEC	HSEX	HRESIDENCE	BICYC	79	101	15,425	21,783	0.71	1.41	77
HPROVINCE	AGEC	HSEX	HRESIDENCE		80	42	5,764	6,864	0.84	1.19	78
HPROVINCE	AGEC	HSEX	MOBILE		81	71	10,679	14,630	0.73	1.37	79
HPROVINCE	AGEC	HSEX	MOBILE	WALL	82	50	6,771	6,862	0.99	1.01	80
HPROVINCE	AGEC	HSEX	MOBILE	WALL	83	137	20,104	25,190	0.80	1.25	81
HPROVINCE	AGEC	ROOMS	SHARETOI		84	136	23,395	23,951	0.98	1.02	82
HPROVINCE	AGEC	ROOMS	SHARETOI		85	71	9,927	10,910	0.91	1.10	83
HPROVINCE	AGEC	ROOMS	HSEX		86	100	15,067	19,943	0.76	1.32	84
HPROVINCE	AGEC	ROOMS	HSEX		87	128	19,486	21,872	0.89	1.12	85
HPROVINCE	AGEC	HSEX	GOAT	FLOOR	88	92	14,246	15,104	0.94	1.06	86
HPROVINCE	AGEC	HSEX	GOAT	FLOOR	89	50	7,774	7,774	1.00	1.00	87
HPROVINCE	AGEC	HSEX	GOAT	WALL	90	153	23,000	24,581	0.94	1.07	88
HPROVINCE	AGEC	HSEX	GOAT	WALL	91	512	79,022	91,955	0.86	1.16	89
HPROVINCE	AGEC	HSEX	HLANGQ	BICYC	92	83	12,352	12,792	0.97	1.04	90
HPROVINCE	AGEC	HSEX	HLANGQ	BICYC	93	70	11,690	11,690	1.00	1.00	91
HPROVINCE	AGEC	HSEX	HLANGQ		94	842	127,795	134,863	0.95	1.06	92
HPROVINCE	AGEC				95	104	16,856	17,430	0.97	1.03	93
HPROVINCE	AGEC	WATSRC	HSEX	FLOOR	96	63	18,664	19,769	0.94	1.06	94
HPROVINCE	AGEC	WATSRC	HSEX	FLOOR	97	50	14,885	17,856	0.83	1.20	95
HPROVINCE	AGEC	WATSRC	HSEX		98	154	46,649	48,410	0.96	1.04	96
HPROVINCE	AGEC	WATSRC			99	90	20,968	26,349	0.80	1.26	97
HPROVINCE	AGEC	HSEX	ECONSUP1_A	HRESIDENCE	100	425	125,236	130,734	0.96	1.04	98
HPROVINCE	AGEC	HSEX	ECONSUP1_A	HRESIDENCE	101	90	22,286	25,172	0.89	1.13	99
HPROVINCE	AGEC	HSEX	ECONSUP1_A		102	118	32,264	36,430	0.89	1.13	100
HPROVINCE	AGEC	HSEX	WATSRC		103	108	36,629	36,629	1.00	1.00	101
HPROVINCE	AGEC	HSEX	WATSRC	TOILET	104	703	196,921	204,556	0.96	1.04	102
HPROVINCE	AGEC	HSEX	WATSRC	TOILET	105	56	13,238	13,238	1.00	1.00	103
HPROVINCE	AGEC				106	78	21,680	21,680	1.00	1.00	104
HPROVINCE	AGEC				107	120	35,874	38,277	0.94	1.07	105
HPROVINCE	HHMEM				108	67	11,863	11,863	1.00	1.00	106
HPROVINCE	HHMEM	FLOOR	CHICKEN		109	49	8,371	9,263	0.90	1.11	107

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjustment factor	Final adjustment cell
HPROVINCE	HHMEM	FLOOR	CHICKEN		110	58	10,800	10,931	0.99	1.01	108
HPROVINCE	HHMEM	FLOOR			111	52	7,911	9,398	0.84	1.19	109
HPROVINCE	HHMEM	HLANGQ	PC	AGEC	112	109	20,503	30,217	0.68	1.47	110
HPROVINCE	HHMEM	HLANGQ	PC	AGEC	113	143	30,967	37,048	0.84	1.20	111
HPROVINCE	HHMEM	HLANGQ	PC	AGEC	114	59	10,735	11,046	0.97	1.03	112
HPROVINCE	HHMEM	HLANGQ	PC	AGEC	115	159	32,534	40,376	0.81	1.24	113
HPROVINCE	HHMEM	HLANGQ	PC		116	84	13,104	13,571	0.97	1.04	114
HPROVINCE	HHMEM	HLANGQ			117	49	9,282	17,786	0.52	1.92	115
HPROVINCE	HHMEM	HLANGQ	CANOE	AGEC	118	290	53,142	68,632	0.77	1.29	116
HPROVINCE	HHMEM	HLANGQ	CANOE	AGEC	119	833	152,266	173,416	0.88	1.14	117
HPROVINCE	HHMEM	HLANGQ	CANOE	AGEC	120	327	61,334	65,688	0.93	1.07	118
HPROVINCE	HHMEM	HLANGQ	CANOE	HSEX	121	57	12,425	12,703	0.98	1.02	119
HPROVINCE	HHMEM	HLANGQ	CANOE	HSEX	122	55	12,291	12,291	1.00	1.00	120
HPROVINCE	HHMEM				123	60	24,553	24,746	0.99	1.01	121
HPROVINCE	HHMEM	HSEX	WMELIG		124	80	29,582	32,304	0.92	1.09	122
HPROVINCE	HHMEM	HSEX	WMELIG	SMARTPH	125	129	82,948	108,344	0.77	1.31	123
HPROVINCE	HHMEM	HSEX	WMELIG	SMARTPH	126	82	24,847	38,553	0.64	1.55	124
HPROVINCE	HHMEM	HSEX	SHARETOI	MNELIG	127	60	28,578	29,177	0.98	1.02	125
HPROVINCE	HHMEM	HSEX	SHARETOI	MNELIG	128	147	76,066	82,765	0.92	1.09	126
HPROVINCE	HHMEM	HSEX	SHARETOI	TABLE	129	46	24,035	31,905	0.75	1.33	127
HPROVINCE	HHMEM	HSEX	SHARETOI	TABLE	130	145	62,993	70,617	0.89	1.12	128
HPROVINCE	HHMEM	AGEC	WMELIG	HRESIDENCE	131	46	38,257	75,452	0.51	1.97	129
HPROVINCE	HHMEM	AGEC	WMELIG	HRESIDENCE	132	68	17,905	23,242	0.77	1.30	130
HPROVINCE	HHMEM	AGEC	WMELIG	MNELIG	133	40	15,987	22,266	0.72	1.39	131
HPROVINCE	HHMEM	AGEC	WMELIG	MNELIG	134	46	26,822	28,510	0.94	1.06	132
HPROVINCE	HHMEM	AGEC	WMELIG	MNELIG	135	34	22,521	31,800	0.71	1.41	133
HPROVINCE	HHMEM	AGEC	SMARTPH		136	67	48,861	50,578	0.97	1.04	134
HPROVINCE	HHMEM	AGEC	SMARTPH		137	59	16,327	20,360	0.80	1.25	135
HPROVINCE	HHMEM	AGEC	GRINDER	COWS	138	53	30,617	46,983	0.65	1.53	136
HPROVINCE	HHMEM	AGEC	GRINDER	COWS	139	97	42,936	50,579	0.85	1.18	137
HPROVINCE	HHMEM	AGEC	GRINDER		140	73	61,657	68,591	0.90	1.11	138
HPROVINCE	HHMEM	AGEC			141	48	28,464	30,815	0.92	1.08	139
HPROVINCE	FUEL				142	57	16,336	16,336	1.00	1.00	140
HPROVINCE	FUEL	AGEC	SOFA	BICYC	143	309	74,143	94,732	0.78	1.28	141
HPROVINCE	FUEL	AGEC	SOFA	BICYC	144	61	15,226	24,073	0.63	1.58	142
HPROVINCE	FUEL	AGEC	SOFA	MNELIG	145	49	13,377	17,781	0.75	1.33	143
HPROVINCE	FUEL	AGEC	SOFA	MNELIG	146	94	26,014	27,501	0.95	1.06	144

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjustment factor	Final adjustment cell
HPROVINCE	FUEL	AGEC	MNELIG	CHICKEN	147	71	16,449	16,813	0.98	1.02	145
HPROVINCE	FUEL	AGEC	MNELIG	CHICKEN	148	136	34,782	34,782	1.00	1.00	146
HPROVINCE	FUEL	AGEC	MNELIG	HRESIDENCE	149	495	119,312	126,015	0.95	1.06	147
HPROVINCE	FUEL	AGEC	MNELIG	HRESIDENCE	150	99	33,964	38,263	0.89	1.13	148
HPROVINCE	FUEL	AGEC	MNELIG	HSEX	151	233	64,457	79,266	0.81	1.23	149
HPROVINCE	FUEL	AGEC	MNELIG	HSEX	152	214	56,935	62,012	0.92	1.09	150
HPROVINCE	FUEL	AGEC	MNELIG		153	63	15,950	15,950	1.00	1.00	151
HPROVINCE	FUEL	AGEC	MNELIG		154	331	84,843	88,239	0.96	1.04	152
HPROVINCE	FUEL	AGEC	LAND		155	52	14,292	14,292	1.00	1.00	153
HPROVINCE	FUEL	AGEC	LAND	PLOUGH	156	363	88,091	92,752	0.95	1.05	154
HPROVINCE	FUEL	AGEC	LAND	PLOUGH	157	97	26,633	30,603	0.87	1.15	155

Table A.2. Blood weights: Nonresponse adjustment cells for male adults aged 15+ years

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjust-ment factor	Final adjust-ment cell
HPROVINCE	IVISITS	HFHIVTSTOFFER	PREPWDTK	AGEC	1	446	271,982	296,081	0.92	1.09	1
HPROVINCE	IVISITS	HFHIVTSTOFFER	PREPWDTK	AGEC	2	51	24,264	24,264	1.00	1.00	2
HPROVINCE	IVISITS	HFHIVTSTOFFER	PREPWDTK	LIFETIMESEX	3	98	66,767	87,111	0.77	1.30	3
HPROVINCE	IVISITS	HFHIVTSTOFFER	PREPWDTK	LIFETIMESEX	4	265	166,987	190,327	0.88	1.14	4
HPROVINCE	IVISITS	HFHIVTSTOFFER	HRESIDENCE	LNGVQX_LNG	5	145	110,057	170,653	0.64	1.55	5
HPROVINCE	IVISITS	HFHIVTSTOFFER	HRESIDENCE	LNGVQX_LNG	6	62	31,803	32,290	0.98	1.02	6
HPROVINCE	IVISITS	HFHIVTSTOFFER	HRESIDENCE	LNGVQX_LNG	7	56	26,667	31,037	0.86	1.16	7
HPROVINCE	IVISITS	HFHIVTSTOFFER	HRESIDENCE	ALCFREQ	8	123	68,458	86,629	0.79	1.27	8
HPROVINCE	IVISITS	HFHIVTSTOFFER	HRESIDENCE	ALCFREQ	9	169	99,975	107,573	0.93	1.08	9
HPROVINCE	IVISITS	HFHIVTSTOFFER	WMELIG	NORMWORK	10	81	50,741	53,426	0.95	1.05	10
HPROVINCE	IVISITS	HFHIVTSTOFFER	WMELIG	NORMWORK	11	65	35,724	44,225	0.81	1.24	11
HPROVINCE	IVISITS	HFHIVTSTOFFER	WMELIG		12	98	64,643	66,401	0.97	1.03	12
HPROVINCE	IVISITS	MCWHOMED			13	104	66,473	84,044	0.79	1.26	13
HPROVINCE	IVISITS	MCWHOMED			14	54	34,221	67,820	0.50	1.98	14
HPROVINCE	I0714A	HRESIDENCE	IVISITS	PAST12MONUM	15	330	211,770	278,438	0.76	1.31	15
HPROVINCE	I0714A	HRESIDENCE	IVISITS	PAST12MONUM	16	115	71,540	80,534	0.89	1.13	16
HPROVINCE	I0714A	HRESIDENCE	IVISITS		17	61	39,493	66,797	0.59	1.69	17
HPROVINCE	I0714A	HRESIDENCE	ROOMS	TV	18	73	23,510	26,010	0.90	1.11	18
HPROVINCE	I0714A	HRESIDENCE	ROOMS	TV	19	53	15,433	15,762	0.98	1.02	19
HPROVINCE	I0714A	HRESIDENCE	ROOMS		20	92	23,534	28,456	0.83	1.21	20
HPROVINCE	I0714A				21	71	32,690	33,620	0.97	1.03	21
HPROVINCE	PAST12MONUM	IVISITS	LITTLEINTEREST	TV	22	472	180,817	187,176	0.97	1.04	22
HPROVINCE	PAST12MONUM	IVISITS	LITTLEINTEREST	TV	23	150	62,582	70,732	0.88	1.13	23
HPROVINCE	PAST12MONUM	IVISITS	LITTLEINTEREST		24	69	30,412	30,412	1.00	1.00	24
HPROVINCE	PAST12MONUM	IVISITS			25	44	14,960	19,003	0.79	1.27	25
HPROVINCE	PAST12MONUM	TABLE			26	54	19,273	19,518	0.99	1.01	26
HPROVINCE	PAST12MONUM	TABLE			27	71	28,831	28,831	1.00	1.00	27
HPROVINCE	TOILET	CONDOMWHERE_E	HHMEM		28	26	22,402	54,206	0.41	2.42	28
HPROVINCE	TOILET	CONDOMWHERE_E	HHMEM		29	75	79,480	119,871	0.66	1.51	28
HPROVINCE	TOILET	CONDOMWHERE_E			30	73	78,929	93,892	0.84	1.19	29
HPROVINCE	TOILET	HFHIVTSTOFFER	HIVTSTRSLT		31	92	83,675	87,707	0.95	1.05	30
HPROVINCE	TOILET	HFHIVTSTOFFER	HIVTSTRSLT		32	100	96,288	111,973	0.86	1.16	31
HPROVINCE	TOILET	HFHIVTSTOFFER	PREPWDTK		33	96	93,990	111,420	0.84	1.19	32
HPROVINCE	TOILET	HFHIVTSTOFFER	PREPWDTK		34	72	66,767	97,674	0.68	1.46	33
HPROVINCE	TOILET	HFHIVTSTOFFER			35	55	54,508	61,197	0.89	1.12	34
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	FIRSTSXAGE		36	156	28,308	30,659	0.92	1.08	35

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjust-ment factor	Final adjust-ment cell
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	FIRSTSXAGE		37	66	11,665	14,841	0.79	1.27	36
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	CHICKEN		38	56	9,749	15,452	0.63	1.59	37
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	CHICKEN		39	67	11,982	15,083	0.79	1.26	38
HPROVINCE	PREPWDTK	HFHIVTSTOFFER			40	77	13,229	15,348	0.86	1.16	39
HPROVINCE	PREPWDTK	HRESIDENCE	CMETHOD_C	FLOOR	41	170	31,905	58,737	0.54	1.84	40
HPROVINCE	PREPWDTK	HRESIDENCE	CMETHOD_C	FLOOR	42	67	12,890	17,370	0.74	1.35	41
HPROVINCE	PREPWDTK	HRESIDENCE	CMETHOD_C		43	47	9,058	10,370	0.87	1.14	42
HPROVINCE	PREPWDTK	HRESIDENCE	CMETHOD_C	HFHIVTSTOFFER	44	57	9,674	13,737	0.70	1.42	43
HPROVINCE	PREPWDTK	HRESIDENCE	CMETHOD_C	HFHIVTSTOFFER	45	42	7,179	15,054	0.48	2.10	43
HPROVINCE	PREPWDTK	HRESIDENCE	SMARTPH		46	40	7,394	9,262	0.80	1.25	44
HPROVINCE	PREPWDTK	HRESIDENCE	SMARTPH		47	49	8,088	8,498	0.95	1.05	45
HPROVINCE	IVISITS	MNELIG	BICYC	WMELIG	48	226	68,012	71,567	0.95	1.05	46
HPROVINCE	IVISITS	MNELIG	BICYC	WMELIG	49	71	20,032	20,032	1.00	1.00	47
HPROVINCE	IVISITS	MNELIG	BICYC	PREPWDTK	50	191	61,341	64,514	0.95	1.05	48
HPROVINCE	IVISITS	MNELIG	BICYC	PREPWDTK	51	120	32,299	37,679	0.86	1.17	49
HPROVINCE	IVISITS	MNELIG			52	109	30,761	30,761	1.00	1.00	50
HPROVINCE	IVISITS				53	111	37,432	45,177	0.83	1.21	51
HPROVINCE	IVISITS	TOILET	ALCFREQ	LITTLEINTEREST	54	441	94,433	102,840	0.92	1.09	52
HPROVINCE	IVISITS	TOILET	ALCFREQ	LITTLEINTEREST	55	106	22,925	23,396	0.98	1.02	53
HPROVINCE	IVISITS	TOILET	ALCFREQ	HFHIVTSTOFFER	56	68	14,766	15,985	0.92	1.08	54
HPROVINCE	IVISITS	TOILET	ALCFREQ	HFHIVTSTOFFER	57	67	13,881	17,500	0.79	1.26	55
HPROVINCE	IVISITS	TOILET			58	79	17,601	17,601	1.00	1.00	56
HPROVINCE	IVISITS	PREPWDTK			59	136	29,617	38,794	0.76	1.31	57
HPROVINCE	IVISITS	PREPWDTK			60	35	6,971	15,677	0.44	2.25	57
HPROVINCE	HIVTSTRSLT				61	70	60,446	70,833	0.85	1.17	58
HPROVINCE	HIVTSTRSLT				62	57	37,883	39,459	0.96	1.04	59
HPROVINCE	HIVTSTRSLT	ADHIVPREVA	HVISITS	ELECTRI	63	107	94,286	106,914	0.88	1.13	60
HPROVINCE	HIVTSTRSLT	ADHIVPREVA	HVISITS	ELECTRI	64	86	26,954	35,708	0.75	1.32	61
HPROVINCE	HIVTSTRSLT	ADHIVPREVA	HVISITS		65	46	27,892	45,447	0.61	1.63	62
HPROVINCE	HIVTSTRSLT	ADHIVPREVA	WATSRC		66	60	21,633	29,095	0.74	1.34	63
HPROVINCE	HIVTSTRSLT	ADHIVPREVA	WATSRC		67	29	30,171	67,647	0.45	2.24	63
HPROVINCE	PREPWDTK	FIRSTSXAGE			68	68	18,748	21,290	0.88	1.14	64
HPROVINCE	PREPWDTK	FIRSTSXAGE	TABLE	ECONSUP1_A	69	152	40,286	41,744	0.97	1.04	65
HPROVINCE	PREPWDTK	FIRSTSXAGE	TABLE	ECONSUP1_A	70	53	14,880	14,880	1.00	1.00	66
HPROVINCE	PREPWDTK	FIRSTSXAGE	TABLE	GRINDER	71	351	112,261	122,437	0.92	1.09	67
HPROVINCE	PREPWDTK	FIRSTSXAGE	TABLE	GRINDER	72	50	13,204	13,416	0.98	1.02	68
HPROVINCE	PREPWDTK	FIRSTSXAGE			73	50	14,744	14,744	1.00	1.00	69

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjust-ment factor	Final adjust-ment cell
HPROVINCE	PREPWDTK	ROOMS	GRINDER	COWS	74	44	13,867	20,711	0.67	1.49	70
HPROVINCE	PREPWDTK	ROOMS	GRINDER	COWS	75	110	30,478	37,486	0.81	1.23	71
HPROVINCE	PREPWDTK	ROOMS	GRINDER		76	54	12,835	13,556	0.95	1.06	72
HPROVINCE	PREPWDTK	ROOMS	LNGVQX_LNG		77	69	22,637	22,637	1.00	1.00	73
HPROVINCE	PREPWDTK	ROOMS	LNGVQX_LNG		78	49	16,401	19,521	0.84	1.19	74

Table A.3. Blood weights: Nonresponse adjustment cells for female adults aged 15+ years

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjustment factor	Final adjustment cell
HPROVINCE	IVISITS	PREPWDTK	HUSLIVEW	WMELIG	1	402	233,308	255,090	0.91	1.09	1
HPROVINCE	IVISITS	PREPWDTK	HUSLIVEW	WMELIG	2	279	160,367	164,232	0.98	1.02	2
HPROVINCE	IVISITS	PREPWDTK	HUSLIVEW	CMETHOD_F	3	431	235,503	255,988	0.92	1.09	3
HPROVINCE	IVISITS	PREPWDTK	HUSLIVEW	CMETHOD_F	4	273	153,148	186,110	0.82	1.22	4
HPROVINCE	IVISITS	PREPWDTK	HUSLIVEW		5	56	30,793	40,943	0.75	1.33	5
HPROVINCE	IVISITS	PREPWDTK	WMELIG	BED	6	38	22,956	35,770	0.64	1.56	6
HPROVINCE	IVISITS	PREPWDTK	WMELIG	BED	7	180	104,065	132,364	0.79	1.27	7
HPROVINCE	IVISITS	PREPWDTK	WMELIG	SHARETOI	8	261	144,030	174,138	0.83	1.21	8
HPROVINCE	IVISITS	PREPWDTK	WMELIG	SHARETOI	9	85	54,230	56,690	0.96	1.05	9
HPROVINCE	IVISITS	MNELIG			10	44	28,956	31,953	0.91	1.10	10
HPROVINCE	IVISITS	MNELIG	HFHIVTSTOFFER		11	50	36,880	43,478	0.85	1.18	11
HPROVINCE	IVISITS	MNELIG	HFHIVTSTOFFER	HRESIDENCE	12	32	21,029	38,166	0.55	1.81	12
HPROVINCE	IVISITS	MNELIG	HFHIVTSTOFFER	HRESIDENCE	13	56	29,337	38,401	0.76	1.31	13
HPROVINCE	IVISITS	BICYC			14	47	32,350	50,268	0.64	1.55	14
HPROVINCE	IVISITS	BICYC			15	25	15,539	43,235	0.36	2.78	14
HPROVINCE	SHARETOI	HFHIVTSTOFFER	PREPWDTK		16	130	63,860	73,461	0.87	1.15	15
HPROVINCE	SHARETOI	HFHIVTSTOFFER	PREPWDTK		17	76	37,742	54,828	0.69	1.45	16
HPROVINCE	SHARETOI	HFHIVTSTOFFER	FIRSTSXAGE		18	31	12,801	29,982	0.43	2.34	17
HPROVINCE	SHARETOI	HFHIVTSTOFFER	FIRSTSXAGE		19	144	72,152	101,543	0.71	1.41	17
HPROVINCE	SHARETOI	HFHIVTSTOFFER	FIRSTSXAGE		20	61	24,078	42,339	0.57	1.76	18
HPROVINCE	SHARETOI	HFHIVTSTOFFER			21	142	66,994	82,245	0.81	1.23	19
HPROVINCE	SHARETOI	HIVPOSPROV			22	62	25,424	25,917	0.98	1.02	20
HPROVINCE	SHARETOI	HIVPOSPROV	WORK12MO	CHICKEN	23	104	57,830	75,391	0.77	1.30	21
HPROVINCE	SHARETOI	HIVPOSPROV	WORK12MO	CHICKEN	24	131	65,528	75,630	0.87	1.15	22
HPROVINCE	SHARETOI	HIVPOSPROV	WORK12MO		25	63	29,206	31,429	0.93	1.08	23
HPROVINCE	MNELIG	HIVTSBP	PREPWDTK	HIVTSTRSLT	26	83	28,391	28,391	1.00	1.00	24
HPROVINCE	MNELIG	HIVTSBP	PREPWDTK	HIVTSTRSLT	27	445	158,503	169,429	0.94	1.07	25
HPROVINCE	MNELIG	HIVTSBP	PREPWDTK	WMELIG	28	127	45,646	55,282	0.83	1.21	26
HPROVINCE	MNELIG	HIVTSBP	PREPWDTK	WMELIG	29	81	30,083	30,762	0.98	1.02	27
HPROVINCE	MNELIG	HIVTSBP	PREPWDTK	WMELIG	30	51	18,728	21,919	0.85	1.17	28
HPROVINCE	MNELIG	HIVTSBP			31	71	27,428	27,428	1.00	1.00	29
HPROVINCE	MNELIG	LIFETIMESEX	ADTPSX		32	51	20,561	20,806	0.99	1.01	30
HPROVINCE	MNELIG	LIFETIMESEX	ADTPSX		33	53	23,265	23,265	1.00	1.00	31
HPROVINCE	MNELIG	LIFETIMESEX			34	71	26,012	27,754	0.94	1.07	32
HPROVINCE	MNELIG				35	84	35,169	35,169	1.00	1.00	33
HPROVINCE	LITTLEINTEREST	HLANGI	LIFETIMESEX	HPVVACC	36	137	104,492	152,026	0.69	1.45	34

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjustment factor	Final adjustment cell
HPROVINCE	LITTLEINTEREST	HLANGI	LIFETIMESEX	HPVVACC	37	43	32,477	63,606	0.51	1.96	35
HPROVINCE	LITTLEINTEREST	HLANGI	LIFETIMESEX	ADTPSX	38	149	109,262	152,696	0.72	1.40	36
HPROVINCE	LITTLEINTEREST	HLANGI	LIFETIMESEX	ADTPSX	39	62	52,270	59,429	0.88	1.14	37
HPROVINCE	LITTLEINTEREST	HLANGI	HFHIVTSTOFFER		40	86	69,879	78,134	0.89	1.12	38
HPROVINCE	LITTLEINTEREST	HLANGI	HFHIVTSTOFFER	CMETHOD_F	41	101	72,758	86,143	0.84	1.18	39
HPROVINCE	LITTLEINTEREST	HLANGI	HFHIVTSTOFFER	CMETHOD_F	42	33	24,431	37,000	0.66	1.51	40
HPROVINCE	LITTLEINTEREST	HLANGI	HFHIVTSTOFFER		43	53	45,172	48,754	0.93	1.08	41
HPROVINCE	LITTLEINTEREST	INTERNET	LIVEB		44	47	40,730	47,567	0.86	1.17	42
HPROVINCE	LITTLEINTEREST	INTERNET	LIVEB		45	96	74,850	75,916	0.99	1.01	43
HPROVINCE	LITTLEINTEREST	INTERNET	LIVEB		46	51	42,305	45,557	0.93	1.08	44
HPROVINCE	LITTLEINTEREST	INTERNET			47	39	30,226	39,402	0.77	1.30	45
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	TABLE	HIVTSTRSLT	48	56	8,121	8,889	0.91	1.09	46
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	TABLE	HIVTSTRSLT	49	115	20,513	26,823	0.76	1.31	47
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	TABLE		50	122	19,913	21,796	0.91	1.09	48
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	HRESIDENCE	LIFETIMESEX	51	108	16,996	28,515	0.60	1.68	49
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	HRESIDENCE	LIFETIMESEX	52	82	14,739	19,119	0.77	1.30	50
HPROVINCE	PREPWDTK	HFHIVTSTOFFER	HRESIDENCE		53	67	9,822	11,128	0.88	1.13	51
HPROVINCE	PREPWDTK	HFHIVTSTOFFER			54	175	27,777	30,709	0.90	1.11	52
HPROVINCE	PREPWDTK	FLOOR	HFHIVTSTOFFER	BICYC	55	82	14,094	25,504	0.55	1.81	53
HPROVINCE	PREPWDTK	FLOOR	HFHIVTSTOFFER	BICYC	56	61	10,027	13,738	0.73	1.37	54
HPROVINCE	PREPWDTK	FLOOR	HFHIVTSTOFFER	CHICKEN	57	48	7,005	19,747	0.35	2.82	55
HPROVINCE	PREPWDTK	FLOOR	HFHIVTSTOFFER	CHICKEN	58	34	5,967	11,569	0.52	1.94	55
HPROVINCE	PREPWDTK	FLOOR	HFHIVTSTOFFER		59	57	9,402	16,013	0.59	1.70	56
HPROVINCE	PREPWDTK	FLOOR			60	118	19,019	25,974	0.73	1.37	57
HPROVINCE	DOGS				61	107	29,954	29,954	1.00	1.00	58
HPROVINCE	DOGS	ADHIVSCHMTG	HIVPOSPROV		62	55	14,809	14,809	1.00	1.00	59
HPROVINCE	DOGS	ADHIVSCHMTG	HIVPOSPROV	IVISITS	63	739	222,360	236,652	0.94	1.06	60
HPROVINCE	DOGS	ADHIVSCHMTG	HIVPOSPROV	IVISITS	64	86	23,060	31,364	0.74	1.36	61
HPROVINCE	DOGS	ADHIVSCHMTG			65	59	16,889	16,889	1.00	1.00	62
HPROVINCE	DOGS	ADHIVSCHMTG			66	44	14,288	17,033	0.84	1.19	63
HPROVINCE	IVISITS	PREPWDTK			67	73	14,605	14,699	0.99	1.01	64
HPROVINCE	IVISITS	PREPWDTK	HLANGR		68	75	14,790	14,790	1.00	1.00	65
HPROVINCE	IVISITS	PREPWDTK	HLANGR	HIVTSTRSLT	69	79	16,574	16,712	0.99	1.01	66
HPROVINCE	IVISITS	PREPWDTK	HLANGR	HIVTSTRSLT	70	435	87,620	94,978	0.92	1.08	67
HPROVINCE	IVISITS	PREPWDTK	SMARTPH	HFHIVTSTOFFER	71	135	35,716	42,517	0.84	1.19	68
HPROVINCE	IVISITS	PREPWDTK	SMARTPH	HFHIVTSTOFFER	72	81	20,110	30,295	0.66	1.51	69
HPROVINCE	IVISITS	PREPWDTK	SMARTPH		73	69	17,752	18,859	0.94	1.06	70

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Initial adjustment cell	Number of respondents	Weighted number of respondents	Weighted number of eligibles	Weighted response rate	Adjustment factor	Final adjustment cell
HPROVINCE	IVISITS	MNELIG	LIFETIMESEX		74	38	7,084	15,508	0.46	2.19	71
HPROVINCE	IVISITS	MNELIG	LIFETIMESEX		75	60	12,641	19,014	0.66	1.50	71
HPROVINCE	IVISITS	MNELIG			76	67	15,052	19,514	0.77	1.30	72
HPROVINCE	TBSYMP12MO	HPVVACC	HVISITS	LIFETIMESEX	77	50	26,622	28,842	0.92	1.08	73
HPROVINCE	TBSYMP12MO	HPVVACC	HVISITS	LIFETIMESEX	78	259	164,460	216,110	0.76	1.31	74
HPROVINCE	TBSYMP12MO	HPVVACC	HVISITS		79	63	34,569	57,200	0.60	1.65	75
HPROVINCE	TBSYMP12MO	HPVVACC	LNGVINT_LNG		80	68	31,398	41,505	0.76	1.32	76
HPROVINCE	TBSYMP12MO	HPVVACC	LNGVINT_LNG		81	43	25,593	58,159	0.44	2.27	76
HPROVINCE	TBSYMP12MO	ELECTRI			82	59	45,912	46,983	0.98	1.02	77
HPROVINCE	TBSYMP12MO	ELECTRI			83	66	20,488	23,642	0.87	1.15	78
HPROVINCE	ELECTRI	HHMEM	HFHIVTSTOFFER	MOBILE	84	73	18,249	22,312	0.82	1.22	79
HPROVINCE	ELECTRI	HHMEM	HFHIVTSTOFFER	MOBILE	85	61	18,492	20,005	0.92	1.08	80
HPROVINCE	ELECTRI	HHMEM	HFHIVTSTOFFER		86	142	38,792	50,646	0.77	1.31	81
HPROVINCE	ELECTRI	HHMEM	HFHIVTSTOFFER		87	61	18,306	19,398	0.94	1.06	82
HPROVINCE	ELECTRI	HHMEM	HVISITS	TABLE	88	347	88,520	91,657	0.97	1.04	83
HPROVINCE	ELECTRI	HHMEM	HVISITS	TABLE	89	507	141,534	154,449	0.92	1.09	84
HPROVINCE	ELECTRI	HHMEM	HVISITS		90	70	18,180	22,646	0.80	1.25	85
HPROVINCE	ELECTRI	LIFETIMESEX			91	60	19,641	20,177	0.97	1.03	86
HPROVINCE	ELECTRI	LIFETIMESEX			92	111	35,697	35,697	1.00	1.00	87
HPROVINCE	TBSYMP12MO	HPVVACC	HVISITS	LIFETIMESEX	78	259	164,460	216,110	0.76	1.31	74
HPROVINCE	TBSYMP12MO	HPVVACC	HVISITS		79	63	34,569	57,200	0.60	1.65	75
HPROVINCE	TBSYMP12MO	HPVVACC	LNGVINT_LNG		80	68	31,398	41,505	0.76	1.32	76
HPROVINCE	TBSYMP12MO	HPVVACC	LNGVINT_LNG		81	43	25,593	58,159	0.44	2.27	76
HPROVINCE	TBSYMP12MO	ELECTRI			82	59	45,912	46,983	0.98	1.02	77
HPROVINCE	TBSYMP12MO	ELECTRI			83	66	20,488	23,642	0.87	1.15	78
HPROVINCE	ELECTRI	HHMEM	HFHIVTSTOFFER	MOBILE	84	73	18,249	22,312	0.82	1.22	79
HPROVINCE	ELECTRI	HHMEM	HFHIVTSTOFFER	MOBILE	85	61	18,492	20,005	0.92	1.08	80
HPROVINCE	ELECTRI	HHMEM	HFHIVTSTOFFER		86	142	38,792	50,646	0.77	1.31	81
HPROVINCE	ELECTRI	HHMEM	HFHIVTSTOFFER		87	61	18,306	19,398	0.94	1.06	82
HPROVINCE	ELECTRI	HHMEM	HVISITS	TABLE	88	347	88,520	91,657	0.97	1.04	83
HPROVINCE	ELECTRI	HHMEM	HVISITS	TABLE	89	507	141,534	154,449	0.92	1.09	84
HPROVINCE	ELECTRI	HHMEM	HVISITS		90	70	18,180	22,646	0.80	1.25	85
HPROVINCE	ELECTRI	LIFETIMESEX			91	60	19,641	20,177	0.97	1.03	86
HPROVINCE	ELECTRI	LIFETIMESEX			92	111	35,697	35,697	1.00	1.00	87