

Sample design for the Fifth Rwanda Demographic and Health Survey in 2014

(RDHS-V)

By Ruilin Ren, PhD

ICF International

I. Introduction

The fifth Rwanda Demographic and Health Survey (RDHS-V) in 2014 follows those implemented in 1992, 2000, 2005 and 2010. A nationally representative sample of about 12,800 households will be selected. All women age 15-49 who are usual residents of the selected households or who slept in the households the night before the survey are eligible for the survey. The survey will result about 13,650 interviews of women age 15-49. As with the prior surveys, the main objectives of the RDHS-V survey are to provide up-to-date information on fertility and childhood mortality levels; fertility preferences; awareness, approval and use of family planning methods; maternal and child health; knowledge and attitudes toward HIV/AIDS and other sexually transmitted infections (STI); prevalence of HIV among the adult population.

Apart from the female survey, a male survey will also be conducted at the same time in a sub-sample consisting of one household in every two selected for the female survey. All men age 15-59 who are usual residents of the selected households or who slept in the households the night before the survey are eligible for the male survey. The survey will collect information on their basic demographic and social status; on their knowledge and use of family planning methods; and on their knowledge and attitudes toward HIV/AIDS and other sexually transmitted infections. The survey will result about 6,300 interviews of men age 15-59. In this sub-sample, all women and men eligible for the individual survey will be asked consent to provide a drop of blood for HIV testing. Also in this sub-sample, all children under 5 will be measured and weighed for the anthropometry indicators; all children under 5 and all women 15-49 will be tested for anemia and malaria. In 50% of the households selected for male survey, all children 0-14 years will be tested for HIV prevalence; in this sub-sample, a module of domestic violence against woman will be applied to only one woman randomly selected from the households with at least one eligible woman. In the other 50% of the households selected for male survey where there is no women domestic violence module, a module of domestic violence against man will be applied to only one man randomly selected from the households with at least one eligible man.

The survey is designed to produce representative estimates for the main demographic and health indicators for the country as a whole, for the urban and rural areas, and for each of the five provinces. For some indicators, representative results may be available for each of the thirty districts, where possible.

II. Sampling frame

The sampling frame used for RDHS-V is the Rwanda Population and Housing Census (RPHC 2012) which was conducted in Rwanda in 2012. The sampling frame is a complete list of Enumeration Areas (EA) covering the whole country, provided by the National Institute of Statistics of Rwanda (NISR), the implementing agency for the RDHS-V. An EA is a natural village, or a part of a village, created for the RPHC 2012 which was served as counting unit for the census. Each EA appears with identification information, administrative belongings and a measure of size which is the number of residential households residing located in the EA. Each EA is also classified into one of the four types of residence, urban, semi-urban, peri-urban and rural. The urban and the semi-urban are grouped together as “urban” areas, and the peri-urban and rural are grouped together as “rural” areas. Each EA also has accompanied cartographical materials which delineate its geographical locations, boundaries, main access and land marks in or outside the EA which helps to identify the EAs.

Rwanda’s administrative units had been reformed in 2006, reduced the number of provinces from 11 to 5 compared to the last population census conducted in 2002. According to the RPHC 2012, Rwanda is divided into provinces; each province is sub-divided into districts; each district into sector, and each sector into cells and each cell into villages. There are 5 provinces, with a total number of 30 districts and 417 sectors. Table 2.1 below shows the residential population distribution after excluding the institutional population, by province, by district within each province, and according to type of residence (urban and rural). Population share goes from 10.8% for Kigali

province, to 24.7% for East province. The population share for districts is quite homogeneous, goes from 2.7% to 5.1%. In Rwanda, 16.6% of the residential population lives in urban areas. The urban percentage of the provinces varies from 76 percent of urban for Kigali province, to 7.4 percent of urban for East province. Table 2.2 below shows the distribution of residential households. There is a slight difference between the two distributions because in general, the urban household size is smaller than the rural household size. As a result, the percentage of urban households is 17.4%. Table 2.3 below shows the distribution of EAs and their average size in number of households after excluding 88 institutional EAs. Among the 16640 EAs, 2554 EAs are in urban areas, and 14086 EAs are in rural areas. The average size of the EAs is 165 households for the urban EAs, and 142 households for the rural EAs, with an overall average of 146 households per EA.

Table 2.1 Population by province and by district within province according to type of residence

Province	District	Population			Population share	
		Urban	Rural	Total	Urban	Total
Kigali City	Nyarugenge	215069	69746	284815	75.5	2.7
	Gasabo	366717	164184	530901	69.1	5.1
	Kicukiro	280361	38752	319113	87.9	3.0
Kigali City Total		862147	272682	1134829	76.0	10.8
South	Nyanza	25419	297969	323388	7.9	3.1
	Gisagara	5014	317789	322803	1.6	3.1
	Nyaruguru	5131	288293	293424	1.7	2.8
	Huye	53101	275504	328605	16.2	3.1
	Nyamagabe	24932	305721	330653	7.5	3.2
	Ruhango	26470	295551	322021	8.2	3.1
	Muhanga	54362	264603	318965	17.0	3.0
	Kamonyi	38767	304025	342792	11.3	3.3
South Total		233196	2349455	2582651	9.0	24.6
West	Karongi	22898	292872	315770	7.3	3.0
	Rutsiro	6736	316515	323251	2.1	3.1
	Rubavu	148368	255910	404278	36.7	3.9
	Nyabihu	40610	254582	295192	13.8	2.8
	Ngororero	12280	322133	334413	3.7	3.2
	Rusizi	63868	340844	404712	15.8	3.9
	Nyamasheke	6199	376939	383138	1.6	3.7
West Total		300959	2159795	2460754	12.2	23.5
North	Rulindo	8706	279746	288452	3.0	2.8
	Gakenke	9367	329219	338586	2.8	3.2
	Musanze	102799	265764	368563	27.9	3.5
	Burera	6240	330215	336455	1.9	3.2
	Gicumbi	23839	361798	385637	6.2	3.7
North Total		150951	1566742	1717693	8.8	16.4
East	Rwamagana	27179	283059	310238	8.8	3.0
	Nyagatare	47888	419056	466944	10.3	4.5
	Gatsibo	23719	400196	423915	5.6	4.0
	Kayonza	37179	309572	346751	10.7	3.3
	Kirehe	10056	330927	340983	2.9	3.3
	Ngoma	15461	323101	338562	4.6	3.2

	Bugesera	29511	333828	363339	8.1	3.5
East Total		190993	2399739	2590732	7.4	24.7
Rwanda		1738246	8748413	10486659	16.6	100.0

*Source: Residential population, 2012 population census, Rwanda

Table 2.2 Distribution of residential households by province and by district within province according to type of residence

Province	District	Household			Household share	
		Urban	Rural	Total	Urban	Total
Kigali City	Nyarugenge	53512	17369	70881	75.5	2.9
	Gasabo	100235	41718	141953	70.6	5.9
	Kicukiro	68538	9980	78518	87.3	3.2
Kigali City Total		222285	69067	291352	76.3	12.0
South	Nyanza	6533	68514	75047	8.7	3.1
	Gisagara	1242	75995	77237	1.6	3.2
	Nyaruguru	1395	59895	61290	2.3	2.5
	Huye	11350	67002	78352	14.5	3.2
	Nyamagabe	4933	70093	75026	6.6	3.1
	Ruhango	6517	70034	76551	8.5	3.2
	Muhanga	10445	63296	73741	14.2	3.0
	Kamonyi	9624	71482	81106	11.9	3.3
South Total		52039	546311	598350	8.7	24.7
West	Karongi	5904	67847	73751	8.0	3.0
	Rutsiro	1457	69813	71270	2.0	2.9
	Rubavu	34345	54702	89047	38.6	3.7
	Nyabihu	8671	57551	66222	13.1	2.7
	Ngororero	3021	76209	79230	3.8	3.3
	Rusizi	13314	70442	83756	15.9	3.5
	Nyamasheke	1389	80914	82303	1.7	3.4
West Total		68101	477478	545579	12.5	22.5
North	Rulindo	2087	65364	67451	3.1	2.8
	Gakenke	2505	77257	79762	3.1	3.3
	Musanze	23262	61520	84782	27.4	3.5
	Burera	1504	72197	73701	2.0	3.0
	Gicumbi	5629	80796	86425	6.5	3.6
North Total		34987	357134	392121	8.9	16.2
East	Rwamagana	6615	67585	74200	8.9	3.1
	Nyagatare	12128	94622	106750	11.4	4.4
	Gatsibo	5877	90254	96131	6.1	4.0
	Kayonza	7433	70735	78168	9.5	3.2
	Kirehe	2359	75331	77690	3.0	3.2
	Ngoma	3360	76585	79945	4.2	3.3
	Bugesera	7238	78237	85475	8.5	3.5
East Total		45010	553349	598359	7.5	24.7
Rwanda		422422	2003339	2425761	17.4	100.0

*Source: Residential households, 2012 population census, Rwanda

Table 2.3 Distribution of EAs and their average size in number of households by province and by district within province according to type of residence

Province	District	Number of EAs			Average EA size		
		Urban	Rural	Total	Urban	Rural	Total
Kigali City	Nyarugenge	396	122	518	135	142	137
	Gasabo	585	262	847	171	159	168
	Kicukiro	473	72	545	145	139	144
Kigali City Total		1454	456	1910	153	151	153
South	Nyanza	36	432	468	181	159	160
	Gisagara	9	533	542	138	143	143
	Nyaruguru	8	391	399	174	153	154
	Huye	64	486	550	177	138	142
	Nyamagabe	31	525	556	159	134	135
	Ruhango	40	511	551	163	137	139
	Muhanga	49	361	410	213	175	180
	Kamonyi	41	386	427	235	185	190
South Total		278	3625	3903	187	151	153
West	Karongi	35	511	546	169	133	135
	Rutsiro	9	482	491	162	145	145
	Rubavu	203	375	578	169	146	154
	Nyabihu	44	445	489	197	129	135
	Ngororero	16	484	500	189	157	158
	Rusizi	83	543	626	160	130	134
	Nyamasheke	8	602	610	174	134	135
	West Total		398	3442	3840	171	139
North	Rulindo	11	492	503	190	133	134
	Gakenke	17	603	620	147	128	129
	Musanze	116	405	521	201	152	163
	Burera	10	582	592	150	124	124
	Gicumbi	34	611	645	166	132	134
North Total		188	2693	2881	186	133	136
East	Rwamagana	39	467	506	170	145	147
	Nyagatare	59	635	694	206	149	154
	Gatsibo	28	643	671	210	140	143
	Kayonza	35	426	461	212	166	170
	Kirehe	17	613	630	139	123	123
	Ngoma	20	510	530	168	150	151
	Bugesera	38	576	614	190	136	139
East Total		236	3870	4106	191	143	146
Rwanda		2554	14086	16640	165	142	146

*Source: 2012 population census excluding 88 institutional EAs

III. Structure of the sample and the sampling procedure

The sample for RDHS 2014 will be a stratified sample selected in two stages from the 2012 census frame. Stratification will be achieved by separating each district into urban and rural areas, the urban and rural areas of each district forms each a sampling stratum. In total, 60 sampling strata have been created. Samples will be selected independently in each sampling stratum, by a two stages selection. Implicit stratification and proportional allocation would be achieved at each of the lower administrative unit levels by sorting the sampling frame within the explicit stratum according to administrative unit in different levels before sample selection and by using a probability proportional to size selection at the first stage of sampling.

In the first stage, 492 EAs will be selected with probability proportional to the EA size and with independent selection in each sampling stratum with the sample allocation given in table 3.1 below. A household listing operation will be carried out in all of the selected EAs before the main survey. The household listing operation consists of visiting each of the 492 selected EAs; to draw a location map and a detailed sketch map; and to record on the household listing forms all residential households found in the EA with the address and the name of the head of the households. The resulting list of households will be served as the sampling frame for the selection of households in the second stage. Some of the selected EAs may be found large in size in the household listing operation. In order to minimize the task of household listing, the selected EAs with an estimated number of households greater than 300 will be segmented. Only one segment will be selected for the survey with probability proportional to the segment size. The methodology and the detailed household listing procedure are addressed in the household listing manual.

At the second stage, a fix number of 26 households will be selected from each selected EA. Table 3.1 below shows the sample allocation of EAs and households. Among the 492 EAs, 113 EAs will be from urban areas and 379 EAs will come from rural areas. The total number of household to be select is 12792, among them, 2938 household will be from urban areas, and 9854 households will be from rural areas. The urban area is slightly oversampled because of the low urbanization in most of the districts, where at least two urban clusters need to be selected. With the request of representative results for some indicators at district level, the total sample size is tight and therefore an equal size allocation was adopted, with a slightly larger sample size for the districts in the province of Kigali City because the low fertility level in Kigali City. In fact, the equal size allocation is not far from the proportional allocation which is the best allocation, because the district sizes are quite homogeneous. On the other hand, the total sample size is already large; any substantial increase in the total sample size in order to provide representative results for most of the indicators at district level will compromise the data quality due to the limited implementing capability. With the current sample size, adequate survey precision at district level will be obtained for women indicators above 15%; and for children (under five) indicators above 20%.

Table 3.2 below shows the sample allocation of expected number of women and men interviews. Table 3.3 below shows the expected number of eligible women and men for HIV testing. The sample calculations were based on the survey results of the RDHS-IV 2010: the average number of women 15-49 per household is 1.10; the average number of men 15-59 per household is 1.02; household response rate is 98%; women individual response rate is 99%; men individual response rate is 99%.

Table 3.1 Sample allocation of EAs and households by province and by district within province according to type of residence

Province	District	Allocation of EAs			Allocation of households		
		Urban	Rural	Total	Urban	Rural	Total
Kigali City	Nyarugenge	15	5	20	390	130	520
	Gasabo	14	6	20	364	156	520
	Kicukiro	17	3	20	442	78	520
Kigali City Total		46	14	60	1196	364	1560
South	Nyanza	2	14	16	52	364	416
	Gisagara	2	14	16	52	364	416
	Nyaruguru	2	14	16	52	364	416
	Huye	3	13	16	78	338	416
	Nyamagabe	2	14	16	52	364	416
	Ruhango	2	14	16	52	364	416

	Muhanga	3	13	16	78	338	416
	Kamonyi	3	13	16	78	338	416
South Total		19	109	128	494	2834	3328
West	Karongi	2	14	16	52	364	416
	Rutsiro	2	14	16	52	364	416
	Rubavu	6	10	16	156	260	416
	Nyabihu	3	13	16	78	338	416
	Ngororero	2	14	16	52	364	416
	Rusizi	3	13	16	78	338	416
	Nyamasheke	2	14	16	52	364	416
West Total		20	92	112	520	2392	2912
North	Rulindo	2	14	16	52	364	416
	Gakenke	2	14	16	52	364	416
	Musanze	5	11	16	130	286	416
	Burera	2	14	16	52	364	416
	Gicumbi	2	14	16	52	364	416
North Total		13	67	80	338	1742	2080
East	Rwamagana	2	14	16	52	364	416
	Nyagatare	3	13	16	78	338	416
	Gatsibo	2	14	16	52	364	416
	Kayonza	2	14	16	52	364	416
	Kirehe	2	14	16	52	364	416
	Ngoma	2	14	16	52	364	416
	Bugesera	2	14	16	52	364	416
East Total		15	97	112	390	2522	2912
Rwanda		113	379	492	2938	9854	12792

Table 3.2 Sample allocation of expected number of interviews by province and by district within province according to type of residence

Province	District	Women 15-49			Men 15-59		
		Urban	Rural	Total	Urban	Rural	Total
Kigali City	Nyarugenge	416	139	555	193	65	258
	Gasabo	389	166	555	180	77	257
	Kicukiro	471	83	554	219	38	257
Kigali City Total		1276	388	1664	592	180	772
South	Nyanza	55	389	444	25	180	205
	Gisagara	55	389	444	25	180	205
	Nyaruguru	55	389	444	25	180	205
	Huye	83	360	443	38	168	206
	Nyamagabe	55	389	444	25	180	205
	Ruhango	55	389	444	25	180	205
	Muhanga	83	360	443	38	168	206
	Kamonyi	83	360	443	38	168	206
South Total		524	3025	3549	239	1404	1643
West	Karongi	55	389	444	25	180	205

	Rutsiro	55	389	444	25	180	205
	Rubavu	166	278	444	77	128	205
	Nyabihu	83	360	443	38	168	206
	Ngororero	55	389	444	25	180	205
	Rusizi	83	360	443	38	168	206
	Nyamasheke	55	389	444	25	180	205
West Total		552	2554	3106	253	1184	1437
North	Rulindo	55	389	444	25	180	205
	Gakenke	55	389	444	25	180	205
	Musanze	139	305	444	65	141	206
	Burera	55	389	444	25	180	205
	Gicumbi	55	389	444	25	180	205
North Total		359	1861	2220	165	861	1026
East	Rwamagana	55	389	444	25	180	205
	Nyagatare	83	360	443	38	168	206
	Gatsibo	55	389	444	25	180	205
	Kayonza	55	389	444	25	180	205
	Kirehe	55	389	444	25	180	205
	Ngoma	55	389	444	25	180	205
	Bugesera	55	389	444	25	180	205
East Total		413	2694	3107	188	1248	1436
Rwanda		3124	10522	13646	1437	4877	6314

* Male survey will be carried out in ½ of the households selected for female survey.

Table 3.3 Expected number of eligible individuals for HIV testing

Province	Women 15-49			Men 15-59			Total
	Urban	Rural	Total	Urban	Rural	Total	
Kigali City	644	196	840	598	182	780	1620
South	265	1528	1792	241	1418	1660	3452
West	279	1290	1569	256	1196	1452	3020
North	181	940	1121	167	870	1036	2158
East	209	1361	1569	190	1261	1451	3020
Rwanda	1578	5314	6892	1452	4926	6378	13270

* HIV testing will be carried out in the sub-sample of households selected for male survey.

IV. Selection probability and sampling weight

Due to the non-proportional allocation of the sample to the different provinces and to their districts and the possible differences in response rates, sampling weights will be required for any analysis using RDHS-V data to ensure the actual representative of the survey results at national level and as well as at domain level. Since the RDHS-V sample is a two-stage stratified cluster sample, sampling weights will be calculated based on sampling probabilities separately for each sampling stage and for each cluster. We use the following notations

- P_{1hi} : first-stage sampling probability of the i^{th} EA in stratum h
 P_{2hi} : second -stage sampling probability within the i^{th} EA (household selection)

Let a_h be the number of EAs selected in stratum h , M_{hi} the total population according to the sampling frame in the i^{th} EA, and $\sum M_{hi}$ the total population in the stratum h . The probability of selecting the i^{th} EA in the RDHS-V sample is calculated as follows:

$$\frac{a_h M_{hi}}{\sum M_{hi}}$$

Let b_{hi} be the proportion of households in the selected segment compared to the total number of households in the EA i in stratum h if the EA is segmented, otherwise $b_{hi} = 1$. Then the probability of selecting EA i in the sample is:

$$P_{1hi} = \frac{a_h M_{hi}}{\sum M_{hi}} \times b_{hi}$$

A RDHS-V cluster is either a EA or a segment of a large EA. Let L_{hi} be the number of households listed in the household listing operation in the cluster i in stratum h , let g_{hi} be the number of households selected in the 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 production of the two stages selection probabilities:

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

The design weight for each household in cluster i of stratum h is the inverse of its overall selection probability:

$$W_{hi} = 1 / P_{hi}$$

A spreadsheet containing all sampling parameters and selection probabilities will be prepared to facilitate the calculation of the design weights. Design weights will be adjusted for household non-response and as well as for individual non-response to get the sampling weights, for women and men surveys respectively. The differences of the household sampling weights and the individual sampling weights are introduced by individual non-response. The final sampling weights will be normalized in order to give the total number of un-weighted cases equal to the total number of weighted cases at national level, for both household weights and individual weights, respectively. The normalized weights are relative weights which are valid for estimating means, proportions and ratios, but not valid for estimating population totals and for pooled data. The sampling weights for HIV testing are calculated in a similar way, but the normalization of the individual sampling weights is different compared to the individual survey weights. The HIV testing weights are normalized for male and female together at national level, in order that the HIV prevalence calculated for male and female together are valid. Sampling weights for the domestic violence surveys will be calculated based on the number of eligible respondents in the households selected for domestic violence module, for male and female surveys respectively. There are a large number of sets of weights will be calculated:

- one set for all households selected for the survey
- one set for women individual survey
- one set for households selected for the male survey
- one set for male individual survey
- one set for women selected for the domestic violence survey
- one set for men selected for the domestic violence survey
- one set for women HIV testing
- one set for male HIV testing
- one set for HIV testing for children 0-14 years

It is important to note that the normalized weights are relative weights which are valid for estimating means, proportions and ratios, but not valid for estimating population totals and for pooled data. Also the number of weighted case by using the normalized weight has no direct relation with the survey precision because it is relative, especially for oversampled areas, the number of weighted cases will be much smaller than the number of un-weighted cases which are directly related to survey precision.

Sampling errors will be calculated for selected indicators for the national sample, for the urban and rural areas separately, and for each of the five provinces.