The World Health Survey (WHS)

Sampling Guidelines for Participating Countries

A. WHO Sampling Policy

This section generally outlines WHO sampling guidelines. Specific aspects of sampling design and possible country-specific exceptions are discussed in detail in the following sections.

- Surveys in the WHS program must employ a **probability sampling** design. This means that every single individual in the sampling frame has a known and non-zero chance of being selected into the survey sample. While a **Single Stage Random Sample** is ideal if feasible, it is recognized that most sites will carry out **Multi-stage Cluster Sampling**.

- The WHS sampling frame should cover 100% of the eligible population in the surveyed country. This means that every eligible person in the country has a chance of being included in the survey sample. It also means that particular ethnic groups or geographical areas may not be excluded from the sampling frame.

- The sample size of the WHS in each country is 5000 persons (exceptions considered on a by-country basis). An adequate number of persons must be drawn from the sampling frame to account for an estimated amount of non-response (refusal to participate, empty houses etc.). **The highest estimate of potential non-response and empty households should be used to ensure that the desired sample size is reached at the end of the survey period**. This is very important because if, at the end of data collection, the required sample size of 5000 has not been reached additional persons must be selected randomly into the survey sample from the sampling frame. This is both costly and technically complicated (if this situation is to occur, consult WHO sampling experts for assistance), and best avoided by proper planning before data collection begins.

- **All steps of sampling, including justification for stratification, cluster sizes, probabilities of selection, weights at each stage of selection, and the computer program used for randomization must be communicated to WHO (see checklist section H)**

B. Stratification

Stratification is the process by which the population is divided into subgroups. *Sampling will then be conducted separately in each subgroup*. Strata or subgroups are chosen because evidence is available that they are related to the outcome (e.g. health, responsiveness, mortality, coverage etc.). The strata chosen will vary by country and reflect local conditions. Some examples of factors that can be stratified on are geography (e.g. North, Central, South), level of urbanization (e.g. urban, rural), socio-economic zones, provinces (especially if health administration is primarily under the jurisdiction of provincial authorities), or presence of health facility in area. **Strata to be used must be identified by each country and the reasons for selection explicitly justified**.

Stratification is strongly recommended at the first stage of sampling. Once the strata have been chosen and justified, all stages of selection will be conducted separately in each stratum. We recommend stratifying on 3-5 factors. It is optimum to have half as many strata (**note** the difference between stratifying variables, which may be such variables as gender, socio-economic status, province/region etc. and strata, which are the combination of variable categories, for example Male, High socio-economic status, Xingtao Province would be a stratum)

Strata should be as homogenous as possible within and as heterogeneous as possible between. This means that strata should be formulated in such a way that individuals belonging to a stratum should be as similar to each other with respect to key variables as possible and as different as possible from individuals belonging to a different stratum. This maximises the efficiency of stratification in reducing sampling variance.

Example 1. Stratification

- Country X has decided to stratify on three factors:
    1) Region: North/ Center /South
    2) Socio-economic status: High/ Low
    3) Presence of health care facility: Yes/ No

- This will result in 12 strata in which sampling needs to be conducted separately:

Strata	Region	SES	Health Facility
1	North	High	Yes
2	North	High	No
3	North	Low	Yes
4	North	Low	No
5	Center	High	Yes
6	Center	High	No
7	Center	Low	Yes
8	Center	Low	No
9	South	High	Yes
10	South	High	No
11	South	Low	Yes
12	South	Low	No

B.1 Why Stratify?

- To reduce sampling error. In a stratified sample, the sampling error depends on population variance within strata but not between.

- Stratification allows subgroups of interest to be included in the sample whereas in a non-stratified sample some may have been left out due to the random nature of the selection process. Since sampling is conducted separately in each stratum, it is possible to ensure that there are a sufficient number of people in each subgroup to allow meaningful analysis.

B.2 Implicit Stratification

Implicit Stratification is the systematic sampling (*see section D.3 Systematic Sampling*) of units from an ordered list. This has the effect of stratification with respect to the variable on which the ordered list is based. Implicit stratification is recommended at the second stage of cluster selection and thereafter. The variable on which the ordered lists are based is most often geography, but may be any relevant health-related variable.

C. Multi-stage Cluster Selection

A cluster is a naturally occurring unit or grouping within the population (e.g. enumeration areas, cities, universities, provinces, hospitals etc.); it is a unit for which the administrative level has clear, non-overlapping boundaries. Cluster sampling is useful because it avoids having to compile exhaustive lists of every single person in the population.

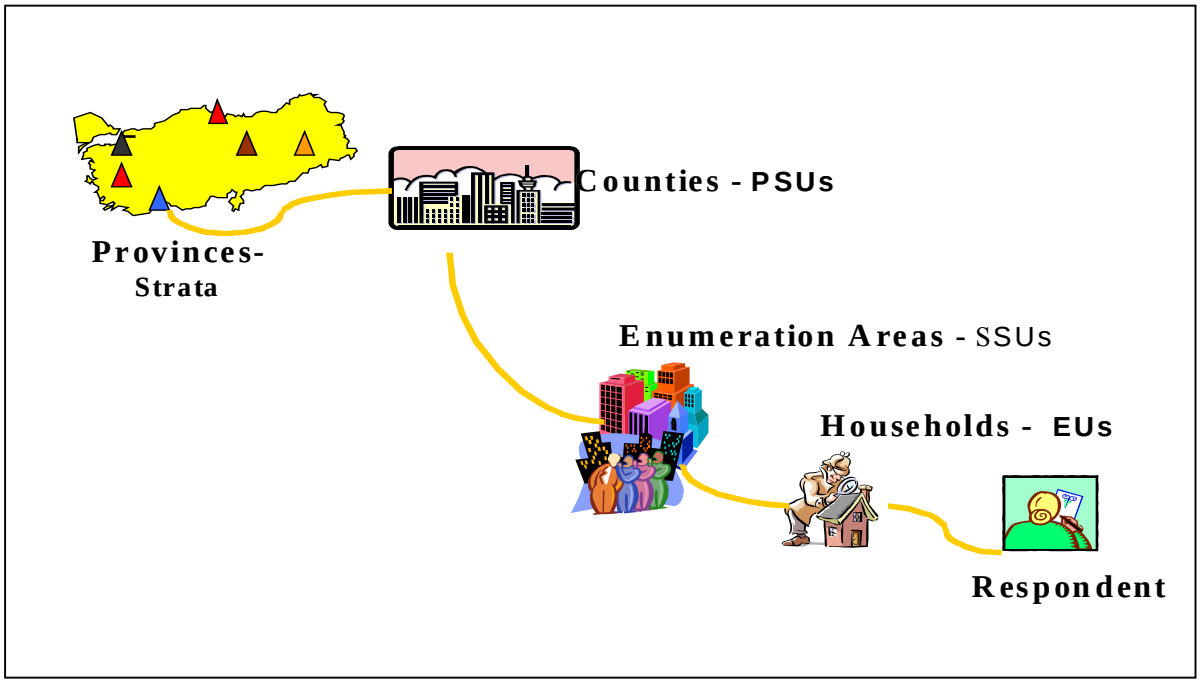
Clusters should be as heterogeneous as possible within and as homogenous as possible between (**note** that this is the opposite criterion as that for strata). Clusters should be as small as possible (i.e. large administrative units such as Provinces or States are not good clusters) but not so small as to be homogenous.

In cluster sampling, a number of clusters are randomly selected from a list of clusters. Then, either all members of the chosen cluster or a random selection from among them are included in the sample. Multi-stage sampling is an extension of cluster sampling where a hierarchy of clusters are chosen going from larger to smaller as illustrated in the following example and figure.

Example 2. Multi-stage stratified cluster sampling

- Strata: Provinces
- Primary Sampling Unit (PSU): Counties
- Secondary Sampling Unit (SSU): Enumeration Areas
- Elementary Unit (EU): Households
- Final Unit: Persons

Figure 1. Multi-stage cluster sampling



In order to carry out multi-stage sampling, one needs to know only the population sizes of the sampling units (*see section D. Probability Sampling*). For the smallest sampling unit above the elementary unit however, the TSU(neighbourhoods) in the above example, **a complete list of all elementary units (households) is needed**; in order to be able to randomly select among all households in the TSU, a list of all those households is required. This information may be available from the most recent population census. If the last census was >3 years ago or the information furnished by it was of poor quality or unreliable, the survey staff will have the task of enumerating all households in the smallest randomly selected sampling unit. It is very important to budget for this step if it is necessary and ensure that all households are properly enumerated in order that a representative sample is obtained.

It is always best to have as many clusters in the PSU as possible. The reason for this is that the fewer the number of respondents in each PSU, the lower will be the clustering effect which increases sample variance and effectively reduces our estimating power. WHO requires an absolute **maximum of 50 respondents per PSU**, and ideally would suggest 20-30. This means that for a sample size of 5000 respondents, 100-200 PSU clusters should be taken into the sample. Calculating that, roughly, one fifth of the total number of PSU clusters in a country will be randomly selected into the survey sample, the sampling frame should consist of 500-1000 PSU clusters, as illustrated by the formulas in example 3.

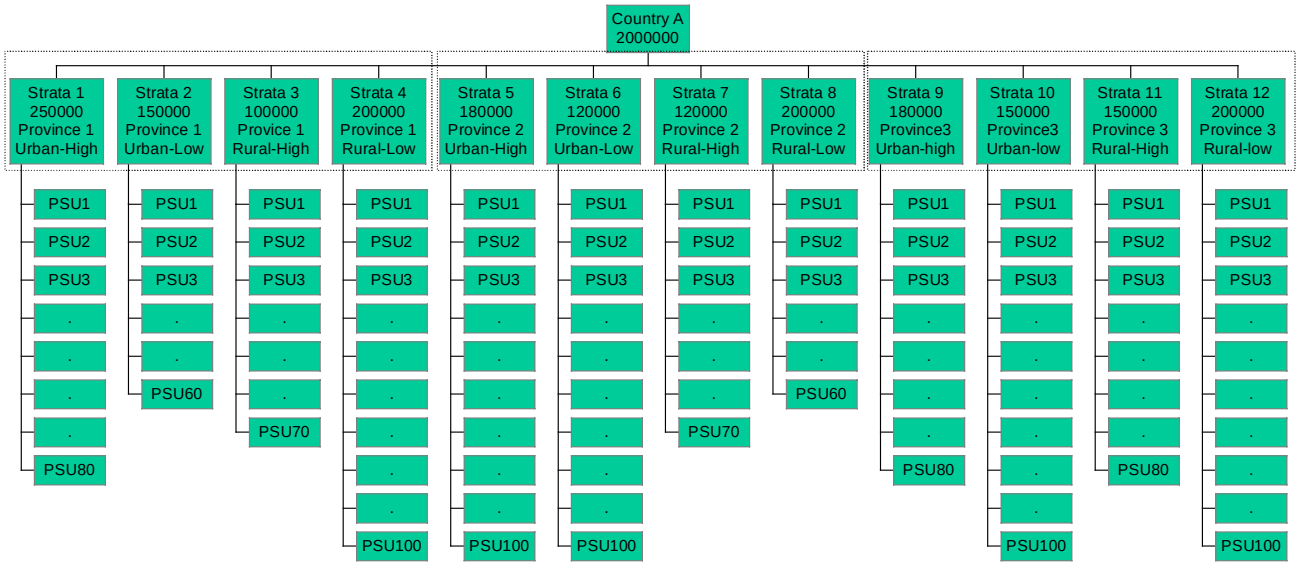
Example 3. Number of PSU clusters in Survey Sampling Frame

- *Minimum Number of PSU Clusters in Survey Sample*
  = Total Sample Size / Subjects per PSU Cluster
  = 5000 / 50
  = 100

- *PSU Clusters in Sampling Frame ≅ PSU Clusters in Survey Sample x 5*
  ≅ 100 x 5
  = 500

Example 4. Strata and PSU selection

- Country A has 3 provinces and a total population of 2 million
- It is decided to stratify by Province (1/2/3), Setting (urban/rural), and by Economic status (high/low)
- Counties are selected as the PSU- there are 1000 counties in the country



- Generally, the number of PSU should be large, the number of PSUs in the sampling frame should be at least 5 times larger than the number of PSUs randomly selected into the survey sample
- Sometimes, PSU sizes in each STRATA varies greatly. In this case, it may be desirable to further stratify PSUs by size, and to select the PSUs with probability proportional to size
- Sometimes the relative size of a PSU is very large (i.e. accounts for >50% of the Strata population). In such cases, we need to further stratify on the PSU in question, a so-called self representative or self-selecting stratum (*see section D.2 Exceptions to Probability Proportional to Size*)

D. Probability Sampling

Probability sampling means that every single individual in the sampling frame has a known and non-zero chance of being selected into the survey sample. **Non-probability methods of sampling such as quota or convenience sampling and random walk, may introduce bias into the survey, will throw survey findings into question, and are not accepted by WHO.**

D.1 Probability Proportional to Size

The probability of selection into the survey sample for each cluster will be proportional to its relative size.

- $\Pr obabilitySelection(clusterA) = Population(clusterA) \Big/ TotalPopulation(allclusters)$

This is called probability proportional to size sampling and is illustrated in the following example.

Example 5. Probability Proportional to Size (PPS) Sampling

- Country X has 11,000,000 inhabitants

- The primary sampling unit is chosen to be Counties, It has been decided that 4 PSUs out of 8 will be chosen for the sample *(note again that the number of PSU clusters in the sampling frame will typically be on the order of 500-1000 out of which 100-500 will be randomized into the survey sample; the small number of PSU clusters in this example is for illustrative purposes)*. The populations and probabilities of selection of each PSU into the sample will be as follows:

	Population Size	Probability of Selection
County 1	900000	900000/11000000 = 0.08
County 2	1200000	1200000/11000000 = 0.11
County 3	650000	650000/11000000 = 0.06
County 4	2125000	2125000/11000000 = 0.19
County 5	1650000	1650000/11000000 = 0.15
County 6	800000	800000/11000000 = 0.07
County 7	3300000	3300000/11000000 = 0.30
County 8	375000	375000/11000000 = 0.03
Sum	**11000000**	**1**

In a purely random selection process each county in the above example would have a 0.125 (1/8) chance of being selected into the survey sample. In PPS random sampling, the probabilities of selection of each cluster are weighted by the relative size of each cluster. These weighted probabilities are entered into a computer program (a PPS algorithm) which then randomly chooses 4 from the 8 PSUs. Note that every cluster (county) in the above example has a known and non-zero probability of being selected to the survey sample and sampling units with larger populations have a greater chance of being included. This same methodology is then applied to each stage of the multi-stage cluster sampling process: SSUs, TSUs etc. are all randomized using probability proportional to size sampling. Note that the probabilities of selection of sampling units at each stage must sum to one.

- **Countries must provide WHO with the population sizes, probabilities of selection and sampling weights of all sampling units for each stage of the sampling process**

Since clusters are often of unequal size, sampling weights are necessary to be able to reconstruct population estimates from our sample estimates.

- $Weight(clusterA) = {}^{1}\!/\!_{\Pr obabilitySelection(clusterA)}$

- The weights of each sampling unit in the example above would be:

	Weight	
County 1:	1 / 0.08 =	12.5
County 2:	1 / 0.11 =	9.1
County 3:	1 / 0.06 =	16.7
County 4:	1 / 0.09 =	5.3
County 5:	1 / 0.15 =	6.7
County 6:	1 / 0.07 =	14.3
County 7:	1 / 0.30 =	3.3
County 8:	1 / 0.03 =	33.3

The weights essentially describe the number of persons in the sampling frame represented by each person in the cluster (i.e. each person in County 1 represents 12.5 people, each person in County 2 represents 9.1 persons etc.). Weights for SSUs, TSUs, etc. are calculated in the same way.

The probability of selection of the elementary unit, the household, is not proportional to the number of people in the household. Rather, the household level weights will be generated at the time of respondent selection within the household (*see section E. Selection of Respondents from Household Roster*). The number of households selected within each chosen sampling unit will be proportional to the total number of households in that sampling unit. All households in each unit will have equal probability of selection as illustrated in the following example.

Example 6. Random Selection of Households

- In country X, employing a 3-stage sampling design (PSUs-Households-Individuals), 10 enumeration areas out of 50 have been selected into the sample as the primary sampling units (PSU). The following describes the number of households in each PSU, the number that will be selected into the sample and the probability of inclusion of each. (*Note again that the number of PSUs in the survey sample should be on the order of 500-1000, and also that the number of households present in enumeration areas, typically, is much smaller*).

EA	Number of Households	Number of Households (EUs) to be Sampled	Probability of Selection of Each Household
1	1800	5000 x (1800/13000) = 692	692 / 1800 = 0.38
2	2050	5000 x (2050/13000) = 788	788/2050 = 0.38
3	600	5000 x (600/13000) = 231	231 / 600 = 0.38
4	1200	5000 x (1200/13000) = 462	462 / 1200 = 0.38
5	2300	5000 x (1300/13000) = 885	885 / 2300 = 0.38
7	750	5000 x (750/13000) = 288	288 / 750 = 0.38
8	1300	5000 x (1300/13000) = 500	500 / 1300 = 0.38
9	300	5000 x (300/13000) = 115	115 / 300 = 0.38
10	1900	5000 x (900/13000) = 731	731 / 1900 = 0.38
11	800	5000 x (800/13000) = 308	308 / 800 = 0.38
Total	**13000**	**5000**	

D.2 Exceptions to Probability Proportional to Size

The following are exceptions to general WHO sampling guidelines. All exceptions must be raised and justified by countries and operationalized in consultation with WHO sampling experts.

Exception (1): Disproportionate Allocation

The first exception concerns allocation of sample sizes to strata. After stratification, the sample size of each strata is chosen proportional to the total size of that strata. There may be cases, however, where disproportionate sampling in each stratum is acceptable. Generally, this will be when the stratum comprises a very small percentage of the population and proportionate sampling would not produce sufficient numbers in that group for analysis. The following is an example of proportionate and disproportionate sample allocation in a hypothetical country.

Example 7. Disproportionate Allocation

- Country X has 1,000,000 inhabitants
- A survey is conducted that stratifies the country by socio-economic status
- In this case it is calculated that a sample size of 250 is insufficient to calculate parameters of interest in the high SES stratum. It is therefore decided to quadruple the high SES sampling fraction relative to the other two strata.

Strata	Population	Proportionate Sample	Disproportionate Sample
High SES	50,000	5,000 x (50,000 / 1,000,000) = **250**	1,000 x (0.87) = **870**
Mid SES	550,000	5,000 x (550,000 / 1,000,000) = **2,750**	2,750 x (0.87) = **2391**
Low SES	400,000	5,000 x (400,000 / 1,000,000) = **2,000**	2,000 x (0.87) = **1739**

- The sample size is quadrupled for the High SES stratum, resulting in 1,000 and, as the Mid SES sample size remains at 2,750 and the Low SES sample size remains at 2000, the total sample size becomes 5,750. In order to keep the desired sample size of 5000, the sample size in each stratum is multiplied by a factor k which is 5000/5750 = 0.87.

Exception (2): Self-selecting Sample

There may be instances where a particular cluster needs to be included into the survey sample and the random chance that it may not be included is not acceptable. For example, a city that comprises 75% of a country's population should generally be part of the survey sample. In such cases a **self-selecting sample** may be appropriate. This means that the cluster in question will have a selection probability of 1 (and therefore a weight of 1). All other clusters at each stage of the sampling procedure will still be selected by probability proportional to size sampling. Though the concept of a self-selecting sample may be confusing, it can be thought of, as stratifying on the factor of the city in question (strata 1- city X; strata 2 - rest of country).

D.3 Systematic Sampling

Systematic sampling is the ordered sampling at fixed intervals from a list, starting from a randomly chosen point. Typically, systematic sampling is not used at the first stage of sampling (selection of PSUs) because it renders the estimation of sampling error difficult. **Systematic sampling is recommended at the SSU, TSU, and household selection stages of sampling.** Systematic sampling may be linear or circular, as illustrated in the following examples.

Example 8. Linear Systematic Sampling

- Block G has been selected into the survey sample at the fourth stage of selection in country X.
- Block G has N= 289 households (which have been enumerated and listed), n= 26 of which will be included into the survey sample
- An interval size of 10 is chosen. Interval size $I \approx N/n$
- A random number generator, which **randomly generates a number between 1 and interval size n=10**, has given the number 8.
- Starting with the 8th house on the list, every tenth house is sampled (8, 18, 28, 38, 48…) until the desired sample size of 26 is reached.

Note that only households 1 to 260 are covered in this example, and the resulting sample will be biased. Hence, circular systematic sampling is preferable, as illustrated in example 9.

Example 9. Circular Systematic Sampling

- Using the same example as above
- A random number generator, which **randomly generates a number between 1 and total number of households N= 289**, has given the number 271.
- Starting with the 271st house on the list, every tenth house is sampled (271, 281, 2, 12, 22…) until the desired sample size of 26 is reached.

Note that after the 281st house, the 2nd house is sampled since in the circular listing of households, the 1st household comes after the 289th.

E. Selection of Households

The Household is a device used to get at the individual. The household is the sampling unit while the individual is the observational unit. While it would be preferable to randomly select from a list of all eligible persons in a country, such lists, with a few exceptions, are not available, so we must employ a final cluster, the household, to get at our observational units.

Households will be selected from lists of dwelling units. **Non-probabilistic methods of household selection such as the random walk are not acceptable**. Such lists are typically available from population registries, household listings, voter lists and census list. As it is essential to include all households in the sampling frame, an assessment of the methodology employed to select households must be made:

-How much has the population changed since these lists were made?
-Completeness of coverage. Are there unregistered populations (e.g. slums)
-Population shifts
-Changes in Registry Quality

Almost all lists will suffer from routine problems. WHO recommends that survey institutions manually enumerate all the households in the sampling units randomly selected into the survey sample. If existing lists or registries will be used, then a detailed analysis of their quality must be made and they must be updated to ensure that there is no exclusion of households from the survey sampling frame.

F. Selection of Individuals from Household Roster

All members of each household selected into the survey sample will be enumerated on the household roster. A member of the household is defined as someone who usually stays in the household, sleeps and shares meals, who has that address as primary place of residence, or who spends more than 6 months a year living there. **Country-specific variations in this standard definition are allowed in consultation with WHO.**

F.1 Kish Tables

The respondent for the survey will be selected among all eligible members of the household using **Kish tables**. Kish tables provide a method by which each eligible person in a household has an equal probability of selection into the survey sample.

- $\Pr obabilitySelection(respondent) = {}^{1}\!\!\diagup\!\!_{NumberPeopleInHousehold}$

- $Weight(respondent) = {}^{NumberPeopleInHousehold}\!\!\diagup\!\!_{1}$

- **It is extremely important for the representativeness of the survey sample and the integrity of the survey that the Kish tables are properly implemented. All interviews where the Kish selection method is not properly implemented will be rejected.**

The Kish technique allows adequate representation for all the persons in the household.

The following is a brief outline of Kish table implementation; sites and interviewers will be given extensive training and training materials in Kish table implementation as well as continuing support from WHO headquarters. The proper implementation of the process will be strictly enforced through the quality-control process.

Once the survey sample has been selected (e.g. 6000 households, accounting for 20% non-response), each Household will be numbered serially (e.g. from 1-6000). There are 8 Kish selection tables (A, B1, B2, C, D E1, E2, F). They will be assigned to each selected household in the following manner:

Household	Kish Table
1	A
2	A
3	B1
4	B2
5	C
6	C
7	D
8	D
9	E1
10	E2
11	F
12	F
13	A
14	A
15	B1
16	B2
17	C
18	C
19	D
20	D
21	E1
22	E2
23	F
24	F
25	A
26	A
27	B1
Etc……..	

- Interviewers will have a list of the households they are to cover and the corresponding Kish tables.

F.2 Household Roster & Non-response

The proper and complete enumeration and description of the entire household is a critical component of the survey process. **The household roster must be completed for all households selected randomly into the survey sample, whether they agreed to participate in the survey or not**. It is only in this way that we can collect the basic information required to estimate the non-response bias in the survey.

The requirement of augmenting the survey sample size to adjust for estimated non-response is necessary to ensure that we have adequate persons in the sample to have the power to make precise estimates. This does not, however, account for the bias that is created by non-response, since non-responders are often different from responders with respect to key variables that are linked to the domains under study in the survey. All effort, therefore, must be made to minimise non-response, and to interview as many people in the survey sample as possible. A detailed discussion of refusal conversion methods, survey awareness raising, and call-backs is found in the WHS Survey Manual.

There are two possible scenarios of non-response:

1) The interviewer completes the household roster and the randomly chosen respondent refuses to participate
2) The interviewer is refused access to the household and is unable to fill in the household roster

In second scenario, sites must ensure that, at least, pages 00.1 and 00.3 of the Coversheet are completed for the household. In addition, if available from census information, the number of adult (18 years of age or older) males and females in the household, and their respective ages should be provided.

It is important to note that the completion of the household roster serves a purpose above and beyond providing a list from which a respondent will be selected. The demographic and other information collected in the household roster and requested from sites serves to calculate the denominators for statistical analysis of the survey data; without the information in the household roster, we would not be able to determine the health-related outcome rates in your country.

G. Domain Coverage

G.1 Demographic Coverage

The WHS will include **all male and female adults (18 years of age and older)** who are not out of the country during the survey period.

It should be noted that this includes the population who may be institutionalized for health reasons at the time of the survey: all persons who would have fit the definition of household member at the time of their institutionalisation are included in the eligible population. If the randomly selected individual is institutionalized short-term (e.g. a 3-day stay at a hospital) the interviewer must return to the household when the individual will have come back to interview him/her. If the randomly selected individual is institutionalized long term (e.g. has been in a nursing home the last 8 years), the interviewer must travel to that institution to interview him/her.

G.2 Geographic Coverage

WHO specifies that the survey sampling frame must cover 100% of the country's eligible population, meaning that the entire national territory must be included. This does not mean that every province or territory need be represented in the survey sample but, rather, that all must have a chance (known probability) of being included in the survey sample.

There may be exceptional circumstances that preclude 100% national coverage. Certain areas in certain countries may be impossible to include due to reasons such as accessibility or conflict. **All such exceptions must be discussed with WHO sampling experts**. If any region must be excluded, it must constitute a coherent area, such as a particular province or region. For example if ¾ of region D in country X is not accessible due to war, the entire region D will be excluded from analysis.

H. Geographic Information Systems (GIS)

Geographical data is increasingly being recognized as an important factor in modelling health and health-related outcomes. Geographic information systems combine positional data such as latitude and longitude with health-related geographic information, such as location of hospitals, bodies of water, roads etc. The WHS will collect geographical information from sampled areas in order to be able to assess and analyze health and health-related outcomes in complex ways. For this purpose, survey supervisors in each country are asked to fill the questionnaire section entitled Sampling and GIS information for each survey respondent as well as providing sampling unit maps and administrative unit details, as explained below.

H.1 GIS Data Collected in the WHS

1. Sampling Unit Names

-The name of all sampling units above the household level that the respondent belongs to.

e.g. Respondent X:
PSU- Boropa (County)
SSU- Tangey (Village)
TSU- Lalli (Neighbourhood)

2. Census Information

-If the sampling was conducted using information from a census, the census unit number (the numeric identification code given to a census unit by the institution that conducted the survey) of the Enumeration Area selected and the date of the census is requested.

3. Sampling Unit Maps

-Maps of all the sampling units randomly selected into the survey sample should be provided

-Maps must include: 1) A minimum of two reference points (latitude/longitude) and/or
2) The delineation (border outline) of the sampling unit

-Three types of maps may be provided:

1) *Digital maps:*
   Computerized maps; may be in ArcView shape or ArcInfo export file.

2) *Census unit delineation maps:*
   Maps representing the extension of the considered census unit (or cluster)

3) *Sketch maps:*
   A representative picture or abstract diagram that distills the key components of a landscape

-It is likely that maps of administrative units within a country are readily available. Since in most cases sampling units will correspond to administrative units (counties, cities, villages etc.), this will pose no problem. Where the sampling unit is not an administrative level unit a linkage between the two should be provided as outlined in #5 below.

4. Setting

-The setting for the smallest sampling unit above the household for a respondent may be *urban*/ *peri-urban*/ *semi-rural*/ *rural*/ *other*. The first four categories may be defined by each site; the definitions used must be communicated to WHO to allow for cross-country comparisons. The *other* category is to allow for sites to use categories beyond the four provided.

5. Administrative level unit linkage

-Administrative units are areas within a country that are under the governance of an administration.

Example 8. Administrative Level Unit Names in the United States

1st Administrative Level Unit Name:	State
2nd Administrative Level Unit Name:	County
3rd Administrative Level Unit Name:	Minor civil division (Town)

-WHO will provide participating countries with a list of administrative level unit names for that country. In many cases we may only be able to provide lists for the 1st and 2nd administrative levels; lists should be checked with local authorities before use. **Sites should verify the information provided and bring the lists up-to-date if necessary, communicating all such changes to WHO**.

-Where the sampling unit chosen at each stage corresponds to an administrative level unit simply indicate this when providing WHO with the list of clusters in the sampling frame and survey sample.

-Where the sampling unit chosen is not an administrative level unit, a linkage between the two needs to be provided, by indicating the smallest administrative level unit that contains the sampling unit, as illustrated below.

Table 1. Linking Sampling Unit and Administrative Level Unit Information

PSU level Name	**Administrative level unit? (Yes/No)**		
1	Yes →	Administrative unit level (1st, 2nd, 3rd 4th)	
	No →	Name of Administrative level unit containing primary sampling unit	Administrative unit level (1st, 2nd, 3rd 4th)
2	Yes →	Administrative unit level (1st, 2nd, 3rd 4th)	
	No →	Name of Administrative level unit containing primary sampling unit	Administrative unit level (1st, 2nd, 3rd 4th)
3	Yes →	Administrative unit level (1st, 2nd, 3rd 4th)	
	No →	Name of Administrative level unit containing primary sampling unit	Administrative unit level (1st, 2nd, 3rd 4th)
.			
.			
.			
SSU level Name	**Administrative level unit? (Yes/No)**		
1	Yes →	Administrative unit level (1st, 2nd, 3rd 4th)	
	No →	Name of Administrative level unit containing primary sampling unit	Administrative unit level (1st, 2nd, 3rd 4th)
2	Yes →	Administrative unit level (1st, 2nd, 3rd 4th)	
	No →	Name of Administrative level unit containing primary sampling unit	Administrative unit level (1st, 2nd, 3rd 4th)
3	Yes →	Administrative unit level (1st, 2nd, 3rd 4th)	
	No →	Name of Administrative level unit containing primary sampling unit	Administrative unit level (1st, 2nd, 3rd 4th)
.			
.			
.			
TSU level Name	**Administrative level unit? (Yes/No)**		
1	Yes →	Administrative unit level (1st, 2nd, 3rd 4th)	
	No →	Name of Administrative level unit containing primary sampling unit	Administrative unit level (1st, 2nd, 3rd 4th)
2	Yes →	Administrative unit level (1st, 2nd, 3rd 4th)	
	No →	Name of Administrative level unit containing primary sampling unit	Administrative unit level (1st, 2nd, 3rd 4th)
3	Yes →	Administrative unit level (1st, 2nd, 3rd 4th)	
	No →	Name of Administrative level unit containing primary sampling unit	Administrative unit level (1st, 2nd, 3rd 4th)
.			
.			
.			

6. Positional Data

-Positional data (Latitude/Longitude/Altitude) is requested for the smallest sampling unit above the household level. This information will be available at the local geographical survey office. **It is very important that the specific location at which the positional measurement is taken (the waypoint) is identified.**

-The GPS data should be collected in WGS84 datum, and in decimal degrees with 5 places of precision (dd.ddddd). In order to maintain consistency we are strongly discouraging the use of other datum and/or coordinate systems. If it is not possible to fulfill this requirement, a clear indication of the datum and coordinate system used must be specified in the questionnaire.

Definitions of the fields to be filled

Latitude:	The angular distance of a point on the earth's surface along a meridian North or South of the equator.
Longitude:	The angular distance of a point on the earth's surface East or West of the Greenwich meridian (Greenwich, England).
Altitude:	Height of a point on the earth's surface above or below (+/-) sea level
Waypoint:	Point on the ground where a Latitude/Longitude/Altitude measurement was taken.

I. Information for WHO Headquarters Checklist

Please ensure that all of the following information has been compiled and sent to WHO headquarters:

1- Strata

- -Names ☐
- -Justification ☐
- -Population size of each strata ☐
- -Number of persons randomly selected into in each strata ☐

2- Sampling Units

- -Number of stages of selection ☐
- -Names of stages of selection ☐
- *-PSU*
  - Complete list of PSU names ☐
  - Complete list of PSU population sizes ☐
  - Complete list of PSU selection probabilities ☐
  - Complete list of PSU sampling weights ☐
  - Complete list of PSUs randomly selected into survey sample ☐
- *-SSU*
  - Complete list of SSU names ☐
  - Complete list of SSU population sizes ☐
  - Complete list of SSU selection probabilities ☐
  - Complete list of SSU sampling weights ☐
  - Complete list of SSUs randomly selected into survey sample ☐
- *-TSU*
  - Complete list of TSU names ☐
  - Complete list of TSU population sizes ☐
  - Complete list of TSU selection probabilities ☐
  - Complete list of TSU sampling weights ☐
  - Complete list of TSUs randomly selected into survey sample ☐
- *-QSU*
  - Complete list of QSU names ☐
  - Complete list of QSU population sizes ☐
  - Complete list of QSU selection probabilities ☐
  - Complete list of QSU sampling weights ☐
  - Complete list of QSUs randomly selected into survey sample ☐

3- Randomization Program

-Name of computer program used to operationalize randomization ☐

4- Census

-Year of census ☐

-Additional information used to update census ☐

-Enumeration area census unit numbers ☐

5- GIS Maps

Primary Sampling Unit

Census unit delineation map ☐

Sketch map ☐

Other (specify) ☐

Secondary Sampling Unit

Census unit delineation map ☐

Sketch map ☐

Other (specify) ☐

Tertiary Sampling Unit

Census unit delineation map ☐

Sketch map ☐

Other (specify) ☐

Quaternary Sampling Unit

Census unit delineation map ☐

Sketch map ☐

Other (specify) ☐