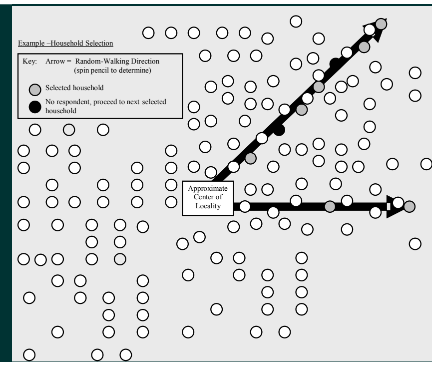
WFP
World Food Programme

Thematic Guidelines

vam
vulnerability analysis and mapping



Sampling Guidelines
for Vulnerability Analysis

Sampling

December 2004

ODAV (VAM) – WFP, Rome

Prepared by Greg A. Collins

For any questions, queries and feedback please contact the following:

Greg Collins, VAM consultant gacollins@ucdavis.edu

Eric Kenefick, Program Advisor, VAM eric.kenefick@wfp.org

Table of Contents

Acronyms and abbreviations

CO	Country Office
EPI	Expanded Program on Immunization
PPS	Probability Proportional to Size
SI	Sampling Interval
SRS	Simple Random Sampling
SYS	Systematic Sampling
UNHCR	United Nations High Commission for Refugees
UNICEF	United Nations Children's Fund
WFP	World Food Programme

Introduction

These guidelines have been designed to assist WFP Country Offices and their partners in choosing appropriate sampling methods for conducting food security and vulnerability studies. Although ideal sampling procedures are widely agreed upon, ideal situations are seldom encountered in the field. Accordingly, the guidance provided in this document was designed with the typical constraints and limitations faced by field staff in mind.

The document has been organized in sections that correspond with the decision-making process involved in developing a sampling strategy. Within each section, detailed guidance and examples are provided.

After describing basic sampling terms and concepts in Section I, Section II presents a decision-tree to assist readers in choosing an appropriate sampling method giving the conditions and objectives of study they wish to undertake. The decision tree asks a series of questions to help field staff identify the most relevant sampling options given the objectives of the assessment and the information available about the population. Once the relevant method options have been identified, proceed to the method sub-sections (simple random sampling, systematic sampling, cluster sampling, two-stage cluster sampling and multi-stage sampling) for more detailed guidance on choosing and applying the appropriate option.

Although the guidelines were designed to cover a wide range of scenarios, it is impossible to predict every constraint and limitation encountered. Additional technical assistance is available through VAM regional and headquarters staff.

Section I - Key terms and concepts

This section introduces key concepts and terms associated with sampling.

1.1 - Sampling

The term sampling refers to the selection of a limited number of individual units of analysis (denoted as n) from a population of interest (denoted as N) with the purpose of inferring something about that population from the individual units selected in the sample. Households are the most common units of analysis in VAM food security assessments[1]. Sampling is used in VAM food security assessments because total enumeration of all of the households in the population (as in a census) is too costly and too time consuming.

> Probability methods are appropriate when the objective of the assessment is to determine the percentage or number of people who are food insecure.

There are two broad categories of sampling relevant to VAM food security assessments: probability sampling (also called formal sampling) and non-probability sampling (also called informal sampling).

1.1.1 - Probability sampling

Probability sampling relies upon probability theory to draw statistical inferences about the population of interest from a randomly selected sample. Because probability sampling employs random selection techniques it is more objective than non-probability sampling. Probability sampling also allows for the degree of error around food security estimates to be quantified.

Example An assessment employing probability sampling methods estimates that 28% (+/- 4 percentage points) households in the peri-urban areas outside of Port au Prince, Haiti consume less than two meals per day. In other words, based on a sample survey, the estimated percentage of households in the peri-urban area outside Port au Prince, Haiti consuming less than two meals per day is between 24% and 32 percent[2].

The types of probability sampling discussed in these guidelines include:

- simple random sampling
- systematic sampling
- cluster sampling
- two-stage cluster sampling
- multi-stage sampling

1.1.2 - Non-probability sampling

Non-probability sampling relies on a more subjective means of inferring something about the population of interest from a sample. Sample households or individuals are selected because there is reason to believe that they 'represent' the population well or that they are well positioned to provide information about the population (as with key informants). Other non-probability methods select sample households or individuals as a matter of convenience. The inherent subjectivity and bias associated with non-probability methods is both its strength and its weakness.

Example To understand the flow of livestock from southern Somalia into Kenya in-depth discussions are held with a few strategically selected traders (purposive, non-probability sampling). In this case, it makes more sense to select individuals who are knowledgeable than to randomly select individuals that may or may not know how cross-border trade networks work.

---

[1] By contrast, nutritional surveys that collect anthropometric data normally treat individuals within households as the unit of analysis. Combined food security and nutritional surveys may use a combination of household and individual level analyses.
[2] This range estimate is known as a *confidence interval* and is discussed in detail in the section entitled Determining the Appropriate Sample Size.

Non-probability sampling methods are appropriate for meeting many of WFP's information needs. Beneficiary Contact Monitoring (BCM) provides a prominent example. However, they lack the necessary objectivity and quantification of error around the estimates that are required to meet the primary objective of most VAM food security assessments: to quantify the percentage of households that are food insecure within defined populations and sub-populations. Therefore, these guidelines focus exclusively on probability (or formal) sampling methods[3].

1.2 - Sampling frames

A sampling frame is an exhaustive list of all sampling units[4] and their physical locations within the population of interest (N). The purpose of constructing a sampling frame is to ensure that each household within the population of interest has an equal or known probability of being randomly selected for inclusion in the sample. Random selection of sampling units from a sampling frame allows for estimates from the sample population (n) to be generalized to the larger population of interest (N) defined by the sampling frame.

> Sampling Frames ensure that every household in the population of interest has an equal chance of being included in the sample.

In practice sampling frames that are 100% complete and accurate do not exist. Recognizing this reality, the sampling frames constructed for VAM food security assessments should strive to be as accurate and complete as possible, but should rely primarily on pre-existing data sources rather than primary data collection[5]. Government census data or demographic data from other surveys are among the most useful data sources for constructing sampling frames.

It is important to be transparent about groups or areas that are intentionally left out of the sampling frame because population (N) level estimates generated by the sample population (n) do not apply to these groups. Security is perhaps the most common reason for intentionally excluding groups or areas. However, in practice, some individual households or villages will be omitted from the sampling frame unintentionally. Although strictly speaking estimates derived from the sample population (n) cannot be used to generalize about these households, a limited number of chance omissions will not undermine the validity of an assessment's findings.

1.3 - Primary and ultimate sampling units

The sampling units listed in the sampling frame are the primary sampling units. In some cases, such as long-term refugee camps or countries in which a detailed census has recently been conducted, a reasonably accurate sampling frame of all households and their locations is available or can easily be constructed. In these cases, households listed in the sampling frame are both the primary sampling units and the desired units of analysis (also known as ultimate sampling units).

> Households are the most common ultimate sampling unit in food security assessments. Villages are the most common primary sampling unit.

However, in many cases a complete list of households for a population of interest is unavailable and would be costly and time consuming to construct. In these cases, the sampling frame is constructed at the lowest aggregation of households for which accurate information on the existence, location, and relative size[6] of aggregates is available. In rural settings, this aggregation is often villages such that an exhaustive list of villages (primary sampling units) within the population of interest can be constructed. In urban settings, neighborhoods or blocks often provide a suitable aggregation of households and can be used when constructing a sampling frame. Households (the most

---

[3] Some VAM food security assessments use a combination of both probability methods and non-probability methods, drawing on the strengths of each for different information needs.
[4] See Primary and Ultimate Sampling Units for a detailed explanation.
[5] In this instance, primary data collection refers to population data collected in the field by WFP for the purpose of constructing a sampling frame. By contrast, secondary data refers to pre-existing data collected for another purpose that can be used to construct a sampling frame.
[6] The utility of size estimates is discussed in detail under Cluster Sampling, Two-Stage Sampling, and Multi-Stage Sampling.

common unit of analysis in VAM food security assessments) remain the ultimate sampling units[7].

Several options exist for choosing households (the ultimate sampling unit) for inclusion in the sample when the primary sampling units are an aggregation of households such as a village or neighborhood/block. The choice of a particular method of household selection is driven by the information available and time/cost constraints. Guidance on choosing an appropriate household selection method is described in detail under each of the five sampling method described in the next section.

1.4 - Stratification or stratified sampling

Stratification or stratified sampling involves dividing the population of interests into sub-groups (e.g. strata) that share something in common based on criteria related to the assessment objectives[8]. Stratification is used when separate food security estimates are desired at a pre-defined, minimum level of precision for each of these sub-groups. When used appropriately, stratification also increases the precision of overall food security estimates for the population of interest.

> Consider stratified sampling when comparing sub-groups within the population of interest is an important objective of the assessment

Stratification by administrative boundaries allows for separate estimates to be generated for disaggregated areas within a population. For example, a national sample may be stratified by district in order to ensure the precision of food insecurity estimates at the district level for comparative purposes.

However, stratification is most effective when it is used to define sub-groups within the population that share characteristics related to food security. Livelihoods and land-use zones are examples. Defining groups in this way serves two functions. First, administrative boundaries rarely correspond with household characteristics that are related food insecurity and estimates for administrative aggregations are likely to mask meaningful differences between sub-groups. Second, defining sub-groups for stratification using criteria that are related to food insecurity improves the precision/accuracy of both sub-group and overall food security estimates[9].

Example The estimated percentage of food insecure households for Garissa, Kenya, a rural district containing both nomadic pastoralists and sedentary farmers, is 35% (+/- 5 percentage points). However, this average at the district level masks the fact that 70% of pastoralists are food insecure and only 10% of sedentary farmers are food insecure.

Stratified sampling requires that each sub-group (stratum) must be mutually exclusive; meaning that every household in the population of interest must be assigned to only one sub-group. The strata should also be collectively exhaustive; meaning every household in the population of interest must belong to a sub-group.

Despite the clear advantages of stratified sample for generating meaningful sub-group estimates and improving overall precision, several practical considerations may limit its use. First, stratifying a population into sub-groups using criteria related to food security requires pre-existing information about those sub-groups. In order to take the sample, the location of the sub-group must be known and households within the sub-group must be identifiable. This is often made difficult by the fact that information is most often found

---

[7] In rare cases it may be necessary to have multiple levels of sampling units. For example, if no information on villages and their location is available, a higher aggregate, such as a district, may be used. In this example, district is the PSU, villages are the secondary sampling unit (SSU) and households (the desired unit of analysis) remain the USU. A more detailed discussion of this issue is provided in the section entitled <u>Multi-Stage Sampling</u>.

[8] The purpose of stratification is to define homogenous sub-groups within a heterogeneous population for comparison and to increase the overall precision of estimates derived from the sample.

[9] Stratification by sub-groups defined by criteria related to food security result in more homogenous groupings in terms of food security outcomes. The result is an increase in the precision/accuracy of estimates for each sub-group and the combined overall estimate for the population by reducing sampling error. By contrast, stratification by administrative boundary is likely to result in heterogeneous groupings similar to the heterogeneity found in the overall population under study.

for administrative aggregations (districts, divisions, provinces, departments, etc.) and different sub-groups defined by livelihoods overlap with one another within administrative boundaries.

Second, each additional sub-group (e.g. stratum) represents an increase in cost and time required to conduct the assessment. Therefore, cost and time constraints will figure heavily into if and how a sample can be stratified. If the sample size required for a district level of estimate at a reasonable level of precision is 200 households, stratifying the district into two livelihood sub-groups would require applying the same sample size to each of the two livelihood groups if the same level of precision was desired for each sub-group (200 * 2 = 400). If the criteria used to define strata results in three sub-groups, the sample size is tripled. For four groups, the sample size is quadrupled and so on. Stratification by two or more criteria results in a minimum of four strata (2 criteria, each defining 2 groups) and will increase an assessment's costs substantially.

Example A food security assessment in Haiti was originally designed to yield district level estimates for four districts (four strata). The estimated sample size required was 400 households per district for a total of 1600 households.

Upon further reflection, the Country Office decided to stratify by major land-use zones within each district (stratifying by two criteria). Land-use maps suggested that two of the districts had four land-use zones and the other two districts had three land-use zones for a total of 14 land-use zones. Rather than apply the sample size of 400 to each zone (n = 5,600), the desired precision of the estimates was relaxed such that t overall sample size required was 2,440 households.

Given these practical limitations, it will not be possible to stratify a sample by every comparison that you wish to make during analysis. But, if a sub-group is well represented in the population it is likely that a sufficient number of households within that sub-group will be randomly selected. As a result, a fairly precise estimate of the food security status of the sub-group can be generated during analysis without pre-stratifying the sample.

For example, almost all VAM food security assessments will compare the percentage of food secure households among female and male headed households. However, few, if any, of these assessments stratify on the basis of the gender of the head of household.

Why? First, in most contexts the gender of the head of household can only be determined by asking the household or a neighboring household, meaning that extensive fieldwork would be required to create separate sampling frames for male and female headed households. Second, although food security comparisons by gender of the head of household are important, they rarely are the primary comparison objective for a VAM food security assessment and the cost associated with adding an additional stratification criterion is usually prohibitive. Third, even if the minimum precision of estimates for female headed households is not pre-determined by stratification, it is likely that the sample will contain a proportion of female headed households similar to that found in the population. Where female headed households represent a significant proportion of all households, the sample size will be large enough to generate food security estimates for this sub-group with reasonable levels of precision.

With these limitations and constraints in mind, stratified sampling should be reserved for those instances when all four of the following criteria are met:

- Sub-group food security estimates are a critical part of the assessment's objectives
- A minimum level of precision around the food security estimates for these sub-groups is required to meet the assessment objectives such that a guaranteed minimum sample size from each sub-group is required
- The predicted sub-group sample size suggests that estimates for sub-groups will not be precise enough to meet assessment objectives.
- Pre-existing information can be used to construct separate sampling frames for each sub-group defined by the stratification criteria.

Section II - Choosing an appropriate sampling method

A variety of probability sampling methods exist to suit different situations encountered in the field. The most commonly used methods during VAM food security assessments involve one or more of the following methods: simple random sampling, systematic sampling, cluster (or area) sampling, two-stage cluster sampling, and, on rare occasion, multi-stage sampling.

> Two-stage cluster sampling is the most frequently used sampling method for food security assessments. However, opportunities to use more cost effective methods such as simple random or systematic sampling are often missed.

The decision tree on the next page asks a series of questions to help identify the appropriate sampling method(s) given the available information and the objectives of the assessment. Once the appropriate method(s) has/have been identified, proceed to the appropriate section(s) for a detailed explanation of when and how to apply a particular method.

Although stratified sampling is often treated as a method unto itself, the choice to stratify or not stratify a sample is in many ways independent of the choice of between the five probability sampling methods above. In other words, stratification can be used in combination with any of the five sampling methods. Accordingly, the first question in the decision tree accesses whether or not the sample will be stratified before moving on to choosing an appropriate sampling method.

2.1 - Simple random sampling

As the name implies, Simple Random Sampling (SRS) is the most straightforward of the probability sampling methods. A simple random sample involves the random selection of households from a complete list of all households within the population of interest (e.g. sampling frame). Households are therefore both the primary and ultimate sampling units. Simple random sampling has a statistical advantage over other sampling methods[1] and requires a smaller sample size (approximately half of the sample size required for cluster or two-stage cluster sampling).

2.1.1 When to apply simple random sampling

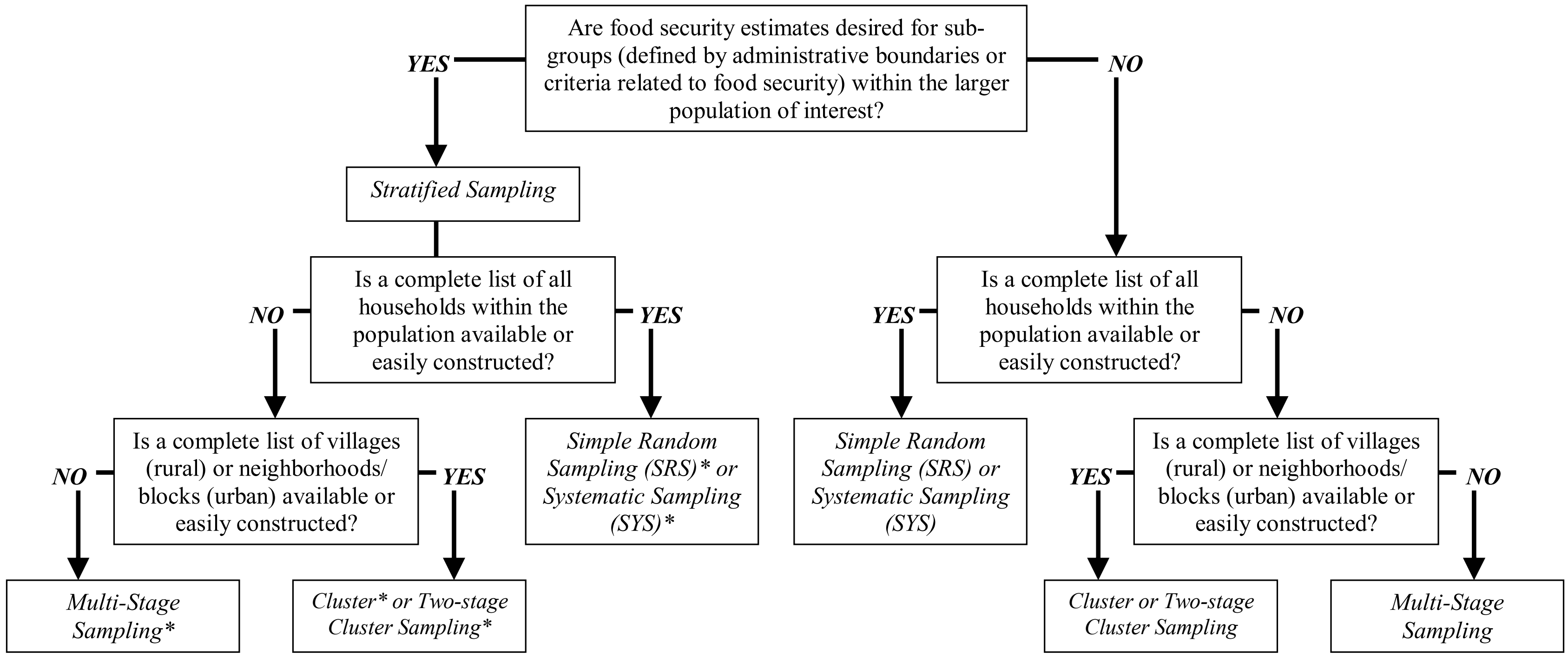
In practice, household level sampling frames are rarely available. However, assessments conducted in long-term refugee camps or areas in which a census has recently been conducted may provide enough information at the household level to construct one.

Despite the statistical advantage and reduced sample size requirements, the existence of a household level sampling frame does not mean that simple random sampling is always the most appropriate method. Because households are selected randomly from the population, the list of households included in the sample can be widely dispersed and may require visiting a large number of villages to collect the sample.

By comparison, cluster and two-stage cluster sampling limit the number of villages to be visited and may present a logistical advantage over simple random sampling. When the area being covered by an assessment is large, cluster or two-stage cluster sampling may be more cost effective despite the larger sample size requirements.

[1] Systematic Sampling shares in this advantage.

Figure 1 - Decision Tree: Choosing an Appropriate Sampling Method



* Apply sampling method and sample size calculation to each sub-group (strata) defined by the stratification criteria

2.1.2 - How to apply simple random sampling

Step 1 - Each household in the sampling frame is assigned a unique number between and the total number of households in the sampling frame. For stratified samples, a separate sampling frame must be developed for each stratum (e.g. sub-groups defined by stratification criteria).

Step 2 - A randomization method is then used to select households for inclusion in the sample[2]. The website http://www.randomizer.org/form.htm provides an easy-to-use random numbers generator.

> ***Example*** Each household in a sampling frame containing 5,500 households was assigned a number (from 1 to 5,500). The random numbers generator form available through randomizer.org was then used to select the 300 randomly selected households for inclusion in the sample (enter values into each field). Suggested default values for format fields are provided in the example form below.

To generate a set of random numbers, simply enter your selections (integer values only):

How many sets of numbers do you want to generate? 1
Help

How many numbers per set? 300
Help

Number range (e.g., 1-50):
From: 1
To: 5500
Help

Do you wish each number in a set to remain unique? Yes
Help

Do you wish to sort your outputted numbers? Yes: Least to Greatest
Help

How do you wish to view your outputted numbers? Place Markers Off
Help

Research Randomizer Results

1 Set of 300 Unique Numbers Per Set
Range: From 1 to 5500 -- Sorted from Least to Greatest

[2] The total number of households to be randomly selected from the sampling frame is determined by the sample size requirements (see Section III).

Set #1:

7, 23, 37, 40, 44, 68, 98, 120, 123, 124, 144, 172, 176, 194, 223, 259, 267, 272, 274, 280, 310, 337, 354, 379, 414, 446, 505, 521, 523, 543, 556, 559, 571, 633, 660, 666, 688, 730, 738, 749, 787, 794, 872, 879, 902, 903, 923, 935, 946, 967, 991, 997, 1019, 1092, 1142, 1153, 1172, 1182, 1202, 1233, 1284, 1289, 1320, 1325, 1336, 1351, 1367, 1416, 1427, 1438, 1453, 1491, 1516, 1541, 1542, 1601, 1639, 1659, 1674, 1690, 1708, 1710, 1715, 1775, 1789, 1810, 1818, 1819, 1849, 1869, 1964, 1968, 1973, 1979, 2019, 2020, 2055, 2059, 2066, 2128, 2135, 2182, 2188, 2200, 2226, 2229, 2275, 2285, 2316, 2320, 2361, 2365, 2425, 2441, 2465, 2477, 2487, 2497, 2499, 2525, 2531, 2546, 2556, 2560, 2563, 2580, 2622, 2640, 2662, 2665, 2677, 2694, 2717, 2761, 2764, 2770, 2779, 2828, 2829, 2834, 2855, 2873, 2912, 2930, 2939, 2985, 2995, 3030, 3032, 3040, 3055, 3061, 3068, 3076, 3097, 3115, 3122, 3161, 3166, 3172, 3186, 3195, 3215, 3217, 3218, 3249, 3260, 3281, 3290, 3345, 3347, 3365, 3368, 3384, 3390, 3399, 3404, 3430, 3444, 3457, 3459, 3462, 3464, 3481, 3484, 3491, 3500, 3519, 3566, 3570, 3579, 3590, 3606, 3651, 3659, 3660, 3670, 3735, 3736, 3743, 3773, 3794, 3795, 3798, 3810, 3832, 3837, 3859, 3863, 3877, 3881, 3896, 3908, 3915, 3946, 3962, 4024, 4030, 4055, 4116, 4118, 4126, 4131, 4135, 4148, 4190, 4230, 4288, 4299, 4319, 4334, 4358, 4365, 4368, 4385, 4445, 4464, 4492, 4516, 4519, 4529, 4537, 4564, 4597, 4598, 4607, 4624, 4625, 4627, 4637, 4649, 4652, 4664, 4671, 4675, 4693, 4721, 4727, 4742, 4836, 4850, 4860, 4865, 4887, 4901, 4934, 4958, 4973, 5017, 5032, 5054, 5068, 5072, 5081, 5088, 5096, 5150, 5175, 5185, 5199, 5203, 5208, 5216, 5250, 5273, 5285, 5287, 5338, 5356, 5357, 5358, 5369, 5382, 5402, 5404, 5410, 5413, 5445, 5478, 5490

Step 3 - The selected households are then noted in the sampling frame.

1	38	75	112	149	186	**223**	260	297
2	39	76	113	150	187	224	261	298
3	**40**	77	114	151	188	225	262	299
4	41	78	115	152	189	226	263	300
5	42	79	116	153	190	227	264	301
6	43	80	117	154	191	228	265	302
7	**44**	81	118	155	192	229	266	303
8	45	82	119	156	193	230	**267**	304
9	46	83	**120**	157	**194**	231	268	305
10	47	84	121	158	195	232	269	306
11	48	85	122	159	196	233	270	307
12	49	86	**123**	160	197	234	271	308
13	50	87	**124**	161	198	235	**272**	309
14	51	88	125	162	199	236	273	**310**
15	52	89	126	163	200	237	**274**	311
16	53	90	127	164	201	238	275	312
17	54	91	128	165	202	239	276	313
18	55	92	129	166	203	240	277	314
19	56	93	130	167	204	241	278	315
20	57	94	131	168	205	242	279	316
21	58	95	132	169	206	243	**280**	317
22	59	96	133	170	207	244	281	318
23	60	97	134	171	208	245	282	319
24	61	**98**	135	**172**	209	246	283	320
25	62	99	136	173	210	247	284	321
26	63	100	137	174	211	248	285	322
27	64	101	138	175	212	249	286	323
28	65	102	139	**176**	213	250	287	324
29	66	103	140	177	214	251	288	325
30	67	104	141	178	215	252	289	326
31	**68**	105	142	179	216	253	290	327
32	69	106	143	180	217	254	291	328
33	70	107	**144**	181	218	255	292	329
34	71	108	145	182	219	256	293	330
35	72	109	146	183	220	257	294	331
36	73	110	147	184	221	258	295	332
37	74	111	148	185	222	**259**	296	333

Households 334 to 5349 are removed for ease of presentation only

5350	5387	5424	5461	5498
5351	5388	5425	5462	5449
5352	5389	5426	5463	5500
5353	5390	5427	5464	
5354	5391	5428	5465	
5355	5392	5429	5466	
5356	5393	5430	5467	
5357	5394	5431	5468	
5358	5395	5432	5469	
5359	5396	5433	5470	
5360	5397	5434	5471	
5361	5398	5435	5472	
5362	5399	5436	5473	
5363	5400	5437	5474	
5364	5401	5438	5475	
5365	**5402**	5439	5476	
5366	5403	5440	5477	
5367	**5404**	5441	**5478**	
5368	5405	5442	5479	
5369	5406	5443	5480	
5370	5407	5444	5481	
5371	5408	**5445**	5482	
5372	5409	5446	5483	
5373	**5410**	5447	5484	
5374	5411	5448	5485	
5375	5412	5449	5486	
5376	**5413**	5450	5487	
5377	5414	5451	5488	
5378	5415	5452	5489	
5379	5416	5453	**5490**	
5380	5417	5454	5491	
5381	5418	5455	5492	
5382	5419	5456	5493	
5383	5420	5457	5494	
5384	5421	5458	5495	
5385	5422	5459	5496	
5386	5423	5460	5497	

Households included in Sample (highlighted)

Step 4 - Next, selected households are mapped to facilitate data collection. Importantly, the data collection team must also have a household replacement strategy for the households in which a) the household cannot be located (inaccurate information in the sampling frame) or b) an appropriate respondent is not available.

Step 5 - Replacement households can be pre-selected prior to data collection using the sampling frame by identifying the next household in the sampling frame as the replacement household. Alternatively, a protocol[3] for replacing households in the field can be agreed upon prior to data collection. Examples include choosing the next closest household or spinning a pencil in front of the absentee household to select a transect line and choosing the first house encountered in that line as the replacement household. The means of household replacement is less important than the uniform application of whatever procedure is chosen.

2.1.2.1 - Example applications of simple random sampling

Western Tanzania - A food security assessment in a Western Tanzania refugee camp housing Congolese refugees requires a sample size of 400 households. A list of all households within the camp is available from UNHCR, along with maps locating each household within a block and each block within the camp.

Each household within the camp is assigned a number between 1 and 5,050 (the total number of households in the camp). A random numbers generator (www.randomizer.org) is used to select four hundred households. The selected households are then mapped. The workload is divided among four data collection teams with each team given a mapped area containing approximately 100 households.

Given the proximity of households to one another within the camp, data collection teams are able to walk between selected households. Households that are non-existent or that do not have a suitable respondent available at the time of data collection are replaced by choosing the closest household to the mapped location of the original household selected.

Southern Malawi - A simple random sample of 300 households from a sampling frame containing 10,000 households throughout southern Malawi resulted in having to visit 200 different villages (100 villages contain only one selected household each and 100 villages contain two selected households each for a total of 200 villages and 300 households).

However, the expense and time associated with driving to 200 villages, many of which are geographically remote, forces the assessment team to reconsider its method choice. A decision is made to use a two-stage cluster sampling method. The change in method requires a doubling of sample size to 600 households, but greatly reduces the number of villages to be visited. At the first stage of selection, 30 villages are selected randomly from a list of all villages within the population of interest. At the second stage, 20 households are selected from the household lists for each of the 30 selected villages (see 2.4 for a detailed explanation of this method).

2.2 - Systematic sampling

Systematic sampling shares the same information requirements as simple random sampling. In contrast to random selection, this method involves the systematic selection of households from a complete list of all households within the population of interest (e.g. sampling frame). Once again, households are both the primary and ultimate sampling units. Like simple random sampling, systematic sampling has a statistical advantage over other sampling methods and requires a smaller sample size (approximately half of the sample size required for cluster or two-stage cluster sampling).

2.2.1 When to apply systematic sampling

In practice, household level sampling frames are rarely available. However, assessments conducted in long-term refugee camps or areas in which a census has recently been conducted may provide enough information at the household level to construct one.

[3] The protocol should be written and provided to each enumerator for reference during data collection.

When the household sampling frame is ordered geographically, systematic sampling will result in a more even geographic distribution of sampled households than simple random sampling. This may prove to be an advantage over simple random sampling in that the workload and areas to be visited will be more evenly spread among multiple data collection teams. However, not all lists are geographically ordered. Care must be taken to assess what patterns, if any, exist in the sampling frame. If the ordered pattern has any relation at all to food security, simple random sampling is a better choice.

Despite the statistical advantage and reduced sample size requirements, the existence of a household level sampling frame does not mean that systematic sampling is always the most appropriate method. Because households are selected systematically from the population, the list of households included in the sample will be even more widely dispersed than for simple random sampling and will require visiting a large number of villages to collect the sample.

By comparison, cluster and two-stage cluster sampling limit the number of villages to be visited and may present a logistical advantage over systematic sampling. When the area being covered by an assessment is large, cluster or two-stage cluster sampling may be more cost effective despite the larger sample size requirements.

2.2.2 - How to apply systematic sampling

Step 1 - As with simple random sampling, each household in the sampling frame is assigned a unique number between 1 and the total number of households in the sampling frame. For stratified samples, a separate sampling frame must be developed for each stratum (e.g. sub-groups defined by stratification criteria).

> ***Example***[4] For a sampling frame containing 1950 households, each household is assigned a number between 1 and 1950 with no household having the same number.

Step 2 - Next, a sampling interval (SI) is derived by dividing the total number of households in the sampling frame by the required sample size[5]. Limit the sampling interval to two decimal places.

> ***Example*** The sampling interval for a systematic sample of 200 households from a sampling frame containing 1950 households is 9.75
>
> SI = 1950/200 = 9.75

Step 3 - After calculating a sampling interval, a random starting household is selected. The website http://www.randomizer.org/form.htm provides an easy-to-use random numbers generator. Choose a random starting household between 1 the sampling interval. When the sampling interval contains a decimal, round down.

> ***Example*** The random numbers generator form available through randomizer.org was used to select one household as the 'random starting household'. The range for selected the starting households is 1 to 9 (e.g. 1 and the last integer contained by the sampling interval). Suggested default values for other fields are provided form below.

---

[4] This example uses small numbers to illustrate the steps involved. In practice, the total number of households in the sampling frame will be much larger.

[5] The total number of households to be systematically selected from the sampling frame is determined by the sample size requirements (see Section 3)

To generate a set of random numbers, simply enter your selections (integer values only):

How many sets of numbers do you want to generate?	1	Help
How many numbers per set?	1	Help
Number range (e.g., 1-50):	From: 1 To: 9	Help
Do you wish each number in a set to remain unique?	Yes	Help
Do you wish to sort your outputted numbers?	Yes: Least to Greatest	Help
How do you wish to view your outputted numbers?	Place Markers Off	Help

Research Randomizer Results

1 Set of 1 Unique Numbers Per Set
Range: From 1 to 9 -- Sorted from Least to Greatest

Job Status: Finished
Bottom of Form

Set #1:
2

Step 4 - The random starting household (2 in the example) is the first household selected. Add the sampling interval (9.75 in the example) to the random starting household to select the second household. Round up if the decimal is 0.5 or greater. Round down if the decimal is less than 0.5.

Example Add 2 + 9.75 = 11.75. Round this number up to 12. Household number 12 is the second household.

Step 5 - The third household is selected by again adding to sampling interval to the sum of the starting household plus the sampling interval. Again round up if < .5 and down if >.5 to select the third household. Repeat until the end of the sampling frame is reached. A mistake has been made if you have reached the end of the sampling frame and do not have the number of households required.

Example Add 11.75 + 9.75 = 21.5. Household number 22 is the third household selected. Add 21.5 + 9.75 = 31.25. Household number 31 (round down since .25 is less than .5) is the fourth household selected…..and so on.

1	38	75	112	149	186	223	260	297
2	39	76	113	150	**187**	224	261	298
3	40	77	114	151	188	225	262	299
4	**41**	78	115	152	189	**226**	263	300
5	42	79	116	153	190	227	264	301
6	43	**80**	117	154	191	228	**265**	302
7	44	81	118	155	192	229	266	303
8	45	82	**119**	156	193	230	267	**304**
9	46	83	120	157	194	231	268	305
10	47	84	121	**158**	195	232	269	306
11	48	85	122	159	196	233	270	307
12	49	86	123	160	**197**	234	271	308
13	50	87	124	161	198	235	272	309
14	**51**	88	125	162	199	**236**	273	310
15	52	89	126	163	200	237	274	311
16	53	**90**	127	164	201	238	**275**	312
17	54	91	128	165	202	239	276	313
18	55	92	**129**	166	203	240	277	**314**
19	56	93	130	167	204	241	278	315
20	57	94	131	**168**	205	242	279	316
21	58	95	132	169	206	243	280	317
22	59	96	133	170	**207**	244	281	318
23	60	97	134	171	208	245	282	319
24	**61**	98	135	172	209	**246**	283	320
25	62	99	136	173	210	247	284	321
26	63	**100**	137	174	211	248	**285**	322
27	64	101	138	175	212	249	286	323
28	65	102	**139**	176	213	250	287	**324**
29	66	103	140	177	214	251	288	325
30	67	104	141	**178**	215	252	289	326
31	68	105	142	179	216	253	290	327
32	69	106	143	180	**217**	254	291	328
33	**70**	107	144	181	218	255	292	329
34	71	108	145	182	219	**256**	293	330
35	72	**109**	146	183	220	257	294	331
36	73	110	147	184	221	258	**295**	332
37	74	111	**148**	185	222	259	296	333

Households 334 to 1799 are removed for ease of presentation only

1800	1837	1874	1911	1948
1801	1838	1875	1912	1949
1802	1839	1876	1913	1950
1803	1840	1877	1914	
1804	1841	1878	1915	
1805	1842	1879	1916	
1806	1843	1880	1917	
1807	1844	1881	1918	
1808	**1845**	1882	1919	
1809	1846	1883	1920	
1810	1847	**1884**	1921	
1811	1848	1885	1922	
1812	1849	1886	**1923**	
1813	1850	1887	1924	
1814	1851	1888	1925	
1815	1852	1889	1926	
1816	1853	1890	1927	
1817	1854	1891	1928	
1818	**1855**	1892	1929	
1819	1856	1893	1930	
1820	1857	**1894**	1931	
1821	1858	1895	1932	
1822	1859	1896	**1933**	
1823	1860	1897	1934	
1824	1861	1898	1935	
1825	1862	1899	1936	
1826	1863	1900	1937	
1827	**1864**	1901	1938	
1828	1865	1902	1939	
1829	1866	**1903**	1940	
1830	1867	1904	1941	
1831	1868	1905	**1942**	
1832	1869	1906	1943	
1833	1870	1907	1944	
1834	1871	1908	1945	
1835	1872	1909	1946	
1836	1873	1910	1947	

Random Start and 1st Household included in Sample (green)
Households included in Sample (yellow)

Step 6 - Selected households are then mapped to facilitate data collection. Importantly, the data collection team must also have a household replacement strategy for the households in which a) the household cannot be located (inaccurate information in the sampling frame) or b) an appropriate respondent is not available.

Step 7 - Replacement households can be pre-selected prior to data collection using the sampling frame by identifying the next household in the sampling frame as the replacement household. Alternatively, a protocol[6] for replacing households in the field can be agreed upon prior to data collection. Options include choosing the next closest household or spinning a pencil in front of the absentee household to select a transect line and choosing the first house encountered in that line as the replacement household. The means of household replacement is less important than the uniform application of whatever procedure is chosen.

2.3 - Cluster sampling

A cluster is simply an aggregation of households that can be clearly and unambiguously defined[7]. For VAM food security assessments in rural areas, villages are the most common cluster used in sampling. For urban studies, blocks or neighborhoods may be more appropriate. Cluster sampling involves selection of a limited number of villages (between 20 and 30) in each strata (non-stratified samples have only one strata). All households within each selected village are then included in the sample.

2.3.1 - When to apply cluster sampling

[6] The protocol should be written and provided to each enumerator for reference during data collection.
[7] FANTA Sampling Guide (Magnani, 1997)

Often, the information needed to construct a list of all households in the population of interest (e.g. household level sampling frame) is unavailable and would be time consuming to construct. Therefore, a sampling frame is constructed at the lowest aggregation of households (often villages, neighborhoods, or blocks) for which information is available.

Even when a household level sampling frame does exist, using a random or systematic sampling method is likely to produce a geographically dispersed sample (see Simple Random and Systematic Sampling). Therefore, a large number of villages may need to be visited to select a relatively small number of households. To reduce the costs and time needed to conduct an assessment, particularly those covering large physical area, a decision may be made to use a cluster sampling. Cluster sampling reduces costs and time needed because it limits the number of villages/neighborhoods/or blocks to be visited. However, there is a cost to doing so. For most assessments the sample size required for a cluster sampling approach will be double that required for a simple random or systematic sample[8].

Cluster sampling involves only one stage of selection (selection of clusters). All households within the selected clusters are then included in the sample. Since a minimum number of clusters is required (normally between 20 and 30), standard cluster sampling only makes sense in assessments where clusters contain a relatively small number of households. Otherwise, the number of households in the sample will greatly outnumbered the number of households required for the sample. Furthermore, cluster sampling works best where clusters are fairly uniform in terms of size. If they are not, managing the workload between data collection teams and ensuring the required sample size is achieved can be problematic.

Example It is determined that the required sample size for an assessment in an urban settlement in Tajikistan is 700 households. A recent mapping exercise by the government provides a list of city blocks and the approximate number of households per block. Although each block is different, on average there are 25 households per block. A cluster sampling approach is used with clusters defined as city blocks. Thirty (30) clusters are randomly selected from the block level sampling frame for an expected sample size of n = 750 (e.g. 25 * 30 = 750).

Example It is determined that the required sample size for an assessment in West Haraghe, a rural district in Ethiopia, is 500 households. Although there has not been a recent census, a reasonably accurate list of all villages and there approximate size is available through the government's statistics department. Villages range in size rather drastically and, on average, contain 150 households. A cluster sampling approach using villages as clusters would require selection of a minimum of 20 clusters. Since this would yield an expected sample size of 3,000 households (in comparison to the 500 required), a decision is made to use a two-stage, cluster sampling approach[9].

As illustrated in the examples, cluster sampling is most useful in urban settings, where aggregations of households such as blocks or neighborhoods contain a relatively small and uniform number of households. It may also be useful in small rural settlements. Multiply the average number of households per cluster (village, neighborhood/block) by 20 (the minimum number of clusters required) to get the expected sample size. Compare this with the required sample size[10]. If the expected sample size is much larger than the required, two-stage cluster sampling is a more appropriate method.

2.3.2 - How to apply cluster sampling

Applying cluster sampling requires two distinct steps to be taken: defining clusters and assembling the sampling frame (step 1), and selecting clusters and household for inclusion in the sample (step 2). Each of these steps involves a number of intermediary steps.

---

[8] This is due to the design effect of using a cluster sampling methodology. This issue is discussed in detail in section 3.
[9] This method is described in 2.4
[10] see Section III

2.3.2.1 Defining clusters and constructing the sampling frame

Step 1a - The first step in cluster sampling is defining the aggregation of households that will be used as 'clusters'. The following criteria are helpful for defining appropriate clusters[11]:

- Aggregations should be pre-existing and recognized. Villages, blocks, neighborhoods, and census blocks are good examples.

- Aggregations used for clusters should be as unrelated to food security as possible. Unlike stratification – in which households were categorized into sub-groups on the basis of criteria related to food security such as livelihoods, land-use zones (e.g. homogeneity) – the aim of clustering is just the opposite (e.g. heterogeneity). Ideally, each cluster should contain households that reflect the diversity (in terms of food security related factors such as livelihoods and land-use) that is found in the entire population of interest. For the majority of VAM food security assessments the use of administrative aggregations as clusters will most closely approximate this ideal.

- Clear physical boundaries exist between clusters to assist in identification during data collection.

- Information on the size of the cluster (households or populations) is available. Where population estimates are unavailable, key informants can be used to provide rough/relative estimates (very large, large, medium, small, very small).

Step 1b – Next, assemble the sampling frame. For stratified samples, a separate sampling frame must be developed for each stratum (e.g. sub-groups defined by stratification criteria). Microsoft Excel or similar spreadsheet software is useful, though a simple table can also be used. In the first column list each cluster. In the second column list the size of the cluster (either population or number of households). If you are using rough estimates from key informants use relative size codes. The table on the right provides example codes.

Cluster Size	Code
Very Large	5
Large	4
Medium	3
Small	2
Very Small	1

Step 1c - Use the third column to list the cumulative size values for all clusters. The cumulative size value for cluster 2 is the sum of clusters 1 and 2. The cumulative size value for cluster 3 is the sum of clusters 1, 2, and 3…..and so on.

Example Sampling Frame with Cluster Population Estimates

CLUSTER	SIZE	CUMM SIZE
A	50	50
B	125	175
C	35	210
D	20	230
E	80	310
F	20	330
G	25	355
H	40	395
I	25	420

Example Sampling Frame with Key Informant Generated Cluster Size Estimates

CLUSTER	SIZE	CUMM SIZE
A	3	3
B	1	4
C	5	9
D	2	11
E	1	12
F	1	13
G	4	17
H	5	22
I	3	25

[11] The first, third, and fourth criterion were adapted from the FANTA Sampling Guide (Magnani, 1997)

2.3.2.2 Selecting Clusters and Households for Inclusion in the Sample

From a technical standpoint, the more clusters the better. But, more clusters mean more villages and, as a result, more expense and time.

Step 2a - The next step is to decide how many clusters will be included in the sample. As indicated above, 20 to 30 clusters per strata are recommended for most settings (non-stratified samples have only one strata).

The recommendation of 30 clusters per strata is somewhat arbitrary, but provides a commonly used and technically sound standard that assessments should attempt to follow. However, choosing the most appropriate number of clusters requires striking a balance between technical and logistic considerations.

A minimum of 20 clusters per strata provides a lower limit for assessments where cost and time considerations are major constraints[12]. Most assessments fall somewhere in between the standard of 30 clusters and this minimum.

Recommended Number of Clusters

Standard	Compromise	Minimum
30	25	20

Step 2b - Since all households within selected clusters are included in cluster sampling, use the average number of households per cluster and the desired number of clusters (from above) to determine the number of enumerators/data collection teams required. Where possible, the number of households per cluster should correspond to the number of interviews that one or two data collection teams of reasonable size (3 to 5 enumerators) can complete in a day[13]. At times, constraints on the number of enumerators and teams available may require using the compromised (25) or minimum (20) number of clusters. However, a serious attempt should be made to find additional enumerators or add data collection days before reducing the number of clusters.

Example The required sample size for an assessment of peri-urban settlements in the capital city of Bangladesh is determined to be 600 households. Maps of blocks containing an average of 26 households each are available through a local NGO working in the area. A pre-test suggests that a team of 4 enumerators can interview approximately 1 block per day (6 interviews per day).

The assessment will employ a total of twenty enumerators (5 teams) with 1 supervisor per team. Although 30 clusters would be ideal, the Country Office has only 5 days to collect the data so that a report will be available for an upcoming assessment mission due to arrive in 2 weeks. Furthermore, government counterparts and local staff are being used to ensure high quality data collection and only 20 are available to participate.

A decision is made to select 25 clusters (one cluster per data collection team per day) of approximately 26 households each (6 interviews per enumerator per day) for a sample size of n = 650.

Twenty-three (23) clusters would yield an expected sample size (n = 598) closer to the required number (n = 600). However, it is possible that the average size of the selected blocks will be slightly smaller than the average for all blocks in the population such that extra clusters are included to ensure at least 600 households are included in the sample.

Step 2c - Clusters are then randomly or systematically selected from the cluster-level sampling frame. Cluster population figures are used to select clusters **probability**

[12] Reducing the number of clusters below 20 requires a technical assessment of the expected inter-cluster heterogeneity and intra-cluster homogeneity and should not be done without appropriate technical guidance. Fewer than 20 clusters may be possible in samples in which stratification produces a large number of sub-groups (e.g. strata are very homogenous on factors related to food security, reducing the range of heterogeneity within and between clusters within a particular strata).
[13] This issue is more pronounced in two-stage, cluster sampling where the number of households per cluster is constant and, therefore, can be managed.

proportional to size (PPS); meaning that larger clusters have a higher probability of selection. As indicated earlier, key informants can be used to provide rough estimates where existing information on cluster size is unavailable.

Probability Proportional to Size (PPS)

The purpose behind selecting clusters 'PPS' is to ensure that each household in the population of interest, whether from a large or small village, has an approximately equal probability of selection. To approximately equate probability of household selection at the second stage, large villages must have a higher probability of selection at the first stage. Selecting clusters without PPS will lead to households having different probabilities of selection. Such samples are non-self-weighting and will complicate analysis (Magnani, 1997).

Example The required sample size for an assessment in urban settlements in Freetown, the capital of Sierra Leone, is 500 households. Information on the location and approximate size of city blocks is available. Blocks will be used as define clusters. A total of 75 blocks are listed in the sampling frame with an average size of 22 households per block. Twenty-five (25) blocks will be chosen out of a total of 75 blocks in the population. Given the average block size, this expected to yield a sample size of 550 households.

Random Selection - Use the random numbers generator (www.randomizer.org) to generate 25 random numbers. Use the cumulative size (CUMM SIZE) to define the number range (in the example 1 to 1667). The numbers generated correspond with numbers in the column CUMM SIZE. The clusters containing each of the cumulative numbers selected are included in the sample.

All households with in selected clusters are included in the sample. Therefore, each cluster can only be selected once. Generate additional random numbers for each duplicate until 25 clusters are selected.

CLUSTER	# OF HH	CUMM SIZE	CLUSTER	# OF HH	CUMM SIZE	CLUSTER	# OF HH	CUMM SIZE
1	15	15	31	30	745	61	40	1365
2	25	40	32	25	770	62	25	1390
3	35	75	33	20	790	63	60	1450
4	20	95	34	10	800	64	15	1465
5	10	105	35	25	825	65	10	1475
6	20	125	36	40	865	66	20	1495
7	25	150	37	20	885	67	10	1505
8	40	190	38	10	895	68	30	1535
9	25	215	39	15	910	69	10	1545
10	20	235	40	45	955	70	25	1570
11	30	265	41	25	980	71	10	1580
12	35	300	42	30	1010	72	35	1615
13	15	315	43	20	1030	73	10	1625
14	10	325	44	10	1040	74	15	1640
15	15	340	45	20	1060	75	27	1667
16	20	360	46	15	1075			
17	15	375	47	25	1100			
18	35	410	48	10	1110			
19	10	420	49	10	1120			
20	60	480	50	15	1135			
21	50	530	51	15	1150			
22	25	555	52	25	1175			
23	30	585	53	10	1185			
24	35	620	54	15	1200			
25	20	640	55	20	1220			
26	20	660	56	20	1240			
27	20	680	57	15	1255			
28	10	690	58	15	1270			
29	15	705	59	20	1290			
30	10	715	60	35	1325			

Selected Cluster
Selected Twice in Random Numbers Selection
Selected as replacements for duplicate Numbers

Systematic Selection – To determine the sampling interval (S.I.), divide the total cumulative size (CUMM SIZE) indicated in the last cluster listed in sampling frame by the number of clusters to be selected (25).

1667/25 = 66.68

Use the random numbers generator to generate one random starting number. The sampling interval defines the number range (1 to 66.68 in the example) from which the random start is selected. The number generated corresponds with the numbers in the column CUMM SIZE (not the cluster number!). The cluster containing the cumulative number selected is the random starting household (cluster 2 in the example).

To select the second cluster, add the sampling interval (66.6) to the cumulative size generated above (66.68 + 37 = 103.68). The cluster containing the product is the second cluster (cluster 5) To select the third cluster, add the sampling interval to the cumulative size used to select the second cluster (66.68 +105.68= 170.36, cluster 8)…..and so on.

All households within selected clusters are included in the sample. Therefore, each cluster can only be selected once. Note the number of duplicate selections. Limit the sampling frame to only those clusters that have not been selected and repeat the steps outlined above (pick a new random start and generate a new sampling interval corresponding to the total cumulative size divided by the number of duplicates). Repeat again (as needed) until the total number clusters required (25 in the example) are selected.

CLUSTER	# OF HH	CUMM SIZE	CLUSTER	# OF HH	CUMM SIZE	CLUSTER	# OF HH	CUMM SIZE
1	15	15	31	30	745	61	40	1365
2	25	40	32	25	770	62	25	1390
3	35	75	33	20	790	63	60	1450
4	20	95	34	10	800	64	15	1465
5	10	105	35	25	825	65	10	1475
6	20	125	36	40	865	66	20	1495
7	25	150	37	20	885	67	10	1505
8	40	190	38	10	895	68	30	1535
9	25	215	39	15	910	69	10	1545
10	20	235	40	45	955	70	25	1570
11	30	265	41	25	980	71	10	1580
12	35	300	42	30	1010	72	35	1615
13	15	315	43	20	1030	73	10	1625
14	10	325	44	10	1040	74	15	1640
15	15	340	45	20	1060	75	27	1667
16	20	360	46	15	1075			
17	15	375	47	25	1100			
18	35	410	48	10	1110			
19	10	420	49	10	1120			
20	60	480	50	15	1135			
21	50	530	51	15	1150			
22	25	555	52	25	1175			
23	30	585	53	10	1185			
24	35	620	54	15	1200			
25	20	640	55	20	1220			
26	20	660	56	20	1240			
27	20	680	57	15	1255			
28	10	690	58	15	1270			
29	15	705	59	20	1290			
30	10	715	60	35	1325			

Random Start (first cluster selected)
Selected Cluster

Step 2d - All households within each selected cluster are included in the sample. If clusters are large this can result far too many households being included in the sample. Managing the data collection workload among different teams can also be made difficult if clusters vary widely in size. In either of these cases, two-stage cluster sampling should be considered as an alternative.

Step 2e - An attempt should be made to return to households that are unavailable (no one home or inappropriate respondent) at the time of initial data collection. However, a small number of absent households will not affect the overall validity of the assessment.

2.4 - Two-stage cluster sampling

In practice, two-stage cluster sampling is far more common than all of the other methods described in these guidelines combined. The combination of minimal information requirements and logistical ease make it particularly well suited to many of the scenarios encountered during VAM food security assessments.

As the name implies, two-stage cluster sampling is a variant of cluster sampling. A cluster is simply an aggregation of households that can be clearly and unambiguously defined[14]. For VAM food security assessments in rural areas, villages are the most common cluster used in sampling. For urban studies, blocks or neighborhoods may be more appropriate. Two-stage cluster sampling involves selection of a limited number of villages (between 20 and 30) in each strata (non-stratified samples have only one strata). Instead of selecting all households in each selected cluster (as for cluster sampling), two-stage cluster sampling uses a second step to select a limited and fixed number of households within each selected cluster.

2.4.1 - When to apply two-stage cluster sampling

The information needed to construct a list of all households in the population of interest (e.g. household level sampling frame) is often unavailable and such a list would be time consuming and expensive to construct. Therefore, a sampling frame is constructed at the lowest aggregation of households (often villages, neighborhoods, or blocks) for which information is available.

Even when a household level sampling frame does exist, using a random or systematic sampling method is likely to produce a geographically dispersed sample (see Simple Random and Systematic Sampling). Therefore, a large number of villages may need to be visited to select a relatively small number of households. To reduce the costs and time needed to conduct an assessment, particularly those covering large physical area, a decision may be made to use a two-stage cluster sampling. Two-stage cluster sampling reduces costs and time needed because it limits the number of villages/neighborhoods/or blocks to be visited and the number of households to be interviewed each village/neighborhood/ or block selected. However, there is a cost to doing so. For most assessments the sample size required for a two-stage cluster sampling approach will be double that required for a simple random or systematic sample[15].

Two-stage cluster sampling is more widely applicable than cluster sampling because it does not require that clusters contain a relatively small and uniform number of households. Therefore, the approach is well suited to rural settlements commonly encountered in VAM food security assessments. Two-stage cluster sampling may also be appropriate in urban settlements where the size of clusters is not conducive to standard cluster sampling (e.g. too large or too variable).

Example It is determined that the minimum required sample size for an assessment in East Haraghe, a rural district in Ethiopia, is 440 households. Although there has not been a recent census, a reasonably accurate list of all villages (150 in total) and there approximate size is available through the government's statistics department. Villages range in size from 20 to 300 households and, on average, contain 150 households. At the first stage of selection, 30 villages are randomly selected for inclusion in the assessment. At the second stage of selection, 15 households are selected within each of the 30 villages for a total sample size of n = 480 (e.g. 30 * 15 = 450).

2.4.2 - How to apply two-stage cluster sampling

---

[14] FANTA Sampling Guide (Magnani, 1997)

[15] This is due to the design effect of using a cluster sampling methodology. This issue is discussed in detail in Section III.

Two-stage cluster sampling requires three distinct steps: defining clusters and constructing the sampling frame (step 1), choosing clusters for inclusion in the sample (step 2), and choosing households from within selected clusters for inclusion in the sample (step 3). As with cluster sampling, each of these steps involves a number of intermediate steps.

2.4.2.1 - Defining Clusters and Constructing the Sampling Frame

Step 1a - the first step in two-stage cluster sampling is defining the aggregation of households that will be used as 'clusters'. The following criteria are helpful for defining appropriate clusters[16]:

- Aggregations should be pre-existing and recognized. Villages, blocks, neighborhoods, and census blocks are good examples.
- Aggregations used for clusters should be as unrelated to food security as possible. Unlike stratification – in which households were categorized into sub-groups on the basis of criteria related to food security such as livelihoods, land-use zones (e.g. homogeneity) – the aim of clustering is just the opposite (e.g. heterogeneity). Ideally, each cluster should contain households that reflect the diversity (in terms of food security related factors such as livelihoods and land-use) that is found in the entire population of interest. For the majority of VAM food security assessments the use of administrative aggregations as clusters will most closely approximate this ideal.
- Clear physical boundaries exist between clusters to assist in identification during data collection.
- Information on the size of the cluster (households or populations) is available. Where population estimates are unavailable, key informants can be used to provide rough/relative estimates (very large, large, medium, small, very small).

Cluster Size	Code
Very Large	5
Large	4
Medium	3
Small	2
Very Small	1

Step 1b - The second step is assembling the sampling frame. For stratified samples, a separate sampling frame must be developed for each stratum (e.g. sub-groups defined by stratification criteria). Microsoft Excel or similar spreadsheet software is useful, though a simple table can also be used. In the first column list each cluster. In the second column list the size of the cluster (either population or number of households). If you are using rough estimates from key informants use relative size codes. The table on the right provides example codes.

Step 1c - Use the third column to list the cumulative size values for all clusters. The cumulative size value for cluster 2 is the sum of clusters 1 and 2. The cumulative size value for cluster 3 is the sum of clusters 1, 2, and 3…..and so on.

Example Sampling Frame with Cluster Population Estimates

CLUSTER	SIZE	CUMM SIZE
A	50	50
B	125	175
C	35	210
D	20	230
E	80	310
F	20	330
G	25	355
H	40	395
I	25	420

Example Sampling Frame with Key Informant Generated Cluster Size Estimates

[16] The first, third, and fourth criterion were adapted from FANTA Sampling Guide (Magnani, 1997)

CLUSTER	SIZE	CUMM SIZE
A	3	3
B	1	4
C	5	9
D	2	11
E	1	12
F	1	13
G	4	17
H	5	22
I	3	25

2.4.2.2 - Selecting Clusters Inclusion in the Sample

> From a technical standpoint, the more clusters the better. But, more clusters mean more villages and, as a result, more expense and time.

Step 2a - The next step is to decide how many clusters will be included in the sample. As indicated above, 20 to 30 clusters per strata are recommended for most settings (non-stratified samples have only one strata). The recommendation of 30 clusters per strata is somewhat arbitrary, but provides a commonly used and technically sound standard that assessments should attempt to follow. However, choosing the most appropriate number of clusters requires striking a balance between technical and logistic considerations. A minimum of 20 clusters per strata provides a lower limit for assessments where cost and time considerations are major constraints[17]. Most assessments fall somewhere in between the standard of 30 clusters and this minimum.

Example A VAM food security assessment in a rural Indian requires a sample size of 300 households in each of 5 strata (sub-groups defined by land-use zones) for a total sample size of n = 1,500. Information from the government allows for the use of villages as clusters. The following options are considered for each of the 5 strata:

30 clusters of 10 households each (n = 300)
25 clusters of 12 households each (n = 300)
20 clusters of 15 households each (n = 300)

Since there are 5 strata, a decision is made to take the minimum acceptable number of clusters to reduce the number of vehicles and other costs associated with the assessment. The total number of clusters/villages to be visited is 100 (20 clusters in each of 5 strata) for a total sample size of n = 1,500 (15 in each cluster).

Step 2b - Use the number of clusters, number of households per cluster, and number of days allotted for data collection to determine the number of enumerators/data collection teams required. Since adding few more households per village is logistically easier than having more villages of smaller size, constraints on the number of enumerators and teams available may suggest using the compromised (25) or minimum (20) number of clusters. However, a serious attempt should be made to find additional enumerators or add data collection days before reducing the number of clusters. A pre-test will help to estimate the number of interviews that a data collection team of reasonable size (3 to 5 enumerators) can complete in a day.

Example (Continuing from the Indian example given above with 20 clusters in each of 5 strata, with 15 households taken per cluster for a total sample size of n = 1500). It is estimated that each enumerator can complete 5 interviews per day. Therefore a team of 3 enumerators and 1 supervisor can complete 1 cluster per day. Fourteen days have been allotted for data

---

[17] Reducing the number of clusters below 20 requires a technical assessment of the expected inter-cluster heterogeneity and intra-cluster homogeneity and should not be done without appropriate technical guidance. Fewer than 20 clusters may be possible in samples in which stratification produces a large number of sub-groups (e.g. strata are very homogenous on factors related to food security, reducing the range of heterogeneity within and between clusters within a particular strata).

collection. Since some travel time between clusters is required, it is estimated that 8 teams will be needed (24 enumerators).

Step 2c - Clusters are then randomly or systematically selected from the cluster-level sampling frame. Cluster population figures are used to select clusters probability proportional to size (PPS); meaning that larger clusters have a higher probability of selection. As indicated earlier, key informants can be used to provide rough estimates where existing information on cluster size is unavailable.

> **Probability Proportional to Size (PPS)**
>
> The purpose behind selecting clusters 'PPS' is to ensure that each household in the population of interest, whether from a large or small village, has an approximately equal probability of selection. To approximately equate probability of household selection at the second stage, large villages must have a higher probability of selection at the first stage. Selecting clusters without PPS will lead to households having different probabilities of selection. Such samples are non-self-weighting and will complicate analysis (Magnani, 1997).

Example The required sample size for an assessment in rural, northern Uganda is 500 households. Information on the location and approximate size of villages is available through the government. A total of 75 villages are listed in the cluster-level sampling frame. Twenty-five (25) villages will be chosen for the sample and twenty (20) households will be taken in each of the selected villages for a total sample size of n = 500.

Random Selection - Use the random numbers generator (www.randomizer.org) to generate 25 random numbers. Use the cumulative size (CUMM SIZE) to define the number range (in the example 1 to 5001). The numbers generated correspond with numbers in the column CUMM SIZE. The clusters containing each of the cumulative numbers selected are included in the sample. If a cluster is selected twice, 40 households will be taken in that cluster (e.g. 2 x 20 hh).

To generate a set of random numbers, simply enter your selections (integer values only):

How many sets of numbers do you want to generate? [1] Help

How many numbers per set? [25] Help

Number range (e.g., 1-50): From: [1] To: [5001] Help

Do you wish each number in a set to remain unique? [Yes] Help

Do you wish to sort your outputted numbers? [Yes: Least to Greatest] Help

How do you wish to view your outputted numbers? [Place Markers Off] Help

Research Randomizer Results

1 Set of 25 Unique Numbers Per Set
Range: From 1 to 5001 -- Sorted from Least to Greatest

Job Status: Finished

1	25	1	5001	Unique	Sorted

Set #1:

192,251,373,5 192, 251, 373, 552, 610, 705, 845, 1228, 1578, 1605, 2259, 2278, 2379, 2636, 3047, 3340, 3478, 3719, 3834, 3910, 4020, 4055, 4244, 4334, 4667

CLUSTER	# OF HH	CUMM SIZE	CLUSTER	# OF HH	CUMM SIZE	CLUSTER	# OF HH	CUMM SIZE
1	45	45	31	90	2235	61	120	4095
2	75	120	32	75	2310	62	75	4170
3	105	225	33	60	2370	63	180	4350
4	60	285	34	30	2400	64	45	4395
5	30	315	35	75	2475	65	30	4425
6	60	375	36	120	2595	66	60	4485
7	75	450	37	60	2655	67	30	4515
8	120	570	38	30	2685	68	90	4605
9	75	645	39	45	2730	69	30	4635
10	60	705	40	135	2865	70	75	4710
11	90	795	41	75	2940	71	30	4740
12	105	900	42	90	3030	72	105	4845
13	45	945	43	60	3090	73	30	4875
14	30	975	44	30	3120	74	45	4920
15	45	1020	45	60	3180	75	81	5001
16	60	1080	46	45	3225			
17	45	1125	47	75	3300			
18	105	1230	48	30	3330			
19	30	1260	49	30	3360			
20	180	1440	50	45	3405			
21	150	1590	51	45	3450			
22	75	1665	52	75	3525			
23	90	1755	53	30	3555			
24	105	1860	54	45	3600			
25	60	1920	55	60	3660			
26	60	1980	56	60	3720			
27	60	2040	57	45	3765			
28	30	2070	58	45	3810			
29	45	2115	59	60	3870			
30	30	2145	60	105	3975			

Selected Cluster
Cluster selected twice (40 households taken instead of 20)

Systematic Selection – To determine the sampling interval (S.I.), divide the total cumulative size (CUMM SIZE) indicated in the last cluster listed in sampling frame by the number of clusters to be selected (25).

Example In the example below there are 5001 total households and the number of clusters required is 25. The sampling interval is therefore 5001/25 = 200.04

Use the random numbers generator to generate one random starting number. The sampling interval defines the number range (1 to 200.04 in the example) from which the random start is selected. The number generated corresponds with the numbers in the column CUMM SIZE. The cluster containing the cumulative number selected is the random starting household.

Example 111 is the randomly selected 'first household' selected from the range 1 – 200 (e.g. range defined by the sampling interval). This CUMM SIZE corresponds with cluster 2 in the example below.

To select the second cluster, add the sampling interval to the cumulative size given by the random start. The cluster containing the product is the second cluster. To select the third cluster, add the sampling interval to the cumulative size used to select the second cluster…..and so on until 25 clusters are selected.

Example Second Household 200.04 + 111 = 311.04 located in cluster 5. Third household 200.04 + 311.04 = 511.08 located in cluster 8.

CLUSTER	# OF HH	CUMM SIZE	CLUSTER	# OF HH	CUMM SIZE	CLUSTER	# OF HH	CUMM SIZE
1	45	45	31	90	2235	61	120	4095
2	75	120	32	75	2310	62	75	4170
3	105	225	33	60	2370	63	180	4350
4	60	285	34	30	2400	64	45	4395
5	30	315	35	75	2475	65	30	4425
6	60	375	36	120	2595	66	60	4485
7	75	450	37	60	2655	67	30	4515
8	120	570	38	30	2685	68	90	4605
9	75	645	39	45	2730	69	30	4635
10	60	705	40	135	2865	70	75	4710
11	90	795	41	75	2940	71	30	4740
12	105	900	42	90	3030	72	105	4845
13	45	945	43	60	3090	73	30	4875
14	30	975	44	30	3120	74	45	4920
15	45	1020	45	60	3180	75	81	5001
16	60	1080	46	45	3225			
17	45	1125	47	75	3300			
18	105	1230	48	30	3330			
19	30	1260	49	30	3360			
20	180	1440	50	45	3405			
21	150	1590	51	45	3450			
22	75	1665	52	75	3525			
23	90	1755	53	30	3555			
24	105	1860	54	45	3600			
25	60	1920	55	60	3660			
26	60	1980	56	60	3720			
27	60	2040	57	45	3765			
28	30	2070	58	45	3810			
29	45	2115	59	60	3870			
30	30	2145	60	105	3975			

Random Start (first cluster selected)
Selected Cluster

2.4.2.3 Selecting Households within Selected Clusters[18]

Three options exist for selecting households within selected clusters. Each option can be applied regardless of whether the clusters were selected randomly or systematically (step 2c in Section 2.4.2.2). The options are listed in order of preference; that is option 1. is preferred over option 2., and option 2 is preferred over option 3. However, the options are listed in reverse order of logistic ease; that is option 3 is cheaper and faster than option 2, and option 2 is cheaper and faster than option 1. Choosing the right method for household selection will vary by assessment. Assessments should strive to use the preferred method (1), choosing options 2 or 3 when required due to logistic, time, and resource constraints.

Option 1 - The most ideal household selection method involves constructing a sampling frame of all households within the selected clusters. Where clusters are small in size this approach is manageable. However, this approach will be costly and time prohibitive when the clusters are large in size. Once the sampling frame has been constructed, follow the guidance given for simple random sampling or systematic sampling for selecting households for inclusion.

Example An assessment is being carried out in rural Bangladesh. Villages will serve as clusters. Thirty (30) villages have been selected for inclusion in the sample in each of two strata for a total of 60 villages. Ten (10) households will be selected in each village for a per strata sample size of n = 300 and a total sample size of n = 600. Upon arrival in each selected village, the data collection team maps the village, giving each household a unique number (no two households can have the same number). In the first cluster there are 35 households, such that households are numbered 1 to 35.

---

[18] This section borrows heavily from the procedures outlined in FANTA Sampling Guide (Magnani, 1997).

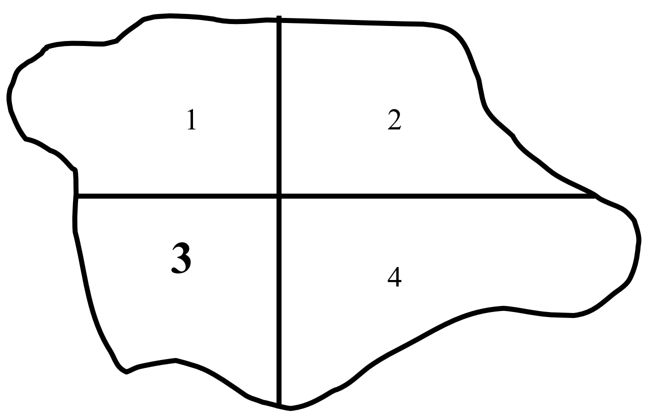
Option 1a - One option is to select households systematically. A sampling interval of 3.5 is calculated (35 divided by 10). Household 2 is selected as the random starting households (chosen between the range of 1 to 3, since 3.5 contains a decimal). The sampling interval of 3.5 is added to the random start to select the second household (5.5, round up to household 6). Add the sampling interval again to get the third households (5.5 + 3.5 = 9) and so on.

Option 1b - A second option is to select households randomly. Write each household number (1 to 35) down on a slip of paper and put them in a hat. Shake the hat and then select 10 slips of paper. The number on the slip of paper corresponds with the household to be interviewed.

> Letting members of the community choose from the hat provides an excellent means of involving the community in the process, helping them to understand the meaning of 'random selection', and avoiding scenarios in which village leaders attempt to dictate which households are interviewed.

Option 2 - When cluster are too large or time constraints prevent using the method outlined above, a method called segmentation can be used. Segmentation requires that a rough map of the cluster exists or can be quickly created. The cluster is then divided into smaller segments, with each segment containing approximately the number of households required from the cluster. The total number of segments in a particular cluster will be equal to an estimate of the total number of households in the cluster by the number of households required. All households within the segment are then included in the sample. Note the actual number of households within the chosen segment may be slightly more or slightly less than the target number of households.

Example An assessment was being carried out in a rural district in Yemen. Hamlets within the district served as clusters. Thirty (30) hamlets were selected and 15 households were selected in each hamlet for a total sample size of n = 450. Upon arrival in each selected hamlet, the data collection team asked two key informants to map the hamlet. The hamlet was then divided into segments with each containing approximately 15 households. The first hamlet selected contained approximately 60 households and was divided into 4 segments (4 x 15 = 60). A random hamlet was selected by numbering 4 slips of paper (1 to 4) and picking one of them from a hat. Segment 3 was chosen and all households within the segment were included in the sample.

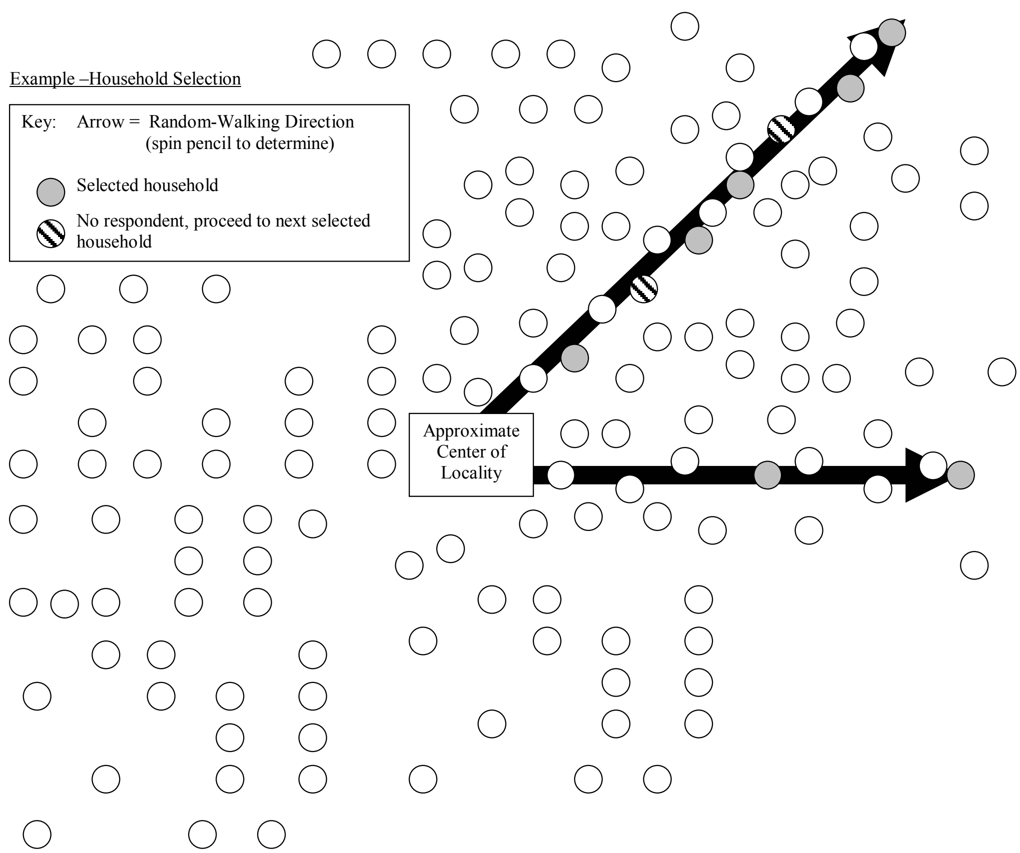


Option 3 - The third option for selecting households is the most rapid, but also the least preferred method. This method is commonly used in Expanded Program on Immunization (EPI) surveys and in UNICEF anthropometric surveys. Once the data collection team arrives in the cluster, the approximate middle of the cluster is identified. A pencil or bottle is spun to select a random walking direction (also called a transect line). The data collection team then counts the number of households encountered along the transect line between the center and the perimeter of the cluster. This number is divided to determine the interval at which households will be selected in the transect line.

When the transect line contains less than the number of households required, all households in the line are included in the sample and the data collection team returns to the center of the cluster to pick a second random walking direction and the process is repeated. If a household without an appropriate respondent is encountered, skip it and proceed to the next selected household. This may require returning to the center and repeating the process as for transects with fewer than the number of required households.

Example An assessment was carried out in Tambura District in Southern Sudan. Villages served as clusters. Thirty (30) villages were selected in each of two livelihood zones, each represented a strata. Seven (7) households were selected in each village for a per strata sample size of n = 210 and an overall sample size of n = 420. Upon arrival in each selected village, the data collection team asked two key informants to help locate the center of the village. A pencil was spun to pick a random walking direction (transect). The number of households encountered when walking from the center of the village to the perimeter was 14. Therefore, every other household was selected for inclusion in the sample.

In two households, an appropriate respondent was unavailable. Therefore, the data collection team was required to repeat the process by returning to the center, picking a transect line, dividing the number of households in that line by 2 (the number of replacement households needed). This resulted in every 4th household in the second transect line being sampled.



2.5 - Multi-Stage Sampling

In the majority of scenarios in which a household level sampling frame is unavailable and expensive to construct, a two-stage cluster sampling methodology can be used. However, there are rare occasions where a multi-stage method may be required.

Multi-stage sampling is simply an extension of the two-stage random sampling (e.g. three or more stages). For example, accurate information may only exist at the division level, necessitating three (or more) sampling stages:

Stage 1 - random or systematic selection of divisions
Stage 2 – random or systematic village selection within selected divisions
Stage 3 – random or systematic household selection within selected villages

The design effect, and therefore sample size requirements, goes up with each additional sampling stage. Before considering the use of multi-stage sampling methods, consult with NGOs, other U.N. agencies, and the VAM regional and headquarters staff to help decide if doing so is necessary.

Section III - Determining the Appropriate Sample Size

The aim of the section is twofold; to provide a basic understanding of the factors to be considered in the calculation of sample size and, more importantly, to provide easy-to-use[1] sample size guidance for common scenarios found in VAM food security assessments. Two different sets of guidance are given for stratified samples (e.g. samples that are designed to ensure comparability between sub-groups) and non-stratified samples.

It should also be noted that the choice of sample size formulas depends on whether the key food security indicator (or indicators) of interests for the assessment is a mean or proportion[2]. A primary objective of most VAM assessments is to estimate the percentage of food insecure households within the population. However, some VAM food security assessments will use indicators expressed as means. The Coping Strategies Index (CSI) provides a notable example.

> There are many misunderstandings concerning sample size. Perhaps the most common has to do with population size. Except where a population is exceptionally small and a 'finite population adjustment' is required, population size has nothing to do with the size of the sample

Ultimately, the choice of sample size is almost always driven by practical limitations on time and resources. However, this does not render the calculation of sample size on the basis of technical factors irrelevant. The sample size calculation provides the ideal sample size required to meet the objectives of the assessment. Knowing this is critical for understanding the consequences of deviating from the ideal due to cost and time constraints and allows for informed choices to be made.

3.1 - Non-stratified samples

3.1.1 Sampling when key indicators are expressed as percentages

The formula for calculating the sample size for assessments with key indicators expressed as percentages is:

$$\mathbf{n = (D)(Z^2 * p * q)/d^2}$$

Where:	n	=	The required minimum sample size
	D	=	Design effect (varies by type of sampling)
	Z	=	The Z-score corresponding to the degree of confidence
	p	=	Estimated proportion of key indicator expressed as a decimal (e.g. 20% = .20)
	q	=	1 – p
	d	=	Minimum desired precision or maximum tolerable error expressed in decimal form (e.g. +/- 10 percentage points = .10).

Taken as a whole the formula can be intimidating, particularly for those who are unfamiliar with mathematical notation. However, taken separately, each parameter in the formula is relatively easy to define and automated sample size calculators are available to perform the computation (an example, website, and instructions for use are provided below). In addition, recommended sample sizes (not requiring computations) are provided for common scenarios encountered in VAM food security studies.

D The design effect for simple random sampling and systematic sampling is equal to 1 (meaning there is no design effect). The design effect for cluster or two-stage cluster sampling is the factor by which the sample size must be increased in order to produce survey estimates with the same precision as a simple random sample[3]. The default value for cluster and two-stage cluster sampling is 2, resulting in a doubling of the sample size requirement. However, it may be possible to reduce

[1] Guidance is provided that does not require users to make the calculation themselves.
[2] The term proportion includes percentages and prevalence.
[3] See FANTA Sampling Guide for a more in-depth discussion (Magnani, 1987).

this value when design effect estimates for the same indicator are available from previous surveys[4].

Z Due to the fact that estimates are based on a sample, rather than total enumeration of the population (as in a census), it is not possible to be 100% confident that the estimate derived from a sample is a true reflection of the population. The conventional degree of confidence for almost all social research is 95%; meaning that if you were to perform the assessment 100 times, 95 of the 100 assessments would yield range estimates known as a confidence intervals (e.g. 20% +/- 5 percentage points) containing the true population proportion. By contrast, 5 of the 100 assessments would yield confidence intervals that do not contain the true population proportion due to chance. The Z-score corresponding with 95% confidence is 1.96.

p An estimate (in decimal form) for the primary food security indicator of interest allows the sample size to be reduced. Where no reasonably accurate estimate can be found, a default value of 50% should be used. This default offers a safe, albeit more expensive, alternative as the value of 50% will yield the largest required sample size.

However, many assessments blindly and inappropriately use this default value without attempting to derive an estimate from pre-existing information. Previous WFP, NGO, and governments assessments often provide estimates of the same or similar indicator (e.g. another food security indicator). Although recent estimates for the same population are desired, it may be necessary to use estimates that are several years old. Taking the time to generate a 'best guess' estimate for the primary indicator of interest is worthwhile and can result in significant savings in time and cost (compare the sample sizes required for different estimates in the table entitled Non-Stratified Sample Size Recommendations).

d The primary technical choice in determining sample size for a non-stratified sample is defining a minimum level of precision (or maximum tolerable error). Precision refers to the degree of error (or confidence interval) around the estimate due to the fact that the estimate is based on a sample.

Example It is estimated that 28% (+/- 5 percentage points) of households in a rural district in Bolivia consume meat less than one time per week. The '+/- 5 percentage points' is the degree of error around the estimate and defines the confidence interval. The point estimate, 28%, reflects the percentage actually found in the sample population. The range or confidence interval of 23% - 33% better reflects the larger population from which the sample was taken[5]. The larger the sample, the more narrow the confidence interval

3.1.2 - Sample size guidance tables

Table 1 depicts the sample size requirements for simple random and systematic samples with various combinations of food security indicator estimates (p and q) and maximum tolerable error/minimum level of precision (d)[6]. Table 2 depicts the sample size requirements for cluster and two-stage sampling with various combinations of food security indicator estimates (p and q) and maximum tolerable error/minimum level of precision (d)[7].

---

[4] Demographic and Health Surveys (DHS) often have estimates of the design effect of two-stage cluster sampling for food security indicators.
[5] As discussed under Z, the convention for confidence intervals is 95%. A comprehensive statement about the estimate given in the example would be 'we are 95% confident that the true proportion of households in X District, Bolivia consuming meat less than one time per week falls between 23% and 33%' or 'it is estimated that the 28% (95% C.I. 23% - 33%) of households in X District, Bolivia consume meat less than one time per week'.
[6] The confidence level (Z) and design effect (D) are held constant at 95% and 1 respectively.
[7] The confidence level (Z) and design effect (D) are held constant at 95% and 2 respectively.

Table 1 – Simple Random and Systematic Sampling

Maximum Tolerable Error (+/-)

Estimate for Key Indicator

	5 pp	10 pp	15 pp
5%	73		
10%	139	35	
15%	196	49	22
20%	246	62	28
25%	289	73	33
30%	323	81	36
35%	350	88	39
40%	369	93	41
45%	381	96	43
50%	385	97	43
55%	381	96	43
60%	369	93	41
65%	350	88	39
70%	323	81	36
75%	289	73	33
80%	246	62	28
85%	196	49	22
90%	139	35	
95%	73		

Table 2 – Cluster and Two-Stage Cluster Sampling

Maximum Tolerable Error (+/-)

Estimate for Key Indicator

	5 pp	10 pp	15 pp
5%	146		
10%	278	70	
15%	392	98	44
20%	492	124	56
25%	578	146	66
30%	646	162	72
35%	700	176	78
40%	738	186	82
45%	762	192	86
50%	770	194	86
55%	762	192	86
60%	738	186	82
65%	700	176	78
70%	646	162	72
75%	578	146	66
80%	492	124	56
85%	392	98	44
90%	278	70	
95%	146		

Example An assessment in West Bank/Gaza will employ a two-stage, cluster sampling method (table Y). An estimate for the key food security indicator is 60% for the population of interest (% in row). The assessment team decides that the estimate for the population should have a degree of error no larger that 5 percentage points (pp in column) in either direction (+/- 5 pp). The required sample size is n = 738.

Maximum Tolerable Error (+/-)

Column 5 pp, row 60%
n = 738

Estimate for Key Indicator

	5 pp	10 pp	15 pp
5%	146		
10%	278	70	
15%	392	98	44
20%	492	124	56
25%	578	146	66
30%	646	162	72
35%	700	176	78
40%	738	186	82
45%	762	192	86
50%	770	194	86
55%	762	192	86
60%	738	186	82
65%	700	176	78
70%	646	162	72
75%	578	146	66
80%	492	124	56
85%	392	98	44
90%	278	70	
95%	146		

3.1.3 - Web-based sample size calculators

A web-based sample size calculator (http://calculators.stat.ucla.edu) can be used for scenarios not contained in tables 1 and 2.

- Select the calculator 'Sample Size Calculator (4th from the top)
- Select 'Proportion'
- Enter the desired value for 'Maximum allowable difference' – this is the same as the maximum tolerable error (usually between .05 and .15)
- Leave default value for 'confidence' (.95)
- Enter the estimated population proportion for the key food security indicator of interest. If no estimate is available use the default value (.50)
- Click 'Submit query'
- Remember that the sample size calculated is for simple random and systematic samples. If you are using cluster or two-stage, cluster sampling you must multiply the sample size by the design effect (default = 2)

Example The assessment will employ a two-stage cluster sampling method (design effect = 2). The estimated population proportion is 35% and the maximum tolerable error is +/- 5 percentage points.

Estimating a Proportion	
Population Size	Infinity
Maximum Allowable Difference	0.05
Confidence	0.95
Population Proportion	0.35
Submit Query	
Required Sample Size	350

3.2 - Stratified Samples

The sample size calculation for stratified samples is slightly different due to the fact that comparisons between sub-groups (strata) are an important part of the objective of the assessment.

3.2.1 - Sampling when key indicators are expressed as percentages

In the formula for non-stratified samples the confidence interval around the estimate derived from sample is defined at 95% to ensure that there is only a 5% probability that the true population proportion falls outside of this confidence interval (Z in the formula for non-stratified samples). For stratified samples this same factor (e.g. statistical confidence) can be described as the confidence with which it is desired to be able to conclude that an observed difference between sub-groups did not occur by chance[8]. In addition, the confidence with which it is desired to be certain of detecting a difference between sub-groups if one actually exits (e.g. statistical power) must also be defined[9].

$$\mathbf{n = D\,[(Z_{alpha} + Z_{beta})^2 * (P_1\,(1\text{-}P_1) + P_2\,(1 - P_2)/(P_2 - P_1)^2]}$$

Where:

n	=	Required minimum sample size per strata (zone)
D	=	Design effect (varies by type of sampling)
$\mathbf{P_1}$	=	Estimated level of an indicator measured as a proportion in decimal form

---

[8] Statistical confidence controls for type I or alpha errors. Alpha is the probability of falsely accepting difference a difference between sub-groups when in fact there is no difference.

[9] Statistical power controls for type II or beta errors. Beta is the probability of falsely accepting no difference between sub-groups when in fact a difference does exist.

P_2 = The estimated level of the same indicator for a comparison sub-group such that the difference between P_2 and P_1 is the minimum difference between sub-groups that the sample is designed to detect.

Z_{alpha} = The Z-score corresponding to the degree of confidence with which it is desired to be able to conclude that an observed difference between strata of size ($P_2 – P_1$) would not have occurred by chance (the level of statistical significance)

Z_{beta} = The Z-score corresponding to the degree of confidence with which it is desired to be certain of detecting a difference between strata of size ($P_2 – P_1$) if one actually exists.

3.2.2 Sample size guidance tables

Tables 3 and 4 provide sample size guidance from common scenarios encountered in VAM food security assessments[10]. To use the tables, locate the percentage corresponding with the estimate for the key food security indicator of interest (% in column). Next, decide on the magnitude of difference you want to be able to detect in percentage points (pp in rows). The options included in the table are 5, 10, 15, and 20 percentage points.

Table 3 – Sample Size for Stratified Samples: Simple Random and Systematic Sampling

Estimate for Key Indicator

Minimum Detectable Difference	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
5 pp	473	724	943	1131	1288	1413	1507	1570	1601	1601	1601	1570	1507	1413	1288	1131	943	724	473
10 pp	160	219	269	313	348	375	395	407	411	407	411	407	395	375	348	313	269	219	160
15 pp	88	113	134	151	165	176	182	186	186	182	186	186	182	176	165	151	134	113	88
20 pp	66	72	82	91	98	103	106	107	106	103	106	107	106	103	98	91	82	72	66

Table 4 – Sample Size for Stratified Samples: Cluster and Two-Stage Cluster Sampling

Estimate for Key Indicator

Minimum Detectable Difference	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
5 pp	946	1448	1886	2262	2576	2826	3014	3140	3202	3202	3202	3140	3014	2826	2576	2262	1886	1448	946
10 pp	320	438	538	626	696	750	790	814	822	814	822	814	790	750	696	626	538	438	320
15 pp	176	226	268	302	330	352	364	372	372	364	372	372	364	352	330	302	268	226	176
20 pp	132	144	164	182	196	206	212	214	212	206	212	214	212	206	196	182	164	144	132

[10] Table 3: The confidence level (Za), power level (Zb) and design effect (D) are held constant at 95%, 80% and 1 respectively. Table 4: The confidence level (Za), power level (Zb) and design effect (D) are held constant at 95%, 80% and 2 respectively.

Example The assessment will employ a simple random sampling method (table 3). The estimated percentage of food insecure households for strata 1 is 20%. You want to be able to detect a difference between strata 1 and other strata when the true difference is +/- 10 percentage points. The required sample size is n = 313 for strata 1 (red). The estimated percentage of food insecure households for strata 2 is 40%. You also want to be able to detect a difference between strata 2 and other strata when the true difference is +/- 10 percentage points. The required sample size if n = 407 for strata 2 (green).

Row 10 pp, Column 20%
n = 313

	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
5 pp	473	724	943	1131	1288	1413	1507	1570	1601	1601	1601	1570	1507	1413	1288	1131	943	724	473
10 pp	160	219	269	313	348	375	395	407	411	407	411	407	395	375	348	313	269	219	160
15 pp	88	113	134	151	165	176	182	186	186	182	186	186	182	176	165	151	134	113	88
20 pp	66	72	82	91	98	103	106	107	106	103	106	107	106	103	98	91	82	72	66

Row 10 pp, Column 40%
n = 407

Example The assessment will employ a two-stage, cluster sampling method (table 4). Strata level (sub-groups defined by stratification criteria) estimates of the percentage of food insecure households are unavailable. But, an overall estimate for the population of interest is available (60%). At minimum, you want to be able to detect a difference between strata when the true difference is +/- 15 percentage points. The required sample size for each strata is n = 372.

Row 15 pp, Column 60%
n = 372

	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
5 pp	946	1448	1886	2262	2576	2826	3014	3140	3202	3202	3202	3140	3014	2826	2576	2262	1886	1448	946
10 pp	320	438	538	626	696	750	790	814	822	814	822	814	790	750	696	626	538	438	320
15 pp	176	226	268	302	330	352	364	372	372	364	372	372	364	352	330	302	268	226	176
20 pp	132	144	164	182	196	206	212	214	212	206	212	214	212	206	196	182	164	144	132

3.2.3 - Web-based sample size calculators

A web-based sample size calculator (http://calculators.stat.ucla.edu) can be used for scenarios not contained in Tables 3 and 4.

- Select the calculator 'Power Calculator'
- Select the button (sample size for a given power) in the row entitled Fisher's Exact Test[11]

[11] Using the Fisher's Exact Test yields the most appropriate sample size regardless of the type of analyses that will be performed (SamplePower Manual, SPSS)

- Enter the estimated value for the key indicator interest in the field P-1
- Enter the value corresponding to the difference you wish to be able to detect in the field P-2. Always choose the value that is closer to 50% from the estimate as it will yield the larger sample size requirement.

> ***Example*** If the estimated value for the key indicator of interest is 20% and you want to detect differences between sub-groups of 10% of more, both 10% and 30% correspond with this difference. Enter 30% for P-2, since it is closer to 50%.

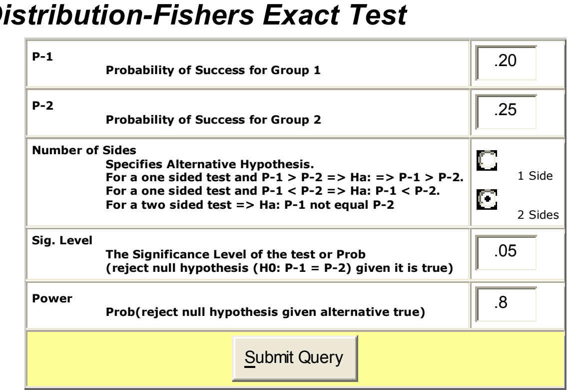
> ***Example*** If the estimated value for the key indicator of interest is 60% and you want to detect differences between sub-groups of 10% of more, both 50% and 70% correspond with this difference. Enter 50% for P-2, since it is closer to 50%.

- Choose two-sided for number of sides (e.g. to capture a difference in either direction[12])
- The default value for 'Sig. Level' is .05 (this corresponds to .95 confidence from the non-stratified calculation)
- The default value for 'Power' is .80
- Click 'Submit query'
- Remember that the sample size calculated is for simple random and systematic samples. If you are using cluster or two-stage, cluster sampling you must multiply the sample size by the design effect (default = 2)

Example An assessment will employ a two-stage cluster sampling method (design effect = 2). The sample is stratified into 3 groups. The estimated population proportion for strata 1 is 25%. The assessment team decides that small differences in the percentage of food secure between strata are not very important for program decision making. Therefore, it is decided that detecting differences of +/- 10 percentage points or more between this and other strata would provide adequate information for comparing strata.

Binomial Power Calculations

Binomial Distribution-Fishers Exact Test



P-1 **Probability of Success for Group 1**	.20
P-2 **Probability of Success for Group 2**	.25
Number of Sides **Specifies Alternative Hypothesis. For a one sided test and P-1 > P-2 => Ha: => P-1 > P-2. For a one sided test and P-1 < P-2 => Ha: P-1 < P-2. For a two sided test => Ha: P-1 not equal P-2**	1 Side / 2 Sides
Sig. Level **The Significance Level of the test or Prob (reject null hypothesis (H0: P-1 = P-2) given it is true)**	.05
Power **Prob(reject null hypothesis given alternative true)**	.8
Submit Query	

[12] The sample size requirement for one-sided tests (e.g. designed to capture differences or change in one direction) are smaller, but inappropriate for most assessments. One-sided tests are more appropriate for program evaluations in which declines are unlikely and improvements are expected.

Binomial Power Calculations

Binomial Distribution-Fishers Exact Test

P-1 **Probability of Success for Group 1**	.20
P-2 **Probability of Success for Group 2**	.30
Number of Sides **Specifies Alternative Hypothesis.** **For a one sided test and P-1 > P-2 => Ha: => P-1 > P-2.** **For a one sided test and P-1 < P-2 => Ha: P-1 < P-2.** **For a two sided test => Ha: P-1 not equal P-2**	1 Side 2 Sides
Sig. Level **The Significance Level of the test or Prob** **(reject null hypothesis (H0: P-1 = P-2) given it is true)**	.05
Power **Prob(reject null hypothesis given alternative true)**	.8
Submit Query	

Result:

N: 313.>

Multiply 313 x the design effect (2.0)

n= 626

Therefore, the sample size required for stratum 1 is 626. Repeat the steps for strata 2 and 3.

Section IV - Two examples from the field

This section outlines two full examples from the field (Haiti and Tanzania) that illustrate the various stages of the sampling decision-making process.

4.1 - Haiti

WFP Haiti conducted a household food security and vulnerability survey in 2004. One purpose of the survey was to quantify the number and severity of food insecure households by deriving prevalence estimates from a sample survey. The Country Office (CO) chose to use probability sampling so that a) statistical inferences could be made from the sample to the larger population from which the sample was taken and b) estimates generated would have a quantifiable degree of error.

4.1.1 - Stratification

Although overall estimates were important, the CO deemed it necessary to also have estimates with a pre-defined level of precision at lower aggregations (stratified sampling). The assessment was originally designed to yield estimates for each of four departments that comprised the population of interest. After some consideration it was decided that estimates for sub-groups defined by land-use zones within these four departments would be more useful for programming purposes. Therefore, the sample was designed to yield estimates with pre-defined levels of precision for the following 14 sub-groups (strata).

Table 6 – Stratification: Land-Use Zones (14) by Department

	Center	North	North-East	West
Cultures agricoles denses	X	X	X	X
Systèmes agroforestiers denses	X	X	X	
Cultures agricoles moyennement denses	X	X	X	X
Savanes / Pâturage avec présence d'autres occupations des sols	X			X
Urbain Discontinu				X

4.1.2 - Sampling Method and Sampling Frame

Accurate information at the household level was unavailable, making simple random sampling and systematic sampling impractical. However, census data from 1996 provided information on the size and location of localities (e.g. villages). A decision was made to use a two-stage cluster sampling method with households as the unit of analysis (ultimate sampling unit) and localities serving as clusters (primary sampling unit).

Localities were then categorized by land-use zone with each village belonging to only one land-use zone (mutually exclusive) and all villages within the population categorized (collectively exhaustive). A list of all localities was constructed for each of the 14 strata identified in table A.

4.1.3 - Sample Size

Department-level estimates of stunting prevalence were used as a basis for calculating the required sample size for strata contained in each of the four departments. The stunting estimates for each department are:

Center	35%
North	25%
Northeast	30%
West	20%

The CO decided that only approximate food security estimates were required and, therefore, the strata level estimates did not need to be very precise. After considering the costs, a decision was made to pre-define the minimum detectable difference between strata at 20 percentage points in either direction. The required sample size for the strata within each department was determined using the table provided in the sampling guidelines (Section 3.2.2, Table 4).

Estimate for Key Indicator

Minimum Detectable Difference

	5%	**10%**	**15%**	**20%**	**25%**	**30%**	**35%**	**40%**	**45%**	**50%**	**55%**	**60%**	**65%**	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**
5 pp	946	1448	1886	2262	2576	2826	3014	3140	3202	3202	3202	3140	3014	2826	2576	2262	1886	1448	946
10 pp	320	438	538	626	696	750	790	814	822	814	822	814	790	750	696	626	538	438	320
15 pp	176	226	268	302	330	352	364	372	372	364	372	372	364	352	330	302	268	226	176
20 pp	132	144	164	182	196	206	212	214	212	206	212	214	212	206	196	182	164	144	132

West Dept → 182 (20 pp, 20%); North Dept → 196 (20 pp, 25%); NE Dept → 206 (20 pp, 30%); Center Dept → 212 (20 pp, 35%)

Center	=	4 strata at 212 each	n = 848
North	=	3 strata at 206 each	n = 618
Northeast	=	3 strata at 196 each	n = 588
West	=	4 strata at 182 each	n = 728
TOTAL	**=**	**14 strata**	**n = 2,783**

4.1.4 - Choosing the Clusters: How Many and Which Ones?

Although 30 clusters within each stratum would be ideal, the number of strata dictates that a compromise be made. To reduce the number of localities to be visited, a decision is made to take 20 clusters in each stratum. Because the sample size required per strata varies by department, the number of households to be taken within each cluster varies by department.

Center	212/20 = 10.6	= 11 hh per cluster
North	206/20 = 10.3	= 11 hh per cluster
Northeast	196/20 = 9.8	= 10 hh per cluster
West	182/20 = 9.1	= 10 hh per cluster
TOTAL		**= 42 hh per cluster**

Within each stratum, clusters were chosen randomly with probability proportional to size (PPS). A random numbers generator available at www.randomizer.org was used to select clusters/localities for inclusion in the sample. The number of households in each locality contained in a stratum were added together to get the cumulative number of households. This number was used to define the range from which random numbers were selected. Next, 20 random numbers were generated (e.g. equal to the number of clusters required). The localities containing these numbers in the column CUMM NUM (e.g. the cumulative number of households) were included in the sample.

Two replacement localities were identified for each of the selected localities in the event that the original selected locality cannot be located. Where such replacements are made, enumeration team supervisors noted the replacement. Replacements localities were the two closest localities within the zone and commune of the original selection.

4.1.5 - Selecting Households within Clusters

Due to the lack of locality maps/locality population figures and time constraints that prevented mapping all households within selected localities, the Expanded Program on Immunization (EPI) approach was used to select households within localities. The following 6 steps are required:

1. In each selected locality, 2 to 3 key informants were asked to identify the approximate center of the locality.

2. Once in the center, the enumeration team supervisor spun a pencil to pick a random-walking direction.

3. Once the direction has been chosen, the supervisor and enumerators walked in a straight line counting the number of households in that line until the end of the boundary is reached. For the majority of villages these tasks can be completed within a reasonable timeframe. However, for the few large (e.g. in excess of 500 households) or geographically dispersed localities that were encountered, 2 to 3 key informants were used to approximate the number of households that would be encountered when walking from the center to the perimeter of the locality.

4. Next the total number of households in the line was divided by the number of households required in the locality (varies by department) to derive a sampling interval. An interval of less than 2 required sampling each household in the random walk-direction. An interval greater than 2, but less than 3 required sampling every other household in the random walk direction. An interval greater than 3, but less than 4 required sampling every third household and so on. An illustration is provided on the next page.

5. When the interval was less than 1 (e.g. fewer households in the random walk direction than were needed for the locality), all households in the random-walk direction were sampled. Then the survey team returned to the center of the locality to pick another random walk direction in order to sample the required number of households. This procedure was repeated until the total number of households required was achieved.

6. When a selected household was unavailable to participant in the survey, enumerators proceeded to the next selected households. On some occasions this replacement strategy required that the enumeration team return to the center to pick a second random walk direction (as for localities in which the initial interval was less than 1).

4.2 - Tanzania

In 2004, the Government of Tanzania restricted refugee access to external markets due to security concerns. The WFP CO suspected that this had a negative impact on the food security status of refugee populations. A decision was made to undertake a food security assessment, using probability sampling methods, in order to quantify the prevalence of food insecurity in the refugee camps.

4.2.1 - Stratification

Market restrictions were unevenly applied across the twelve refugee camps located in western Tanzania (e.g. the population of interest). To assess the effect of market access of food security status, the CO divided the population of interests into two strata according to market access[1] and a separate sample was taken from each (Table 7).

Table 7 – Stratification Criteria

Strata 1 – Good market access	Strata 2 – Poor market access
1. Lukole A	1. Mtabila 2
2. Lukole B	2. Muyovozi
3. Nduta	3. Nyarungusu
4. Kanembwa	4. Lugufu 1
5. Mtendeli	5. Lugufu 2
6. Karago	
7. Mtabila 1	

[1] Each camp was classified into one of four categories by WFP program staff during the CSI Training Workshop in May, 2004: Very Good = External markets with good supply, Good = internal markets with good supply, Poor = internal markets with limited supply, Very Poor = no markets. These categories were collapsed into 2 categories, good and poor, for use in stratifying the sample.

4.2.2 - Sampling Method and Sampling Frame

Relatively complete lists of refugee households and their address (block and household numbers) were available for each camp through UNHCR (household level sampling frame). This list was used to construct two sampling frames, one for each 'market access' strata.

To ensure that each camp was included in the sample population, systematic sampling was used. Therefore, households were both the primary and ultimate sampling units.

4.2.3 - Sample Size

No food security estimates were available for the refugee population. Therefore, the default value of 50% was used in calculating the sample size required from each strata. For programming purposes a difference in the prevalence of food insecure households of less than 10 percentage points between strata was deemed marginal. Therefore, 10 percentage points was defined as the minimum difference to be detected. The required sample size for each stratum (n = 407) was determined using the table provided in the sampling guidelines systematic sampling for a stratified sample (Section 3.2.2, Table 3). The total sample size was 814

Estimate for Key Indicator

Minimum Detectable Difference	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
5 pp	473	724	943	1131	1288	1413	1507	1570	1601	1601	1601	1570	1507	1413	1288	1131	943	724	473
10 pp	160	219	269	313	348	375	395	407	411	407	411	407	395	375	348	313	269	219	160
15 pp	88	113	134	151	165	176	182	186	186	182	186	186	182	176	165	151	134	113	88
20 pp	66	72	82	91	98	103	106	107	106	103	106	107	106	103	98	91	82	72	66

Per strata sample size
n = 407

4.2.4 - Selecting Households

For each stratum, the total number of households was divided by 407 to derive a sampling interval (S.I.). A random starting household was chosen between 1 and the S.I. to select the first household for inclusion in the sample. The second household was selected by adding the S.I. to the random starting household. The third household was selected by adding the S.I. to the sum of the S.I. and the random start…..and so on until 407 households were selected in each stratum.

A protocol was developed for replacing households in which an appropriate respondent was unavailable. The desired respondent for the questionnaire was the head of household; defined as the primary decision maker within the household concerning food and income use decisions. When this person was unavailable, the spouse of the head of household was interviewed. If the spouse was unavailable, any other adult age 16 or above in the household was interviewed. If no respondents meeting these criteria were available the household was replaced by selecting the next closest plot in any direction as described in the survey protocol.

Annex I – References and additional resources

General Sampling Guidance

Sampling Guide. FANTA. Magnani, Robert, 1997.

Constructing Samples for Characterizing Household Food Security and for Monitoring and Evaluating Food Security Interventions: Theoretical Concerns and Practical Guidelines. IFPRI Technical Guide #8. Carletto, 1999.

Sample Size Calculators (on-line)

UCLA Department of Statistics (http://calculators.stat.ucla.edu)

- Non-stratified samples: (http://calculators.stat.ucla.edu/sampsize.php)
- Stratified samples: (http://calculators.stat.ucla.edu/powercalc)

University of Calgary (http://www.health.ucalgary.ca/~rollin/stats/ssize/)

Random Numbers Generators (on-line)

Randomizer (www.randomzer.org)

Random (www.random.org)

Food Security Indicators

Food Security Indicators and Framework for use in the Monitoring and Evaluation of Food Aid Programs. FANTA. Riely, Mock, Cogill, Bailey, and Kenefick, 1996.