

SMART

**Standardized Monitoring and Assessment
for Relief and Transitions**

Manual 2.0

2017

© 2017. All rights reserved. This document is copyrighted by SMART, Action Against Hunger Canada, and the Technical Advisory Group. It may be reproduced as a complete document, without charge or permission provided that the original source is acknowledged and copyright retained by the authors. It can be freely used in national protocols by governments, international agencies and humanitarian charities involved with nutrition and mortality assessments provided that it is not abstracted, changed or edited without the changes being agreed by the copyright holders (which will not be withheld if the changes are deemed to be safe and give reliable data). For copying in other circumstances the prior written permission must be obtained from the copyright holders/ authors. The document or any part of the document cannot be reproduced in any format for sale, used for commercial purposes or referred to in advertising literature to promote any products. Permission should be obtained by SMART, Action Against Hunger Canada, or the Technical Advisory Group if it is used to train fee-paying students.

Where this document or sections are reproduced in derivative publications the following should be inserted in the acknowledgements: "Reproduced with permission from copyrighted material of SMART, Action Against Hunger Canada and the Technical Advisory Group."

Table of Contents

TABLE OF CONTENTS	3
LIST OF FIGURES AND TABLES	5
INTRODUCTION	7
1. PLANNING A SURVEY	8
DETERMINE THE RELEVANCE AND TIMING OF DOING A SURVEY	8
GATHER EXISTING AND CONTEXTUAL INFORMATION FROM LOCAL PARTNERS AND AUTHORITIES	9
DEFINE SURVEY OBJECTIVES.....	9
DEFINE GEOGRAPHIC AREA AND POPULATION GROUP TO BE SURVEYED	10
DECIDE WHAT ADDITIONAL INFORMATION TO COLLECT	11
2. SURVEY TEAMS & TRAINING	12
TRAINING OF SURVEY TEAMS.....	12
STANDARDISATION TEST.....	14
STANDARDISATION TEST PROCEDURE	16
INTERPRETATION OF THE STANDARDISATION TEST REPORT.....	20
PILOT TESTING.....	24
SURVEY MANAGEMENT.....	25
3. SAMPLING METHODS AND SAMPLE SIZE CALCULATION	27
WHAT IS SAMPLING?	27
SAMPLING METHODS	29
SAMPLE SIZE CALCULATION.....	35
SECOND STAGE SAMPLING	45
SPECIAL CASES	49
4. ANTHROPOMETRY	53
WHY DO A NUTRITION SURVEY?	53
POPULATIONS FOR ANTHROPOMETRIC SURVEYS: 6–59-MONTH-OLD CHILDREN.....	53
WHEN TO MEASURE THE NUTRITIONAL STATUS OF PEOPLE OVER AGE FIVE.....	53
NUTRITION INDICES AND INDICATORS	54
THE REFERENCE POPULATION CURVES	55
EXPRESSION OF NUTRITION INDICES	56
DEFINITIONS OF ACUTE MALNUTRITION IN CHILDREN AGED 6–59 MONTHS	57
NUTRITIONAL MEASUREMENTS	59
5. PLAUSIBILITY CHECKS	64
INTRODUCTION	64
SUMMARY TABLE OF THE ANTHROPOMETRY PLAUSIBILITY CHECK.....	65
SECTIONS OF THE ANTHROPOMETRY PLAUSIBILITY CHECK REPORT.....	67
SECTIONS OF MORTALITY AND DEMOGRAPHY PLAUSIBILITY CHECK REPORT	87
CONCLUSION ON THE PLAUSIBILITY CHECKS	88
6. MORTALITY AND DEMOGRAPHY	90
HOW TO DETERMINE THE NEED TO CONDUCT A MORTALITY AND DEMOGRAPHY SURVEY?	91
HOW SHOULD THE DATA BE COLLECTED?	92
MORTALITY AND DEMOGRAPHY SURVEY OBJECTIVES	93
CRUDE DEATH RATE (CDR)	95
UNDER-FIVE DEATH RATE (U5DR) AND UNDER-FIVE MORTALITY RATE (U5MR)	97
SAMPLING DESIGN.....	98
DECIDING ON THE LENGTH OF THE RECALL PERIOD	98
SAMPLE SIZE CALCULATION — CDR.....	101

SAMPLE SIZE CALCULATION — U5DR	106
ENUMERATING HOUSEHOLD MEMBERS	107
MORTALITY AND DEMOGRAPHY QUESTIONNAIRES	108
7. INTERPRETATION OF RESULTS	116
ANNEX 1: POSSIBLE BIASES FOR SMART SURVEYS AND RECOMMENDATIONS ON HOW TO AVOID THEM.....	117
ANNEX 2: SMART METHODOLOGY – CHRONIC MALNUTRITION AS THE MAIN INDICATOR. 120	
ANNEX 3: MORTALITY & DEMOGRAPHY SURVEY TRAINING SCHEDULE EXAMPLE	121
ANNEX 4: EPI INFO/ENA MANUAL.....	123
2. SOFTWARE INSTALLATION	124
4. MAKING A SURVEY WITH THE QUESTIONNAIRE TEMPLATE OF EPIINFO/ENA	129
5. IMPORTING SURVEY DATA FROM ANOTHER SOFTWARE	131
6. GENERATING THE ANTHROPOMETRIC INDICATORS OF A SURVEY	133
7. AUTOMATIC GENERATION OF THE DATA QUALITY REPORT.....	136
8. AUTOMATIC GENERATION OF THE ANTHROPOMETRY GRAPHS AND SURVEY REPORT	137
9. SOFTWARE SETTINGS	139
10. SURVEY PLANNING AND SAMPLE SIZE CALCULATION	140
11. ANTHROPOMETRY EXERCISE FOR TRAINING IN EPI INFO/ENA	140
12. DATA ANALYSIS	141
ANNEX 5: REFERENCES.....	144

List of Figures and Tables

Figure 1: Practicing Weight, Height and MUAC Measurements.....	13
Figure 2: Standardisation Test Forms Templates.	18
Figure 3: Example of TEM output from ENA in the new Standardisation Test report.	21
Figure 4: Use of the two Outputs of Accuracy.....	23
Figure 5: Acceptable Limits for Standardisation Test Measurements.	23
Figure 6: Camp Stankovic II, Skopje. Macedonia, 1999.	27
Figure 7: General Types of Sampling.	28
Figure 8: ENA for SMART Random Number Generator.	29
Figure 9: List of Random Numbers in a Word Document Generated by ENA.	30
Figure 10: Example of Geometric Setting of Dwellings (adapted from MSF).	32
Figure 11: Example of Villages of Different Sizes.	33
Figure 12: ENA table for cluster selection.....	34
Figure 13: Precision and accuracy.....	36
Figure 14: Effect of Changing the Estimated Prevalence (assuming 95% CI, +/- .05, large population).....	38
Figure 15: Effect of Changing the Desired Precision (assuming 95% CI, 50% prevalence, large population) (source: CDC).	39
Figure 16: ENA functionality for assigning clusters.	44
Figure 17: Segmentation into Unequal Parts.	46
Figure 18: Decision Tree for Household Selection at the Last Stage of Sampling.	48
Figure 19: Example of a Cluster Control Form.	50
Figure 20: Reference curve used to derive a z-score.....	56
Figure 21: How to Measure the Length of a Child Aged Less than 2 years of age.	61
Figure 22: How to measure the height of a child of more than 87cm.....	62
Figure 23: How to Check for Oedema.	63
Figure 24: Summary Table of the Anthropometry Plausibility Check in ENA for SMART.	66
Figure 25: Example of SMART vs. WHO Flags.....	69
Figure 26: Location of Flagged Values.	70
Figure 27: Statistical Evaluation of Sex and Age Ratios Table.	72
Figure 28: Expected Proportions of Children by Age Group.	72
Figure 29: Evaluation of Standard Deviation Table.	76
Figure 30: Distribution of WHZ with Good Standard Deviation (mean -1.17, standard deviation 1.11).	76
Figure 31: Effects of a "Melting" Curve on the Prevalence of GAM with a Large Standard Deviation (mean -1.12, standard deviation 1.55).	77
Figure 32: Examples of Distribution.	79
Figure 33: Index of Dispersion as Reported in the ENA for SMART Software.	81
Figure 34: Distribution of Cases of GAM per Cluster.....	82
Figure 35: Time Analysis on the SD for WHZ.	84
Figure 36: How to Generate the Mortality and Demography Plausibility Check Report.	87
Figure 37: Household Members During the Recall Period (Adapted from Woodruff, 2002).	107
Figure 38: Demography and Mortality Questionnaire.	110
Table 1: Standardisation Test Equipment and Materials.	15
Table 2: Units of analysis for anthropometric measurement, new Standardisation Test report.	20
Table 3: Recommended Options if Enumerators fail the Standardisation Test.	24
Table 4: Example of Cluster Selection by PPS.	34
Table 5: Effect of Changing Precision on the Sample Size.	39
Table 6: Variation of Precision Needed with Different Malnutrition Prevalences.	39
Table 7: Effect of Changing Design Effect on Sample Size (assuming a 3% precision).	41
Table 8: Definitions of Acute Malnutrition Using WFH and/or Oedema in Children aged 6–59 months.....	57

Table 9: Percentage of Body Weight Accounted for by Nutritional Oedema.	63
Table 10: Sample Size where one Erroneous Flag (Real Value) is Expected with Various SMART Flag Exclusion Ranges and Various Standard Deviations (SD) of Z-score Distribution.....	68
Table 11: The Number of GAM or MAM Cases Expected in a Survey for each SAM Case (oedema excluded), depending on survey z-score mean and SD.	82
Table 12: Advantages and Disadvantages of Household Mortality and Demography Surveys.	93
Table 13: Calculation of Person-Years and Person-Days Lived in a Population Numbering 500 on January 1 and 600 on December 31.	96
Table 14: Selecting a Recall Period for a Mortality and Demography Survey.	98
Table 15: Advantages and Disadvantages of Shorter and Longer Recall Periods.	100
Table 16: Mortality Benchmarks for Defining Crisis Situations.	102
Table 17: The Effect of Changing Recall Period in Mortality & Demography Survey Sample Size Calculations.	104
Table 18: Calculating Sample Sizes for Mortality & Demography Surveys (CDR).	105
Table 19: Calculating Sample Sizes for Mortality & Demography Surveys (U5DR).	106

Introduction

SMART is a simplified and standardised household-level survey methodology used to understand the severity and magnitude of a public health situation. It is applicable in all contexts, including humanitarian crises, development settings, and displaced populations.

The basic indicators for assessing the *severity* of a crisis are the **mortality**, or death rate, and the **nutritional status** of the population. These are both estimated by conducting a survey of the affected population.

To know the *magnitude* of the problem we also need to know the population size and, if possible, the demographic characteristics of the population. The scale and type of intervention will depend upon the magnitude of the crisis rather than simply on the prevalence of malnutrition. For example, a high proportion of malnourished in a small population is normally of less magnitude than a lower proportion of malnourished in a large population.

In order to understand the reasons for the crisis and to plan and implement appropriate relief, the following must be considered:

- The usual situation for that population;
- The evolution of the changes;
- The context in which the emergency has arisen.

There are many sources of information that are relevant in putting a crisis in context and that may affect the types of responses that are appropriate. Context-specific information is also required for more stable populations in development settings, where an indication of chronic malnutrition may be of importance. In order to be useful, however, the information has to be relatively easy to collect, reliable and accurate.

It is not difficult to conduct a survey, but there are a number of critical points that have to be correct in order for the results to be valid. It does require planning, training, supervision of staff, interaction with the community and a basic understanding of the concepts of epidemiology and statistics.

This manual is designed to provide agencies with the basic tools to collect those data necessary for planning direct interventions in an emergency or development setting, or for advocacy purposes.¹ It should be used in conjunction with the accompanying software, Emergency Nutrition Assessment (ENA) for SMART, which is freely available from www.smartmethodology.org.

¹ Collecting and analysing data for advocacy, policy-making and other such purposes is also necessary; methods for collecting data for these purposes may be different from the methods advocated in this manual. Because this manual is designed for use by field staff without epidemiological knowledge or experience, it is limited to the methods that yield reliable information for programming.

1. Planning a Survey

The steps to consider when planning cross-sectional field surveys are as follows:

1. Determine the relevance and timing of doing a survey;
2. Gather existing and contextual information from local partners and authorities;
3. Define survey objectives;
4. Define the geographic area and population group(s) to be surveyed;
5. Decide what additional information to collect.

Determine the relevance and timing of doing a survey

The decision to undertake an assessment is usually made in conjunction with the government, partner agencies and donors. It is always important to share information about when and where you plan to undertake a survey to prevent unnecessary repetition or overlap by different agencies. Surveys are usually much more informative if they are coordinated so that data from several agencies, geographic areas or population groups can be examined together to give a wider perspective on the situation.

Conducting an assessment is expensive and time-consuming, so before starting an assessment you should consider the following points:

- *Are the results crucial for decision-making?* If a population's needs are obvious, immediate programme implementation takes priority over doing a survey, and the survey should be deferred. If large numbers of malnourished children are present at a centre, implementation of relief programmes should not be delayed until a survey is conducted. The Sphere minimum standards require that a nutrition assessment is conducted when a targeted feeding programme is implemented (The Sphere Project, 2004); this does not mean that the programme should be delayed until after the survey is complete if the need is clear. Where such programmes exist, periodic surveys should be conducted. If another agency has recently carried out a nutrition assessment in the same area, then that data should be used rather than repeating the survey.
- *Will it be possible to implement evidence-based recommendations?* There is little point of doing a survey if it is already known that a response will not be possible. If the agency cannot itself implement a programme where needed, the results must be useful in advocating for a response.
- *Is the affected population accessible?* Insecurity or geographical constraints may result in limited access to the population of interest. If this is extreme, a survey cannot be conducted.

The exact dates of the assessment should be chosen with the help of community leaders and local authorities in order to avoid market days, local celebrations, food distribution days, vaccination campaigns, or other times when people are likely to be away from home. Roads may be impassable during the rainy season. In agricultural areas, women may be in the fields for most of the day during ground preparation, planting, or harvesting. Healthy children are more likely to accompany adults to the market or the fields and are less likely to be at home than ill or malnourished children. The survey results could be wrong if only children who were at home at the time of the survey team's visit are sampled. Wherever possible, the community leaders should inform the villages chosen to be surveyed in advance.

There needs to be sufficient time allocated for preparation and literature review, training, pilot testing, community mobilisation, data collection, analysis and reporting.

Gather existing and contextual information from local partners and authorities

Before starting the survey, it is important to learn as much about the population as possible from existing sources. These include population characteristics and figures, previous surveys and assessments, health statistics, situation reports (security and political situation), maps, and anthropological, ethnic and linguistic information. Only after these data are gathered can a judgement be made about any extra information that should be collected.

It is absolutely essential to meet the community leaders and local authorities before starting a survey. The meetings should at least cover the following points:

- **Discuss with the community about the objectives of the survey.** If the population does not understand why you are doing an assessment they may not co-operate during the survey.
- **Obtain a map of the area in order to plan the survey.** Use this map during the discussions with the local authorities and community leaders.
- Obtain detailed information on population figures (particularly at the village or camp level).
- Obtain information on security and access to the prospective survey area.
- **Obtain letters of permission from the local authorities (in the local language),** addressed to the district or village leaders, stating that you will be visiting. The letters should explain why you are conducting an assessment and ask for the population's co-operation.
- Agree upon the dates of the survey with the community and local authorities.
- **Agree how the results will be used.** In particular, realistically discuss the prospects for intervention, the steps that will be taken and types of programmes that are likely to be implemented if the situation is found to be as poor as expected. Do not make promises that may not be fulfilled.

Define survey objectives

Clear objectives make it much easier for your team, the population, and donors to understand why the survey is being conducted. This should be clearly stated at the outset.

Emergency nutrition assessments are usually conducted to assess the *severity* of the situation by quantifying the acute malnutrition and mortality in a given population at a defined point in time. This is done by estimating the prevalence of wasting and oedema in children aged 6-59 months and the death rate of the entire population. With an estimate of population size, the proportions of malnourished and the death rate give an estimate of the absolute number of malnourished there are in the community and how many have died in the recent past. These figures indicate the *magnitude* of the problem. The estimates, together with previous surveys and contextual data, also indicate the *urgency* of the situation and how it may evolve in the future.

Where the survey is undertaken during a stable period, the data can be used to establish a baseline, from which future changes can be monitored over time.

Undertaking a nutrition and mortality and demography survey provides an opportunity to collect additional information that can be critical in deciding which interventions are most important. Immunisation and nutrition programme coverage, vitamin A, iodine, anaemia or other micronutrient deficiency, disease morbidity, trauma experience, cause of death, demographics, migration and many other variables can all be important.

It is critical to understand that each additional piece of data collected degrades the accuracy of the whole dataset and prolongs and complicates the survey. Thus, any additional information to be collected should be clearly stated and justified in the objectives and have a realistic prospect of leading to a meaningful intervention.

Consideration has to be given to whether the information could be collected more efficiently in other ways (for example from health clinics, sentinel sites or a surveillance system), or whether it would be better to conduct a separate survey to collect the supplementary information. If additional information is to be included in the survey it must be quickly and reliably obtainable during a short visit to the household.

Define geographic area and population group to be surveyed

Geographic area

In designing the survey, the area and population to be surveyed should be carefully defined. The report should contain a map of the area. Many agencies do a survey confined to the area in which they intend to implement a programme. They have normally chosen this area because it is thought to be most needy. This decision is usually made after a rapid assessment, interviews with key informants, migrants and refugees, by determining the origin and history of severely malnourished patients attending clinics or hospitals, and by looking for indications of increased mortality in the population. These data are used to justify the survey. The survey is often the last step before implementation and is used to persuade donors of the severity and urgency of the situation in a particular area. Data from such a survey **cannot** be extrapolated to indicate the severity of problems in other areas because the area has been chosen on the basis of an expectation that it is particularly affected.

In many cases, the area chosen will correspond to one or more administrative areas (for example, a district). A survey should be conducted in an area where the population is expected to have a similar nutritional and mortality situation. If an area is assessed which has two or more very different agro-ecological zones, the results will be an average of the two zones and will not give an appropriate perspective of either zone. Such heterogeneity can be resolved by doing separate assessments, although this usually increases the cost². In general, urban and rural areas, and refugee/IDP and resident populations should be assessed separately.

Frequently, there are areas which cannot be accessed because of insecurity; perhaps part of a district or other administrative zone. These areas need to be defined before the survey, clearly marked upon the map and reported as having been excluded from the survey. Populations living in highly insecure areas normally have a worse nutritional status and higher mortality than those living in more secure areas; nevertheless, it is unlikely that a programme can be implemented successfully in areas that cannot be surveyed.

Measurements can be made on new arrivals from insecure areas. Although such data give a valuable indication of the situation in the insecure area, they do not constitute a survey and should not be reported as such. Arrivals are often better off than those who have not been able to migrate from an area of insecurity. However, this should not be assumed as they may have left the area because, unlike those remaining, they are unable to sustain themselves, or have been rejected by the rest of the population. Often, relief programmes have to take account of such migration from insecure areas which have not been, and cannot be, surveyed.

² It may not increase the cost appreciably. This is because a much larger sample size is needed where there is heterogeneity. Sometimes two separate surveys can be conducted in two areas, each of which is homogeneous, with the same overall number of subjects from one large survey from a heterogeneous population. This is addressed in the section of the manual dealing with "design effects".

Population groups

Most commonly, anthropometric measurements and oedema assessments are made among children between 6 and 59 months of age, and a crude death rate (CDR) is assessed for the entire population (*all* deaths within a defined period of time). The 6-59 month-old child is considered to be the most sensitive to acute nutritional stress. This age group is chosen, therefore, to give an indication of the severity of the situation in the whole population. Furthermore, there are often baseline data for this age group, considerable experience in conducting surveys of their nutritional status and defined criteria for interpretation. However, in some situations it may be appropriate to include other age groups, such as less than six month-old infants, adolescents, adults or the elderly if it is suspected that their nutritional status differs significantly from that of the 6-59 month-old child. Although other age groups do not need to be surveyed, it is crucial to emphasize that limiting the survey to the 6-59 months age group cannot be used to justify confining interventions to this age group. If a survey has to be made for each age group before it receives help, the surveys themselves would become extremely cumbersome. Every malnourished individual should be eligible for relief.

Decide what additional information to collect

The data collected must correspond to the survey's objectives to inform subsequent actions and programming decision-making. When discussing the relevance of adding indicators to your survey, it is important to limit the number of additional information to not over-burden the survey and affect data quality. The two main public health indicators automatically analysed by ENA for SMART are the following:

Children's nutritional data

To estimate the prevalence of acute malnutrition in children less than 5 years of age (generally 6-59 months), the following data should always be collected:

1. Age, in months (from a known date of birth or based on an estimate derived from a calendar of local events);
2. Sex;
3. Weight, in kilograms (to the nearest 100 g);
4. Height, in centimetres (to the nearest millimetre);
5. Presence or absence of oedema.

Mortality data

To estimate the mortality rate, the following information needs to be collected:

1. The total number of people (including the age and sex of each individual) currently in the household;
2. The number of people who were in the household at the start of the recall period;
3. The number of deaths during the recall period, differentiating between traumatic and non-traumatic deaths;
4. The number of births during the recall period;
5. The number of people who have left the household during the recall period;
6. The number of people who joined the household during the recall period.

Other child data that are often collected (depending upon specific survey objectives) includes the Mid Upper Arm Circumference (MUAC), measles immunization (and possibly BCG) status, micronutrient supplementation status, particularly vitamin A, nutrition programme coverage, and morbidity information.

2. Survey Teams & Training

For the nutrition and mortality components, each survey team generally consists of a minimum of three people. Two make the anthropometric measurements and one records the data and serves as the team leader. The team leader is responsible for the quality and reliability of the data collected. The same team members can sometimes take both the anthropometry and conduct the mortality interview. However, it is usually better to have a fourth team member who interviews the head of household to collect the census and mortality data.

In order to make implementation of the survey more efficient and rapid, a respected member of the community should be asked to participate with each team. The member will know the village and be able to guide the team, locating households. The community member can more easily learn the whereabouts of absent households, and perform the often vital function of translator. If for some reason the community member does not speak the local dialect of the survey team, it may be necessary to include an additional translator on the team.

Team members do not have to be health professionals. In fact, anyone from the community can be selected and trained. They need to be fit, as there is usually a lot of walking. They should have a relatively high level of education, as they will need to read and write fluently, count accurately and ideally speak the local language. The gender composition of the team should also be considered when hiring enumerators. Depending on the cultural context, it may be necessary to have at least one woman or at least one man on each survey team. Most importantly, team members should be friendly, personable, eager to learn and hard-working.

Two to six teams may be needed depending upon the number of households to be visited, the time allocated to complete the survey and the size and the accessibility of the area covered. Although it is faster to have more teams, the quality of the data deteriorates. It is also much more difficult to train, supervise, provide transport and equipment, and organise a large number of teams. All team members and their equipment should be able to fit comfortably in the available transport.

Training of survey teams

Adequate training of the survey team members before the survey is critical. All scheduled training must be completed prior to data collection, and every team member should undergo exactly the same training, whatever their former experience, to ensure the standardisation of methods. During the survey, the survey manager must continually reinforce good practice, identify and correct errors, and prevent declining measurement standards. Before undertaking a standardisation test, the survey manager needs to make sure that the enumerators (both new and experienced) are ready and well-trained and understand their roles and responsibilities.

Training the teams is one of the most important aspects of doing a survey. It should not be rushed or assumed, as inaccurate measurements can have a very large effect upon the prevalence of malnutrition reported by a survey. The enumerators' training must be theoretical and practical, even if they are experienced. The training usually takes a minimum of five days and should include the following:

Theoretical training

Team members have to understand the principles of doing a survey, the reasons why the survey is being done, and the likely interventions that could result depending upon the findings. They have to be comfortable with this knowledge and not confused. They have to be able to explain and answer questions from community leaders, fathers, and mothers. They have to be sufficiently well-trained to be confident when they go into the community. They must also learn to select the houses, talk to the mothers, make the specific measurements, and record the results. This portion of the training should focus on:

- The objectives of the survey;
- A clear explanation of roles and responsibilities of each team member, the team leader, and the survey supervisor (they should also be given a written “job description”);
- An explanation of the sampling method that stresses the reasoning behind and importance of each child and household member having an equal chance of being selected (including households without children for the mortality and demography survey). The idea of random selection is sometimes difficult to grasp. Games, such as selecting numbers from a bag, should be used to illustrate those principles;
- An explanation of, and training on all questions included in the questionnaire to be used during field work;
- Instruction on administering the mortality questionnaire followed by practical exercises administering the questionnaire. In practicing, have teams interview several male- and female-headed households. As there is likely to be a “learning” effect by the respondent, each team should be the first to visit two of the practice households. Compare results to ensure teams obtain identical results when visiting the same household.

Practical anthropometry training

This consists of demonstrations of all anthropometric measurements teams will be expected to perform during that are planned for the survey. It is recommended to use a doll, or ideally a well-behaved child between 3-5 years old, to show how to measure weight, height and MUAC (if included) properly (Figure 1). From the outset, it is important that the measurers are not rounding the values inappropriately, and that there are not large numbers of unlikely values. The data they collect will be analysed in this way, and that any shortcuts taken are likely to become apparent. It is recommended to spend at least half a day practicing anthropometric measurements so that:

- Measurers should become comfortable with preparing a space for taking measures, turning on and off scales, assembling and disassembling the height boards and proper methods for reading and reporting measures;
- Measurers should practice before the standardisation test;
- Some corrections related to the standard protocols of measurements can be provided by the trainers and the other enumerators.

Figure 1: Practicing Weight, Height and MUAC Measurements.



Teams can also visit a therapeutic feeding center, hospital, or clinic to see cases with oedema and practice checking for oedema with actual cases. However, it should be noted that most cases encountered during a survey will have mild oedema, and the teams should not be trained exclusively on cases with gross oedema. It is very common for teams to make errors in assessing oedema if they have only practiced on normal children, none of whom have oedema.

Standardisation Test

A Standardisation Test is a practical assessment of enumerators' measurement skills. Conducting a Standardisation Test for anthropometric measurements is a fundamental step in the training of enumerators for an anthropometric survey, and is mandatory before each survey. It allows trainers to objectively evaluate the quality (precision and accuracy) of the measurements taken by each enumerator. Performing a Standardisation Test is therefore essential to assess, measure, and address every enumerator's strengths and weaknesses.

The Standardisation Test consists of each enumerator measuring at least ten different children twice. An interval of time is taken between the first and second round of measurements. The children must be healthy and of the targeted age group for the anthropometric survey (generally 6-59 months), with at least two children under two years of age.

The Standardisation Test evaluates:

- The variance between an enumerator's two measurements of a single child in order to measure the enumerators' precision;
- The mean of both measurements compared to the reference (supervisor's measurements) in order to assess the accuracy (or Technical Error of Measurement) of the enumerators' measurements.

The ENA for SMART software automatically analyses each enumerator's scores for each measurement undertaken in the Standardisation Test (weight, height, and MUAC, if included in the survey). Each enumerator will be assigned a score (Good, Acceptable, Poor or Reject) on both precision and accuracy for each type of measurement (e.g., weight, height, and MUAC). This score will be used to evaluate the enumerator's performance and to correct any mistakes at the training stage.

Location

The location identified for the test must be sufficiently spacious and shaded, if outdoors. The distance between each child must be large enough so that each enumerator cannot hear the results of the others, with ideally at least three meters between each station. It is possible to conduct the Standardisation Test in a village, an orphanage, or the courtyard of a school. It is best to have an open area where one supervisor can see the activities of all trainees at once to reduce the risk of cheating. Discuss with trainees before the exercise the implications of cheating during the standardisation exercise.

If working in a school, nursery or orphanage, be careful not to disrupt normal activities. It is important to obtain authorisation from those responsible for the selected location before undertaking the Standardisation Test.

Recruitment of children

When recruiting children, it is important to inform the mothers or caretakers that the Standardisation Test is long and tedious, and that children will get very tired. It is recommended to explain through visuals the procedure of the Standardisation Test, what is expected of them and the children, and that everyone is clear about the incentive(s) offered.

Equipment

All measurement equipment must be calibrated and in good working order the day before the Standardisation Test (Table 1). It is important to use the same equipment as the one planned for the survey.

At the beginning of the Standardisation Test, each team of two persons is given a clipboard, a pen, and one standardisation test form. The trainee who is taking the measures (measurer) must identify him or herself on the form by writing his or her name and ID. The name and ID of the assistant can also be recorded. For example, if there are 15 trainees, they should be numbered one through 15. Enumerators should never have access to more than one form at a time. Towels, water, and soap must be available to clean the equipment in the case children urinate or vomit.

Table 1: Standardisation Test Equipment and Materials for 10 children.

Item	Quantity	Specifications
Height Boards	12	✓ Wood Height boards should be minimum 130 cm. ✓ Measuring tape on only one side of height board. ✓ 2 extra in case
Scales	12	✓ Same type of scale as will be used in the survey (SECA/electronic scales are recommended over Salter scales). ✓ At least two extra sets of batteries if using SECA/ electronic scales (in addition to working batteries in scales). ✓ 2 extra ones
MUAC Tapes	20	✓ New MUAC tapes (not bent, torn or worn). ✓ 10 extra MUAC tapes
Clipboards	1/team	
Standardisation Test Forms	2/team	✓ One form per round of measurements (two rounds total). Each team should only have access to one form at a time.
Pens	2/team	

Roles

The survey manager assigns roles: (1) measurer or (2) assistant, then pair up measurers and assistants into teams. Enumerators are grouped in pairs because it is necessary to have an assistant when conducting anthropometric measurements. It is possible to assign somebody whose role is to **take care of the mothers or caretakers**. This person can explain the procedure to the mother or caretaker, to reassure them and their child.

A pair should also be chosen for the reference measurements (measurer and assistant). This pairing is often made up of the supervisors. The **choice of the reference measurer is crucial**. If the reference team is not experienced in anthropometric measurements, the test results will not be reliable.

Set up

The ten stations (one for every child measured) should be at least three meters apart. The distance between each child should be enough to prevent the trainees from seeing or hearing each other's results. Each station should include:

- The station number taped to the wall;

- A child with their mother or caretaker;
- A chair;
- Anthropometric equipment (scale, height board, and MUAC tape if included in the survey).

Once mothers or caretakers and their children enter the room, assign each pair to a station. With permission, write the number of the station on the child's hand (this helps with identification in case the child moves around). The mothers or caretakers and children remain at their assigned station for the duration of the test.

Each team (measurer and assistant) begins at the station that corresponds to their team ID. Team members will rotate, but the equipment will remain with the child so that each child is always measured with the same equipment; it is the team that is being tested, not the equipment. The supervisor must verify the positioning of the equipment and the adjustment to zero before the Standardisation Test.

Children

A **minimum of ten children** must be measured. Children identified for testing should be between six and 59 months (or of the age group targeted for the anthropometric survey). It is important that the enumerators are tested with two or three children less than 24 months old (or measuring less than 87 cm, if height is used as inclusion criteria for the survey), and the remaining children more than 24 months old (or more than 87 cm) to practice the height measurement in both recumbent and standing position.

Children identified for testing must be in good health. Be careful not to recruit sick or malnourished children. It is good practice to recruit 12 children in case a child falls ill or cannot be measured for any other reason. Trainers must assess the health status of children before the beginning of the test, and replace those who are not in good health.

This test is often long and children get tired quickly. Advise the youngest children's mothers to breastfeed their child, if possible, to help calm them.

Timeframe

Regardless of the scale of the survey, it is important to dedicate several sessions or days to complete the test for all trainees. Additional days may be required if the results of the Standardisation Test are unacceptable.

Standardisation Test Procedure

First Round of Measurements

Each child to be measured will be given an ID number starting with one (1). If there are no more than ten measurers in the Standardisation Test, each team begins the test at the station (child) with the ID number corresponding to their team number. Measurements are recorded on the Standardisation Test form in the row corresponding to the child's ID.

When a team arrives at a station, the measurer and assistant introduces themselves to both mother and caretaker and to the child, and then determine the child's age. Children under two years old (or less than 87 cm) are measured in a recumbent position (length) and children two years old and older (or 87 cm and over) are measured standing up (height). It is important to clarify to enumerators if age or height is being used as inclusion criteria for the survey (and also for the Standardisation Test). The most common is age inclusion.

The mothers or caretakers and their children stay at the same station during the whole test. Every child participating in the test is given an ID number that facilitates their identification on the Standardisation Test forms. Make sure that the mother keeps the paper on which her child's number is written. Often, identification can be made easier if the child's number is written on their hand (always ask permission first).

The measurer then proceeds to measure the child (weight, height/length, and MUAC, if included in the survey) with the assistant's help, and then moves on to the next child. The measurer takes their first round of measurements for all ten children, with the help of the assistant. **Only the pairs of measurer and assistant rotate from one child to the next;** the children and their mothers or caregivers stay put at their designated station.

Each team hands in their Standardisation Test form to the supervisor at the end of the first round, which is when the team has finished measuring all ten children once.

Break

Once all the teams have measured all ten children once, everyone takes a short break. Provide water and light snacks for enumerators, mothers or caretakers, and children. Small toys such as balloons may help calm children in the second round. Depending on the context, a numbers game may sometimes be necessary to prevent enumerators from remembering their first measurements. It is forbidden for enumerators to write or copy measurements onto any other paper other than the Standardisation Test form.

Second Round of Measurements

The supervisor hands out a second Standardisation Test form to each team (teams do not have access to their form from the first round), and the whole procedure is repeated for a second time. Example: In round one, Enumerator A takes the measurements of Child 1 through Child 10, assisted by Enumerator B. There is a break, then Enumerator A again takes measurements of Child 1 through Child 10, for a second time, still with Enumerator B as their assistant (round two).

In the above scenario, only the measurer (not the assistant) will be analysed by the ENA software. However, if there are no predefined roles for the survey prior to the Standardisation Test, then each enumerator should be evaluated. In this case, there are two options:

- Option 1: Each member of the pair must take the measurements in turn and complete their own Standardisation Test form. Once each member has taken the measurements of one child, the pair moves on to the next child.
- Option 2: A second Standardisation Test using different children should be performed, with the assistants now as measurers (see *Timeframe* section above). Roles in the survey should then be assigned based on objective results.

Important: No communication is allowed between measuring teams (participants) during the Standardisation Test, and it is forbidden for participants to copy measurements onto any other paper other than the Standardisation Test form. The supervisor should correct any major errors during the Standardisation Test, and make notes on any other errors to be discussed later on with the team members. Corrections and advice will be given after the test analysis. Observations generally contribute to the interpretation of the results and allow the measurers to better focus their practice after the test.

The Standardisation Test forms should look like those presented in Figure 2.

Figure 2: Standardisation Test Forms Templates.

Measurer ID #.....		Assistant ID #.....		Date:/...../.....	
1st measurement :					
# Child	Date of Birth	Sex	Weight (kg)	Height (cm)	MUAC (mm)
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Measurer ID #.....		Assistant ID #.....		Date:/...../.....	
2nd measurement :					
# Child	Date of Birth	Sex	Weight (kg)	Height (cm)	MUAC (mm)
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

End of the Standardisation Test

At the end of the Standardisation Test, it is very important to thank everyone and offer the promised incentives or gifts to the mothers or caretakers (e.g. soap) and the children (e.g. fruit, clothes, toys). Choice of appropriate incentive(s) should be discussed with the local partner.

Data Entry

Data from the Standardisation Test is entered into the ENA for SMART software to generate the Standardisation Test report (see next section). Only the results of the **same** Standardisation Test carried out on the **same** children can be directly compared. In other words, each Standardisation Test data is analysed on its own, and enumerators are compared on their respective scores (“Good”, “Acceptable”, “Poor”, or “Reject”) for each type of measurement (weight, height, and MUAC measurements, if included in the survey).

Summary of the Standardization Test

Set up

1. Re-explain the Standardisation Test procedure to the enumerators (explanations should first have been given before the day of the test, during the theoretical part of the training).
2. When the mothers or caretakers and their children arrive, remind them of the Standardisation Test procedure, and emphasise that it might be tiring for the children (explanations should also have been provided beforehand).
3. Assign an ID number to each child, from one (1) to (10) ten: the number will also correspond to the number of their assigned station.
4. Ensure that all equipment is ready and that stations are far enough from each other (3 meters) for work to be done comfortably, and for communication between teams of enumerators to be avoided. At each clearly numbered station there should be: a child with their mother or caretaker, a chair, a height board, a scale, and a MUAC tape (if MUAC measurements are included in the survey).
5. Assign an ID number to each survey team. The supervisor's ID is zero (0).

First round of measurements

6. Hand out the Standardisation Test forms to the teams for the 1st round of measurements.
7. Assign each team a different child to start with.
8. Give the signal for all teams to begin the 1st round of measurements.
9. All teams (including the reference team of supervisors) carefully measure all ten children once: when measurements on the first child are finished, teams move on to the next child, and so on.
10. At the end of the first series of measurements, collect the forms from the teams: make sure that enumerators do not copy the first round measurements. Any use of notebooks during the test is prohibited.
11. Allow for a break for everyone, including mothers and caretakers and their children.

Second round of measurements

12. Hand out new forms to the teams for the 2nd round of measurements.
13. Give the signal for all teams to begin the 2nd round of measurements.
14. All teams (including the reference team of supervisors) carefully measure all ten children a second time: enumerators repeat step 9.
15. At the end of the 2nd round, collect the forms from the teams and check that they are properly filled out.

Wrap-up

16. Thank all the mothers or caretakers and their children, and offer them the promised incentives or gifts.
17. Enter Standardisation Test data in the ENA for SMART software to generate the Standardisation Test report.

Interpretation of the Standardisation Test Report

Under the *Training* tab, ENA Software produces two different Standardisation Test reports: the 'Report from previous ENA versions' and the 'New Report'. The 'New Report' is more rigorous, and applies the standard method used by anthropometrists with results presented in recognizable units. Therefore, it is recommended to use the more rigorous analyses of the New Report when evaluating enumerators in order to ensure the collection of high quality survey data (measurements).

Based on the methods defined by Ulijaszek and Kerr (1999) and used in the training of the anthropometrists for the *WHO Multicentre Growth Reference Study* (MGRS) for Child Growth Standards, the calculations of the new report analyses the team as a whole and looks at the difference between each enumerators' own measurements, with the results presented in the original unit of measurement (Table 2).

Table 2: Units of analysis for anthropometric measurement, new Standardisation Test report.

Anthropometric Measurement	Unit of Analysis
Weight	kg
Height	cm
MUAC	mm

First Columns of the New Report

The analysis is divided by measurement: weight, height and MUAC (if included in the survey).

The columns *mean*, *standard deviation (SD)* and *maximum (max)* can be useful to highlight major errors, generally made in data entry (e.g. MUAC measurements entered in mm for some enumerators and in cm for others).

Subjects

The number of children (subjects) that the supervisor and each enumerator measured. **The number of subjects should always be ten, measured twice (shown as "10"): this is a minimum.** Looking at team evaluations, "5x10" would mean five enumerators measuring ten children.

Mean

This column shows the mean value of all 20 measurements taken (ten children measured twice). The information is presented in the same unit as the measurement. The actual mean value is not as important as the variance (difference) of this value between enumerators.

Standard Deviation (SD)

The associated standard deviation for the supervisor and each enumerator. This information is presented in the same unit as the measurement.

Maximum (max)

The largest difference between any first and second measurements taken from one child. This information is presented in the same unit as the measurement. Use the paper version of the standardisation test results to identify the child, as this information is not presented in the report.

Precision

Enumerators are evaluated based on how consistent their first measurement is with their second measurement. The New Report has three measures of precision: Technical Error of Measurement (TEM), % TEM, and R (Coefficient of Reliability).

Technical Error of Measurement (TEM) – Individual

TEM is most important measure of Precision in the New Standardisation Test Report.

TEM stands for the Technical Error of Measurement which relates to the difference between the first and second measurements for each child. It is analogous to the standard deviation (comparing SD and TEM column gives the proportion of variation due to measurement error).

TEM is based on the following formula:

$$\text{TEM} = \sqrt{(\Sigma D^2)/2N}$$

N = number of children measured (e.g. 2N would be equal to 20 if 10 children were measure twice), and D = difference between the first and second measurements.

TEM is calculated for each variable (height, weight, MUAC) for all ten children measured. Analysis is done on an individual basis (enumerator) and at team level. There are different cut-offs for each type of measurement (height, weight and MUAC). For example, in Figure 2 below, the TEM of Enumerator 8 is very high for weight. Before drawing conclusions on an individual enumerator's TEM, it is important to first double check for recording or data entry errors.

Figure 3: Example of TEM output from ENA in the new Standardisation Test report.

Standardisation test results: Weight (kg)	Precision
	Technical error of measurement TEM (kg)
Supervisor	0.1
Enumerator 1	0.1
Enumerator 2	0.2
Enumerator 3	0.1
Enumerator 4	0.1
Enumerator 5	0.1
Enumerator 6	0.1
Enumerator 7	0.1
Enumerator 8	1.2
Enumerator 9	0.1
Enumerator 10	0.1

Technical Error of Measurement (TEM) – Team Analyses

The main criteria for evaluating precision in enumerators is their individual TEM. However, the team TEM analyses give an indication of consistency of measurements within and between all the enumerators.

The team TEM analyses are:

- **Enum inter 1st:** Comparing between enumerators. Difference (variance) between all ten first round measurements for all enumerators.
- **Enum inter 2nd:** Comparing between enumerators. Difference (variance) between all ten second round measurements between the enumerators.
- **Inter enum+sup:** Comparing between everyone who took measurements (enumerators and supervisor). Difference (variance) between all first and second measurements for each enumerator PLUS difference (variance) between all first and second measurements of the supervisor (10+1 measurements).

- **Total intra+inter:** Comparing within each enumerator (intra) and comparing between all enumerators (inter). Difference (variance) between each enumerator to themselves (difference between their first and second measurements on each child), plus difference (variance) between all first measurements and difference (variance) between all second measurements .
- **Total+sup:** Same as above (TOTAL intra+inter), but also includes the supervisor's measurements.

The main team TEM analysis to look at (and which has suggested cut-offs) is *Total intra+inter*.

Percentage (%) Technical Error of Measurement (TEM)

The parameter of %TEM is a unit-less measurement that allows the precision of the different measures taken (weight, height, MUAC) to be compared to each other.

The value will vary with the size of the actual measurement. Percentage TEM is calculated as a fraction of the TEM over the value of the measurement taken. For example, if the weight of one child is 8 kg and that of another child is 20 kg, and both were measured with a TEM of 0.5 kg, the %TEM will be 6.25% and 2.5%, respectively.

R (Coefficient of Reliability)

The R-value gives the proportion of between subject variance in the measurement which is free from measurement error. The Coefficient of Reliability is the most widely-used measure of precision in population studies. It is based on a scale of 0 to 1, where 1 is perfect. For example, if R=0.70, this means that 70% of the differences in measurement are due to actual differences in the size of children and 30% of the differences are due to the enumerator measurement error.

In general, R-value should be close to 95%. With good instruments, especially with digital scales for weight (e.g., SECA scales), R-value will be close to 99.9%. Lower R-values are usually seen for height and MUAC.

Accuracy

Accuracy refers to how close the enumerators' measurements are to those of the reference value.

Two outputs are provided for accuracy in the New Report:

1. Bias from super (supervisor).
2. Bias from median.

The first measure (Bias from supervisor) is the most important.

If the supervisor TEM value is "Good" or "Acceptable" for a particular measurement, individual results are compared to **Bias from the supervisor** (the mean of the twenty measurements taken by the enumerator is compared to the mean of the twenty measurements taken by the supervisor).

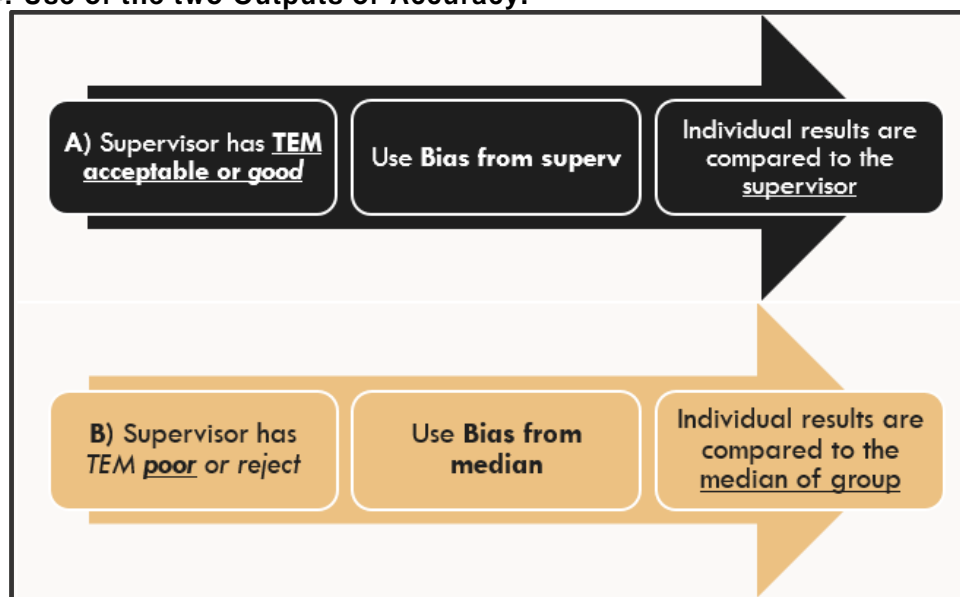
However, the New Report provides an alternative to interpreting the accuracy of enumerator measurements: Bias from median.

If the supervisor TEM value is "Poor" or "Reject" for one of the anthropometric indicators, individual results are compared to the group median – **Bias from median** (the mean of the twenty measurements taken by the enumerator is compared to the median of all measurements taken by all enumerators and the supervisor) for that indicator.

The median is used instead of the mean because it is more stable. If most measures are grouped but have one outlier, that outlier/erroneous measures affects the mean.

Figure 4 summarises when to use which output of accuracy.

Figure 4: Use of the two Outputs of Accuracy.



Bias: Additional Analyses

The additional analyses for bias give an indication of the accuracy of the measurements within and between all the enumerators:

- **Inter-enumerator** compares the bias from the mean of enumerators, without including the mean of the supervisor.
- **Total error** compares both the inter-enumerator (team) and intra-enumerator (within each enumerator) variation for all of the enumerators.
- **Inter enum+sup** compares the mean for all enumerators for all rounds PLUS mean values of supervisor.

Acceptable Limits

When assessing the quality of each enumerator’s measurements in terms of precision and bias, the table of acceptable limits is used (Figure 5). Determined on the basis of the Review Article (Ulijaszek and Kerr, 1999) and the feasibility for field teams from the SMART Expert team, these acceptable limits are used to interpret TEM for each enumerator and the team, as well as Bias (from supervisor or median, depending on TEM of the supervisor).

Figure 5: Acceptable Limits for Standardisation Test Measurements.

Suggested cut-off points for acceptability of measurements					
Parameter		MUAC mm	Weight Kg	Height cm	
Individual TEM (intra)	good	<1.0	<0.04	<0.4	
	acceptable	<1.3	<0.10	<0.6	
	poor	<2.1	<0.21	<1.2	
	reject	≥2.1	≥0.21	≥1.2	
Team TEM (intra+inter) and Total	good	<1.3	<0.10	<0.5	
	acceptable	<2.1	<0.21	<1.0	
	poor	<3.0	<0.24	<1.5	
	reject	≥3.0	≥0.24	≥1.5	
R value	good	>99	>99	>99	
	acceptable	>95	>95	>95	
	poor	>90	>90	>90	
	reject	≤90	≤90	≤90	
Bias	good	<1	<0.04	<0.4	
	From sup if good outcome, otherwise	acceptable	<2	<0.10	<0.6
	poor	<3	<0.21	<1.4	
	from median	reject	≥3	≥0.21	≥1.4

What to do if some Enumerators fail the Standardisation Test?

The table below shows the different recommended options to improve the anthropometric measurements when some measurers do not pass the Standardisation Test (Table 3). Each option offers suggestions of activities to undertake.

Among the options suggested below, the best one to consider is a combination of the first and second. However, the third option remains acceptable and is often selected during rapid nutrition surveys, which are generally constrained by time and budget.

Table 3: Recommended Options if Enumerators fail the Standardisation Test.

Options	Rationale
1. Do more practice in the field	If time allows it, re-organise a practical training on taking anthropometric measurements (as described above), only for measurers who had a low score on precision or accuracy for weight, height, or MUAC.
2. Re-do the Standardisation Test	Once the practical training is done, a second Standardisation Test is undertaken only for the measurers who failed the first test. The measurers who passed the first test can then be involved in the organisation and supervision.
3. Create pairs of measurers	When building the teams, the survey supervisors is be careful to pair measurers whose strengths are complementary: each team should be composed of a “strong” measurer paired with a “weaker” one during the pilot test and the first days of the survey.
4. Conduct the field test	The pilot test may be considered as a field practice, and used to reinforce and improve the quality of the weakest measurers’ anthropometric techniques. The survey supervisors must focus their attention on the measurers who did not pass the test during the pilot test and during the first days of the survey.
5. Reassign roles	If a member of the team cannot take the measurements appropriately after re-training, they should be replaced or given another task in the survey which does not require taking measurements.

Pilot Testing

The pilot test is a practical exercise and not confined to the classroom. This step is essential for the teams to feel confident when they begin conducting the actual survey. It takes place after the teams are able to take accurate and precise measurements, have “passed” the Standardisation Test, and have formed teams that have practiced working together. For pilot-testing, teams go to a convenient local village or community that has *not* been chosen to be sampled and to contain a cluster. Community members need to be sensitised so that they can inform household members that the survey teams will be visiting the village for this exercise. All teams work within the same village, but each team should have different randomly selected households from the village to visit (otherwise it is likely that the selection was not random).

During the pilot test, the teams will demonstrate the following:

1. Sampling procedures: the team will practice selecting the houses that will form the cluster, selecting the first house, the respondent(s), and children of the right age.
2. Data collection: within a household, teams will practice approaching respondents and explaining the purpose of the survey, taking and recording the measurements correctly, and administering the mortality questionnaire, if included under field

conditions. This means that the team members will practice working as a team, each with his or her assigned tasks, and interacting amicably and effectively with the respondents.

3. Their ability to organize, transport, and care for the equipment.

The practice data from the pilot test of each of the teams is entered into the ENA for SMART software and analysed. Each team's results will be slightly different; this is used as a practical demonstration of the effect of sampling error and the importance of taking a random sample.

At the end of the pilot test, the team members, team leader(s), and survey supervisor should be confident that each team can undertake the survey accurately, and should know how long it takes to complete a survey of each household³. This information allows the supervisor to calculate how many households and children can be expected to be completed each day during the real survey. If the time required to complete one household is excessive, any additional information being collected will need to be reconsidered.

Survey Management

The supervisor or team leader⁴ has the overall responsibility for training team members, visiting teams in the field, ensuring that households are selected properly, that the equipment is checked and calibrated each morning during the survey, and that measurements are taken and recorded accurately. Unexpected problems nearly always arise during a survey, and the supervisor is responsible for deciding how to overcome them. Each problem encountered and decision made must be promptly recorded and included in the final report. The survey supervisor is also responsible for overseeing data entry into the ENA for SMART software, and for the analysis and report writing.

It is also the duty of the survey supervisor to regularly supervise teams in the field. It is particularly important to check cases of oedema, as there are often no cases of oedema seen during the training and some team members may therefore be prone to mistaking a fat child for one with oedema (particularly with younger children). The supervisor should note teams that report a lot of oedema and visit some of these children to verify their status. The supervisor should also confirm measles and death cases.

Where possible, the survey supervisor should organize an evening "wrap-up" session with all the teams together to discuss any problems that have arisen during the day⁵.

Before leaving the field (or at the end of the day, if the team remains in the field), each team leader should review and sign all forms to ensure that no pieces of data have been left out. If there were people absent from the house during the day, the team should return to the household at least once before leaving the area.

It is very important not to overwork survey teams. There is a lot of walking involved in carrying out a survey, and when people are tired, they may make mistakes or fail to include more distant houses selected for the survey. The supervisor must make sure the team has enough time to take appropriate rest periods and has refreshments with them.

Each evening, or during the next day while the teams are in the field, the supervisor should arrange for data to be entered into the ENA for SMART software. Recording errors, unlikely results and other problems with the data may become clear at this stage. Depending on the settings in the "Variable View" at the bottom of the *Data Entry Anthropometry* Tab, the ENA software can automatically flag any values beyond the acceptable range set as the data are

³ The teams normally take longer during the pilot test than after they become used to working as a team.

⁴ Communication with teams in the field is often very difficult, and in these circumstances each team leader must be sufficiently trained to be able to take decisions independently.

⁵ This may not be possible if the survey area is large, so that teams are widely separated and remain in the field for several days.

entered. In the morning before the teams set out for the day, there should be a short feedback session. If any team is getting a large number of “flagged” results, the supervisor should accompany that team the next day. If the results are very different from those obtained by the other teams, it may be necessary to repeat the cluster from the day before.

Team leaders and the survey supervisor should record all important points in a notebook as soon as possible (e.g. during breaks, or at base in the evening), including observations, ideas, problems, actions taken to address these problems and the reasoning behind any decisions taken. Each note should be labelled with the date, location, and names of relevant people.

Apart from the evening and morning meetings, survey team members should be encouraged to regularly discuss their experiences and findings together. This often brings out important points and sometimes shows where survey method need to be modified.

If possible, at each household the team leader should calculate or look up in a table the percentage weight-for-height z-score for each child and classify the child’s nutrition status. When a malnourished child is identified he or she should be referred, on the spot, to the nearest health or nutrition facility. Ideally, this will be a therapeutic or supplementary feeding programme. If these are not available, the supervisor should urge the parents to take the child to the nearest health facility, providing a referral slip upon which the name, height, weight, weight-for-height z-score, (MUAC if collected), and diagnosis is written.

3. Sampling Methods and Sample Size Calculation

What is Sampling?

When we want to collect data on a population, there are different ways to proceed. In some cases, it is possible to measure the *entire* population to get an accurate picture of the nutritional status or mortality rates for the population at hand. This is called an exhaustive survey. These cases are however *very rare* as they are feasible only for target populations that are geographically concentrated and not exceeding 1,000 households, as seen in camps or very small areas (Figure 6). For larger populations, an exhaustive survey will be long, expensive, difficult, and unnecessary. In addition, the results of such a survey cannot be extrapolated to other regions.

Figure 6: Camp Stankovic II, Skopje, Macedonia, 1999.



Example: An organization would like to know the malnutrition rate in a small camp. There are between 2,000 and 3,000 people in this camp. The camp is composed of roughly 400 to 600 children less than five years old. In this case, *all* children can be measured. The results will therefore represent the exact nutritional situation of the children in the camp (True Population Value).

When it is impossible to collect data on the whole population because it is large and/or geographically dispersed, we would choose a sub-group of this population, which is called a **sample**. This sub-group, or sample, can be chosen following two broad groups of methods: probability and non-probability⁶ sampling methods.

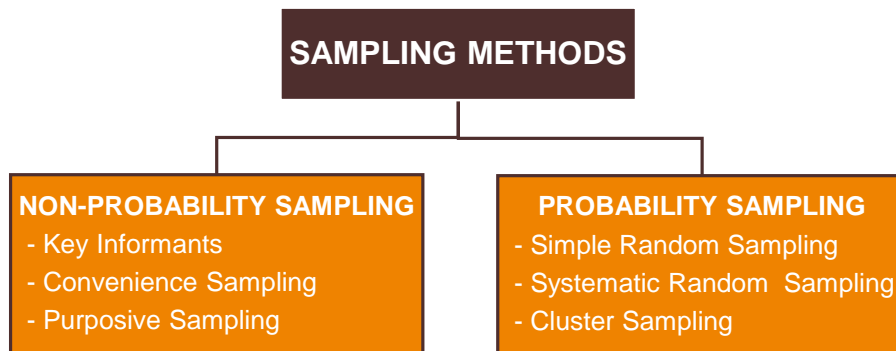
Screenings of children measured in health centers, markets, or other places where people gather are not representative. Children that visit health centers are not the same as those that do not. Well children can go to the market, run around the village, and follow the assessment team, while sick children are more likely to be in the house. The results of measuring such groups should never be called a survey (it can be called a series), and no decisions should ever be based upon such a “grab” sample. Similarly, clusters can never be selected on the basis of convenience. If we were only to select households in villages that are near to the road, the survey would be biased. Villages near roads have better access to transport and may be wealthier than more remote villages. In each of these examples, all the households and children in the community do not have an equal chance of being selected. Merely going to a clinic, running around the village, or living near a road or in the center of a village should not make anyone more likely to be selected.

Figure 7 shows the sampling methods that belong to each of those two groups. An example of the non-probability sampling would be for surveyors to go to a village and call all mothers to bring their 6-59 month-old children to the health center and collect data only on those children. We can easily assume that some mothers might only bring sick children, thinking that they will receive free treatment in the health center. On the other hand, we can also assume that mothers might bring only healthy children to avoid being judged by others as being “bad mothers”, and we can continue with other assumptions regarding the “type” of children that will be taken to that health center for data collection on that day (sex/age of the children, access to the health center, etc.) Therefore, to avoid these problems, we generally choose a **representative sample** using a probability sampling method⁷, which will be covered in this SMART manual.

⁶ In general, non-probability sampling is often used for collecting qualitative information.

⁷ Different probability sampling methods will be discussed into more details in the next section.

Figure 7: General Types of Sampling.



Principles of Sampling

A representative sample needs to have the **same characteristics** as the target population. For example, if our target population are children 6-59 months living in a pastoralist population, our sample needs to contain children 6-59 months who are living in this same population. The distribution of age and sex and other characteristics in our sample should be very similar to the distribution seen in the target population. Having a representative sample also means that:

1. Each individual or sampling unit in the population has a **known, non-zero chance or probability of being selected**.
2. The selection of one individual is **independent** from the selection of another.

Since we are only collecting data on a sub-group of the population when sampling, it is important to remember that the result obtained will only be an **estimate** of the indicator that needed to be measured. As discussed above, to get the True Population Value, it would be necessary to conduct an exhaustive survey.

Definition of Concepts

It is important to understand the terminology used when talking about sampling. Below are some of those concepts with their definition:

Sampling universe: Population from which we are sampling.

Sampling frame: Description of the sampling universe, usually in the form of the *list of sampling units* (for example, villages, households or individuals). Sometimes, it may be outdated or otherwise not accurate, and thus would not provide an accurate description of the sampling universe (census data not recent, recent population movements, etc.).

Sampling unit: Unit selected during the process of sampling. If you are selecting districts during the first stage of cluster sampling, the sampling unit (also called *primary sampling unit*) at the first sampling stage is therefore the district. If you select households from a list of all households in the population, the sampling unit is in this case the household.

Basic sampling unit or elementary unit: Sampling unit selected at the last stage of sampling. In a multi-stage cluster survey (see page 32), if you first select villages and then select households within those selected villages, the basic sampling unit would be the household.

Respondent: Person who's responding to your questionnaires in the field.

Survey subject: Entity or person from whom you're collecting data.

Sampling Methods

Three methods of sampling will be covered in this manual:

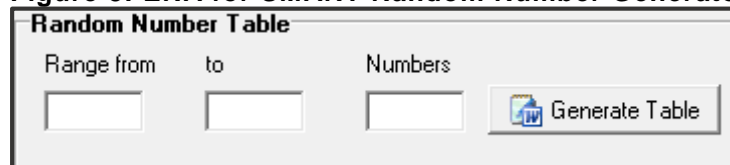
1. Simple random sampling;
2. Systematic random sampling;
3. Cluster sampling.

The sampling method is selected based on the spatial distribution of households and **population size**. The basic principle for selecting households to visit is that each individual in the target population must have a known (and preferably equal) chance of being selected for the survey.

Simple Random Sampling

Simple random sampling can be used for small populations that contain more than 1,000 sampling units (or households). This method requires a complete and updated list of sampling units, which may be difficult to obtain in emergency situations. Since in most contexts the number of basic sampling units (BSU) is large, simple random sampling procedure can be conducted by numbering each basic sampling unit and then choosing the desired number of units randomly from a random number table (Figure 8). Measures will then be taken based on these units only.

Figure 8: ENA for SMART Random Number Generator.



The image shows a software interface for generating a random number table. It consists of a window titled "Random Number Table". Inside the window, there are three input fields: "Range from", "to", and "Numbers". To the right of the "Numbers" field is a button labeled "Generate Table" with a small icon of a document and a plus sign.

Two scenarios are possible for simple random sampling depending on the demographic data available:

- **List of target population is available:** If there is an updated list of 6-59 month-old children, for example, the number of children needed for the sample can be drawn randomly from this list (as in scenario 1 below).
- OR
- **List of households is available:** If there is only a list of all households within the area to survey, the desired number of households (calculated by ENA) can be drawn randomly from this list. Once those households are selected, all children present in those households will be measured. This same procedure as described in scenario 1 can be applied if the basic sampling units are households instead of children.

Scenario 1: Simple random sampling using a list of children.

You are requested to conduct a nutrition survey in a refugee camp that contains about 10,000 people. All births are registered, enabling you to have the list of all children under five years old, with their birth dates. This list contains 1,700 children and after completing your calculations, you find that your sample needs to include 450 children.

Here are the steps to follow, using the *Random Number Table* in ENA for SMART to select your basic sampling units (which are children in this case):

1. Number each of the 1,700 children of the camp from 1 to 1,700.
2. Enter 1 in the box named *Range from* and 1,700 in the box named *To* (Figure 8).
3. Enter the number of children you need for your sample (450) in the box named *Numbers*.
4. Click the *Generate Table* button.
5. A word document will open displaying the 450 numbers selected randomly (Figure 9).

Figure 9: List of Random Numbers in a Word Document Generated by ENA.

Random Number table

Range: 1 to 1700, Number: 450

501	1601	793	458	327	962	1094	631	714	893	1070	1175	328
	614	370	655	1665	1068	1630	1452	1296	1468	767	635	195
	1165	432	1480	58	1057	832	149	1167	915	687	1038	844
	1305	785	856	422	518	1275	897	207	1143	1276	309	
	1419	542	1307	84	530	1124	980	316	951	212	96	899
	1110	1351	1523	963	49	917	1080	1283	821	1206	1033	686
	961	1321	968	54	467	889	129	707	417	1072	1511	
	1389	1406	1357	1641	221	1273	1231	1415	1568	463	982	652
	1045	36	1393	682	469	1519	799	588	1312	1269	329	375
	425	1285	1372	921	336	1085	118	1228	937	255	1464	715
	1022	911	201	1073	1136	1442	209	1002	1337	537	28	
	1508	292	1092	884	337	1056	122	284	1477	1483	741	972
	48	668	1542	854	1335	1404	1618	552	1505	382	353	
	1409	38	916	1016	385	50	883	909	1075	754	788	571
	281	332	364	227	568	1140	1358	1431	305	490	175	
	1166	1086	271	975	269	220	1501	1199	150	1156	324	925
	1220	1597	1054	298	746	75	1041	737	434	431	713	
	1346	459	1173	1646	1007	1586	335	1532	833	1350	683	
	1338	1553	148	76	791	981	869	513	1293	166	823	
	1084	541	1018	251	1226	811	1557	12	438	1593	1650	386
	1420	1079	850	1554	1582	875	855	922	738	749	630	
	1142	814	812	1446	1069	842	1592	236	989	894	1652	789
	1194	1121	532	1635	1394	1052	141	1238	1191	514	1189	
	1640	214	545	941	1541	393	176	165	794	1416	1602	934
	1334	990	273	1144	1685	1364	771	732	1455	1014	317	
	1599	1280	816	1513	611	152	1161	1255	245	843	1319	493
	1546	830	449	242	618	465	510	1438	632	716	1210	733
	1529	388	1376	43	958	585	295	834	1318	564	290	
	1521	681	1326	1010	1549	330	1234	523	1382	847	1405	
	1388	1527	246	1264	1183	147	1197	1090	579	312	366	
	1506	155	1526	413	878	873	171	1700	360	1628	831	283
	258	684	1663	638	1115	594	1088	359	1669	426	294	
	1632	678	404	1425	1678	11	1065	1380	1185	164	1179	
	1411	218	440	208	1373	1625	1550	137	1106	555	127	
	1062	252	1563	1695	549	906	931	15	498	135	759	
	1243	1251	593	1270	740	1135	1174	278	964	1006	502	
	1224	1414	352	1096	64	380	172	1130	1648	1699	1082	177
	696	1298	563	1507	1677	282	706	1324	650	1432	235	63
	750	1277	1049	428	971	1384	675	699	519	1498	703	
	1611											

Systematic Random Sampling

Systematic random sampling is based on selection of units situated at a certain predetermined interval called the **sampling interval**. It is applicable for small scale surveys (approximately 1,000 to 5,000 BSU) and one of its main advantages is that it can also be used without having a list of basic sampling units, as in situations where dwellings are well organized in rows, blocks, or along a river or main road, for example.

Sampling Interval Calculation

$$\text{Sampling interval} = \frac{\text{Total number of basic sampling units (BSU) in the population}}{\text{Number of sampling units needed for the sample}}$$

Example:

You need to conduct a mortality survey in a camp which contains 2,400 households. According to your calculations, your sample should contain 300 households. The sampling interval you will be using is:

$$\text{Sampling interval} = \frac{2400 \text{ households in total population}}{300 \text{ households for your sample}} = 8$$

Procedure for Systematic Random Sampling:

1. After calculating your sampling interval, you are ready to select your first BSU.
2. Choose a random number⁸ between 1 and the sampling interval you calculated.
3. Starting at number 1 of your list or survey area, count the BSUs until you get to the one that has the number you have picked randomly. For example, if you have picked number five, start counting from the start of the list/camp until you get to the fifth BSU. This will be the first BSU in your sample.
4. To choose the second BSU, add the sampling interval to the number of the first BSU. In the example above, your second BSU to survey will be 5+8 (if 8 is the sampling interval), which results in 13. Therefore, after coming out of BSU five, you keep walking and counting BSUs until you get to number 13 and this will be the second BSU of your sample.
5. Subsequent households are selected in the same manner, by adding the sampling interval to the number found for the previous BSU.

In some contexts, the calculated sampling interval may be a decimal number. In that case, the procedure to find the numbers of households to visit is slightly different. The rounding will be done after adding the sampling interval to the decimal number found in the previous step.

Example: Sampling interval is a decimal number.

The village of Kawar is composed of 3,400 households (HH) numbered sequentially from 1 to 3,400. You need a sample of 250 households.

$$\text{The sampling interval: } s = 3400 / 250 = 13.6$$

You ask someone to randomly choose a number between 1 and 13 and they choose **11**. This number is the equivalent to the first household to survey, and the following are chosen as follows:

Table: Calculation of HH numbers using a decimal sampling interval

HH Rank	Calculations	Total	HH number
1 st household	11	11	11
2 nd household	11+13.6	24.6	25
3 rd household	24.6+13.6	38.2	38
4 th household	38.2+13.6	51.8	52
5 th household	51.8+13.6	65.4	65
6 th household	65.4+13.6	79	79
7 th household	79+13.6	92.6	93
8 th household	92.6+13.6	106.2	106
9 th household	106.2+13.6	119.8	120
Etc.	Etc.	Etc.	Etc.

Three scenarios are possible for systematic random sampling depending on the demographic data available:

- **List of target population available** (scenario 1 below): A sample of individuals from the target population can be drawn directly from this list using the ENA for SMART random number table function.

OR

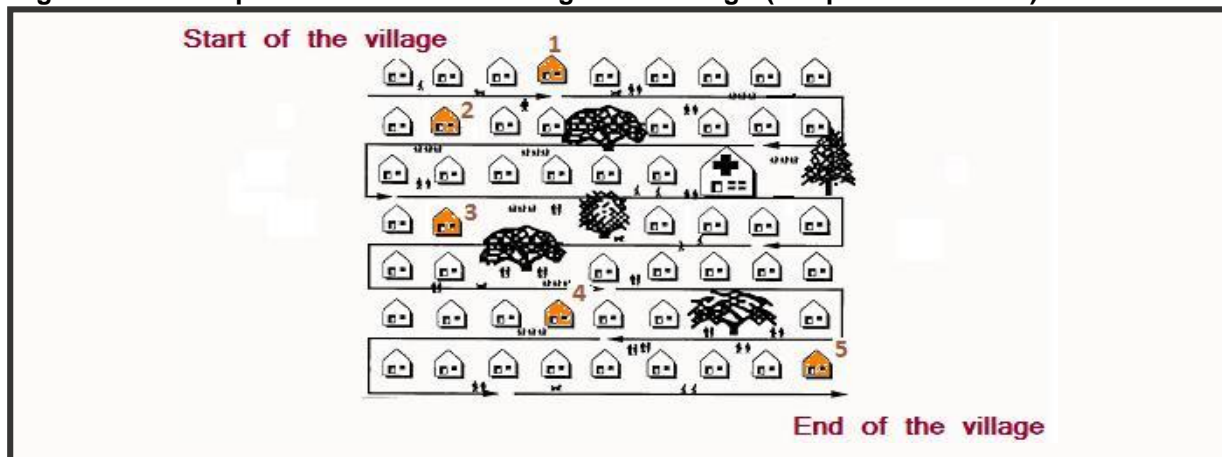
⁸ This number can be chosen by picking a number out of bag or hat, or using a random number table.

- **List of households available** (same as scenario 1): A sample of households can be drawn randomly from this list. Once those households are selected, all children in those households will be measured.

OR

- **No list is available but households are arranged in an easy geometric pattern** (scenario 2 below): The survey team needs to know the boundaries of their survey area, and determine what could be considered as its start point and end point and which direction to follow when walking or driving through the area (Figure 10).

Figure 10: Example of Geometric Setting of Dwellings (adapted from MSF).



Scenario 1: Systematic random sampling using a list of children.

You want to conduct a nutrition survey in a camp which has a list of 1,500 children between six and 59 months. Your sample size is 300 children.

1. First, make sure the list is not ordered in a way that would introduce a bias in the selection (e.g. according to age, mother's level of education, etc.).
2. Calculate your sampling interval: $1500 \div 300 = 5$.
3. Choose a random number between 1 and 5. Let's say you picked the number 3.
4. Count from the top of the list down to child number 3 on the list. This child will be the first child in your sample.
5. The second child in the sample will be: $3+5 = 8$; the eight child on the list.
6. The third will be $8 + 5 = 13$, and so on, until you get the 300 children needed for your sample.

Scenario 2: Systematic random sampling without a list.

You want to conduct a nutrition survey in an area that has a relatively small population, which is arranged in rows (as seen in Figure 10). Your basic sampling units are households. The area contains 3,000 households, but you only need 250 for your sample.

1. First, determine the start and end of this survey area and the direction you will be following.
2. Calculate your sampling interval: $3000 \div 250 = 12$.
3. Choose a random number between 1 and 12. Let's say you picked the number 4.
4. From the start of the survey area, count dwellings until you get to the fourth one. This will be your first HH in the sample (see Figure 10).
5. To find the second HH in your sample, add 4 to the sampling interval (12): $4+12 = 16$. Your second HH of the sample will be the sixteenth in the village.
6. Continue with this same process until you get the 250 HH needed for your sample.

Cluster Sampling

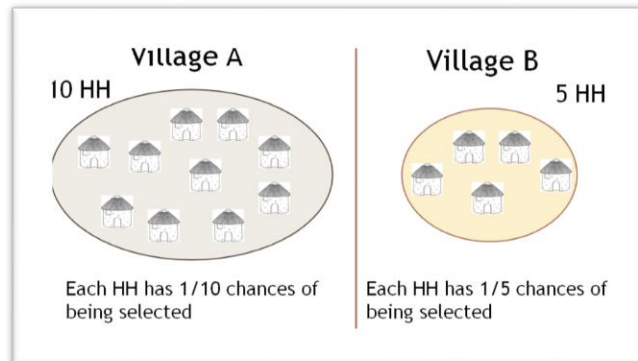
In cases where your survey area is too large (a country, province, or district for example) and/or the population is dispersed, cluster sampling can be used. This method is the **most frequently used in the field**. In cluster sampling, basic sampling units are selected within groups named clusters (villages, administrative areas, camps, etc.) The objective of this method is to choose a limited number of smaller geographic areas in which simple or systematic random sampling can be conducted. It is therefore a **multi-stage sampling method**. Very often, it is completed in two stages:

- **First stage: Random selection of clusters.** The entire population of interest is divided into small distinct geographic areas, such as villages. It is necessary to find an approximate size of the population for each village (in this example). At this stage, the primary sampling unit is the village. Then, clusters will be assigned randomly to villages using the ENA for SMART software.
- **Second stage: Random selection of households within clusters.** Households are chosen randomly within each cluster using simple or systematic random sampling.

This section will discuss in further details the **first stage of sampling**.

Large survey areas often contain geographical units or settlements of various sizes. Therefore, it is important that each individual in each of those settlements has **an equal chance** of being selected whether he or she lives in a large or small settlement. To illustrate this, imagine that you have to survey an area which contains two villages, A and B (Figure 11). Village A contains 10 HH and village B has 5 HH: village B is half the size of village A. Let's assume that we choose randomly one of the two villages by tossing a coin, for example. In village B, each HH has double chances of being selected than in a HH in village A (1/5 compared to 1/10). Therefore, chances of HH in each village to be selected will not be equal by just tossing a coin.

Figure 11: Example of Villages of Different Sizes.



To account for this, the selection of clusters is done by the ENA for SMART software using **probability proportional to population size (PPS)** method. In PPS, larger settlements have a higher chance of being selected as clusters compared to smaller settlements because the probability of selection is proportional to the population size of the settlement.

How does ENA assign clusters randomly using PPS?

- As shown in Table 4 below, ENA calculates cumulative population sizes: i.e. for each geographic unit, the cumulative population size will be the size of the population for that unit, plus the sum of all the units which come before it on the list.
- A range of household numbers is also assigned for each geographical unit according to its cumulated population, as shown in the fourth column of Table 4⁹.

⁹ In this example, we ended up having a lot of clusters in the same villages because the number of villages was limited. In real situations, there may be villages that will have no clusters assigned, while others will have only one or two, and yet others very large that might have three or more.

- ENA then calculates the sampling interval, which is the total population of all the geographic units divided by the number of clusters needed. If the total population is 4,200 and 30 clusters are needed; then, the sampling interval will be $4200/30 = 140$.
- Sampling will begin at a randomly selected starting point: therefore, we choose a random number as our starting point between one and the sampling interval (140). The geographic unit where this number lies will be the cluster number one. Let's assume that the random start number is 100. This falls into village one.
- The second cluster will be: $100+140 = 240$, which also falls into the range defined for village one (1 to 500): this will be cluster two.
- For cluster three, it will be: $240+140 = 380$, which is again in village one. However, if we add again the sampling interval to 380, you will get 520, which falls into village two. ENA continues this way automatically until it assigns the 30 clusters (Figure 12).

Table 4: Example of Cluster Selection by PPS.

VILLAGES	Estimated Total Population	Cumulative Population	N° Allocated	CLUSTERS
N° 1	500	500	1 - 500	1; 2; 3
N° 2	400	900	501 - 900	4; 5; 6;
N° 3	160	1060	901 - 1060	7
N° 4	650	1710	1061 - 1710	8; 9; 10; 11; 12
N° 5	520	2230	1711 - 2230	13; 14; 15; 16
N°6	640	2870	2231 - 2870	17; 18; 19; 20
N°7	700	3570	2871 - 3570	21; 22; 23; 24; 25
N°8	100	3670	3571 - 3670	26
N°9	470	4140	3671 - 4140	27; 28; 29
N°10	60	4200	4141 - 4200	30

Figure 12: ENA table for cluster selection.

The screenshot shows a software window titled "Table for Cluster sampling". At the top, "Population Size:" is set to 491971. Below this is a table with three columns: "Geographical unit", "Population size", and "Cluster". The table lists several geographical units with their respective population sizes and the clusters assigned to them. The first row, BUTAGANZWA, is highlighted with a black border. Below the table, there is a small note: "*RC: Reserve Cluster".

Geographical unit	Population size	Cluster
BUTAGANZWA	52746	1,2,3,4
GAHOMBO	41851	5,6,7
GATARA	66963	8,9,10,11,1...
KABARORE	35187	14,15
KAYANZA	86154	16,17,18,1...
MATONGO	34551	23,24,25
MUHANGA	74124	26,27,28,2...
MURUTA	27955	31,32
RANGO	72440	33,34,35,3...

Once the villages that contain clusters are identified, second-stage sampling will be conducted in each village to randomly select households to be a part of the sample. This will be discussed in more detail on Page 44 of this manual.

Sample Size Calculation

Sample size expressed in terms of children vs. households

SMART Recommendation

There are two ways to express the final sample size required, either with a fixed number of households or with a quota, that is, a fixed number of children.

The quota sampling method means that teams keep selecting households in the cluster until a specific number of children are reached, regardless of the number of households visited. For example, if each cluster should contain 20 children, teams will be selecting as many households as necessary in the village to reach that number of children by the end of the day, and sometimes, they have to visit a neighbouring village to complete their cluster.

In contrast, when using the fixed household method, only a certain number of households, selected randomly in the field, will be visited. At the end of the survey, some clusters will have more children than others and the total should not differ significantly from what was planned. Therefore, survey teams should be given **only the target number of households** per cluster (and not the number of children per cluster) in order to avoid confusion and unnecessary stress and errors.

SMART recommends using the fixed number of households method for three main reasons:

1. **Household selection technique:** SMART recommends using simple or systematic random sampling methods to choose households **within the clusters** (in the second stage of sampling), since they are better than modified EPI in terms of representativeness of the sample and introduce less bias. These two methods are based on the selection of households either from a list (simple) or with a sampling interval (systematic). Therefore, when using one of these two sampling techniques, it is more logical to have a fixed number of households as a target to reach for each cluster. In other words, since it is only possible to estimate the approximate number of eligible children per household prior to data collection (which might not reflect the actual number found in selected houses), it will be impossible to know in advance the required number of households to select to obtain the exact number of children¹⁰;
2. **Inclusion of HH with no children:** When survey teams have a target number of children to reach, they may have a tendency to skip households that do not have children. Other indicators collected during the same survey (such as mortality, water and sanitation, food security, etc.) are measured at household level. Measuring these indicators only in households with small children and excluding all other households will create a serious bias;
3. **Sample size calculation may be done not only for anthropometry, but also for mortality and other indicators measured at household level:** To compare and reconcile these sample size requirement, both sample sizes (for example, anthropometry and mortality sample sizes) should be expressed in the same units (i.e. households).

Conversion of Sample Size

In the planning tab, the ENA for SMART software will automatically convert the number of children found for the sample size into a number of household, following the formula below:

⁵ Quota sampling was more adapted to the EPI method for household selection since it was not necessary to determine the number of households to visit in advance as it is the case for simple or systematic sampling methods. However, SMART does not recommend the use of EPI as a preferred method.

Equation 1: Conversion of sample size from number of children into number of households.

$$N_{HH} = \frac{N_{Children}}{(HH\ size \times \% \ of \ Under\ 5 \times 0.9)}$$

Where: N_{HH} = sample size in terms of households
 $N_{Children}$ = sample size in terms of children
 $HH\ size$ = average household size
 $\% \ of \ under\ 5$ = proportion of under 5 years-old children in the population
 0.9 = fraction of 6-59 months children within the under 5 age category

Note that once N_{HH} is found, ENA adjusts it using the non-response rate (see Page 41 for more details).

Example:

If $N_{Children} = 500$, average HH size = 6.2, and $\% \ of \ under\ 5 = 15\%$ (that is, 0.15 when expressed as a proportion),

Then the sample size in HH is $N_{HH} = 500 / (6.2 \times 0.15 \times 0.9) = \underline{598\ HH}$.

Accuracy, Precision and Confidence Interval

As discussed before, using a sample within a total population will only provide an *estimate* of the indicator of interest, and not the True Population Value. Therefore, as portrayed in Figure 13, we need to consider:

1. How close this estimate is likely to be to the true value (level of **accuracy**);
2. How similar will be the results if several surveys are conducted with the same methodology (level of **precision**).

Figure 13: Precision and accuracy.



Accuracy

Accuracy measures the **validity** of the estimate. Lack of accuracy is related to **bias**, which reflects the difference between the sample estimate and the True Population Value due to error in measurement, selection of a non-representative sample, or factors other than sample size. **We cannot control bias once it is present.** However, we can try to reduce its occurrence by standardising survey procedures, providing adequate training to every member of the survey team, and closely supervising all survey work.

Precision

Precision measures the **consistency** of the results and is related to **sampling error**, which is the difference between the sample estimate and the True Population Value due to measuring a sample instead of the whole population. **We can control sampling error by controlling sample size.** In essence, sampling error is the measure of uncertainty stemming from the fact that our estimate is derived from measuring only a relatively small proportion of the target

population, and therefore it will never be exactly the same as the True Population Value, even if the sample is selected randomly and representatively. The larger the proportion of the target population that is measured, the lower this uncertainty becomes. Therefore, the **higher the sample size, the higher the precision**.

However, a larger sample size increases the precision of the results but does not guarantee the absence of bias. When the sample size is very large, quality control becomes difficult because of the high number of teams to train and supervise, and there may be a higher risk of bias.

Remember. Preventing bias is critical: never try to achieve higher precision at the expense of introducing bias. It might be better to have a smaller sample size with less precision but much less bias. Annex 1 describes a number of potential biases to consider and recommendations on how to avoid them.

Confidence Interval and Level of Confidence

As the ENA for SMART software calculates the needed sample size, it will determine by convention a 95% confidence interval (CI). This means that in absence of bias, once the estimate is generated, one can be 95% sure¹¹ (level of confidence) that the True Population Value of the measured indicator is within the limits of the confidence interval calculated. The more representative is the sample and the more accurate are the measurements, the more certain one can be about the absence of bias and of having the true population value for the indicator within the confidence interval.

Parameters for Sample Size Calculation

The calculation of sample size depends on the sampling design and the indicator.

In most cases, to calculate sample size for **simple or systematic random sampling**, two numbers are required:

1. The **expected prevalence** of the indicator you are measuring;
2. The **desired precision** that needs to be achieved to meaningfully interpret the estimate.

Equation 2: Sample size formula for anthropometry in systematic random sampling.

$$n = \left[z^2 \times \frac{p \times q}{d^2} \right]$$

Where: n = sample size
z = linked to 95% confidence interval (use 1.96)
p = expected prevalence (as fraction of 1)
q = 1- p (expected non-prevalence)
d = relative desired precision

In cluster sampling, the sample size calculated for simple or systematic random sampling is multiplied by a factor called the *design effect* to account for the heterogeneity between clusters with regard to the measured indicator (see Page 39). Therefore, to calculate the sample size for a **cluster survey**, a third number is required:

3. The **expected design effect**.

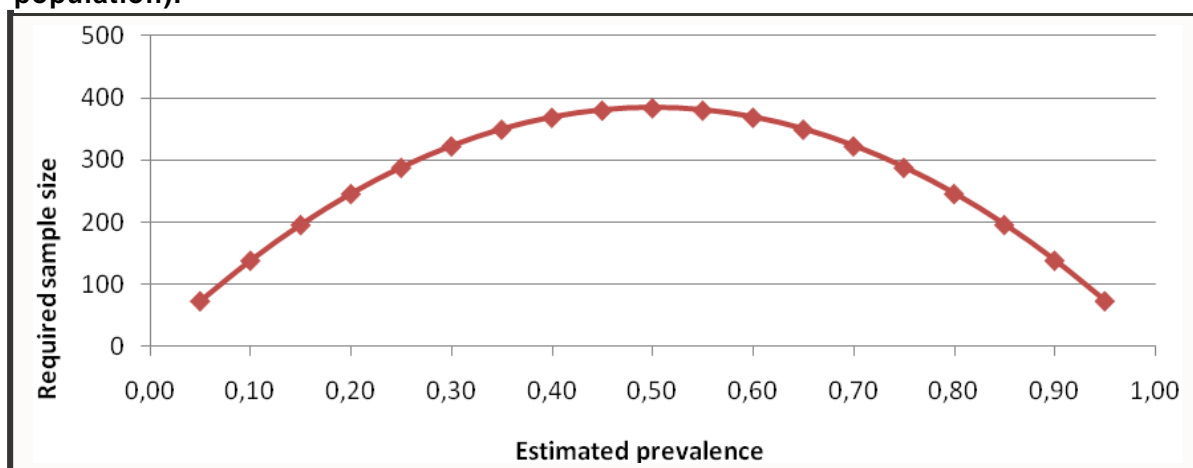
The following sections will provide guidance on how to best determine these numbers.

⁷ The level of confidence used in ENA is always 95%.

Estimated Prevalence

Information to determine the estimated prevalence can be obtained from various sources such as previous surveys, surveillance data, or rapid assessment results. Health workers can also provide interesting information about whether they see more thin children than in the past. If there is a range of possible values, it is always better to choose the one which is closer to 50%, which is the most conservative value (producing the largest sample size possible given the other parameters are held constant), as shown in Figure 14.

Figure 14: Effect of Changing the Estimated Prevalence (assuming 95% CI, +/- .05, large population).



To make the best possible “guesstimate” of expected prevalence:

1. Look at results of previous surveys conducted in this area or in similar areas, and consider confidence intervals reported in these surveys. If no surveys exist, try to estimate the prevalence using data from rapid assessments, anecdotal reports, feeding programme admissions’ trends, etc.
2. Adjust previous estimates for seasonality, and according to your understanding of if and how the situation has changed since the last surveys were conducted.
3. After going through points (1) and (2), you may come up with the range of values where you think the current prevalence might be. To be on the safe side, use the higher limit of this range in the formula for sample size calculation. For example, if the current prevalence is likely to be between 15% and 19%, use 19% for your calculations.

Precision

There is no standard precision to use when planning a survey. Desired precision depends on the objectives of the survey, estimated prevalence or rate, and resources available. Other things kept equal, the higher the desired precision¹², the larger the sample size. This effect can be seen in Equation 2, where squared precision is in the denominator, and therefore is inversely quadratically proportional to sample size. In other terms, to decrease the width of a confidence interval two times, the sample size needs to be increased 4 times.

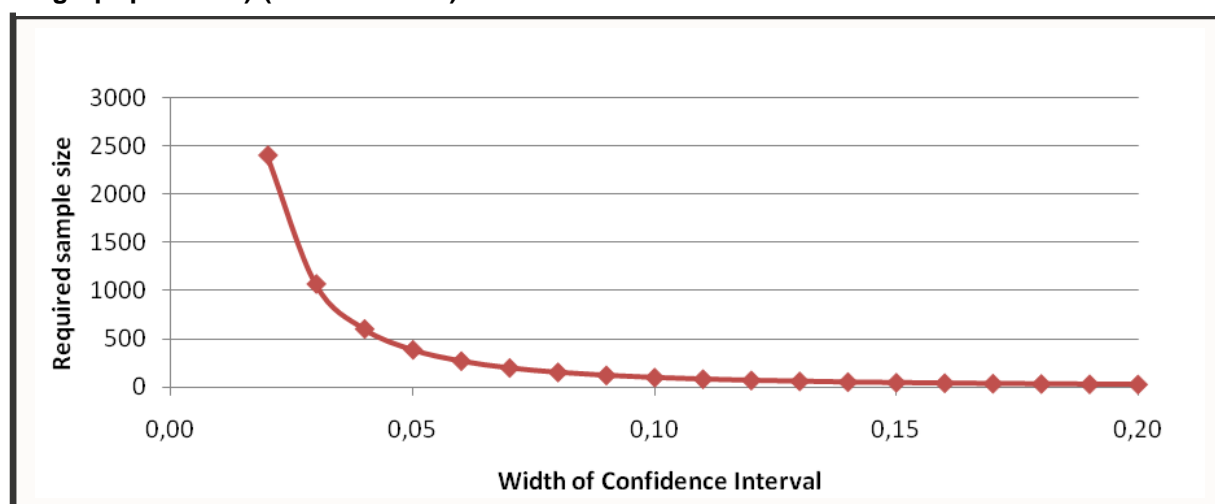
Consequently, logistical or financial resources might also be taken into consideration when calculating sample size. If a sample size of 384 households is required with a precision of +/- 3% (estimated prevalence: 10%), increasing this precision to +/- 2% will result in a sample size of 864 households (2.25 times larger). Much more resources will then be needed to survey that many households (Table 5 and Figure 15).

⁸ Note that a higher precision for measurements means a smaller value associated to that precision. For example, changing from a precision of 10% to 5%, the value is decreased, but this means achieving a higher precision.

Table 5: Effect of Changing Precision on the Sample Size.

Malnutrition Prevalence (%)	Desired Precision (\pm %)	Sample Size (Children)
10	1.5	1537
10	2.0	864
10	2.5	553
10	3.0	384

Figure 15: Effect of Changing the Desired Precision (assuming 95% CI, 50% prevalence, large population) (source: CDC).



In general, to meaningfully interpret Global Acute Malnutrition (GAM) estimate in a survey, a precision of \pm 3% should be sufficient in most cases, even if expected prevalence is low. If expected prevalence of GAM is higher, for example 15-20%, the precision of \pm 4-5% would likely be sufficient. As shown in Table 5, requiring confidence interval to be narrower than \pm 2.5% would result in very large sample sizes that would be very difficult to achieve, and in most cases would be unnecessary. If the estimated prevalence is very high, precision does not need to be very high, since the limits of the confidence interval might not make a big difference in decision-making. For example, if the prevalence of GAM is of 45%, having a precision of \pm 10% might be enough, because the confidence interval will be from 35% to 55% and at that stage, it should not make much difference on whether or not to intervene (Table 6)¹³.

Table 6: Variation of Precision Needed with Different Malnutrition Prevalences.

Malnutrition Prevalence (%)	Desired Precision (\pm %)	Sample Size (Children)
5	3.0	203
10	3.5	282
15	4.0	306
20	4.5	304
30	5.0	323
40	6.0	256

¹³ Values in tables 6 and 7 are only suggestions. Depending on the context, you may want to choose a precision that is more adapted to your survey.

Precision may need to be higher if survey results are to be used to monitor the situation on a yearly basis, or if you are trying to detect statistical difference between two survey estimates. In that case, a different formula for survey size calculation is used.

If chronic malnutrition is the main indicator for the assessment, the reader should refer to Annex 2 for further details on its sample size calculation.

Design Effect

The design effect (DEFF) is a “correction factor” to account for the heterogeneity between clusters with regard to the measured indicator. Therefore, it is only used to determine sample size in cluster sampling (Equation 3).

Equation 3: Sample size formula for anthropometry in cluster design.

$$n = \left[t^2 \times \frac{p \times q}{d^2} \right] \times DEFF$$

Where: n = sample size
 t = linked to 95% confidence interval for cluster sampling (2.045)
 p = expected prevalence (fraction of 1)
 q = 1- p (expected non-prevalence)
 d = relative desired precision
 DEFF = Design Effect

If the prevalence of GAM among the clusters in your survey is very different, the number of GAM cases per cluster will likely vary widely, meaning that the GAM distribution is heterogeneous, and the design effect is likely to be high. On the other hand, if the prevalence of GAM in all clusters is about the same, the number of GAM cases per cluster will also be similar, meaning that GAM distribution in the survey area is fairly homogenous, and the design effect is likely to be low.

Example: A measles outbreak in one or two clusters may increase prevalence of malnutrition in these clusters compared to others with no outbreak. This high cluster prevalence is not representative of the whole area but will increase the general prevalence estimate, which will not reflect the reality of the majority of clusters. To compensate for this, we will increase our sample size by using a correction factor called design effect.

The design effect to use when planning the survey can be taken from previous surveys in the same area, if there is no reason to think that there has been any change that might have increased or decreased the heterogeneity among clusters.

Generally, if there is no previous information about design effect, 1.5 can be used as a default for GAM.

The design effect will depend on the prevalence and the size of the clusters:

1. The higher the expected prevalence, the higher would be design effect. For example, if expected prevalence is around 10%, expected DEFF may be 1.5, whereas if expected prevalence is around 25-30%, expected DEFF would increase to 1.7-1.8.
2. The smaller the number of children per cluster, the smaller the DEFF. For example, if 15 children are measured per cluster, the DEFF may be 1.5, whereas 25-30 children are to be measured per cluster, expected DEFF would increase to 1.7-1.8.

Remember. The design effect is a multiplying factor (i.e. sample size is directly proportional to the design effect), so if the EFF is increased two times, the calculated sample size will be two times higher (Table 7).

Table 7: Effect of Changing Design Effect on Sample Size (assuming a 3% precision).

Malnutrition Prevalence (%)	Desired Effect (\pm %)	Sample Size (Children)
10	1.0	418
10	1.5	625
10	2.0	836
10	3.0	1,255

The 3 components above: estimated prevalence, desired precision and design effect allow the ENA for SMART software to calculate sample size in terms of number children. To **convert this number into number of households**, more elements need to be determined, as described below.

Demographic Data

Two types of demographic data are needed to convert the number of children into the number of households:

1. **Average household size;**
2. **Proportion of children under five years old** in the population.

This information can be obtained from a census or a past mortality survey that was conducted in the same area.

In many contexts of developing countries, the proportion of children under five years old does not reach 20%. Over-estimating this proportion and/or the number of persons per household will result in having a small sample size in terms of households (other parameters kept equal), and therefore in a smaller than expected final sample size in terms of children.

Even if the target population is children between six and 59 months, only the under-five proportion in the ENA for SMART software and it will make the calculations automatically assuming that 90% of the children under five years old are aged between six and 59 months.

Household Definition

Definition of household is a key part of the survey planning since most of the time; households will be the basic sampling unit at the last stage of sampling. There is no clear consensus about a definition. It is recommended to use the definition that is used by the country's departments. This definition needs to be agreed upon **before** the survey starts.

The most frequently used definition is: "**People who slept in the house last night and ate from the same cooking pot.**"

In country where polygamy exists, a clear definition of a household is even more crucial. By having a specific definition, you'll be able to determine how many households there are in a polygamous family. For example, if all wives cook together, eat together and live in the same compound, this would be one household. However, if each wife has her own kitchen and prepares food for her own children, those would be separate households.

Non-Response Rate

Non-Response Rate (NRR) accounts for households that could be either absent or unreachable, refuse to be surveyed, or any other reason that prevent survey teams from surveying a selected household. The ENA for SMART software uses this rate to inflate sample size using the following formula:

Equation 4: Non-response rate calculation

$$Final\ N = \frac{Number\ of\ HH\ needed}{1 - NRR}$$

Example: If according to the estimated prevalence, desired precision, and design effect, number of persons per household and percentage of children under five years old, the sample size is of 695 households as the sample size, and the non-response rate is expected to be about 5%, the final sample size calculated by the ENA for SMART software should be:

$$Final\ N = 695 / (1 - 0.05) = 732\ HH$$

Correction for Small Population Size

As may be noted, the sample size formula does not include the size of the population of the sampling universe. However, if the target population is small, a smaller sample size would be needed to achieve the required precision. If the target population is below approximately 10,000, the box “Correction Small Population Size” should be checked in the ENA for SMART software’s *Planning* tab in order to adjust the sample size.

The ENA for SMART software calculates the target population (for example children 6-59 months) from the total number entered in the cluster selection table and the percentage of children under five years old entered into the calculator. For example, if the total population size in the cluster selection window is 40,000, and the percentage of children under five years old is 15%, the ENA for SMART software would assume that there are $40,000 * 0.15 * 0.9 = 5,400$ children aged 6-59 months in this sampling universe, and use this number for adjustment for small population size.

Combined Surveys

In some contexts, a mortality and anthropometric survey might be conducted simultaneously in the same population. The ENA for SMART software can be used to calculate sample size for both¹⁴. However, in most cases, sample sizes for anthropometry and mortality will not be the same.

There are 2 scenarios possible:

- **Small difference between the two sample sizes:** The total number of households to survey for both components (anthropometry and mortality) will be determined by the higher sample size found through calculations. For example, if sample size is 497 households for anthropometry and 425 households for mortality, the higher value (497 HH) will be considered as the sample size to survey for both indicators. This will add about two to three households to each cluster in a survey containing 30 clusters.
- **Large difference between the 2 sample sizes:** Every indicator should be surveyed with more or less the sample size calculated for that indicator; it will therefore be necessary to

¹⁴ For more information about sample size calculation for mortality, the reader should refer to mortality module in the SMART manual or the SMART Standardized Training Package on www.smartmethodology.org.

set an interval at which teams will be collecting data for that indicator that requires smaller sample size. For example, if the sample size for anthropometry is 200 households and the one for mortality is 400 households, teams will be conducting anthropometry measurements only in children found in every second household selected in each cluster.

Scenarios to Avoid in Combined Surveys

1. **Measuring one indicator only in some of the clusters, and measuring both indicators in other clusters.** This would create a selection bias because the assignment of clusters at the planning stage is done according to the number of clusters needed for the whole survey. By doing an extra selection within those clusters, not every individual in the population will have had the same chance of being selected for that second indicator.
2. **Measuring both indicators only in the first households of every cluster when doing systematic random sampling.** This might also create a selection bias as it alters representativeness of the sample if the last households in the village are different from the first ones (access to road, market, etc.).

Determining the Number of Clusters

Once the sample size has been calculated in terms of households, the number of households to include in each cluster will need to be determined in order to find the number of clusters that will be visited. This is done by taking into consideration the reality of the team work in the field during data collection. This allows survey managers not to overload teams by requesting them to survey too many households per day.

If we suppose that one cluster will be completed each day by each team, the calculation for the number of households to include in each cluster is based on:

1. Travel time;
2. Number of work hours per day;
3. Number of hours spent on the field (excluding transportation);
4. Number of hours spent on surveying in the households and getting from one selected household to the next (excluding break times and time spent to introduce the teams and select the households.).

However, there is a minimum of clusters that should be included in each survey for the survey to be considered valid¹⁵: 25 clusters are considered a minimum, but normally, nutrition surveys include at least 30 or more clusters. Therefore, in most situations, it is strongly recommended to **consider 30 clusters as a default**. If there is a lot of heterogeneity in the distribution of the measured indicator, consider including more clusters and measuring fewer children per cluster to decrease DEFF.

Assigning Clusters

After defining your geographical area at the planning stage and taking into account travel times, security issues, and any other factor that could influence the ability of the survey teams to get to the cluster site, the survey manager needs to choose the smallest administrative unit that has population data and a known name.

At this stage, each geographical unit should have at least the number of household required to complete a cluster. Otherwise, it should be paired with the neighbouring unit when constructing your sampling frame for cluster selection.

¹⁵ Binkin N. Rapid nutrition surveys: how many clusters are enough? *Disasters* 16(2): 97-103.

After that, the sampling frame is pasted into the cluster selection window on the *Planning* tab of the ENA for SMART software. The sampling frame is a list of all geographic units (or settlements) with the population size of each geographical unit (expressed in terms of total number of people, not households). The order of settlements on the list is not important, it can be sorted alphabetically or geographically. It is very important that all settlements in the survey area be included. If some areas are omitted at this step, they will not be part of the surveyed population and this should be clearly described in the report. The number of clusters to select from this sampling frame is then entered and the button “Assign Clusters” should be pressed

Figure 16: ENA functionality for assigning clusters.

The image shows a software interface element with a light gray background. On the left, the text "Number of Cluster" is displayed. To its right is a white input box containing the number "38". Further to the right is a rectangular button with the text "Assign Cluster" inside it.

(Figure 16). The software will automatically select the areas or “villages” that contain a cluster. This is assigned randomly using PPS procedure previously described.

Small Villages

If a village does not have enough households for one cluster, it should be **combined** with the closest one. If this combination of villages is chosen by the ENA for SMART software to contain a cluster, allocate households proportionally among the two villages.

Example:

Village A has 15 households; village B has 60 households. Since village B is 4 times bigger than village A, we will allocate 4 times more households to village B. If the cluster should contain 20 HH:

- Village A: 4 HH
- Village B: 16 HH

Large Villages

If a very large village is chosen to contain only one cluster, use the segmentation technique as explained on Page 45 of this paper.

If a large village is chosen to contain two clusters or more, two options are possible:

- **Segmentation:** Divide the village into two or more segments and assign a cluster randomly to each segment if segments are equal, or using PPS if segments are unequal.
- **Conduct the clusters one after the other:** However, in order to have a proper data analysis, each of those two clusters needs to have a different number. Therefore, teams have to change the cluster number on the questionnaires after finishing each one.

Reserve Clusters

As discussed earlier, the sampling frame should only include sampling units that are accessible and can be visited if selected. However, in some cases surveyors could be prevented from accessing previously selected clusters after the survey starts. Clusters might be impossible to survey because of insecurity, accessibility, refusal of the community or leaders, etc. Examples can be a conflict that starts in one village that contains an original cluster, or a flood might occur and a cluster becomes inaccessible.

At the planning stage, when assigning clusters, the ENA for SMART software automatically chooses additional clusters called *reserve clusters* or *replacement clusters*. The number of reserve clusters (RC) selected by the ENA for SMART software is as follows:

If 25-29 clusters are included in the survey → 3 RC; if 30-39 clusters → 4 RC; if 40-49 clusters → 5 RC, etc.

Reserve clusters should only be used if 10% or more of original clusters were impossible to reach during the survey. For example, if 30 clusters were planned to be surveyed, but during data collection three or more clusters became impossible to survey, the replacement clusters can be included.

All reserve clusters should be surveyed if 10% or more of the planned clusters have not been reached. For example, if the target is to survey 40 clusters and but only 36 clusters were accessible, all replacement clusters should be included in the survey, even if there are five of them.

Second Stage Sampling

If cluster sampling has been selected at the planning stage, the ENA for SMART software will randomly select villages to contain each cluster. This is called the first stage of sampling. The next step is to randomly select basic sampling units within those clusters, which is called the **second stage (or last stage) sampling**.

Preparation for household selection

As mentioned above, one of the reasons to use cluster sampling is to divide the survey area into smaller geographical units where simple or systematic random sampling will be feasible. However, in some cases, even villages that are chosen to contain a cluster might be too large and in most cases, no list of households is available at the village level. Therefore, when arriving to the village, some preparatory work needs to be done before being able to select the actual households that will be included in the survey.

List of Households Available

After introducing the team and explaining the objectives of the survey to the local authorities in the village, it is important to get an updated list of the households for that village. This list should not be based on some socio-economic indicators that would exclude part of the populations in the village: for example, some lists only include mothers who have children and are followed by the female health workers. Using only this list will introduce an important selection bias. If the list is not updated or complete, the team leader can use the local knowledge to do so. It is very important that no household is excluded from that list.

Listing Households

It is very important to explain the definition of households to the local authorities when asking them for lists. Very often, people might confuse families and households. Hence, survey teams should make sure that village leader and people present during those preparatory steps understand the definition of household that was used in the planning stage of the survey.

List of Households Not Available

- **Drawing a Map.** If there is no list in the village, the survey team should try to make one with the local authorities.

The easiest procedure to make this list is to:

1. Ask the village leader or the most knowledgeable person present to draw the boundaries of the village on a paper.
2. Identify the natural and community landmarks on the map (e.g., rivers, roads, church or mosque, school, health center, market, etc.). This will allow the team to define internal boundaries and consider separate sections of the village for the next step.
3. Ask the village leader to list households that live in each section limited by specific landmarks. For example, survey team can ask the village leader to list all households that live in the area between the market and the river. Go through all sections of the village writing the head of household's name and giving each household an ID.

4. If the village is not too large, walk around with the village leader and confirm the number of households for each section.

- **Numbering the households.** In some cases, villages are not very large but village leaders are still unable to list all households. In such situations, survey team members can walk around the village and identify all households, by writing a number (starting at one to the total number of households in the village) with a chalk on their door, for example.

Segmentation

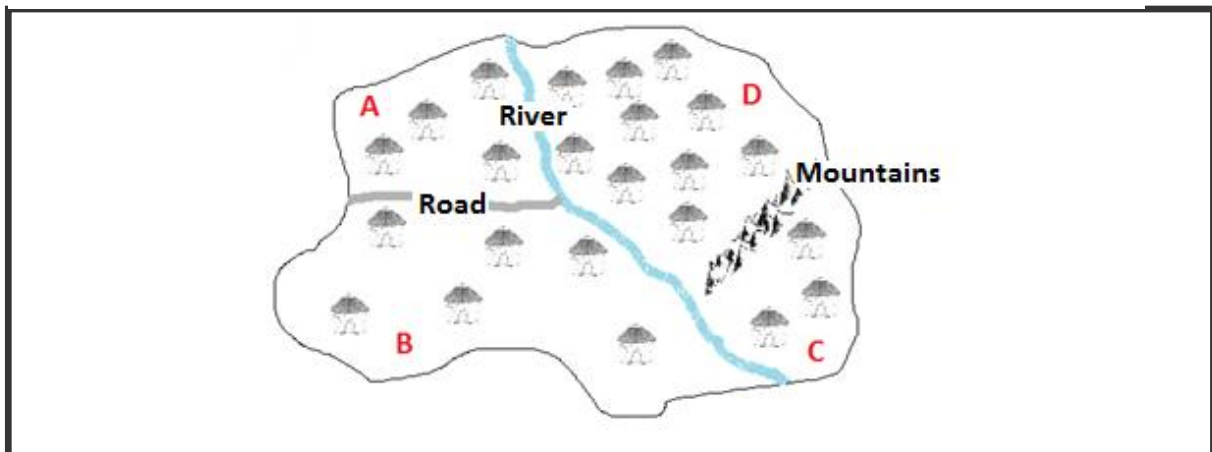
In some cases, villages selected randomly to contain a cluster might be very large or have very dispersed households and sample selection can then become very tedious; teams will have long distances to walk and not enough time to complete one cluster per day. In those scenarios (approximately more than 100 households in the village), segmentation can be used in order to reduce the area that will be covered by the survey teams.

The objective of this procedure is to divide the village into smaller segments and choose one segment randomly to include the cluster. This division can be done based on existing administrative units (neighbourhoods, etc.), natural landmarks (river, road, mountains, etc.) or public places (market, schools, churches, mosques, temples, etc.).

Segmentation can be done into equal or unequal parts.

- **Segmentation into equal parts:** If the village can be divided into two or more approximately equal parts each containing less than 100 households, the survey team leader can write the name of those parts on pieces of paper that he or she folds and put into a bag or hat, and have the village leader or his or her representative choose one paper randomly. Team will therefore go to that part of the village to conduct the survey for that cluster.
- **Segmentation into unequal parts.** In some cases, it might be impossible to divide the village into equal parts, as shown on Figure 17. Natural landmarks should therefore be used to help divide the village into separate and clearly defined segments.

Figure 17: Segmentation into Unequal Parts.



Once those segments are defined with an approximate population size, one segment will be selected randomly using PPS, as shown in the example on the next page.

Example: Selecting a segment using PPS.

You're conducting a survey where each cluster should contain 15 households. **Error! eference source not found.** Figure 17 shows one of the villages selected to contain a cluster. It has segments that have unequal number of households.

- A = 70 HH
- B = 100 HH
- C = 30 HH
- D = 190 HH

TOTAL = 390 HH

To choose one segment randomly using the PPS, we have to create a table with population intervals for each segment (Table below):

Table: PPS for segment selection

Segment	Population (HH)	Cumulative population	Intervals
A	70	70	1-70
B	100	170	71-170
C	30	200	171-200
D	190	390	201-390

After filling out the cumulative population numbers and setting intervals for each segment, pick a three-digit number from the random number table between one and the cumulative total of all the segments.

The segment containing the selected number will be the one to survey.

In this example, a three-digit number must be picked from: 001 – 390: e.g. we picked 167. This number is within the segment B. Survey will therefore be conducted in segment B.

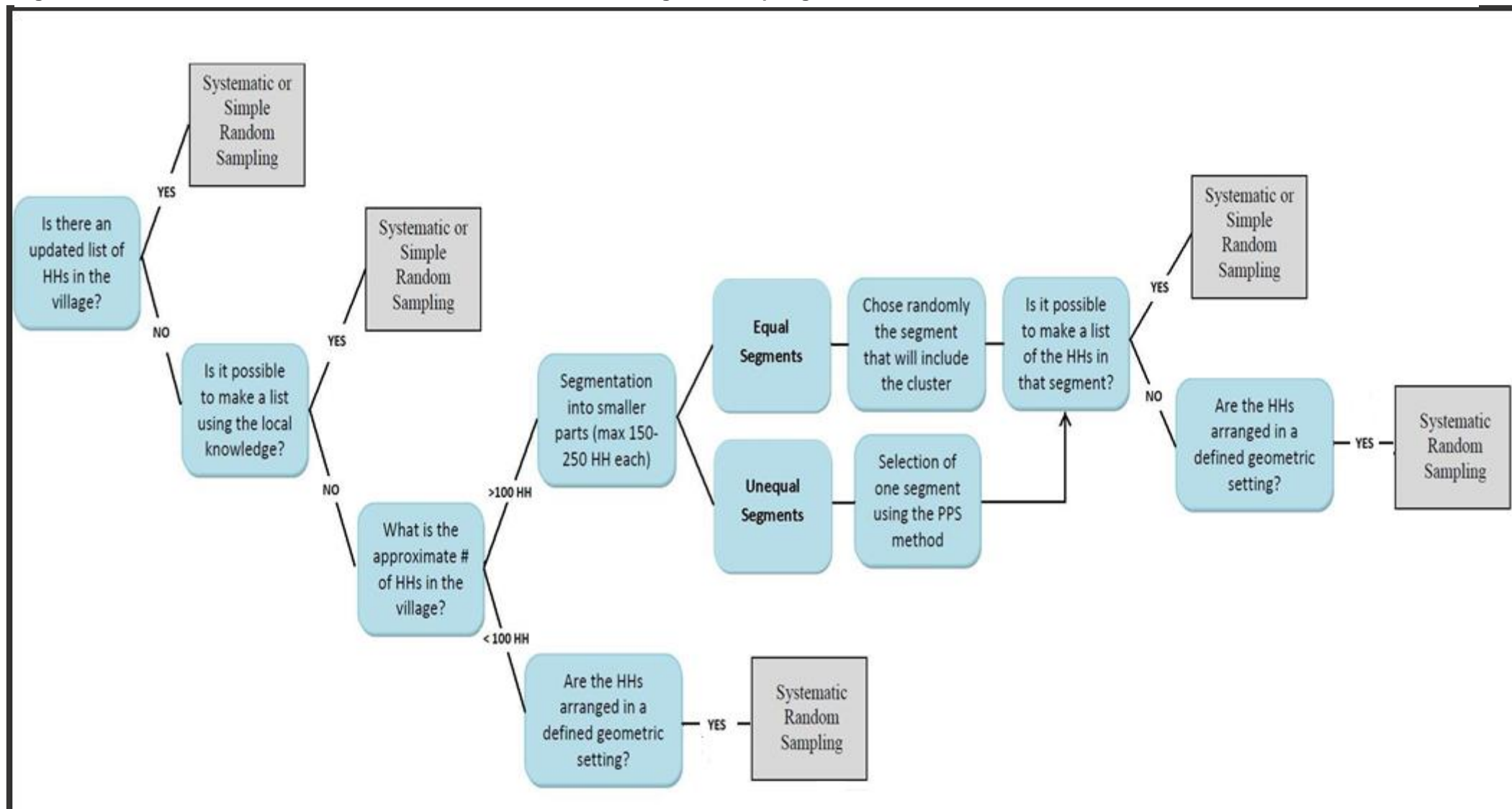
Household Selection Techniques

Once a segment is selected, selection of households in this segment can start using one of the most recommended methods: simple or systematic random sampling. Another method used in the past is the modified EPI. The choice of the method depends on:

- The population size of the village or segment;
- The distribution of households;
- The possibility to get or build an updated list of households.

Figure 18 shows the different pathways to help decide which household selection technique to use depending on the context.

Figure 18: Decision Tree for Household Selection at the Last Stage of Sampling.



Simple random sampling

As previously discussed, simple random sampling can be conducted when an updated and exhaustive list of households is either available or possible to make. Picking numbers from a hat or bag or using a random number table can allow choosing the number of households needed for the cluster.

Systematic random sampling

Systematic random sampling has also been discussed previously and relies on the calculation of a sampling interval that will allow the survey team to find the first household of the cluster, and select the following households until the required number is reached.

Modified EPI

The Expanded Programme on Immunization used to recommend this procedure in the late 1970s to select samples in surveys measuring immunisation coverage.

Modified EPI Limitations

This method should not be used because it induces certain biases in the sampling:

- Since we always select the closest household on the right of the previous one, or the one that has the closest door, it creates a bias because the selection of every household is **not independent** from the selection of the previous one (and this is one important criterion for probability sampling);
- Also, by choosing households by proximity, we are more likely to stay within the same area of the village although households close to each other are more likely to have the same characteristics; hence, you might not be getting a representative sample of your cluster.

Special Cases

In the field, survey teams might encounter various difficulties, some of which are addressed in this section. Teams should therefore use a tool such as a cluster control form (Figure 19) to summarise what has been done in each cluster, keep track of which household refused to participate, which had eligible household members absent and need to be re-visited, etc.

Polygamous families and compounds

As mentioned earlier, a clear definition of household has to be agreed upon before starting the survey. This will allow determining the number of households within a compound or a polygamous family.

If a compound or a polygamous family contains more than one household, each household should be included separately in the list for household selection.

If this information was not communicated when the list of households was constructed with the village leader, and the team finds out about it once in the compound, random selection can be done among those households to choose the one that will be included in the sample. Even if this step will facilitate selection for the teams, this scenario is not ideal since it adds an extra stage of sampling and hence, decreases the chances of each household within a compound of being selected compared to the rest of the sample.

Figure 19: Example of a Cluster Control Form.

CLUSTER CONTROL FORM								
Province: _____ District: _____ Village/Camp: _____ Cluster #: _____ Team number: _____								
Date of sampling: ____ / ____ / ____								
HH no	Head of HH name	First Visit Outcome 1 = completed 2 = partly completed 3 = refused 4 = absent*	Number of eligible children	Number of eligible children measured	Household needs to be revisited Yes or No	Household re-visited Yes or No	Second Visit Outcome (If necessary) 1 = completed 2 = partly completed 3 = refused 4 = absent	Comments
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								

* If household or eligible child are absent, team should re-visit the household once before leaving the village to conduct the interview and/or measure the child.

Impossible to visit a selected household

Many reasons can prevent a team from visiting a household that was selected during the sampling procedure: refusal, fear of dog, access, conflicts with village leader, etc. In this case, the teams should visit the next selected household according to the sampling procedure that was used, **without replacing this household**. The final cluster size of that day will then be decreased by the number of households that were impossible to visit, but this was accounted for in the planning stage by inflating the sample size with the non-response rate.

No children in the household

If the team arrives at a household that only has adults, they should still complete the rest of the questionnaire with a respondent and then move to the next household according to the sampling method chosen. **This household should not be replaced**. It is normal that not all households in the sample contain children and this was also planned for by entering the proportion of expected children in the population. Excluding households without children from selection will introduce serious selection bias in measuring household-level and other non-child variables (e.g., mortality, haemoglobin in women, access to safe water, etc.).

Absent household

Survey team can arrive at a household and find no one because all members are absent. To simplify the procedures, an absent household is defined as the household where some members have slept there the previous night but were absent on the day of the survey.

Survey teams can ask neighbours why the house is empty, but **should not** fill out household survey questionnaire by interviewing neighbours about the absent household. Afterwards, they could move to the following household according to the sampling procedure. Survey teams should return to absent households before leaving the village, to see if residents are back. If not, this should be reported on the questionnaire. **Absent households should not be replaced**. If more than 5% of the households in a cluster are not found, the team should revisit the area at another time to see if they can complete the sample.

Abandoned house (not a household)

An abandoned house is either a house that had no one living in it for a long period, or one that is empty but does not fall in the absent household category (inhabitants left more than one day ago and will not be back at the end of the day, etc.)

Those households should not be included in the list of households used for household selection at the preparatory steps when survey team arrives to the village. **If such empty house is encountered during systematic selection, it should be ignored, and only inhabited houses should be considered in selection procedure.**

Absent child

If a child lives in the household but is not present at the time of the survey, this should be recorded on the household questionnaire and on the cluster control sheet when the household is visited; the rest of his or her information (age, sex, feeding practices, immunisations, etc.) and the rest of the questionnaire can still be filled out with the respondent, such as his or her mother or caregiver.

The team should tell the mother that they will return later on that day to measure the child. If the team returns later but child is not present, this should be recorded as such in the questionnaire and on cluster control sheet, and total number of absentees has to be included in the final report. **Absent child in selected household should not be replaced.**

Child with disability

If the target population of your survey is 6-59 month-old children, those with a disability falling into that age group are eligible for your survey. However, some disabilities might not allow you to take all anthropometric measurements needed or might lead to a biased measure. For example, the weight of a child missing a limb will not be very meaningful when comparing it with the standard population.

All other data that is not influenced by the disability should be collected such as sex, age, oedema (if the child has both feet), etc.

Child in a center or a clinic

When a team finds that one of the eligible children is in a feeding center or a clinic, they should first try to find out the type of disease: malnutrition or not. This child should then be recorded as absent with the reason the absence. The team should attempt to find and measure the child in a feeding centre if it is nearby, or re-visit the household later in the day if the child returns home in the evening.

If a lot of children are being reported as being in the health centre or rehabilitation centre, this may introduce important selection bias. This should be described and discussed in the discussion part of the final report.

Population scattered over a large area

Population scattering is common in pastoral areas. The population numbers are usually very inaccurate and it takes a longer time to travel between sites and find the subjects. If the settlements are small, it is essential to have more clusters, with fewer children per cluster, to ensure there will be a sufficient number of children at each site.

In some situations where the population is very spread out, you may deliberately choose to undertake a survey without sampling certain sections of the population. You might choose to sample only the population in larger settlements, which would save you the time and money needed to get into the bush to the more scattered population. This would mean you would not include population estimates from the scattered area when you are originally selecting your sample. As a result, you would have no accurate data about the nutritional status of children or the mortality rate in the population excluded from the survey. It is critical that the reader of

the report not be misled: there must be a clear description of the exclusion in the report. The decision to exclude or include sections of the population, such as small groups of scattered nomads whose current whereabouts is unclear, will depend upon anecdotal reports of whether these groups are more or less affected than those that are more easily accessible.

Population is very mobile

If you are attempting to undertake a survey in an area of nomads where the population frequently moves large distances, it is likely that you may travel to an area to find that there is no one there or nearby. If you suspect this might happen, you should select some extra clusters before you start the survey. This way, if one cluster is deserted you can replace it with another.

It is sometimes more appropriate to list the names of the nomadic groups, clans, or extended families themselves instead of villages or other fixed settlements. The clans to be sampled are then selected proportional to population size and the teams set out to find the named groups within their migratory range. However, sampling of nomads is a highly specialised topic. Either the advice of an experienced epidemiologist should be sought, or nomadism should be cited as a reason for exclusion. Occasionally in nomadic areas with a very sparse population, it is not possible to do a representative survey.

4. Anthropometry

Why do a Nutrition Survey?

Whether due to starvation, malabsorption, loss of appetite or psychological causes, children who have not taken a sufficient amount of food do not grow; and under more severe circumstances, they lose weight. Decreased growth rate is assessed by comparing the ratio of a child's height to weight to a reference standard for the child's age. For an individual, these measurements are used to decide whether the person is admitted to a supplementary feeding programme or treated for severe malnutrition. At the population level, the same measurements are used in the survey to estimate what proportion of a population as a whole is moderately or severely malnourished.

Malnutrition in the context of this manual takes three forms: 1) failure to grow results in height **stunting**; 2) loss of body tissue results in **wasting**, and 3) accumulation of fluid results in nutritional oedema (also called **kwashiorkor**, or hunger oedema). The prevalence of each of these is assessed during a nutrition survey by recording age, measuring weight and height, and examining for oedema.

Other forms of malnutrition, such as micronutrient deficiency, are not usually assessed during a nutrition and mortality and demography survey, even though they may be very important causes of morbidity and mortality. Most micronutrient deficiency diseases do not cause stunting or wasting, and their prevalence cannot be determined from anthropometric measurements.¹⁶

Populations for Anthropometric Surveys: 6–59-month-old Children

In emergencies, wasting among children aged between six and 59 months is used as a proxy indicator for the general health and wellbeing of the entire community. This assumes that children aged between six and 59 months are the most vulnerable group in the society, at least as vulnerable as each of the other age groups. This is usually, but not always, true. In practice, this group is much easier to measure than other population groups. Young children are generally at home, the parents are usually concerned about their children and willing for them to be measured, and they are not embarrassed by (nor are there as many cultural restrictions about) taking off their clothes. Also, the equipment needed is not as cumbersome as that for older age groups.

There are a few other very basic reasons why children aged between six and 59 months are a good group to survey. First of all, policymakers are used to seeing and acting upon this type of data. There is a lot of experience with surveys of this age group, affording those using the data to make decisions the opportunity to compare the new survey with previous surveys. Furthermore, there is not yet international agreement on the anthropometric indicators and cut-off points used to assess acute malnutrition in adolescents, adults, pregnant and lactating mothers, and older people. It must be reiterated that surveys of children aged between six and 59 months are used to indicate the situation of the whole population and not just young children, and restricting data collection to this group should in no way be understood as justification for confining relief to them.

When to Measure the Nutritional Status of People over age five

Surveys including other age groups are more complex and require greater technical expertise than for children aged between six and 59 months. However, it may be appropriate

¹⁶ An important corollary of this is that if the result of the anthropometric survey does not give rise to concern, there could still be major undetected micronutrient deficiency in the population that is an underlying cause of illness and death.

to assess the nutritional status of other age groups in the following circumstances:

1. When there is a relative increase in the crude death rate (CDR) compared to zero to five years old death rates. The 0-5 death rate is generally about twice the CDR. A disproportionate increase in the CDR suggests that there is a particular problem in older age groups so that the 6-59 months age group is no longer a good indicator of the stress of the whole population.
2. When there is reasonable doubt that the nutrition status of young children reflects the whole population's nutritional situation. For example, in populations where cultural traditions give preference to the feeding of young children, or when there is a high prevalence of HIV, older adults may be more severely affected.
3. When many adults or older children present themselves to selective feeding programmes or health facilities with malnutrition.
4. When credible anecdotal reports of frequent adult or adolescent malnutrition are received.
5. When the data are required as an advocacy tool to persuade policymakers to address the needs of other age groups. Ideally, this should not be necessary.

The methods for sampling and measuring other age groups are the same as those for the 6-59 month age group described in this manual. Infants less than six-month old can be included in the survey, but there are particular difficulties related to the accuracy and precision of the measurements.

Nutrition Indices and Indicators

To determine the nutrition status of an individual, the weight, height, age, and presence of oedema are recorded. The relationship of these measurements to each other is compared to international reference standards. The nutrition surveys are designed with respect to three indices: height-for-age (HFA), weight-for-height (WFH), and weight-for-age (WFA).

Growing children get taller and the height of a child in relation to a "standard" child of the same age gives an indication of whether the growth has been normal or not. This index of growth is called *height-for-age*. Children who have a low HFA are referred to as stunted. Growth is a relatively slow process, and if a child of normal height stops growing it takes a long time for that child to fall below the cut-off point for stunting.¹⁷ For this reason, HFA is often used to indicate long-standing or chronic malnutrition. If the circumstances that led to stunting are in the past, it is possible that the current growth rate is actually normal (although this is unusual without a change in the family circumstances). Stunting may also be due to intrauterine growth retardation followed by normal postnatal growth.

A child getting taller will also gain weight if body proportions remain normal. A thin child will weigh less than a normal child of the same height. *Weight-for-height* is a measure of how thin (or overweight) the child is. Because weight gain or loss is much more responsive to the present situation, WFH is usually taken to reflect recent nutritional conditions. Being excessively thin is called wasting. It is also often termed "acute malnutrition," although individual children may have been thin for a long time. An advantage of using WFH to assess the nutritional state is that it does not involve age; in many poor populations, age is not known and is difficult to estimate reliably, especially in emergency situations.

Neither stunted nor wasted children weigh as much as normal children of the same age.

Weight-for-age is thus a composite index, which reflects both wasting and stunting, or any combination of both. In practice, about 80% of the variation in WFA is related to stunting and

¹⁷ A child who is 100% of normal growth who falters to 70% of normal will take up to half his life to fall below the usual cut-off point and be labeled as moderately stunted. Thus, a 1-year-old child who is gaining height at 70% of normal will not be designated as stunted for six months.

about 20% to wasting. It is *not* a good indication of recent nutritional stress. It is used because it is an easy measurement to take in practice, and can be used to follow individual children longitudinally in the community.

Mid-upper arm circumference (MUAC) is also sometimes measured. It directly assesses the amount of soft tissue in the arm and is another measure of thinness, like WFH. Although it is easier to measure MUAC than WFH, it is more difficult to make a precise measurement and it is not standardized for age. Nevertheless, MUAC is the index to use in the community (for screening) to identify individual children in need of referral for further assessment or treatment. Because MUAC is used in this way in the community, it is useful to know the relationship between WFH and MUAC in a particular community to establish a full nutrition programme, including screening. This is why MUAC is sometimes included in the data collected in a survey. MUAC data are often not reported or emphasized in a report, and decisions are not usually based upon these data alone.

WFH, HFA, and WFA are calculated for individuals and groups using the ENA for SMART software¹⁸. Users of this manual are not expected to have to calculate these values without the aid of a computer.

The Reference Population Curves

To assess malnutrition as determined by WFH, WFA, and HFA, individual measurements are compared to an international reference standard. The WHO Multicentre Growth Reference Study (MGRS) was undertaken between 1997 and 2003 to generate new growth curves for assessing the growth and development of infants and young children around the world¹⁹. The MGRS collected primary growth data and related information from approximately 8,500 children from widely different ethnic backgrounds and cultural settings (Brazil, Ghana, India, Norway, Oman, and the USA)¹⁹. These new growth curves (known as WHO 2006 Standards) are expected to provide a single international standard that represents the best description of physiological growth for all children from birth to five years of age and to establish the breastfed infant as the normative model for growth and development²⁰. Each survey team should have a copy of the WHO 2006 Reference Standards tables so that during the survey they can identify children who need immediate referral to a nutrition or health facility.

Z-scores and percentage of median produce slightly different estimates of the prevalence of wasting²¹. The z-score is said to be more statistically valid than the percentage of the median, and has become the standard index used in nutrition surveys. Nonetheless, there are several reasons why percentage of the median should also be included in the annex of the survey report:

- It is a better predictor of mortality (the outcome of dominant interest) than z-score.
- It is used for the admission of patients to feeding programmes, because it is a better predictor of death and directs resources where they are most efficiently used²². It is used for the admission of adolescents where WHZ cannot be used.

¹⁸ The software “Epi-Info” can also be used to calculate the nutritional variables.

¹⁹ World Health Organization. The WHO Multicentre Growth Reference Study (MGRS). 2006.
<http://www.who.int/childgrowth/mgrs/en/>

²⁰ World Health Organization. The WHO Child Growth Standards. 2006.
<http://www.who.int/childgrowth/en/>

²¹ There will always be a higher prevalence of malnutrition using WHZ than WHM. The cut-off lines cross each other at the height of a normal six-month-old child.

²² The nutritional status of adolescents can also be expressed as WHM, but not as WHZ. Using WHZ for admission leads to many more admissions than WHM; the additional admissions generally do not need to be in therapeutic feeding programmes.

- It is easier for lay people to understand and for survey managers to explain because it does not require an understanding of statistical concepts.
- It is easier to calculate with a simple calculator.

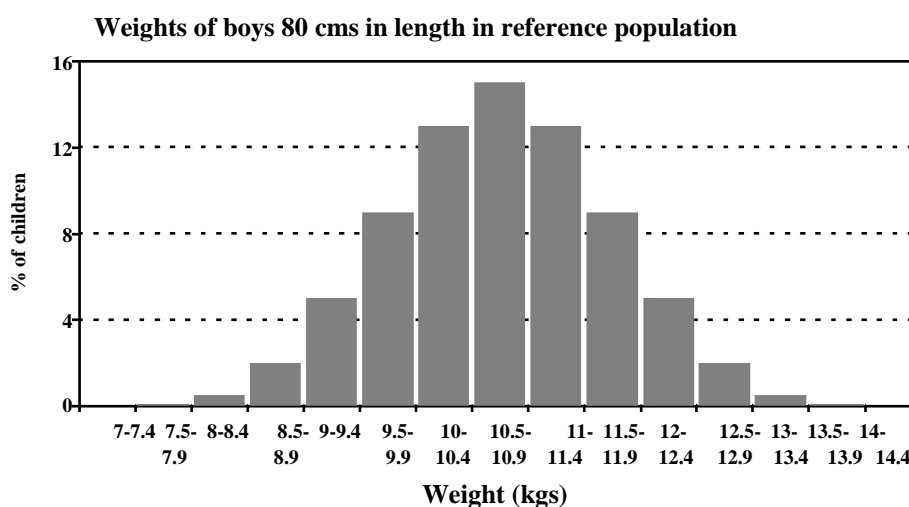
Nutritional surveys should report the prevalence of oedema and of wasting (WFH) in terms of z-score in the body of the report and include the results based on the percentage of the median as an annex.

Expression of Nutrition Indices

Anthropometric indices are generally expressed as z-scores derived from the reference standard. Details on the percentage of the median being no longer utilised can be found in SMART Manual Version 1 (April 2006).

A z-score is a measure of how far a child is from the median WFH of the reference (often written WHZ). In the reference population, all children of the same height are distributed about the median weight, some heavier and some lighter. For each height group, there is a standard deviation among the children of the reference population. This standard deviation is expressed as a certain number of kilograms at each height. The z-score of a child being measured is the number of standard deviations (of the reference population) the child is away from the median weight of the reference population in that age group. This is illustrated in figure 20 below:

Figure 20: Reference curve used to derive a z-score.



The WHZ is based upon the child's weight, the median weight of children of the same height and sex in the reference population, and the standard deviation of the distribution of weights in the reference population for children of the same height and sex.

Formula for WHZ calculation

$$\text{Weight/Height index Z-score} = \frac{\text{weight of the child} - \text{median weight of the reference population}}{\text{standard deviation of the weight in the reference population}}$$

These calculations should all be done by computer using the ENA for SMART software, but it is useful to understand the basis for the calculation.

Example z-score calculation

In a nutrition survey, a male child of 84 cm tall weighs 9.9 kg. The reference median weight for boys of height 84 cm is 11.7 kg. The standard deviation from the reference distribution for boys of height 84 cm is 0.908 kg. The z-score for this child is then as follows:

$$\text{Weight/Height index Z-score} = \frac{9.9 - 11.7}{0.908} = -1.98$$

Definitions of Acute Malnutrition in Children aged 6–59 Months

WFH is the criterion used to assess moderate and severe wasting, monitor changes in the nutrition status of the population, and make decisions on admission and discharge of individuals to and from feeding programmes.

Oedema

The presence of oedema should also be assessed during the survey. Pitting oedema on both feet (bilateral oedema) is the sign of kwashiorkor. In an emergency context, any person with bilateral oedema has severe malnutrition²³ and is classified as severely malnourished even if the WFH z-score or percent of median is normal.

Moderate, severe, and global malnutrition

Individuals are classified as normal if they have no oedema and have a WFH equal to or above -2 z-scores (or 80% of the median). If they are less than -2 z-scores (or below 80% of the median) and equal to or above -3 z-scores (or 70% of the median), they have moderate acute wasting. If they are below -3 z-scores (or 70% of the median), or if they have oedema on both feet, they have severe acute malnutrition, which is sometimes referred to as SAM (Table 8)²⁴.

Table 8: Definitions of Acute Malnutrition Using WFH and/or Oedema in Children aged 6–59 months.

Acute malnutrition using WFH	Acute malnutrition using z-scores	Acute malnutrition using oedema
Severe	< -3 z-scores	Yes/no
	> -3 z-scores	Yes
Moderate	< -2 z-scores to ≥ -3 z-scores	No
Global	< -2 z-scores	Yes/no

Global acute malnutrition (GAM) is the term used to include all children with moderate wasting, severe wasting or oedema, or any combination of these conditions.

The user of this manual will not have to make these calculations: they are done automatically using the ENA for SMART software. GAM and SAM should be presented as prevalences

²³ There are other causes of bilateral edema such as heart failure, kidney disease (nephrotic syndrome), thiamine deficiency, and pre-eclampsia in pregnant women. However, in an emergency context, most bilateral edema, especially in children, is due to kwashiorkor.

²⁴ Note that the terms “severe wasting” and “severe acute malnutrition” are not synonyms. A child with severe acute malnutrition is either severely wasted, oedematous, or both. Severe acute malnutrition is the sum of severe wasting and oedema.

expressed as a percentage of the population based on the GAM and SAM derived from the z-score calculations.

Example of SAM and GAM calculations

A group of 905 children was measured in a survey. None of the children had oedema. Fifteen children had $WHZ < -3$ z-scores and 45 had $WHZ < -2$ z-scores and ≥ -3 z-scores.

Prevalence of severe acute malnutrition (SAM) = number of severely malnourished children divided by total number of children multiplied by 100 = $(15)/905 \times 100 = 1.7\%$

Prevalence of moderate acute malnutrition = number of moderately malnourished children divided by the total number of children multiplied by 100 = $(45)/905 \times 100 = 5.0\%$

Prevalence of global acute malnutrition (GAM) = prevalence of severe acute malnutrition plus the prevalence of moderate acute malnutrition = $1.7\% + 5.0\% = 6.7\%$

SAM/GAM scenario

Another group of 910 children was measured in a survey. Six had oedema, and of these six, one had a WFH z- score < -3.0 , two had WFH z-scores between -3.0 and -2.0 , and three had WFH z-scores > -2.0 . Overall, 17 children had $WHZ < -3$ z-scores, and 55 had $WHZ < -2$ z-scores and ≥ -3 z-scores.

Prevalence of severe acute malnutrition (SAM) = number of severely malnourished children divided by total number of children multiplied by 100 = $(17+6-1) / 910 \times 100 = 2.4\%$

One of the oedematous children was also severely wasted. The child cannot be counted twice. The oedematous are added to the severely wasted, and the number is reduced by the number of children who have both oedema and severe wasting.

Prevalence of moderate acute malnutrition = number of moderately malnourished children divided by the total number of children multiplied by 100 = $(55 - 2)/910 \times 100 = 5.8\%$

Two of the moderately wasted children had oedema: they have already been included in the severely malnourished category. The children cannot be counted in both categories. The moderately malnourished are the total with moderate wasting minus those moderately wasted children who had oedema and were therefore counted with the severely malnourished.

Prevalence of global acute malnutrition (GAM) = $2.4\% + 5.8\% = 8.2\%$

Why focus on z-scores to estimate acute malnutrition?

Z-scores and percentage of median produce slightly different estimates of the prevalence of wasting²⁵. The z-score is said to be more statistically valid than the percentage of the median, and has become the standard index used in nutrition surveys.

²⁵ There will always be a higher prevalence of malnutrition using WHZ than WHM. The cut-off lines cross each other at the height of a normal 6-month-old child.

Nonetheless, there are several reasons why percentage of the median should also be reported:

- It is a better predictor of mortality (the outcome of dominant interest) than z-score.
- It is used for the admission of patients to feeding programmes, because it is a better predictor of death and directs resources where they are most efficiently used²⁶. It is used for the admission of adolescents where WHZ cannot be used.
- It is easier for lay people to understand and for survey managers to explain because it does not require an understanding of statistical concepts.
- It is easier to calculate with a simple calculator.

Nutritional surveys should always report the prevalence of oedema and of wasting (WFH) in terms of z-score. For comparison, results based on the percentage of the median for SAM and GAM can be included in the annex of the report.

Chronic malnutrition in children

The long time scale over which HFA changes makes it less useful for deciding when to intervene in an emergency. It is useful, however, for long-term planning and policy development. Although at an individual level stunting develops slowly, the degree of stunting can change within a few months when averaged over an entire population. Incorrect age data makes HFA information misleading, and reliable age data can be difficult and time-consuming to obtain. For this reason, age data are generally gathered in an emergency survey to determine if the child is appropriately included in the sample (i.e., is the child probably between six and 59 months?) or whether the sample is biased toward a particular age group, rather than to obtain accurate information about stunting.

If chronic malnutrition is the main indicator of the assessment, guidance is provided in Annex 2.

Nutritional Measurements

The caretaker should not be interviewed for any other part (i.e. Mortality and Demography) of the survey while their children's measurements are being taken. If this is attempted, it will be much more difficult to measure the children, and the caretaker will be distracted by the manipulation of their children, even if they do not cry. The caretaker will thus give less attention to providing full and accurate answers. Furthermore, if the team is too large, the caretaker and children are more likely to be intimidated or distracted.

Weighing equipment

The scale has to be light and robust. It is strongly recommended to use electronic scales over 25kg hanging spring scale marked out in increments of 0.1kg with weighing pants (known as Salter scales). Using electronic scales such as the Uniscale improves significantly the quality of weight measurements. Bathroom scales are not sufficiently accurate.

Equipment for measuring height and length

The measuring board should be at least 130cm long and made of hardwood with a hard water-resistant finish. Choice of wood is important; as the boards should be light enough to easily carry from house to house and should not warp in the rainy season²⁷. The board should have two tape measures attached to it, one on each side, and they should be marked out in 0.1cm

²⁶ The nutritional status of adolescents can also be expressed as WHM, but not as WHZ. Using WHZ for admission leads to many more admissions than WHM; and the additional admissions generally do not need to be in therapeutic feeding programmes.

²⁷ Aluminum should *not* be used, because it can get very hot in the sun, and aluminum has a sufficiently high specific heat to burn children. The tapes should not be of metal for the same reason. A wooden board is most comfortable for the child. Rigid plastic boards are acceptable but expensive.

increments. The foot-piece must be easily adjustable, remaining perpendicular to the board²⁸. The board should be easily set upright to measure height; the head piece of the length board becomes the base when the board is set upright. It must be large enough for a child to stand on it and to stabilise the whole board when it is set upright. Measuring boards are usually made by local carpenters, but there should be at least one commercial board to be used as a carpenter's template to standardise the locally made boards. All the boards are standardised with a broom handle or dowel that has been cut to measure exactly 100cm.

Children's length is measured if a child is less than two years of age or ill and cannot stand for height (see MEASURE variable in the ENA for SMART software). The height measurement is taken on all children aged two years or more.

Details on recommended anthropometric equipment can be found on the SMART Methodology website²⁹.

Estimating age

In many rural areas of the developing world, the age of children is not known. In general, the younger the child is, the more accurately you can estimate the month of birth. The following methods may be helpful if the mother does not know the child's birthday.

1. The mother may have the child's immunisation card, road-to-health card, or other written document showing the child's age or date of birth. Always ask to see the child's immunisation card.
2. If the age of a neighbor's child is known, ask whether their child was born before or after the selected child.
3. Use a "local-events calendar," which shows dates on which important events took place during the past five years. It can show local holidays, hailstorms, the opening of a nearby school or clinic and political elections, etc. Ask the mother whether the child was born before or after a certain event. In addition, the local calendar can include agricultural events that occur at the same time each year. These events can help identify which month the child was born in.

Weight

Weight should be measured to the nearest 100g (0.1kg). The scales should be checked for accuracy before and after each day's measurements, using the same known weights. Each team must use the same standard weights at the base. Standard weights do not need to be carried in the field, but the scale should measure the same in the morning and evening when the team returns from the field.

Height and length

Children's height should be measured to the nearest 0.1cm. Children aged less than two years of age are measured lying down on a horizontal measuring board. Children above 85cm are measured standing up.³⁰ Figure 21 shows how to measure the length of a child less than 85cm.

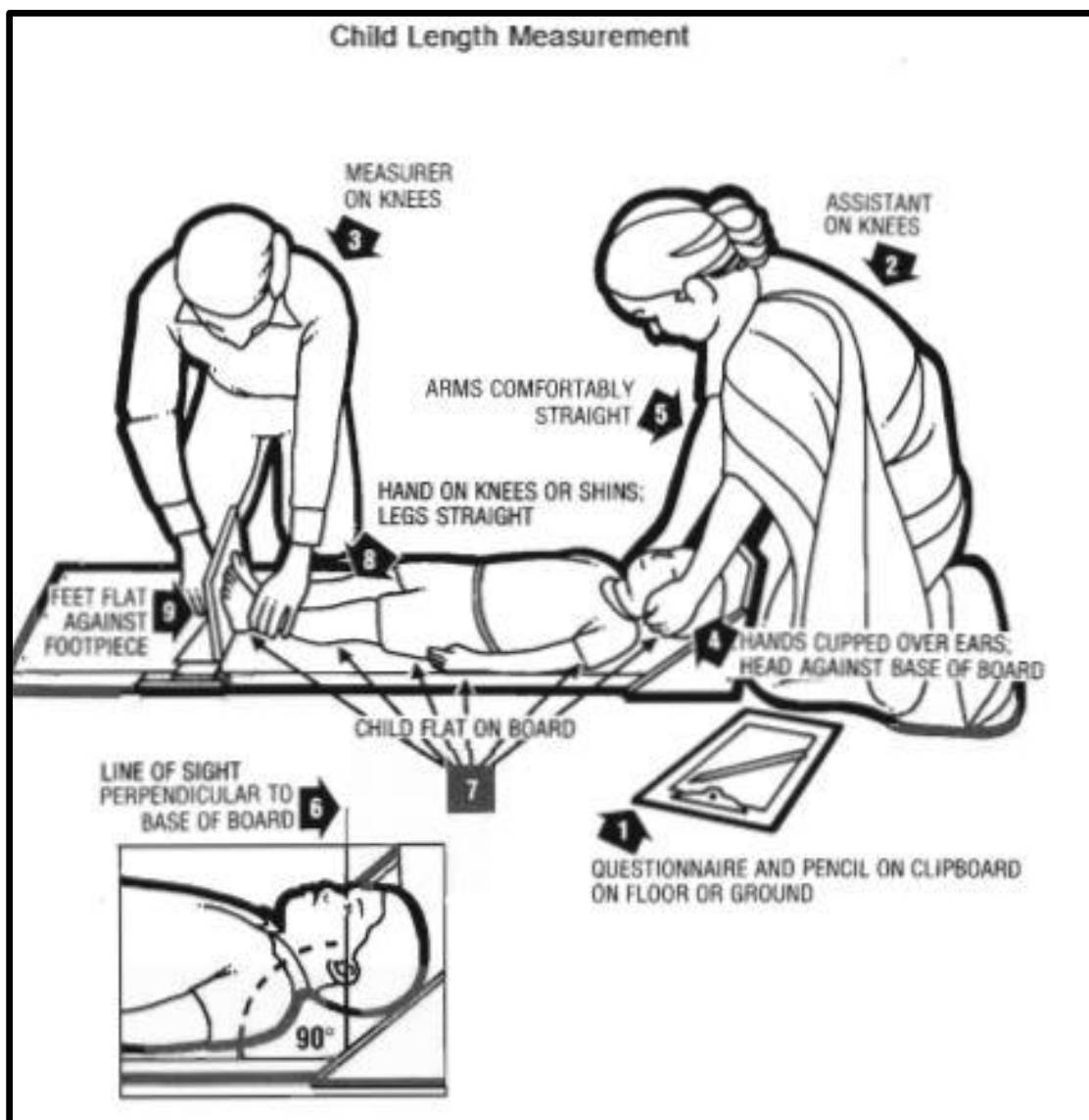
²⁸ It is best lubricated by rubbing with a candle.

²⁹ SMART Website. Recommended survey equipment. 2017.

<http://smartmethodology.org/resources/recommend-survey-equipment/>

³⁰ Some manuals suggest an age cut-off for taking length or height. This should never be used in an emergency survey. The cut-off is only based upon whether the child is above or below 85cm, no matter the age of the child.

Figure 21: How to Measure the Length of a Child Aged Less than 2 years of age.³¹



There are several steps:

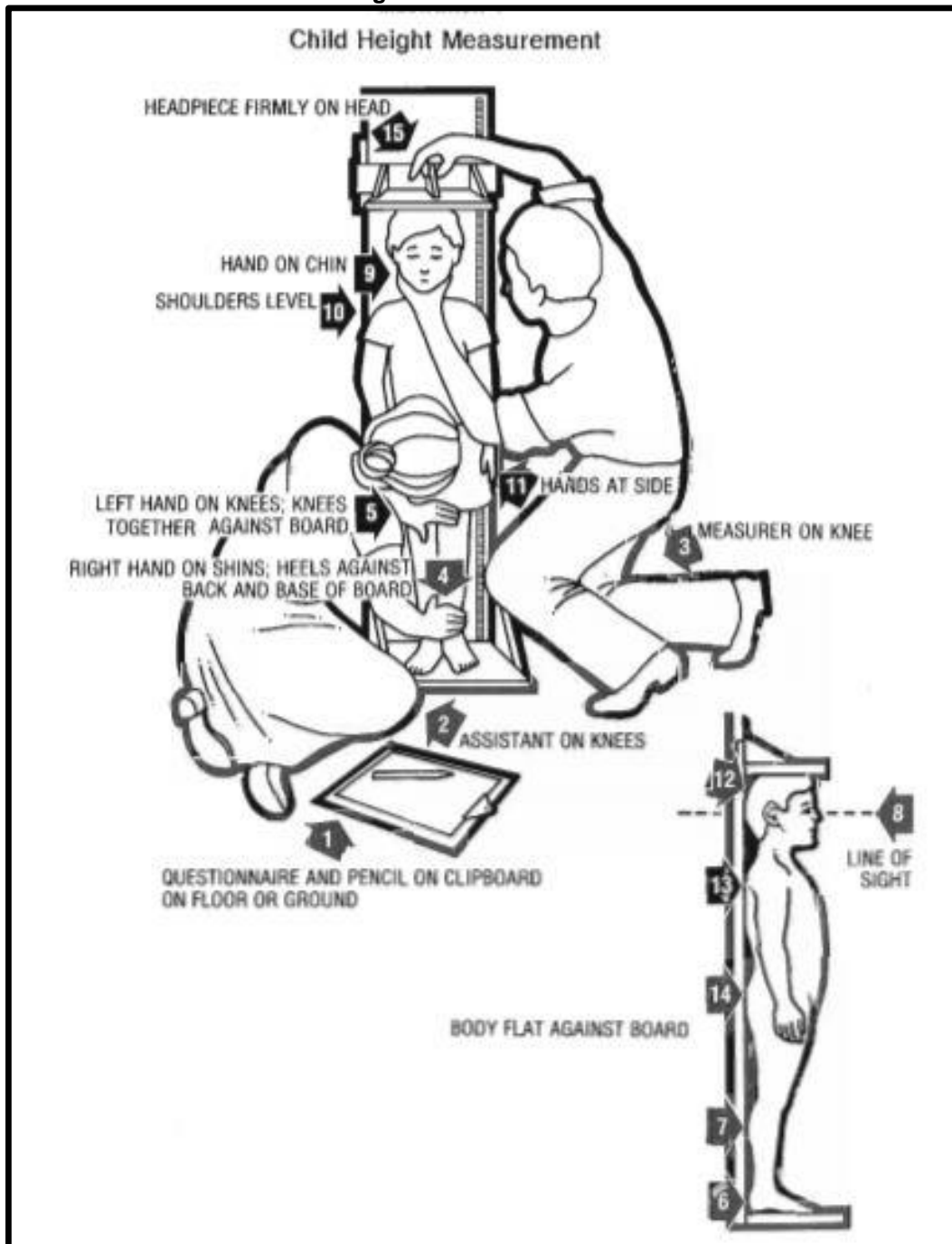
1. Explain the procedure to the child's mother or caregiver.
2. Remove the child's shoes and any hair ornament or top knot on the top of the child's head.
3. Place the child gently onto the board on his or her back, with the head against the fixed vertical part and the soles of the feet near the cursor or moving part. The child should lie straight in the middle of the board, looking directly up.
4. The assistant should hold the child's head firmly against the base of the board.
5. The measurer places one hand on the knees (to keep the legs straight), places the child's feet flat against the cursor with the other hand, and pushes the cursor against the feet firmly but gently.
6. The measurer reads and announces the length to the nearest 0.1cm.
7. The assistant repeats the measurement out loud and records it on the datasheet.

In some cultures, parents may be unhappy about measuring a child lying down, as such a

³¹ Shorr I. *How to weigh and to measure children: Assessing the nutritional status of young children in household surveys*, New York: United Nations, 1986.

procedure is used to measure bodies for a coffin. In these populations, the measurement technique is at best seen as bad luck, and it is particularly important to have local people from the same culture as the main people on the team. It is also crucial that community leaders, religious authorities, and other influential people understand why the measurements are being made and give the teams explicit authority to make the length measurements. The mother must understand this and give her permission specifically. In these circumstances, children who can stand can have their height measured instead of length, and a correction made to the measurement. Again, the height of a child is measured for children aged two years or older. Figure 22 indicates how this is done.

Figure 22: How to measure the height of a child of more than 87cm.³²



³² Shorr I. How to weigh and to measure children: Assessing the nutritional status of young children in household surveys, New York: United Nations, 1986.

Oedema

Oedema is the retention of water in the tissues of the body. To diagnose oedema, moderate finger pressure is applied just above the ankle on the inside of the leg where the shin bone is below the skin, or on the tops of the feet (Figure 23). The pressure is kept for about three seconds (if the enumerator counts “one thousand and one, one thousand and two, one thousand and three” in English, pronouncing the words carefully, this takes about three seconds). If there is oedema, an impression remains for some time (at least a few seconds) where the oedema fluid has been pressed out of the tissue. The child should only be recorded as oedematous if **both** feet clearly have oedema.

Figure 23: How to Check for Oedema.



It can be quite painful if the enumerator presses hard on the skin. Hard pressure is *not* required to test for oedema. The team should practice on each other and if anyone finds it painful or uncomfortable then the team member is pressing excessively. Oedema should be tested for after weight and height/length are measured.

Most people, including doctors, overestimate the amount of oedema in the body. The oedema can be graded as absent, mild (+), moderate (+ +) or severe (+ + +):

+	mild: both feet/ankles
+ +	moderate: both feet, plus lower legs, hands, or lower arms
+ + +	severe: generalized oedema including both feet, legs, hands, arms, and face

If the degree of oedema is recorded (graded 0 to 3), then different adjustments are made by the ENA for SMART software as indicated in Table 9³³.

Table 9: Percentage of Body Weight Accounted for by Nutritional Oedema.

Degree	Oedema as % body weight	Correction factor applied
No Oedema	0.00	weight * 1.0000
Mild (+)	2.68	weight * 0.9732
Moderate (+ +)	4.31	weight * 0.9569
Severe (+ + +)	8.38	weight * 0.9162
Weighted Mean	3.60	weight * 0.964

³³ Golden MH. The Clinical Assessment of Oedema: Implications for Feeding the Malnourished Child. *European Journal of Clinical Nutrition* 1989; 43: 581–82.

Grellety Y. Management of Severe Malnutrition in Africa (PhD Thesis). Scotland: University of Aberdeen. 2000.

5. Plausibility Checks

Introduction

What is the plausibility check and why is it important?

Surveys are expensive and are used to plan interventions. It is vital that the results of a survey are as accurate as possible; otherwise resources will be wasted and those in need may not be identified. There are many ways in which a survey can be biased or inaccurate and give erroneous results; quality control of the data is critical if these problems are to be avoided. It is not a substitute for proper selection of staff, training, planning and supervision, but it does give an indication whether these have been successful and that the data can be relied upon. The SMART plausibility check is the tool that analyses the overall quality of the survey data. It has been developed by examination of well conducted small-scale surveys and theoretical considerations. Refined over years of use in the field, the plausibility check report shows the distribution of the sample against that expected if the subjects are properly sampled, the amount of missing and implausible data, and a series of statistical tests examining the internal structure of the survey data against that which would be expected to occur in a well conducted survey. Nutrition surveys must always include the plausibility check or summary table of plausibility scores by strata.

The quality of survey data can be influenced by multiple factors, one of which is bias. Bias is a systematic error in either selection or measurement of survey subjects. Bias can be avoided with proper training of data collection teams, proper data collection methods and correct anthropometric measurement practices. The ability of the teams to conduct accurate and precise measurements should be tested at the end of the training by performing the standardisation test.

There are two types of bias:

1. **Selection bias** – A systematic or non-random error in selecting participants to be included within a study. In the case of nutritional surveys, selection bias could occur if data collectors preferentially chose to include households that were easier to access or closer geographically within a town (as opposed to using random sampling methods). The extent of this bias cannot be ascertained from the plausibility test. However, selection bias can also occur if the selected sample is not representative of the population sampled by preferentially taking older or younger children or selecting a higher proportion of boys or girls than occurs naturally in the population: this bias is tested in the plausibility test. There might also be a bias introduced by omitting eligible subjects: the plausibility test also examines the number of missing data.
2. **Measurement bias** – A systematic or non-random error in classifying the individual subjects. In the case of nutritional surveys, measurement bias could stem from improper use of measurement tools when conducting anthropometric measurements, or an improper translation of the field questionnaire. This is addressed in the plausibility test by comparing the data obtained by the different teams to determine if one team is consistently getting higher or lower values than the other teams. If it is thought that one of the teams is systematically over or under measuring the subjects, a selection of the subjects can be re-measured and the new measurements compared with the original measurements by the spreadsheet in the training module. This extra work is preferable to either excluding the biased team's data, rejecting the survey, or reporting erroneous results.

Survey data can also be affected by error excessive random error in the measurements, which is a different condition. Random error may not affect the mean value from all the subjects but has the effect of widening the spread of the data (increasing the standard deviation), so that

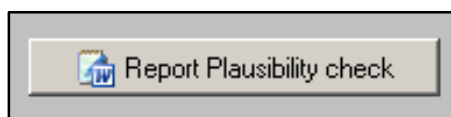
more children fall into the tails of the distribution and the numbers of both undernourished and over-nourished subjects reported is inflated.

The plausibility test uses the default values in the *Options* tab of the ENA for SMART software to compare the actual data with the expected data. These should not normally be changed unless the administrator has good evidence that the population being surveyed does not conform to the demographic profile typical of such populations. The width of the range of excluded values using the SMART flags can also be changed (but should always be within the range of 3.0 to 3.5 Z-scores). It is very important that any changes to these parameters in the *Options* tab be made before the survey is undertaken. After the data has been collected and analysis is undertaken, the *Options* tab should not be changed.

Analysis of data quality using the plausibility check can be undertaken both during and after data collection to ensure that proper survey data collection procedures have been diligently followed.

Underlying principles behind the statistical tests of the plausibility check

The Emergency Nutrition Assessment (ENA) software for SMART is the only software that offers the plausibility check. It is freely available from the SMART Methodology website³⁴. There are quite frequent updates, so it is important to download and install the latest version prior to each survey. In the *Data Entry Anthropometry* tab of the ENA software for SMART, the 'Report Plausibility Check' button generates the plausibility check report that opens in Rich Text Format.



In this chapter, explanations on the logic behind the statistical tests and a step-by-step approach on how to interpret the different sections of the plausibility check are provided.

Summary Table of the Anthropometry Plausibility Check

Layout

The layout of the plausibility check summary table (Figure 24) consists of the following sections:

1. **Name** of survey (Countrycode_town_date_NGO.as – e.g., Afg018_kabul_0614_ngo.as) and details of the reference population (default = WHO 2006 Child Growth Standards).
2. **“Criteria”** refer to the ten statistical tests that evaluate data quality and the presence of bias.
3. **“Flags”** refer to outliers, and whether or not SMART flagged data points are included (*Incl*) or excluded (*Excl*) in the analysis of each of the ten statistical tests.
4. **“Unit”** refers to the unit of measurement of each test.
5. **“Excellent, Good, Acceptable and Problematic”** give the ranges of acceptability for the score of each test.
6. **“Score”** refers to the number of penalty points for each test depending on which of the acceptability ranges the survey data falls in.
7. **“Overall score”** refers to the total amount of penalty points for the survey, and whether or not this is acceptable.

³⁴<http://www.smartmethodology.org/>

Figure 24: Summary Table of the Anthropometry Plausibility Check in ENA for SMART.

1 **Plausibility check for: _Anthro demo.as**

Standard/Reference used for z-score calculation: WHO standards 2006
 (If it is not mentioned, flagged data is included in the evaluation. Some parts of this plausibility report are more for advanced users and can be skipped for a standard evaluation)

Overall data quality

2

Criteria	Flags*	Unit	3	4	5	6	Score
Flagged data (% of out of range subjects)	Incl	%	0-2.5 0	>2.5-5.0 5	>5.0-7.5 10	>7.5 20	5 (2.7 %)
Overall Sex ratio (Significant chi square)	Incl	p	>0.1 0	>0.05 2	>0.001 4	<=0.001 10	0 (p=0.418)
Age ratio(6-29 vs 30-59) (Significant chi square)	Incl	p	>0.1 0	>0.05 2	>0.001 4	<=0.001 10	0 (p=0.802)
Dig pref score - weight	Incl	#	0-7 0	8-12 2	13-20 4	> 20 10	0 (4)
Dig pref score - height	Incl	#	0-7 0	8-12 2	13-20 4	> 20 10	2 (11)
Dig pref score - MUAC	Incl	#	0-7 0	8-12 2	13-20 4	> 20 10	2 (11)
Standard Dev WHZ .	Excl	SD	<1.1 and 0	<1.15 and 5	<1.20 and 10	>=1.20 or <=0.80 20	5 (1.11)
Skewness WHZ	Excl	#	<±0.2 0	<±0.4 1	<±0.6 3	>=±0.6 5	1 (-0.25)
Kurtosis WHZ	Excl	#	<±0.2 0	<±0.4 1	<±0.6 3	>=±0.6 5	1 (-0.25)
Poisson dist WHZ-2	Excl	p	>0.05 0	>0.01 1	>0.001 3	<=0.001 5	0 (p=0.239)
OVERALL SCORE WHZ =			0-9	10-14	15-24	>25	16

7 **The overall score of this survey is 16, this is acceptable.**

Criteria of the summary table

Each of the ten statistical tests (criteria) evaluates an aspect of data quality and presence of bias based on the demographic distribution of the sample (overall sex ratio and overall age distribution), the precision of the anthropometric measurements and the Weight-for-Height (WHZ) distribution. The overall summary table does not include evaluation of Height-for-Age (HAZ), Weight-for-Age (WAZ) or MUAC distributions or mortality data evaluation. The body of the report contains much more detailed analysis of all the anthropometric indices and their distributions, as well as calculation of the differences between the prevalence obtained by calculation (probit analysis) and by counting. The data are then analysed by each team to examine whether there is a bias and by the order within the clusters.

Interpreting the overall score of the summary table

Penalty points awarded are based on acceptable ranges generated for each of the test criteria and are scaled according to the overall effect of an error in the parameter on the survey result.

These penalty points are meant to highlight areas of the data set where survey managers should focus their detective efforts; each individual criterion should be examined in more detail in order to see why these penalty points were awarded, whether they can be justified based on a particular field context, whether the problem lies with one of the teams only, and whether the errors have a significant effect upon the reported prevalence. Only after careful consideration of each criterion along with the survey context can it be determined whether a survey with a high penalty score may actually be acceptable, or whether a survey with a low penalty score requires further investigation.

Key Message

The overall score of the summary table should be used as an indication for further scrutiny of the data and highlight the areas that may need more detailed analysis, and **not** used by itself as the primary criterion to validate or invalidate a survey's results!

Sections of the Anthropometry Plausibility Check Report

Details of the statistical tests from the summary table

❖ Criterion 1: Flagged Data.

Flagged data (% of out of range subjects)	Incl	%	0-2.5	>2.5-5.0	>5.0-7.5	>7.5	
			0	5	10	20	5 (2.7 %)

This test calculates the proportion (%) of flagged data present in the survey data. Flags refer to outliers, extreme values that are so far from the mean that they are unlikely to be correct measurements. The errors are always identified based on SMART exclusion cut-offs which are more stringent than other ways of flagging erroneous data; however, on average, about one in 1,000 measurements excluded may be removed incorrectly (this will have a negligible effect upon the results – there is no penalty until there are 25 flagged data per 1000 measurements). The WHO exclusion cut-offs are based upon data points that are so extreme as to be biologically impossible (see *Options* tab in the ENA for SMART software); less extreme errors are not excluded using WHO flags. It should be noted that with perfect quality of measurements, the results with no flags, WHO flags and SMART flags should all be identical or almost identical. The magnitude of the difference between the three is an indication of the quality of the data collected.

SMART flags (cut-offs) are based on statistical plausibility, and exclude all values outside of ± 3 Z-scores from the mean of the surveyed population. In measurements from a surveyed population form a normal distribution (bell-shaped curve) with a SD equal to 1, based on statistical principles, 99.7% of observations should lie within ± 3 Z-scores from the mean. The upper and lower cut-off points for SMART flags will be different in each survey since the survey mean is used as the reference point for SMART flags (surveyed children are compared to their own population when using SMART flags).

Table 10 below lists the average number of subjects that have to be measured before one subject would be erroneously excluded from the data set when SMART flags are imposed on the data -- depending on exclusion range from the survey mean (varying from ± 3 to ± 5) and the standard deviation of the sample Z-score (varying from 0.8 to 1.2).

Table 10: Sample Size where one Erroneous Flag (Real Value) is Expected with Various SMART Flag Exclusion Ranges and Various Standard Deviations (SD) of Z-score Distribution.

Flag CUT-OFF	SD=0.8 z-score	SD=0.9 z-score	SD= 1.0 z-score	SD=1.1 z-score	SD=1.2 z-score
± 3.00	>10,000	>2,000	741	313	161
± 3.05	>10,000	>2,000	874	360	181
± 3.10	>10,000	>2,000	1,033	414	204
± 3.15	>10,000	>2,000	1,225	478	231
± 3.20	>10,000	>5,000	1,455	552	261
± 3.25	>10,000	>5,000	1,733	639	296
± 3.30	>50,000	>5,000	>2,000	741	336
± 3.35	>50,000	>10,000	>2,000	861	381
± 3.40	>50,000	>10,000	>2,000	1,002	434
± 3.45	>100,000	>10,000	>2,000	1,169	495
<u>MAX ± -3.50</u>	>100,000	>10,000	>2,000	1,367	565
± 3.55	>200,000	>10,000	>5,000	1,600	647
± 3.60	>200,000	>10,000	>5,000	1,878	741
± 3.65	>300,000	>10,000	>5,000	>2,000	850
± 3.70	>500,000	>50,000	>5,000	>2,000	977
± 3.75	>500,000	>50,000	>10,000	>2,000	1,125
± 3.80	>500,000	>50,000	>10,000	>2,000	1,297
± 3.85	>1,000,000	>100,000	>10,000	>2,000	1,498
± 3.90	>1,000,000	>100,000	>10,000	>5,000	1,733
± 3.95	>1,000,000	>100,000	>10,000	>5,000	>2,000
± 4.00	>1,000,000	>200,000	>10,000	>5,000	>2,000
<u>WHO ± 5.00</u>	>1,000,000	>1,000,000	>1,000,000	>300,000	>50,000

It is clear that as the standard deviation of the distribution increases more subjects will be erroneously excluded from the analysis when SMART flags are applied. Normally, we expect the SD of the sample to be around 1.0 in which case 1: 741 would be erroneously excluded if the cut-off point for SMART flags is set at ± 3.0 Z. If the SD is less than 1.0, then it is unlikely that any subjects would be erroneously excluded. With an SD of 1.2, many more children would be excluded – but SDs approaching 1.2 are only likely to occur if random errors of measurement are substantial (i.e. the quality of measurements is seriously compromised). Note that even a very small increase in either the bounds of the cut-off point or the SD of the distribution can have a dramatic effect on the number of potential erroneous exclusions.

It is important to note that the plausibility check should always be generated using exclusion range ± 3 Z-scores to make sure that quality criteria cut-offs are the same across all surveys evaluated by the ENA for SMART software. When reporting survey results, exclusion range may be relaxed to be wider than ± 3 only if the overall quality of the survey is high based on the plausibility check (e.g. if the percentage of flags is below 1% and the observed SD after

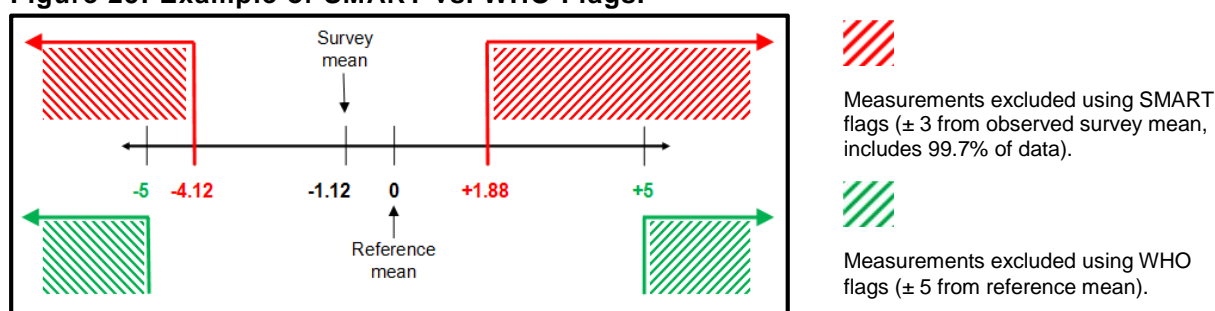
exclusion of SMART flags is below 1.1). In such situations when data quality is high, then the exclusion range can be relaxed to ± 3.5 Z-scores. As seen in the table above, it is very unlikely that real values would be excluded from the analysis if SD is 1.1 or less.

The exclusion range used to produce survey estimates should always be specified in the survey report. Whatever the bounds of the cut-off entered into the *Options* tab, there will be a relationship between the SD of the survey and the number of flags identified.

The cut-off point used for the WHO flags Weight-for-Height is ± 5.0 Z. It is clear from the last row of the table that when using WHO flags the chances of excluding a true value are negligible, even with quite a large SD for the distribution; on the other hand it is very likely that using these flags will result in inclusion of erroneous (statistically implausible) values in the analysis.

In an example in Figure 25 below, the survey mean is -1.12 and therefore the range of acceptability based on SMART is (-4.12; +1.88). The analysis of the plausibility check (each test) will be performed on the values that lie inside this range. However, data (measurements) that lie outside the SMART flag cut-offs should never be manually removed from the dataset (i.e., from the *Data Entry Anthropometry module* of ENA): they will be automatically excluded during the analysis depending upon the flags chosen and the range set in the *Options* tab. Anyone who has access to the original dataset must be able to re-examine all of the data (measurements), including those that are excluded when using SMART flags. **Note:** if outlying data are removed in an effort to improve a poorly collected data the distribution of the data then becomes abnormal with truncated tails and this becomes obvious when other tests (e.g. kurtosis) are considered.

Figure 25: Example of SMART vs. WHO Flags.



WHO flags (cut-offs) are based on biological plausibility, and exclude all values outside of ± 5 Z-scores from the mean of the reference population, which is always 0 Z. The upper and lower ranges for WHO flags will be the same for each survey (+5; -5) since the same reference mean of zero is used (surveyed children are compared to the WHO 2006 Child Growth Standards³⁵ when using WHO flags).

Interpretation

A high proportion of flagged data usually means that measurements collected by at least one (or more) survey teams have been poorly taken or recorded. Up to a maximum of twenty penalty points (for datasets having >7.5% flagged data) are awarded for this test. In the example summary table, 2.7% of the dataset is classified as being flagged using SMART flags, and consequently 5 penalty points have been assigned for this criterion.

Details of the location (line number in the data table, ID number if available) for each of the flagged values based on the cut-offs set in the *Options* tab (default ± 3 Z) from the surveyed population's mean are provided for the following anthropometric indices: Weight-for-Height (WHZ), Height-for-Age (HAZ), and Weight-for-Age (WAZ) (Figure 26).

³⁵<http://www.who.int/childgrowth>

Figure 26: Location of Flagged Values.

Anthropometric Indices likely to be in error (-3 to 3 for WHZ, -3 to 3 for HAZ, -3 to 3 for WAZ, from observed mean - chosen in Options panel - these values will be flagged and should be excluded from analysis for a nutrition survey in emergencies. For other surveys this might not be the best procedure e.g. when the percentage of overweight children has to be calculated):

(...)

Line=187/ID=35: HAZ (-5.463), Age may be incorrect
 Line=218/ID=: WHZ (-4.296), Weight may be incorrect
 Line=232/ID=: WHZ (-8.451), HAZ (3.048), WAZ (-5.455)
 Line=263/ID=: WHZ (1.931), Height may be incorrect
 Line=265/ID=: WHZ (-6.048), HAZ (7.650), Height may be incorrect
 Line=285/ID=: WHZ (2.564), WAZ (1.404), Weight may be incorrect
 Line=304/ID=99: HAZ (5.787), WAZ (1.515), Age may be incorrect
 Line=311/ID=: WHZ (2.841), Weight may be incorrect
 Line=334/ID=: HAZ (1.481), WAZ (1.011), Age may be incorrect
 Line=341/ID=: WHZ (-4.552), Weight may be incorrect
 Line=348/ID=: HAZ (1.600), WAZ (1.346), Age may be incorrect
 Line=442/ID=: HAZ (1.301), Age may be incorrect
 Line=443/ID=: HAZ (2.743), WAZ (1.655), Age may be incorrect

Percentage of values flagged with SMART flags: WHZ: 2.7 %, HAZ: 9.3 %, WAZ: 4.1 %

Each of the records listed in the flagged data section should be double-checked with the original copy of the paper questionnaire, if applicable, to ensure there are no data entry mistakes.

❖ **Criteria 2 & 3: Overall Sex Ratio and Overall Age Distribution.**

Overall Sex ratio (Significant chi square)	Incl	p	>0.1	>0.05	>0.001	<=0.001	0 (p=0.418)
			0	2	4	10	
Age ratio(6-29 vs 30-59) (Significant chi square)	Incl	p	>0.1	>0.05	>0.001	<=0.001	0 (p=0.802)
			0	2	4	10	

These two tests assess the representativeness of the survey sample (selection bias) with respect to the expected age/sex distribution of the child population. The default values in the *Options* tab have been set assuming equal numbers of boys and girls, a steady population size and birth rate, a 0-5 year-old death rate of about 1:10,000/d, which is typical of most developing countries, and an exponential reduction in mortality with age from zero to five years. The default values should be used unless there has been a major change in birth rate (baby boom) or the 0-5 death rate has been substantially greater or less than 1:10,000/d (or the crude death rate has deviated greatly from 0.5:10,000/d) during at least most of the preceding five years (a change in the previous six months should not lead to a change in the default values).

Flagged values are included in this test in order to identify if different proportions in terms of sex and age are observed in the entire sample, irrespective whether or not measurement errors have been made. If these tests are significant (the Chi-square test shows a p-value <0.05), this may indicate either 1) a sampling problem or 2) a substantial deviation from normal demographics for this population (birth, death and differential migration rates).

❖ Age ratio

Age ratio of 6-29 months to 30-59 months: 0.83 (The value should be around 0.85).:
p-value = 0.802 (as expected)

The expected age ratio of children six to 29 months of age to children 30 to 59 months of age should be around 0.85. This value represents the expected ratio of children in each age range based upon research compiled by Michael Golden, as explained in Criteria 2 & 3 above. This ratio takes the expected number of children from six to 29 months of age and divides it by the sum of the expected number of children 30 to 59 months of age based on the number of months³⁶ (see *Options* tab of the ENA for SMART software). If the age ratio is far from 0.85, then there was an imbalanced representation of younger (six to 29 months) children versus older children (30 to 59 months) in the dataset. The overall age distribution statistical test provides further details on which age category was misrepresented.

The “Statistical evaluation of sex and age ratios” table (figure 27) consists of the following sections:

1. **Age categories** with the same age ranges as seen in the *Options* tab of ENA software.
2. **Number of months** in each age category (six or 12) and the total months (54).
3. **Observed number/expected number of children in the whole survey**, by sex (boys, girls). The total expected number is based on the assumption that the sex ratio should be approximately 1:1 (50% male - 50% female)³⁷. In the survey example below where a total of 441 children were observed, 220.5 children for each sex would be expected; however, 229 boys and 212 girls were actually sampled (measured).
4. **Observed number/expected number of children in each age category**, by sex and overall (total). The expected number of children in each age category is based on the actual number of measured children in the survey (for further explanation see next paragraph on calculation for expected proportions). In the survey example above 53.1 boys were expected in the age category of six to 17 months; however, only 39 boys were actually observed (measured) in this age category. Overall 102.3 children six to 17 months were expected in this age category, but only 80 were actually sampled (measured).
5. **Ratio of the observed /expected number** provided in brackets (). This ratio should always be close to 1.
6. **Ratio of boys/girls** provided for each age category and overall (total) which should be close to 1.

³⁶Age ratio of 6-29 months/30-59 months = (proportion 6-17*12 months/1000 + proportion 18-29*12 months/1000) / (proportion 30-41*12 months/1000 + proportion 42-53*12 months/1000 + proportion 54-59 *6 months/1000) thus:

Age ratio of 6-29 months/30-59 months = (12 + 11.7) / (11.34 + 11.16 + 5.52) = 0.8458 ≈ **0.85**.

³⁷At birth, the sex ratio is such that there are slightly more boys than girls– 51 to 53%. There is a higher mortality for boys so that by the time the children reach six months, the ratio falls to about 50-52% boys. Within the age groups, the sample sizes are not sufficiently large to give values that are significantly different from equality.

Figure 27: Statistical Evaluation of Sex and Age Ratios Table.

Statistical evaluation of sex and age ratios (using Chi squared statistic):

Age cat.	mo.	4 boys	girls	total	ratio boys/girls
6 to 17	12	39/53.1 (0.7)	41/49.2 (0.8)	80/102.3 (0.8)	0.95
18 to 29	12	70/51.8 (1.4)	50/48.0 (1.0)	120/99.8 (1.2)	1.40
30 to 41	12	61/50.2 (1.2)	48/46.5 (1.0)	109/96.7 (1.1)	1.27
42 to 53	12	37/49.4 (0.7)	48/45.7 (1.0)	85/95.2 (0.9)	0.77
54 to 59	6	22/24.4 (0.9)	25/22.6 (1.1)	47/47.1 (1.0)	0.88
6 to 59	54	229/220.5 (1.0)	212/220.5 (1.0)		1.08

1 2 3 5 6

The data are expressed as observed number/expected number (ratio of obs/expect)

Calculations for expected proportions by age category

The following values (figure 28) are the expected proportions of children in each age range based upon research compiled by Michael Golden (as seen in the *Options* tab of the ENA for SMART software):

Figure 28: Expected Proportions of Children by Age Group.

Age groupes (months)

Expected proportion of children for plausibility check of sampling

	male	female
6 - 17	1000	1000
18 - 29	975	975
30 - 41	945	945
42 - 53	930	930
54 - 59	920	920

In order to derive the expected number of children for each age category for a specific survey based on the proportions in the *Options* tab (Figure 28), the product of the expected proportion of children in each age category and the number of months for that age category is divided by the total (for all age categories).

Example

The total number of children is $(1 \cdot 12) + (0.975 \cdot 12) + (0.945 \cdot 12) + (0.930 \cdot 12) + (0.920 \cdot 6) = 51.72$

6 – 17 months: $1.00 \cdot 12 / \text{Total} (\approx 23.2\% (0.232 \text{ if expected as a simple proportion}))$

18 – 29 months: $0.975 \cdot 12 / 51.72 \approx 22.6\% (0.226)$

30 – 41 months: $0.945 \cdot 12 / 51.72 \approx 21.9\% (0.219)$

42 – 53 months: $0.930 \cdot 12 / 51.72 \approx 21.6\% (0.216)$

54 – 59 months: $0.920 \cdot 6 / 51.72 \approx 10.7\% (0.107)$

Based on the example above, the total number of children measured was 441 (229 boys and 221 girls). Therefore, the expected proportions for this survey are the following:

6 – 17 months: $0.232 \times 441 = 102.3$ expected children for that age category, and for each sex: 53.1 (0.232×229) for boys, 49.2 (0.232×212) for girls.

18 – 29 months: $0.226 \times 441 = 99.8$ expected children for that age category, and for each sex: 51.8 (0.226×229) for boys, 48.0 (0.226×212) for girls.

30 – 41 months: $0.219 \times 441 = 96.7$ expected children for that age category, and for each sex: 50.2 (0.219×229) for boys, 46.5 (0.219×212) for girls.

42 – 53 months: $0.216 \times 441 = 95.158 \approx 95.2$ expected children for that age category, and for each sex: 49.4 (0.216×229) for boys, 45.7 (0.216×212) for girls.

54 – 59 months: $0.107 \times 441 = 47.067 \approx 47.1$ expected children for that age category, and for each sex: 24.4 (0.107×229) for boys, 22.6 (0.107×212) for girls.

Interpretation

The overall sex ratio and expected age ratio of children six to 29 to 30 to 59 months of age are evaluated using a Chi-squared test. If either of the chi-squared tests for overall sex ratio or expected ratio of younger (six to 29 months) versus older (30 to 59 months) children is significant ($p < 0.05$), then the variation in the sex ratio and/or expected age ratio is not due to random differences in sampling, and potential selection bias took place during the survey. Teams should be interviewed to try to determine the cause of this deviation from the expected ratio. In the examples below, neither the sex ratio ($p = 0.418$) nor the age ratio of six to 29 months to 30 to 59 months ($p = 0.802$) show significant differences.

Overall sex ratio: p-value = 0.418 (boys and girls equally represented)

**Age ratio of 6-29 months to 30-59 months: 0.83 (The value should be around 0.85).:
p-value = 0.802 (as expected)**

Four additional Chi-squared test results below the sex/age table provide further information on sex ratio and age distribution of the sample:

- Overall age distribution based on the observed number/expected number of children in each age category;
- Overall age distribution for boys and for girls - whether the observed boys or girls age distribution matches with the expected age distribution, and;
- Overall sex/age distribution - whether the overall age and sex distribution matches with the expected values).

**Overall age distribution: p-value = 0.020 (significant difference)
Overall age distribution for boys: p-value = 0.003 (significant difference)
Overall age distribution for girls: p-value = 0.762 (as expected)
Overall sex/age distribution: p-value = 0.001 (significant difference)**

❖ **Criteria 4, 5 & 6: Digit Preference (weight, height, mid-upper arm circumference - MUAC).**

Dig pref score - weight	Incl	#	0-7	8-12	13-20	> 20	
			0	2	4	10	0 (4)
Dig pref score - height	Incl	#	0-7	8-12	13-20	> 20	
			0	2	4	10	2 (11)
Dig pref score - MUAC	Incl	#	0-7	8-12	13-20	> 20	
			0	2	4	10	2 (11)

A common mistake when recording measurements is to round to the nearest whole number, or as seen with weight and height, round the first decimal point to zero or five. The digit preference test assesses the last digits of each anthropometric measurement to see if there has been rounding (digit preference). Note that flagged data are included in this test in order to assess the entire dataset.

Excess digit preference can have a large effect upon the result, but a small amount of rounding will not affect the survey conclusions, even though it may introduce noise to the data.

Weight and height measurements are taken to one decimal point, and MUAC measurements are taken to the nearest millimetre. A random or roughly even distribution is expected for the first decimal of weight and height measurements (kg and cm, respectively), and the last whole number of MUAC measurements (mm). In order to identify any rounding of measurements, the digit preference test is performed on each basic measurement. The test used is derived from the WHO MONICA³⁸ study of blood pressure. The formula de-sensitises the test so that trivial degrees of digit preference do not alter the results of the survey. If the numbers are truly random then each terminal digit (from 0 to 9) should occur in approximately one tenth of the observations.

In the plausibility test output each hash tag represents the proportion of measurements recorded with that particular number as the terminal digit. If there has been no rounding, then there should be a similar number of hash tags for each of the ten digits (0 through 9) since the probability of having each of the last (terminal) digits is equal.

Digit preference Weight:	
Digit .0	: #####
Digit .1	: #####
Digit .2	: #####
Digit .3	: #####
Digit .4	: #####
Digit .5	: #####
Digit .6	: #####
Digit .7	: #####
Digit .8	: #####
Digit .9	: #####
Digit preference score: 4 (0-7 excellent, 8-12 good, 13-20 acceptable and > 20 problematic)	
p-value for chi2: 0.637	

³⁸<http://www.ktl.fi/publications/monica/bp/bpqa.htm>

A simple Chi-squared test of the observed frequencies against the expected frequencies is performed to evaluate this criterion, and a Digit Preference Score (DPS) is then computed using the following formula:

$$\text{DPS} = 100 * (\chi^2 / (\text{df} * \text{N}))^{1/2}$$

Where **N** is the number of observations (subjects in the survey), χ^2 is the Chi-square statistic for the test of homogeneity of the terminal digits, and **df** are the degrees of freedom (i.e., df = 9 because there are 10 possible terminal digits). The DPS ranges from 0 to 100. It is low when there is no digit preference and high when the digit preference becomes large enough to affect the result of the survey. Proper training in anthropometric measurements is essential in eliminating high digit preference scores.

Interpretation

A high score often results from inadequate training and supervision. The DPS for weight should be close to zero when using digital scales if the child remains reasonably still during the measurement. It may be higher using hanging scales. The DPS for height and MUAC are usually higher than for weight. If there is an excess of numbers other than 0 and 5, then it is possible that the data have been constructed fictitiously (particularly if there is an excess of numbers such as 6 or 2). In this case a runs test can be conducted (to be added to the analysis and output shortly) in which the order of terminal digits is analysed to determine if there is a paucity of consecutive terminal numbers being identical. This test is used by actuaries to look for fictitious data. Penalty points for DPS can be avoided if the training, standardisation test, and supervision in the field are properly conducted.

❖ **Criterion 7: Standard Deviation.**

Standard Dev WHZ	Excl	SD	<1.1	<1.15	<1.20	>=1.20	
-			and	and	and	or	
-	Excl	SD	>0.9	>0.85	>0.80	<=0.80	
			0	5	10	20	5 (1.11)

The standard deviation (SD) for WHZ criteria and the test for flagged data (criterion 1) are the two most important statistical tests in the plausibility check, and both carry a maximum of 20 penalty points.

The standard deviation increases with random error in the measurements, that is, if the measurements are taken without due care and attention, if the child is upset, struggling or apprehensive, and if the teams are tired, rushed, inexperienced, have poor eyesight or are improperly trained and supervised. The test for SD provides information about the quality of actual measurements that have been taken. When there is random error imposed upon test data the distribution widens, the mean remains approximately the same, and there are more children in the tails of the distribution inflating both the prevalence of malnutrition and overweight.

Thus, when SD is high, the survey results (e.g. prevalence of GAM) may be overestimated. Flagged data are excluded for this test because big mistakes (flags) have already been penalised for in the flagged data test (criterion 1). If flagged data were included in this test, then the SD would always be large if there are many flags. As flags (outliers) are excluded, SD of the distribution decreases.

The “Evaluation of Standard Deviation” table (Figure 29) consists of the following sections:

- 1. Anthropometric index for which z-scores are evaluated.** This criterion evaluates the prevalence of GAM solely based on wasting (WHZ). Children with oedema are excluded from this analysis. Similar calculations are produced for both HAZ and WAZ in the plausibility check report, but are not included in the evaluation of the survey quality and are not assigned penalty points in the summary table at the present time.
- 2. Exclusion criteria.** The “No exclusion” column is based on the analysis of the entire dataset without exclusion of flagged cases; “Exclusion from reference mean (WHO flags)”

excludes from analysis all WHZ values outside of ± 5 Z-scores from the mean of the reference population (0 Z). Exclusion ranges for WHO flags are slightly different for HAZ (+/-6) and WAZ (-5 to +6). “*Exclusion from observed mean (SMART flags)*” excludes from analysis all values outside ± 3 Z-scores (default – see above for effect of changing the SMART flag options) from the mean of the surveyed population; in this case the numerical flag boundaries are different for each survey depending upon the survey’s mean value.

3. **Standard Deviation (SD)** for each of the exclusion criteria. The range of acceptable SD is between 0.8 and 1.2. Penalty points for this criterion are based on SD for WHZ excluding SMART flags (see summary table); the other columns are given for comparison and to demonstrate the effect of data cleaning. Very well conducted surveys should have approximately the same prevalence for each of the 3 columns.
4. **Prevalence observed** counts the number of cases that fall under WHZ < -2 divided by the total number of children for each of the exclusion criteria.
5. **Prevalence calculated.** The ENA for SMART software takes the observed mean and standard deviation (current SD and SD = 1Z) based on the survey data itself and, using the probit function, assumes a normal distribution and calculates the proportion of the area-under-the-curve < -2 WHZ (to the left of < -2 Z). Each curve will be different because each mean and SD will be different depending on the survey data.

For “*Prevalence calculated with current SD*”:

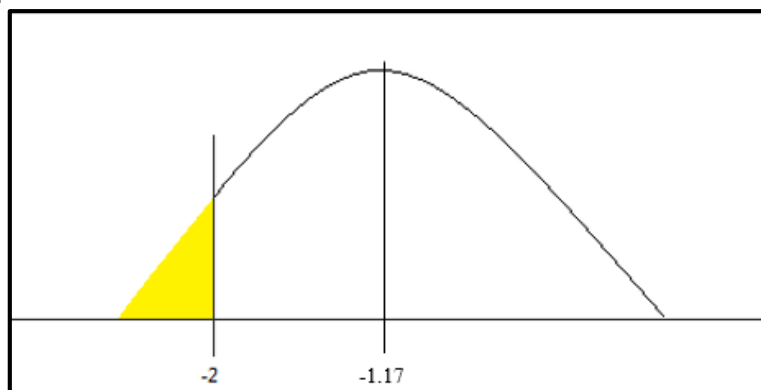
- a. No exclusion curve: mean of -1.12 (see above, or also *Results* tab) with SD 1.55, calculated prevalence of wasting is 28.5%.
- b. WHO Flags: mean of -1.15 with SD 1.20, calculated prevalence of wasting is 24.0%.
- c. SMART Flags: mean of -1.17 with SD 1.11, calculated prevalence of wasting is 22.8%.

Figure 29: Evaluation of Standard Deviation Table.

1	2	no exclusion	exclusion from reference mean (WHO flags)	exclusion from observed mean (SMART flags)
3	WHZ			
3	Standard Deviation SD: (The SD should be between 0.8 and 1.2)	1.55	1.20	1.11
4	Prevalence (< -2) observed:	23.9%	23.6%	23.4%
5	calculated with current SD:	28.5%	24.0%	22.8%
	calculated with a SD of 1:	18.9%	19.7%	20.3%

In the example illustrated in Figure 30, 22.8% of the curve is to the left of -2 WHZ (prevalence of GAM calculated with current SD is 22.8%).

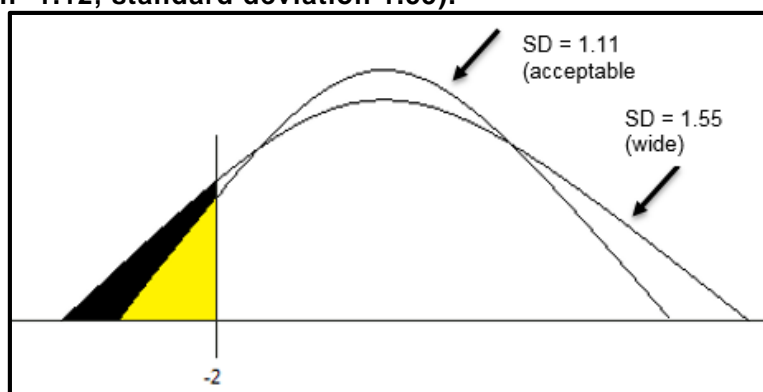
Figure 30: Distribution of WHZ with Good Standard Deviation (mean -1.17, standard deviation 1.11).



Standard deviation is a measure of the dispersion of data around the mean, and is the most important statistic that tells about the quality of the data. Based on the WHO Technical Report³⁹, the standard deviation for Weight-for-Height should be less than 1.1 in high quality data. In a similar study, data from over 200 small-scale surveys was nearly always normally distributed and had standard deviations between 0.8 and 1.2⁴⁰. Both of these analyses were based on the NCHS Standards. The range of acceptable standard deviations for SMART surveys is 0.8 to 1.2.

As data quality decreases (i.e. mistakes in measurements are introduced), the distribution of data around the mean “melts” (gets wider), and SD increases (see Figure 31). If the SD for WHZ is very high, then the prevalence of GAM may be overestimated due to the fact that a larger proportion of the area under the curve would fall below WHZ < -2.

Figure 31: Effects of a "Melting" Curve on the Prevalence of GAM with a Large Standard Deviation (mean -1.12, standard deviation 1.55).



In this example (Figure 31), 28.5% of the curve (black and yellow areas combined) is to the left of -2 WHZ - prevalence of GAM calculated with current SD is 28.5%. Prevalence is likely overestimated due to high standard deviation.

As mentioned above, the acceptable range for standard deviation is between 0.8 and 1.2 (1.0 is ideal). If the SD excluding SMART flags is above 1.2, there is concern that the observed prevalence is an over-estimation of the true prevalence, and that the actual prevalence of wasting would be closer to the “Prevalence calculated with a SD of 1”. “Prevalence calculated with SD of 1” assumes that the data are normally distributed, and a distribution is drawn using the survey mean and a standard deviation of one. When SD = 1, the curves for no exclusion, WHO flags and SMART flags will all be the same shape; however, they will be shifted slightly to the left or right depending on the mean (their central point).

Interpretation

In instances where the SD of “Exclusion from Observed mean (SMART Flags)” is **higher** than 1.2, the data has too many mistakes and the prevalence observed is most likely an over-estimation: it should not be presented as representing the true situation in the survey area. Because the mean does not change substantially with random errors, and is thus a much more robust (reliable) statistic of the nutritional state of the population, if the SD of the survey is excessive, the data should not be abandoned or only the conventional analysis reported. We strongly suggest that the calculated prevalence (with an SD of one) is reported in the final survey report, with adequate explanation of the reasons why this is deemed to be a more reliable estimate than the count of children below the cut-off point or using the calculated prevalence with an abnormally wide SD. The counted prevalence can also be reported as a comparison and to illustrate the uncertainty of the estimation.

³⁹ World Health Organization. 1995. [Physical Status: the use and interpretation of anthropometry](#).

⁴⁰ Michael Golden and Yvonne Grellety, 2002. [Population Nutritional Status During Famine](#) This study was based on NCHS Standards http://www.nutrisurvey.de/ena_beta/documents.htm.

If the SD of “*Exclusion from Observed mean (SMART Flags)*” is less than 1.2, then the data are of sufficient quality to report the “*Observed prevalence*”.

With both a Rapid SMART assessment, where there are relatively few children surveyed, and where the prevalence of malnutrition is low with a conventional SMART survey, there are likely to be very few children identified as malnourished or severely malnourished. For example, if 300 children are included in the survey and the prevalence of SAM is 0.3%, then on average one would find only one SAM child in the whole sample. If no children were found, then the prevalence would be zero, and if two children were identified, then the prevalence would double to 0.6%. In other words, when few children with malnutrition are identified, the estimate changes dramatically with each individual discovered with SAM. If this is the case, then it is recommended that ONLY the calculated prevalence with an SD of 1.0Z is reported. This approach will give an estimate of both GAM and SAM, even if no children with SAM are found within the particular sample that has been obtained in this survey and the measurements have been made rapidly or imprecisely under difficult field conditions.

❖ **Criterion 8: Skewness.**

Skewness	WHZ	Excl	#	<±0.2	<±0.4	<±0.6	>=±0.6	
				0	1	3	5	1 (-0.25)

Both skewness (criterion 8) and kurtosis (criterion 9) assess the overall shape of the distribution of the dataset. Flagged data are excluded for these two tests. The tests assess how close the distribution is to a normal Gaussian distribution (a bell-shaped curve).

Skewness is a measure of the direction and degree of any asymmetry in the data. For skewness, a value close to zero means that the data are symmetrically distributed about the mean. A positive value indicates skewness to the right – the longer tail (more values) is the right of the curve. A negative value indicates skewness to the left --the longer tail is to the left. The absolute value of the test statistic for skewness increases as the distribution becomes more and more skewed.

Interpretation

Skewed data are not necessarily due to poor quality of data collection. Skewness can be generated if there are subgroups within a population that are sufficiently different from the rest of the population to form a distinct subgroup. However, the usual reason is that one of the teams has a systematic bias that make them different from the other teams – that team may be consistently over or under estimating the weight, height or MUAC, and their measurements are then heaped onto one side of the distribution generated by the other teams. If a skewed distribution is found, then the analysis by team should be examined to see if the data from one of the teams is consistently different from the rest of the survey data. To confirm this, the data should be re-analysed omitting each team’s data in turn to see whether the skewness disappears. Where this is the case, then that team’s data needs to be excluded from the survey and the reason for the exclusion and the lower than expected sample size carefully reported. If this is observed early on in the survey, the cluster can be re-surveyed and the aberrant team re-trained. Other types of data, such as haemoglobin level, are skewed in most populations and alternative tests need to be applied for such data: this is not the case with anthropometric data so far examined in non-obese populations.

If there is no consistent difference between teams, then there may be distinct subgroups within a surveyed population that should have been identified during the planning phase of the survey, and consequently surveyed separately. An extreme form of skewness or kurtosis is sometimes found when there are distinct sub-populations. The graph of the distribution should be observed to see if there are then two peaks or a flat top to the distribution, indicating two overlapping normal distribution curves.

❖ **Criteria 9: Kurtosis**

Kurtosis	WHZ	Excl	#	<±0.2	<±0.4	<±0.6	>=±0.6	
				0	1	3	5	1 (-0.25)

Kurtosis measures the shape of the distribution, looking at how many data points (measurements) are in the tails versus in the body (or shoulders) of the distribution. Kurtosis assesses the relative size of the body versus the tails of the distribution (whether the distribution is flat topped with large shoulders, or alternatively like a “Mexican hat” with the shoulders of the distribution missing, a peaked center, and long tails). A distribution that is symmetrical is not necessarily normal.

Interpretation

A positive value for kurtosis indicates a large number of outlying (or extreme) values (tails). Outlying values can occur from errors in: reading scales or height boards, recording measurements, entering data or transferring data between softwares. If there are large numbers of flags, it is likely that there will be a high positive kurtosis. Kurtosis will decrease (become more negative) when flags are excluded from analysis (as the size of the tails decrease). A large negative value for kurtosis usually indicates that data have been “over-cleaned” and real data has been removed from the sample, that the data are fictitious, that pressure has been applied to report a low prevalence of malnutrition or an improvement from a previous survey, or that the field teams have failed to record measurements that they thought may be extreme and thus avoid incurring a penalty for excessive flags. As with skewness, a negative kurtosis can occur when several of the teams are consistently different from each other giving a relatively flat top to the distribution – this should be examined by systematically re-analysing the data with each team excluded in turn to determine whether exclusion of one of the teams removes the abnormal kurtosis.

The Shapiro-Wilk test

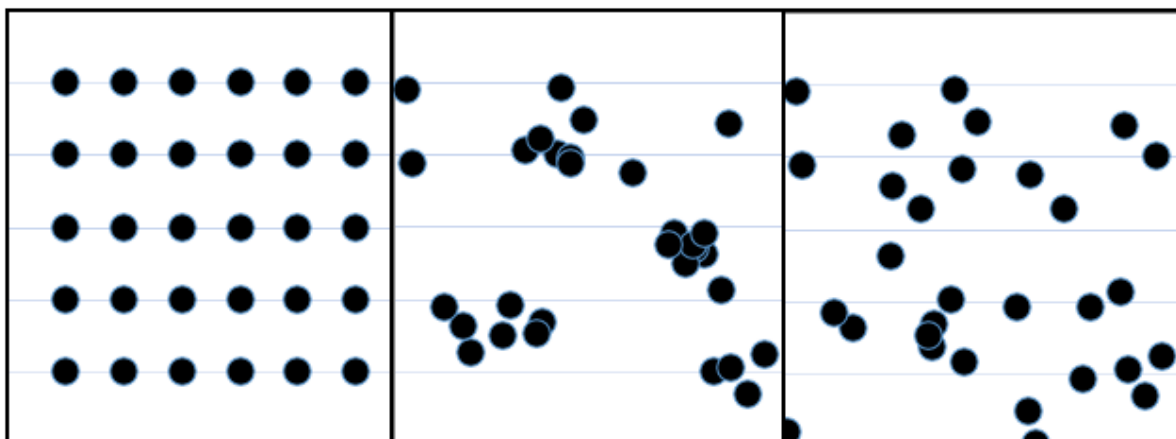
The Shapiro-Wilk test is a statistical test that examines the overall distribution against a normal distribution. It is slightly more stringent than the moments of kurtosis and skewness. The Shapiro-Wilk test is interpreted in the same way as kurtosis and skewness. It is not included in the summary table and penalty points are not assigned for this test.

❖ **Criterion 10: Poisson Distribution and Index of Dispersion.**

Poisson dist	WHZ-2	Excl	p	>0.05	>0.01	>0.001	<=0.001	
				0	1	3	5	0 (p=0.239)

The three boxes below (Figure 32) illustrate a uniform distribution, a “clumped” or pocketed distribution of 30 points, and a random distribution respectively in the same area.

Figure 32: Examples of Distribution.



Let us assume that these are malnourished children distributed in the community. They will be present in the clusters that are included in the sample. It is of interest to discern whether these children occur in the various clusters uniformly, so that the cases are spread evenly throughout the clusters with the same number of malnourished children in each cluster. Alternatively, they can occur at random so that the chance of a child being malnourished is the same for each cluster, and if the sampling was repeated with different households selected in the clusters then different numbers of malnourished children would be found in a cluster when it is re-sampled. The third possibility is that malnourished children occur much more frequently in some clusters than in others, so that the cases are pocketed in certain clusters with other clusters having fewer than expected numbers of malnourished children: in this case, if the clusters were re-sampled, the same clusters would again show higher numbers of malnourished children than the others.

Theoretically, if the malnourished children occur at random in the survey area, then the number of clusters without any malnutrition, with one child, two children, and three children with malnutrition (for instance) should follow a Poisson distribution. This is not the distribution of children but the *distribution of the clusters* themselves that have different numbers of cases. Thus, each cluster has the same chance of having a child with malnutrition; conversely, a malnourished child could occur in any cluster. Some clusters in the survey will then have no malnourished children by chance, some will have one, and some will have two or more, purely by chance.

In this analysis, the numbers of clusters with zero, one, two, or three, etc. children in them are counted and compared with the Poisson distribution. If this follows the Poisson distribution, it means that within the survey area the malnourished children are present at random throughout the whole area. This is what is generally found in small-scale surveys and in individual strata of larger surveys for moderately and severely wasted children. If this is not the case, then there may have been a sampling problem; alternatively, this can be a true reflection of the situation in the sampled area, particularly if the area comprises several livelihood zones, ethnic groups, or residents and displaced persons. Thus, it is possible that there are subgroups within the sampled area and in planning they should have been surveyed/sampled/stratified separately. If the distribution is Poisson and there are some clusters which seem to have a high number of cases, this does NOT indicate that the affected cluster is particularly at risk of malnutrition: it indicates that this is a chance finding and that if the survey was repeated then that cluster would probably have fewer cases and another cluster would then have more cases identified.

Where there are pockets of malnutrition, the distribution will not be random, and will follow what is usually a “negative binomial” distribution. This simply means that there are pockets of high levels of malnutrition within the area, and that there are other areas where there are fewer than the expected number of cases. If there are pockets, then another indication of this is an increase in the design effect reported.

In general, it is found that kwashiorkor cases do not conform to a random distribution and that oedematous cases will be clumped into a few clusters, while most others will have no cases at all. Because SAM includes cases of oedema as well as cases with severe wasting, the distribution of total SAM cases (and perhaps GAM as well) may not follow a Poisson distribution if there are kwashiorkor cases.

A uniform distribution is almost never observed. If there is a uniform distribution of cases it is an indication that the survey data have been manipulated in some way to give a similar number of cases in each cluster.

The Poisson distribution is shown in the graphs. Another way of expressing the results of the analysis is by use of the Index of Dispersion (ID). The ID is defined as the ratio of the variance to the mean. A perfectly uniform distribution will have an ID of 0.00 – such perfect symmetry is never seen. An ID that follows the Poisson distribution is 1.00 and a clustered distribution has an ID significantly higher than one.

The ENA for SMART software compares the ID statistically with the theoretical Poisson distribution (Figure 33). The higher the number of clusters the more discriminatory is the statistic.

Figure 33: Index of Dispersion as Reported in the ENA for SMART Software.

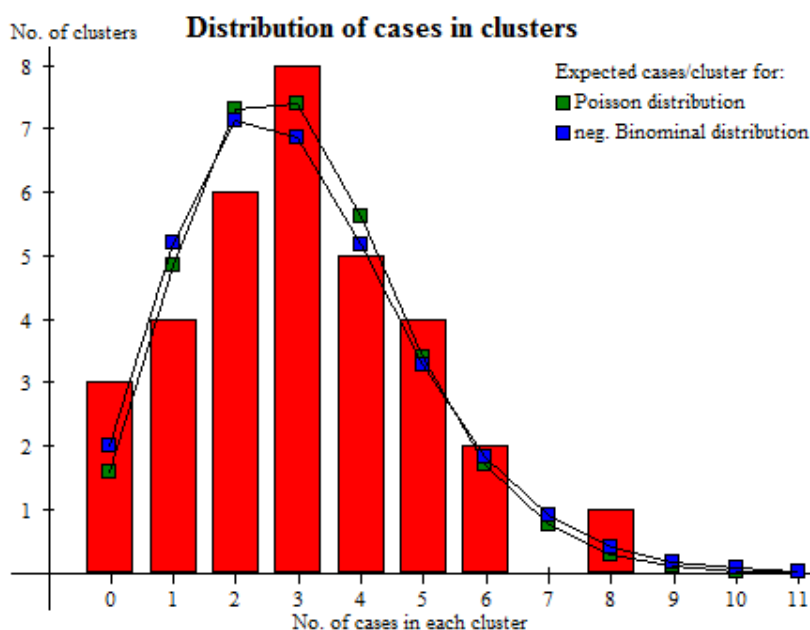
Test if cases are randomly distributed or aggregated over the clusters by calculation of the Index of Dispersion (ID) and comparison with the Poisson distribution for:		
WHZ < -2:	ID=1.16	(p=0.239)
WHZ < -3:	ID=1.33	(p=0.099)
GAM:	ID=1.16	(p=0.239)
SAM:	ID=1.33	(p=0.099)
HAZ < -2:	ID=1.79	(p=0.004)
HAZ < -3:	ID=1.65	(p=0.012)
WAZ < -2:	ID=1.99	(p=0.001)
WAZ < -3:	ID=2.07	(p=0.000)

If the ID is not significant ($p > 0.05$), it means that the distribution conforms to a Poisson distribution. An ID significantly less than 1.0 indicates that the distribution tends towards uniformity; this is nearly always an indication of improper data manipulation, perhaps by inappropriate cleaning or selection of cases, or entirely fictitious data. If an ID is significantly higher than 1 then there is evidence of either pocketing or perhaps that the team that surveyed those particular clusters has a systematic bias in its measurements to identify very few or excess cases. This can occur with oedema cases if the teams have only been trained on non-oedematous subjects and have not been trained with real cases of oedema in a hospital so that they are unsure of whether a child has or does not have oedema (therefore some teams may over-diagnose and some teams under-diagnose oedema, creating artificial “pockets” of cases); nevertheless, it is usually found that oedematous cases occur in clusters and the ID for oedema is significantly greater than 1.

In the example above (Figure 33), HAZ and WAZ do not follow the Poisson distribution, but WHZ, GAM and SAM occur randomly within in the survey area. Perhaps, one of the teams is having difficulty in ascertaining the age of the children; this should now be looked at by team to see if there is one team that is taking the age incorrectly. It would be unusual for GAM to be randomly distributed but not HAZ or WAZ. Alternatively, this could occur if there were clusters that had been disproportionately affected in the past, but not at present, because HAZ and WAZ are particularly affected by chronic and historical deprivation in contrast with acute malnutrition which reflects the recent past. The survey manager, by examining the data, should be able to make this distinction and explain the discrepancy in the report.

The distribution of cases of GAM per cluster (Figure 34) should be further investigated by looking at the *Results* tab of the ENA for SMART software, by selecting *Cluster* and *Weight/Height* under “*Graphs*”, and *Distribution WHZ <-2* from the dropdown menu.

Figure 34: Distribution of Cases of GAM per Cluster.



❖ **SAM/MAM Ratio.**

Normally there is a fixed relationship between moderate and severe wasting depending upon the degree of malnutrition within the community. If it is found that there is an excess of severe wasting (SAM without oedema) over moderate plus severe wasting (GAM) or MAM, then this is an indication that the measurements have been taken poorly. Table 11 shows the number of GAM or MAM cases that would be expected in a survey for each SAM case. It is often assumed that this ratio is constant, but this is not the case. In a normal, non-malnourished population one would expect about 16 moderately wasted cases for every severe case. If the population mean WHZ is -0.5 Z-score then for every one severe case there will be ten moderate cases; this is what is normally observed under field conditions. As the population deteriorates to a mean of -1.0 Z there will be six MAM cases for each SAM case. A population mean value of -2.0 Z is very rare, as it means that the prevalence of GAM will be 50% of the children – in these dire situations, there will only be two MAM cases for each SAM case. Examination of this ratio is another way of assessing the quality of the survey data. The ratio has important implications on the effort that is directed to treating each degree of wasting, and hence the design of implantation programmes. If the ratio differs markedly from those shown in Table 11, then the data reported in the survey is suspect.

Table 11: The Number of GAM or MAM Cases Expected in a Survey for each SAM Case (oedema excluded), depending on survey z-score mean and SD.

WHZ	GAM/SAM SD 0.8	GAM/SAM SD 1.0	GAM/SAM SD 1.2	MAM/SAM SD 0.8	MAM/SAM SD 1.0	MAM/SAM SD 1.2
0.0	70.2	16.9	7.7	69.2	15.9	6.7
-0.1	60.7	15.4	7.2	59.7	14.4	6.2
-0.2	52.5	14.1	6.8	51.5	13.1	5.8
-0.3	45.5	12.9	6.4	44.5	11.9	5.4
-0.4	39.4	11.8	6.0	38.4	10.8	5.0
-0.5	34.2	10.8	5.7	33.2	9.8	4.7

WHZ	GAM/SAM	GAM/SAM	GAM/SAM	MAM/SAM	MAM/SAM	MAM/SAM
Mean Z	SD 0.8	SD 1.0	SD 1.2	SD 0.8	SD 1.0	SD 1.2
-0.6	29.7	9.9	5.3	28.7	8.9	4.3
-0.7	25.8	9.0	5.0	24.8	8.0	4.0
-0.8	22.4	8.3	4.8	21.4	7.3	3.8
-0.9	19.5	7.6	4.5	18.5	6.6	3.5
-1.0	17.0	7.0	4.2	16.0	6.0	3.2
-1.1	14.8	6.4	4.0	13.8	5.4	3.0
-1.2	13.0	5.9	3.8	12.0	4.9	2.8
-1.3	11.4	5.4	3.6	10.4	4.4	2.6
-1.4	10.0	5.0	3.4	9.0	4.0	2.4
-1.5	8.8	4.6	3.2	7.8	3.6	2.2
-1.6	7.7	4.3	3.0	6.7	3.3	2.0
-1.7	6.8	3.9	2.9	5.8	2.9	1.9
-1.8	6.0	3.7	2.7	5.0	2.7	1.7
-1.9	5.3	3.4	2.6	4.3	2.4	1.6
-2.0	4.7	3.2	2.5	3.7	2.2	1.5

❖ **Analysis by Team.**

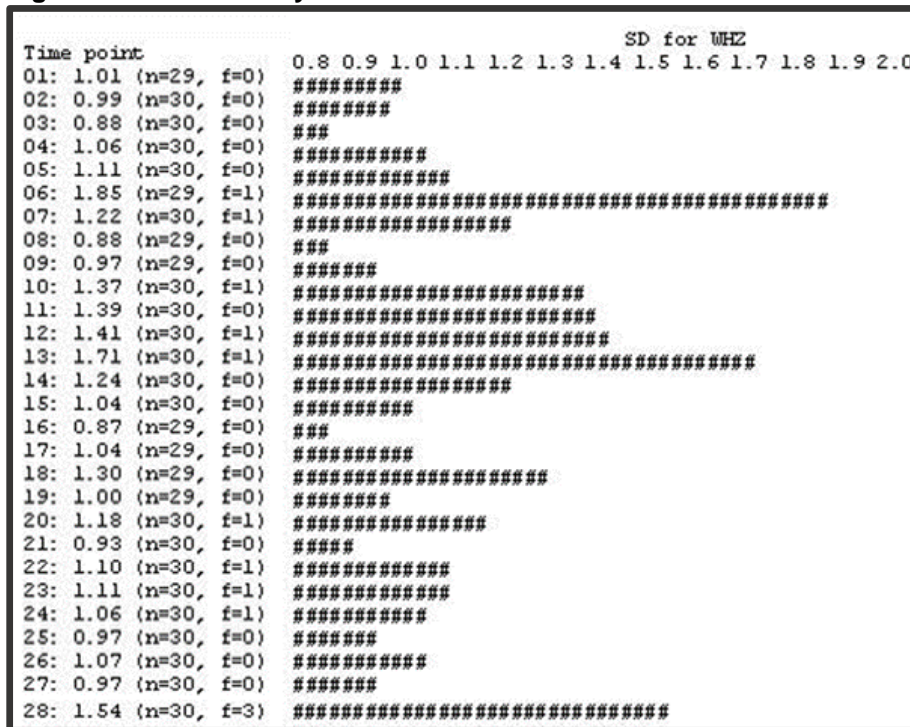
When there are problems with the data, this is often due to one or more of the teams being improperly trained or supervised so that this team is contributing aberrant data and affecting the overall quality of the data in the survey. In this section of the plausibility check report, some of the quality checks used for the whole survey are repeated for each team separately. This section should be examined whenever a problem with the data is identified or anticipated. The team results should be compared with each other.

Time when the measurements are taken

Surveys are difficult. The teams become tired when they have to travel, walk, interview and measure in the community, sometimes under dangerous conditions. The question arises whether stress and overwork affect the quality of data to the extent that can affect the overall results of the survey. Normally, the teams are fresh for the first few cases measured and become progressively tired until the lunch break – and then the same pattern repeats itself in the afternoon until there is a rush to complete the cluster and return to base.

As the quality of data deteriorates, random errors increase and the SD of the data increases. Therefore, we can use this phenomenon to examine the variation in data quality by the order in which the subjects are measured. We assume that the data are recorded in the order in which the subjects were measured. Say that there are 30 clusters; in this analysis we take the first subject to be measured in each cluster and compute the SD of these 30 subjects, one from each of the clusters; the SD of the measurements for the second 30 subjects is then computed, and so on until all the subjects in the survey have been entered into the analysis. The output is expressed graphically. A typical output is shown in Figure 35.

Figure 35: Time Analysis on the SD for WHZ.



The first column is the order of the subjects within the clusters. The next columns list the SD, the number of subjects (n=xx), and the number of SMART flags (f=x). The hashtags then give a graphic depiction of the SDs.

As one subject is taken from each cluster so that the contribution of each cluster, team and geographic area is the same for each time point as if there is no effect of the order in which the children are measured. The results shown are typical. The first five subjects measured seem to have a low SD – but the sixth subjects have a very wide SD: this corresponds to the mid-morning break. After the break, the next three subjects are reasonably measured, but then there is a steady deterioration in data quality until lunch time. In the afternoon, the quality was variable until the last cluster when there was a dramatic increase in SD and three flags were identified (excluded from the computation of the SD). The team were finishing up and preparing to return to base.

This pattern is repeated in many surveys and should give the survey planners insight into why the data quality is less than desirable. It is clear that these teams were having difficulty in completing a cluster each day and were getting hungry, thirsty, or tired after having seen several children. This may have compromised the quality of the data from this survey. This survey had an overall SD of 1.16 Z, which would attract six penalty points in the overall quality summary. Nevertheless, this survey is accepted as high quality. The survey planners should still preferably have planned clusters of a smaller size so that the team would have had time to comfortably complete all the required measurements on survey subjects, as well as travel to and from the cluster, liaise with the village elders, and identify the sampling frame within a day. This analysis also demonstrates why the addition of extra questions to a SMART survey can seriously compromise data quality.

If there are large variations in SD between consecutive children's measurements, it indicates that the survey has not been planned appropriately and this can account for poor quality data.

It is important to note that this analysis is possible and meaningful only if child ID are recorded, if the order of child ID in each cluster is reflective of the order in which these children were measured as the day in the field progressed, and if the child ID in each cluster was started from 1 and ended at the last number of the child measured in that cluster. Nevertheless,

because only one subject is contributed to each calculation from each cluster, the teams, geographic area and any other heterogeneity between the clusters should be the same for each of the calculations, and all the SDs should be approximately the same. If the order of children in some clusters is changed, more than one cluster is examined per day, or one cluster takes more than one day to complete, the differences may not be interpretable in terms of time of day, and any differences may be masked, but if there is a major difference between the children analysed with the same ID number, then there is a problem with the quality of the data which needs to be investigated.

Additional characteristics of the dataset

Missing Data.

Missing data:

WEIGHT: Line=323/ID=

The plausibility check report will show the exact line and ID (when available) of any missing data on sex, age, height or weight. Missing data on other variables are not listed. The missing entry can be found by looking at the data set in the *Data Entry Anthropometry* tab of the ENA for SMART software.

Duplicate Entries.

There were no duplicate entries detected.

Exact duplication of entire rows as entered in the *Data Entry Anthropometry* tab of the ENA for SMART software will be identified (only when values for all variables in two or more rows are identical). If two children have identical anthropometric variables (sex, age, weight, height, edema) but have even one different identifying variable (cluster, household, team, ID or date) they do not show up as duplicates in the plausibility check report.

Percentage of Children with no Exact Birthday.

Percentage of children with no exact birthday: 87 %

The number of children in a dataset without an exact birthdate is an important indication of the quality of the age data. In settings where birth certificates or remembering exact birthdates are not common, age estimates based on local events calendar are less precise than exact birthdays. This will affect the interpretation of the Height-for-Age and Weight-for-Age indicators.

Age Distribution.

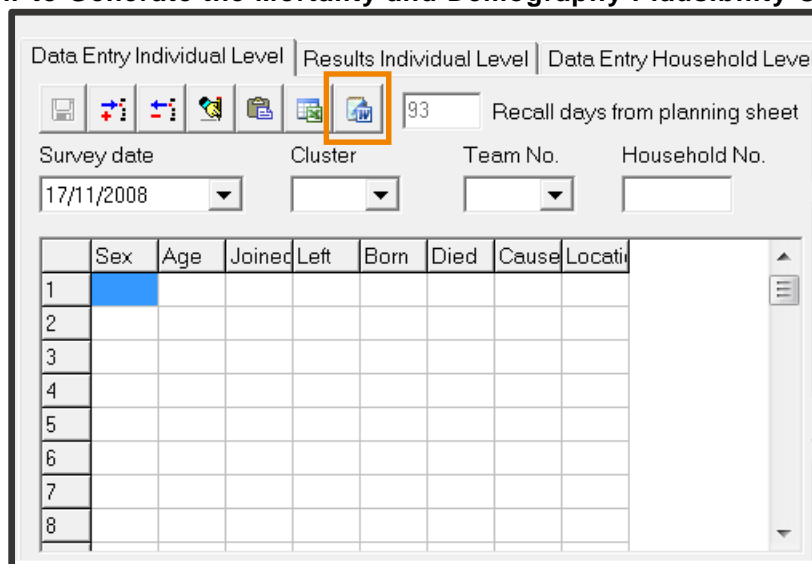


Age distribution provides a visual indication on whether any age heaping (i.e. whether the ages were rounded to whole or half years) occurred in the dataset. Each hashtag (#) represents one survey subject's reported age. With a representative sample that has ages based upon birth documents or a properly administered events calendar, a random distribution of ages should be seen across all the months. When ages are estimated, a distribution with peaks at full years (12, 24, 36, 48 months) and sometimes at half years (18, 30, 42, 54 months) can often occur. A lot of age heaping highlights the poor administration of local events calendar by the survey teams.

Sections of Mortality and Demography Plausibility Check Report

Given recent emergencies requiring quality mortality and demographic data, preliminary logic checks have been incorporated in the ENA for SMART software available on the SMART Methodology website⁴¹. On the *Data Entry Individual Level* tab of Death Rates in the ENA for SMART software, the “*Report Plausibility Check*” button as seen below (Figure 36) generates the mortality and demography plausibility check report that opens in Rich Text Format (also available in *Extras* under “*Plausibility check Mortality and Demography*”).

Figure 36: How to Generate the Mortality and Demography Plausibility Check Report.



The plausibility check for Mortality and Demography focuses on the following tests:

1. **Missing or wrong data** details line numbers of the households in the data table that have missing information on the cluster number, household ID, or team ID, as well as on sex or age of at least one member in the household. This section also highlights the line numbers of the households where a child is born during the recall period and recorded age of that child is more than zero. The age is indicated in completed years, therefore if recall period is shorter than one year, all children born within the recall period cannot be older than 1 year of age.
2. **Duplicate data** highlights which lines has identical information for the following:
 - All identifier variables (cluster, household and team ID);
 - Household members data;
 - Identifier variables (cluster, household and team) and household members' data.
3. **Unusual values – Births and Deaths** indicates unusual events -- which household lines have more than one birth in a household or more than two deaths in a household during the recall period.
4. **Households per cluster** shows the total number of households per cluster, number of households with children under five years of age, and the percentage of households with children among all households surveyed in each cluster. This may detect if the survey teams were only surveying households with children under five and skipping the households without children. It may also highlight the clusters with unusually high non-response, and the clusters where the survey teams did not follow the sampling procedures if the number of households surveyed is above the number of households they were instructed to survey per cluster.
5. **Age distribution** provides a distribution of ages in years of all survey subjects listed in survey households. This provides visual indication on whether any age heaping occurred.

⁴¹<http://www.smartmethodology.org/>

Each hashtag (#) represents one survey subject's reported age for all individuals surveyed, including those who were born, died, joined and left during the recall period. With a representative sample that has ages based upon birth documents or a properly administered events calendar, a random distribution of ages should be seen across all the years. A lot of age heaping highlights the poor administration of local events calendar or age estimation by the survey teams.

6. **Statistical evaluation of sex and age ratios (using Chi-squared statistic)** analyses the observed number/expected number for the overall sex ratio and age distribution using a Chi-squared test for the following categories:
 - **Sex ratio** looks at if there is equal representation of boys (M) versus girls (F) for children under five (U5) and children under ten years (U10) of age;
 - **Age distribution:** analyses whether the age distribution of the sample for children 0-23 (0-<2 years) versus 24-59 months (2-<5 years) of age; and children 0-59 (0-<5 years) versus 60-119 months (5-<10 years) of age is plausible (follows expected demographic patterns: ratio of 0-<2/2-<5 is expected to be close to 0.7, and ratio of 0-<5/5-<10 is expected to be close to 1.1)
 - **Sex/age distribution** provides further information on whether the overall age and sex distribution matches the expected values for two sets of four age/sex categories: M 0-<2, M 2-<5, F 0-<2, F 2-<5 for U5 children, and M 0-<5, M 5-<10, F 0-<5, F 5-<10 for U10 children.
7. **Poisson Distributions** uses a similar logic as in anthropometry plausibility check described previously. Instead of using graphs, the results of this analysis are shown by the use of the Index of Dispersion (ID). As a reminder, the ID is defined as the ratio of the variance to the mean. A perfectly uniform distribution will have an ID of 0.00 – such perfect symmetry is never seen. An ID that follows the Poisson distribution is 1.00, and a clustered distribution has an ID significantly higher than one. If the index of dispersion is not significant ($p > 0.05$), it means that the distribution conforms to a Poisson distribution, meaning the cases appear to be randomly distributed among the clusters. An ID significantly less than 1.0 indicates that the distribution tends towards uniformity: this is nearly always an indication of improper data manipulation, perhaps by inappropriate cleaning or selection of cases, or because of entirely fictitious data. If an ID is significantly higher than 1, then there is evidence of either pocketing or perhaps that the team that surveyed those particular clusters has a systematic bias in its measurements to identify very few or excess cases. The Poisson distributions are analysed similarly to those for anthropometry yet based on the following categories:
 - Distribution of household (HH) size;
 - Distribution of under-five children (U5) per household - this test is run on all households and not on households within each cluster;
 - Distribution of deaths per cluster;
 - Distribution of U5 deaths per cluster;
 - Distribution of births per cluster.

Conclusion on the Plausibility Checks

Plausibility checks are an innovative and very important tool for survey managers to assess the overall quality of their survey data and to identify mistakes or problem areas. For Anthropometry, the score provided by the summary table highlights the type of problems that may have occurred during sampling and data collection, requiring further investigation of contextual data via discussions with the field supervisors or from other in-country partners. Although the results of the plausibility checks cannot be used to explain why children are malnourished or why certain indicators are as reported, they are tools that allow those that use the results of a survey to have confidence in its results. All operations have a “motor” and a “sensory” nervous system – there cannot be appropriate intervention without reliable data.

The plausibility checks allow readers of the report and those evaluating the actual data itself to rely upon the data without reservation if the tests are all passed.

Plausibility checks should always be generated and examined before analysis and interpretation of the actual results takes place. The data should not be used for decision-making if minimal level of confidence in survey results is not achieved. A full copy of the plausibility checks must always be included as one of the appendices in the survey report, and if there are more than one strata for Anthropometry, a summary table should be included with the strata-level results of the plausibility check.

6. Mortality and Demography

In most developed countries, death rates are monitored by civil registration systems that record vital statistics like birth, death, marriage and divorce. In many developing countries, however, vital statistics systems are of poor quality or lacking altogether. A study in 2007 found that each year, nearly 50 million newborn children are not registered and only a third of countries outside of Europe and North America have the capacity to obtain usable mortality statistics⁴².

Even if they were once functioning, vital statistics systems often are disrupted in situations of civil conflict or natural disaster. In some cases, where displacement occurs over a longer term — in refugee camps or internally displaced persons (IDP) settlements, for example — it may be possible to establish registration systems to record vital events and other health indicators. The United Nations High Commissioner for Refugees (UNHCR) supports health information systems in 80 refugee camps in 16 countries, which record monthly reports of live births and deaths by age and cause⁴³. In order to follow morbidity and mortality prospectively and allow more timely response to changes, a surveillance system to register communicable diseases and deaths should be set up as early as possible in humanitarian emergencies⁴⁴. When these services are not available, public health workers have to rely on cross-sectional surveys to estimate death rates retrospectively.

While mortality is the final and most definitive health outcome in almost any setting, understanding — and estimating — mortality in a given situation is even more important for several reasons:

- Knowing how many people have died (or, more accurately, the rate at which people have died within a given time period) may help to determine how urgent is the need for humanitarian action (including intervention and/or advocacy).
- Knowing something about who has died—in terms of age, sex, cause of death, and other characteristics—helps to determine who is at greatest risk in the surviving population.
- Mortality rates (particularly of children under five) have proven to be sensitive tracking indicators of overall population health and security and thus may be useful in monitoring changes in health over time.

An organisation seeking data on mortality should begin by addressing a few basic questions:

- Why does the programme need mortality data?
- What kinds of data are needed?
- Who is best able to collect these data?
- How should the data be collected in a given context?

The kind of decisions that need to be made will dictate not only the kind of information to be sought but the level of precision that is desired. Context and constraints will dictate whether what is desired can be achieved. Time may require that a programme decision be made quickly; just as resources, security and other constraints may limit the precision of the information obtained.

⁴² Setel PW, Macfarlane SB, Szreter S, Mikkelsen L, Jha P, Stout S, AbouZahr C et al. A scandal of invisibility: making everyone count by counting everyone. *Lancet* 2007; 370: 1569-1577

⁴³ United Nations Refugee Agency (UNHCR). *Health Information System, Quarterly Bulletin* 2008; 3. (UNHCR; <http://www.unhcr.org/protect/4614bce84.html> accessed on 1 June 2009).

⁴⁴ Checchi F, Roberts L. Documenting mortality in crises: what keeps us from doing better? *PLoS Med* 2008; 5(7): e146.

How to Determine the need to Conduct a Mortality and Demography Survey?

Information on mortality can have multiple functions. In the first instance, it may be used to identify immediate and severe health risks requiring urgent action during a crisis: immunisation campaigns, emergency feeding, or security interventions, for example. More generally, mortality data — particularly the crude death rate and the under-five death rate — can help to benchmark the severity of a crisis, to provide an empirical basis for advocacy purposes, to be able to compare one crisis situation to another, and to establish a baseline for evaluating impact of programme interventions⁴⁵. Whether the reasons are for short-term, local interventions or longer-term, broader advocacy purposes or to complement a surveillance system, organisations should be able to articulate a clear purpose for gathering mortality data. If they cannot, it is likely that they will not properly invest resources in data collection or have any meaningful use for the results.

It is important to consider the following contextual factors when deciding on whether mortality data should be collected:

Type of emergency

Disasters often are classified by speed of onset, with two main types: sudden-onset and gradual-onset. It is also possible to make distinctions between natural and human-made causes⁴⁶. Common types of natural disasters include earthquakes, tsunamis, floods, droughts, famines and epidemics, while human-made disasters include conflicts, technological and industrial disasters.

Phase of emergency

The phases of a disaster vary according to the type of disaster, although generally three phases can be distinguished: a pre-emergency phase, an emergency phase (often separated into an acute and chronic phase), and a post-emergency phase that may be marked by gradual recovery or may transition back into crisis. In the case of disasters like earthquakes and hurricanes, the pre-emergency and emergency phases may last only a matter of hours or days. Famines and complex emergencies, on the other hand, evolve over months, even years, and the recovery phase may be still more protracted. Epidemic diseases can have a rapid onset (cholera or diarrheal disease) or may take years to develop into a crisis (HIV/AIDS)⁴⁷.

Level of insecurity

Conflict, violence and insecurity not only create dangers for the population of concern, causing some to hide or to flee, but impose security risks and severe logistical constraints on programme staff.

⁴⁵ Working Group for Mortality Estimation in Emergencies. Wanted: studies on mortality estimation methods for humanitarian emergencies, suggestions for future research. *Emerging Themes in Epidemiology* 2007, 4:9. See also: Checchi F, Roberts L. Interpreting and using mortality data in humanitarian emergencies; A primer for non-epidemiologists. *Humanitarian Practice Network Paper* 2005. Prudhon C, Spiegel PB. A review of methodology and analysis of nutrition and mortality & demography surveys conducted in humanitarian emergencies from October 1993 to April 2004. *Emerging Themes in Epidemiology* 2007, 4:10.

⁴⁶ Tailhades M, Toole M. Disasters: What are the needs? How can they be assessed? *Tropical Doctor* 1991; 21: 18-23. See also Noji E. *The Public Health Consequences of Disaster*, New York and Oxford: Oxford University Press, 1997.

⁴⁷ CERTI. Crisis and Transition Tool Kit: Strategic Framework for Crisis Prevention, Mitigation and Transition, 2000.

Level of volatility

In their acute phase, emergencies are inherently unstable. Neither calm nor chaos is entirely predictable, making planning a challenge. Camps and settlements can be full one week and half-empty the next, rendering even an accurate estimate quickly obsolete.

Geography and settlement patterns

Populations of concern may be clustered in one place or widely scattered across a broad area. They may be settled discretely or mixed in among other populations. Combined with severe weather patterns, geographical features like mountains and rivers can hinder or prevent access.

Most mortality estimates focus on crude death rate and under-five death rate but, in some instances, either more or less may be needed. In the very early days of a crisis, when the situation is most dynamic, rough approximations of the number and key causes of death (specifically, traumatic versus non-traumatic deaths) may be all that is necessary or possible to obtain. As a crisis stabilises somewhat, perhaps settling into a chronic and/or transitional phase, organisations may seek more precise and detailed estimates of age-specific, sex-specific and cause-specific death rates (traumatic versus non-traumatic deaths), or they may seek to know more about how death rates may be associated with migration patterns, ethnic or religious background or household composition, in order to better pinpoint risk and target interventions.

It is strongly recommended that before gathering any mortality data, organisations should investigate what sort of data may already exist and who has collected it. These include:

- Census data and vital statistics;
- Data from local health offices and facilities;
- Community surveys (health, nutrition, livelihood, etc.);
- Demographic and Health Surveys (DHS);
- Demographic Surveillance Sites (DSS);
- Complex Emergency Database (CE-DAT) and Emergency Events Database (EM-DAT) run by the Center for Research on the Epidemiology of Disasters (CRED);
- Relief Web, run by the United Nations Office for the Coordination of Humanitarian Affairs (UN-OCHA).

Once this external inventory has been carried out, organisations should conduct an internal inventory to identify what resources it has (or lacks) to collect mortality data. This should include:

- Financial resources;
- Human resources (for administration, field work, data entry and analysis, and dissemination of results);
- Logistical support (including vehicles, radios, phones, computers, etc.).

How should the data be collected?

A *survey* is defined as “a detailed study or inspection that gathers information through observations, questionnaires, etc, and analyzes it.”⁴⁸ A *sample* is “a subset of the population that is used to gain information about the entire population. A sample in this sense is a model of the population. A good sample will represent the population well.”⁴⁹ A sample survey, then,

⁴⁸ *Webster's New World Dictionary*, Fourth Edition, IDG Books Worldwide, 2001.

⁴⁹ Henry G. *Practical Sampling*, Applied Social Research Methods Series, Volume 21. Newbury Park: SAGE Publications, 1990, p. 11.

gathers data on a sub-set of a population through observations, questionnaires and other methods and analyses this data in order to measure something of interest about a population.

Characteristics

- Short-term (often one-time) collection of data from a population sample based on questionnaires and other measurements.
- Measures retrospectively cumulative incidence of mortality over a recall period.
- *Numerator*: deaths over recall period; *denominator*: “person-time” of exposure (measured directly or approximated as mid-interval sample population).

Table 12: Advantages and Disadvantages of Household Mortality and Demography Surveys.

Advantages	Disadvantages:
Numerator and denominator are obtained from the same source (cross-sectional data from sample households)	Surveys cannot follow trends in mortality incidence or prevalence rates unless they are repeated.
Probability-based sample and more detailed questionnaire allows for estimation of prevalence and identification of risk factors	Surveys based on probability sampling designs involve sometimes complicated decisions about sample size, design and post-survey analysis.

Household mortality and demography surveys provide a methodology for obtaining often detailed (though retrospective) information on mortality events and the demographics of the population in the households exposed to the risk of those events. They do require time and resources to plan and implement as well as time and resources to analyse findings and disseminate results.

If surveys are repeated periodically, they can be a part of a surveillance system. Surveillance systems, once in place, can provide timely reporting on mortality events to allow for prompt identification of and response to outbreaks, sanitation problems, nutritional deficiencies, and other mortality risks. Surveillance systems require cooperation from a variety of different actors — health facilities, relief agencies, possibly government entities and, of course, the affected communities—and thus can be both difficult to set up and difficult to maintain. Ideally, retrospective surveys should complement a functioning surveillance system to verify surveillance data, and answer questions that the surveillance system is not providing or about areas that the surveillance system is not covering.

Mortality and Demography Survey Objectives

The most common form of household mortality and demography survey used in disasters is a retrospective study (looking backward in time over a specified interval) employing a probability sample design. A retrospective survey measures mortality over some defined period by identifying deaths during that period to a defined population (usually a sample of households), and estimating the exposure to mortality risk of all members of the households who were alive and living in the household at any time during the recall period.

Although retrospective sample surveys are a common method to estimate mortality rates in crisis settings, they can be challenging to design and implement. A study of 23 surveys

conducted during the 1991-1992 famine in Somalia found extensive differences and inconsistencies in the surveys, including the following:⁵⁰

- Study objectives were unclear or unspecified;
- Survey methods were not documented (only 12 were deemed reproducible);
- Units of measurement and construction of denominators in rate calculations were inconsistent;
- There was considerable variation in the length of the recall period chosen (ranging from 1 month to 12 months), and;
- None of the studies reported confidence intervals around the point estimates of the mortality rates.

Ten years later, a review of 158 mortality and demography surveys conducted in humanitarian emergencies in 17 countries between October 1993 and April 2004 found that only four (3.2%) met the study's criteria for quality (compared to 35.3% of nutrition surveys reviewed). Moreover, while the study found the proportion of nutrition surveys meeting criteria for quality "rose significantly from 1993 to 2004, there was no improvement for mortality and demography surveys during this period."⁵¹ As the authors of the 1994 Somalia study noted:

Definition of specific objectives is a critical element of a public health study. The objectives define, in general terms, what needs to be learned and why. They target the direction of the study, help define the methods, dictate the information to be collected, and promote appropriate use of the data once the study has been completed. Lack of stated objectives may lead to failure to collect necessary information or to the collection of irrelevant information, increasing the expenditure of resources without influencing either public health policy or programme management decisions.⁵²

The objectives for conducting a mortality and demography study in a crisis setting may vary, but it is likely that they involve one (or possibly both) of two main purposes:

1. To benchmark the scale and severity of the crisis in order to promote awareness and influence public policy;
2. To identify proximate mortality risks in a given population in order to target interventions, evaluate programme effectiveness, and make other programme management decisions.

While these two objectives could certainly overlap, the first (which could be called the advocacy objective) may argue for a broader perspective on the population at risk across time and place and may also dictate that more extensive information be collected. The second objective (which could be called the evaluation objective) may argue for a more intensive, epidemiological perspective on population risk, concentrating on more recent and local mortality events as these are the ones that can be targeted for intervention and evaluation.

The first step in conducting a mortality and demography survey is to decide on the objective(s) of the survey. Write them down, share them with stakeholders (including the affected communities), and continue to review them as planning moves forward. **If an organisation cannot articulate clear objectives for conducting a mortality and demography survey, it is a good indication that one is not necessary and, if conducted, may not be useful.**

⁵⁰ Boss L, Toole M, Yip R. Assessments of mortality, morbidity and nutritional status in Somalia during the 1991-1992 famine: recommendations for standardization of methods, *JAMA*, 1994; 272 (5). Of the 23 surveys, 16 focused specifically on mortality.

⁵¹ Prudhon and Spiegel, 2007.

⁵² Boss et al, 1994.

Crude Death Rate (CDR)

In estimating mortality, probably the first element to consider is the Crude Death Rate (CDR). In the context of the SMART mortality module and the ENA for SMART software, although the Crude Death Rate and the Crude Mortality Rate are calculated the same way and the terms may be used virtually interchangeably, for the death rate of the entire population, we will use the term CDR.

A rate measuring population change *during or within* a specified time interval, a mortality rate — whether it is a crude, age-specific or cause-specific rate — measures the average risk of death during a given interval. More precisely, the CDR is the number of people who die in a given time interval, divided by the population at mid-point in the interval:

$$\text{CDR} = \frac{\text{Number of deaths}}{\text{Population at mid-interval}} \times K$$

The CDR reflects the risk for the entire population of dying during a given time period (K). The numerator is the total number of deaths of people of all ages who were alive at some time during the interval. The denominator is “person-time” lived in the interval, and specifies the total population exposed to the risk of that event, as approximated by the mid-interval population, on the assumption that all deaths, births, and movements in or out are evenly distributed across the interval.

In mortality estimation, the exposure to risk in the denominator is expressed by the concept of “person-time” lived, which is the product of the number of persons multiplied by the number of days, weeks, months or years that each person was exposed to the risk of the event. The use of a mid-interval population figure as the denominator is based on a simplifying assumption that births, deaths, as well as in-migration and out-migration, are evenly distributed throughout the interval. Thus, the mid-interval population captures half of all “entries” and half of all “exits” during the full interval. Another way of looking at it is that each birth and death (also in-migration and out-migration if this is being measured) contributes half an interval-length of exposure per person. If the mid-interval population is used, the rate is defined in terms of death per interval, for example per day, per month or per year. Exposure time, however, can be defined in terms of any units; in the context of humanitarian emergencies, epidemiologists have tended to express rates in terms of deaths per 10,000 per day.

To examine the issue of “person-time” lived, consider the example on the next page (Table 13) of a hypothetical village (or refugee camp) with a population of 500 persons on January 1. As the example shows, the population grew from 500 to 600 in the space of one year, with 16 births, 21 deaths, 125 arrivals and 20 departures during the interval. In order to calculate the CDR, the numerator is 21 deaths. The question is: what is the denominator? Using the person-day measurement gives the denominator as 215,607, so the CDR = $(21/215,607) \times 10,000 = 0.97$ per 10,000 per day. An up-to-date and complete vital registration system could provide the data needed for precise calculations of “person-time” of exposure using individual dates of birth, death, in-migration and out-migration.

Table 13: Calculation of Person-Years and Person-Days Lived in a Population Numbering 500 on January 1 and 600 on December 31.

(1) Number of persons	(2) Events and dates in interval	(3) Days lived in the interval by each person	(4) Number of person-days lived in interval:(1)*(3)
500	Alive on Jan. 1		
482	Alive from Jan. 1 to Dec. 31	365	175,930
1	Born Jan. 11	354	354
1	Born Jan. 11, died Nov. 9	302	302
52	Arrived Feb. 19	315	16,380
1	Born Feb. 21, died Apr. 27	65	65
1	Born Mar. 6, died Mar. 31	25	25
2	Died Mar. 14	73	146
37	Arrived on Mar. 19	286	10,582
2	Died Apr. 8	98	196
4	Arrived on Apr. 18	258	1,032
1	Died Jun. 1	152	152
5	Died Jun. 5	156	780
1	Born Jun. 7	207	207
14	Departed Jun. 15	166	2,324
1	Died Jun. 30	181	181
1	Born Jul. 26	158	158
14	Arrived on Aug. 15	138	1,932
1	Born Aug. 24	129	129
1	Died Aug. 26	180	180
1	Born Sep. 13, died Nov. 13	61	61
6	Departed on Sep 7	250	1,500
1	Born Oct. 1	91	91
2	Born Oct. 7	85	170
1	Died Oct. 9	255	255
1	Born Oct. 19	80	80
18	Arrived Oct. 20	72	1,296
1	Born Nov. 18	49	49
1	Born Nov. 30, died Nov. 30	1	1
3	Died Dec. 3	337	1,011
2	Born Dec. 12	19	38
		<i>Total</i>	<i>Total</i>
600	Alive on December 31	4,908	215,607

Source: Adapted from Palmore and Gardner, 1994.

To calculate crude mortality and under-five mortality, as done by the ENA for SMART software, is to take all people alive at the end of the interval (600) and, if using the person-day measurement, to multiply that number by the number of days in the interval ($600 * 365 = 219,000$). From this number is subtracted one half ($\frac{1}{2}$) of the person-days contributed by births and in-migrations (on the simplifying assumption that these persons were in the household for half the interval), and from the same number is added one half ($\frac{1}{2}$) of the person-days contributed by deaths and out-migrants (on the assumption that these persons, though not in the household at the end of the interval, were in the household for half the interval).

Thus, 219,000
 minus $\frac{1}{2}$ births = $219,000 - 2,920 = 216,080$
 plus $\frac{1}{2}$ deaths = $216,080 + 3,832 = 219,912$
 minus $\frac{1}{2}$ in-migrations = $219,912 - 22,812 = 197,100$
 plus $\frac{1}{2}$ out-migrations = $197,100 + 3,650 = 200,750$

The CDR resulting from this approach is $(21/200,750) * 10,000 = 1.05$ deaths per 10,000 per day. The reason it is slightly higher than the rate calculated using actual dates of birth, death,

arrivals, and departures is due to the fact that more than two-thirds of in-migrants arrived in the first quarter of the interval, thus the simplifying assumption that all in-migrants were in the household for only half the interval underestimates the “person-time” of exposure they contributed. By subtracting too many person-days from the calculation, the estimated mortality rate is thus biased upward (note that, using the actual “person-time” calculated from individual dates of birth, death, in-migration, and out-migration yielded a denominator of 592, while a calculation based on a simplifying assumption that the rate of population change was constant across the interval yielded a denominator of 550). The opposite effect could occur, of course, if, for example, significant out-migration had occurred early in the interval or large-scale in-migration late in the interval. Although this example is using patterns of migration into and out of a camp population, similar effects can occur within households.

We might note that, while the absolute difference in mortality rates is not large (0.97 deaths per 10,000 per day compared to 1.05 deaths per 10,000 per day), one falls just below the “emergency” threshold of 1 death per 10,000 per day and the other is just above it). The point here is that, although birth rates are not likely to change much during short periods of time, mortality, and certainly migration patterns, could fluctuate significantly. Any measurement that simplifies assumptions about these patterns risks making errors in estimation. Nevertheless, the numbers of births and in- and out-migrations generally have a small impact on the death rate estimate in comparison to the number of deaths (the numerator), which has a significant impact on the overall estimate.

Under-Five Death Rate (U5DR) and Under-Five Mortality Rate (U5MR)

In contrast to the CDR (measuring deaths at all ages in a population) and the Infant Mortality Rate (measuring deaths below the age of one), the SMART mortality estimation focuses on age-specific death rates at various intervals (fixed at every 5 years or custom intervals), though they can be calculated using broader age-intervals: adults 60 and over, working-age population between 15 and 64, etc. Other than the CDR, the most common measurement of mortality in crisis or transitional settings is deaths of children under five.

In measuring mortality of children under five, it is critical to use the correct terminology to express mortality risk. In the SMART mortality module and the ENA for SMART software, the Under-Five Death Rate (U5DR) can be calculated. In larger surveys, such as the DHS or MICS, the Under Five Mortality Rate (U5MR) is the commonly used indicator by governments, United Nations agencies and research organisations for estimating national mortality rates and trends.

Unlike CDR and CMR, which are calculated the same way and which terms may be used virtually interchangeably, **U5MR and U5DR are calculated using different methods and express different ways of measuring mortality risk.** To prevent confusion, it is important to maintain the clear distinction between the two different measures.

Under-Five Mortality Rate (U5MR)

It is the probability of dying before the exact age of five, expressed per 1,000 live births. U5MR cannot be calculated directly from data on births and deaths by age in a single year because the death of a four-year-old occurs to a child born four to five years before the one-year time interval. U5MR can be calculated from birth history data or by using indirect methods.

Under Five Death Rate (U5DR)

What the SMART Methodology uses.

It is calculated by dividing deaths of children under age five in an interval by the exposure time (in person-years or person-days) of children under age five during the interval.

Thus,

$$\text{U5DR} = \frac{\text{Number of deaths to children} < 5}{\text{Mid-interval population of children} < 5} \times K$$

In humanitarian responses and emergency conditions, the U5DR is used as an age-specific death rate (expressed in deaths per 10,000 per day). Significant literature has been published about mortality in humanitarian emergencies using this indicator. **To maintain clarity with mortality indicators, we insist that mortality estimates for children under five following SMART methods produced are labelled as U5DR.**

Sampling Design

After a study objective has been established for the survey, the next task is to identify the most appropriate sample design. Sampling methods commonly are broken down into two general types: probability and non-probability samples. Probability samples “have the distinguishing characteristic that each unit in the population has a known, non-zero probability of being selected”, while non-probability samples “are selected based on the judgment of the researchers to achieve particular objectives.”⁵³ Probability samples employ some form of random selection mechanism in order to control for subjective bias. Non-probability samples do not randomly select participants — subjectivity is either tolerated or, in some cases, intended in the method. Probability samples using the ENA for SMART software enable researchers to measure the uncertainty (sometimes referred to as confidence limits) in making inferences from sample data to the population of interest. With non-probability samples, uncertainty (as a statistical matter) cannot be measured.

In general, a probability sample will be the appropriate choice assuming that one study objective is to estimate a CDR within measurable confidence limits. For further details on which sampling method to use, the reader should refer to the Sampling Chapter for more information.

Deciding on the length of the Recall Period

Because a mortality rate is a measure of average risk over a period, a retrospective survey must establish a recall period or interval within which deaths are to be counted. Consider the example in Table 14.

Table 14: Selecting a Recall Period for a Mortality and Demography Survey.

You need to assess mortality in a rural population.												
Events which likely have had an impact on mortality include:												
<ul style="list-style-type: none"> ▪ Seasonal food shortages from May to October, but particularly from August to October; ▪ Outbreaks of malaria in the season of heavy rains in January and February ; ▪ Flare-up in ethnic conflict in March and April, leading to ; ▪ Population displacement in April through June. 												
Event	Time											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Food shortage						X	X	X	XX	XX	XX	
Malaria outbreak		X	X									
Conflict			X	XX								
Displacement				X	XX	X						

Adapted from Save the Children, *Nutrition assessment Manual*, London: SCF; 2004.

⁵³ Henry, 1990.

Assume you decide to conduct the mortality and demography survey in October:

- If you use a recall period of three months, you will obtain mortality data that is affected by the worst of the food shortages.
- If you use a recall period of six months, you will obtain mortality data that is affected by food shortages (moderate and severe) and by effects of displacement.
- If you use a recall period of nine months, you will obtain mortality data that is affected by food shortages, conflict, displacement, and malaria outbreak.
- If you use a recall period of 12 months, you will obtain mortality data affected by all of the above, plus possible food shortages of the year before (if these are chronic).

The decision about how long a recall period should be in a retrospective mortality and demography survey depends, first and foremost, on the objectives of the survey. If the objective of the survey is to assess the proximate risk of mortality in order to target interventions and evaluate their impact, it is likely that a shorter recall period, approximately three months, will be more appropriate⁵⁴. If the objective of the survey is to document mortality not only for decision-making in a specific situation but for broader advocacy and awareness, then longer recall periods, perhaps six to 12 months, will be more appropriate. In some special cases, the recall period may only focus on a matter of a few weeks (tsunami mortality, for example), or it may encompass several years (a retrospective mortality and demography survey of North Koreans in 1999 covered a four-year recall period in order to capture mortality risk in 1996 and 1997 — believed to represent the peak years of famine-related mortality — and the two years on either side)⁵⁵. While there are other factors that will affect the decision about length of the recall period (these are explored below), survey objectives should be predominant.

The selection of a particular recall period will affect mortality rates in terms of:

Levels of mortality

Other things being equal, it is likely that mortality rates would be higher in the survey with a three-month recall period than in the survey with a six-month period.

Causes of mortality

Shorter recall periods tend to be most affected by most proximate and highest risk factors for mortality, while longer recall periods will be more heterogeneous. Differential causes will have differential impacts on age-specific death rates (conflict tends to pose greater risk to young adult males, while malnutrition and malaria likely will be of greatest risk to young children).

Locations of mortality

Displacement of populations during a recall period will mean that mortality rates will be affected by where people were living at their place of origin, what risks they faced in flight (in transit), and what risks they face in their place of (at least temporary) settlement. As not all populations in the survey have experienced the same displacement (and some not at all, or not recently), these effects will vary.

Shorter recall periods are defined here as typically ranging between three to six months. Advantages of a shorter recall period include the greater focus provided on recent and proximate mortality events, reduced recall bias, and the likelihood that there has been less dynamic change in household composition over place and time. Disadvantages include the

⁵⁴ Prudhon and Speigel, 2007, analysed 158 mortality and demography surveys and found that recall periods varied between one and 12 months, with a median of three months and a mean of 4.1 months.

⁵⁵ Doocy S, Rofi A, Moodie C, Spring, CE, Bradley S, Burnham G, Robinson C. Tsunami mortality in Aceh Province, Indonesia. *Bulletin of the World Health Organization* 2007; 85(2): 273-278. See also Robinson C, Lee MK, Hill K, Burnham G. Mortality in North Korean migrant households: A retrospective study. *Lancet* 1999; 354.

fact that they will miss mortality events outside the recall period that may have had a significant impact on the population of interest, and that they are more sensitive to “calendar error” (meaning that while respondents may have trouble distinguishing a death that occurred 13 weeks prior to the interview from one that occurred 12 weeks previously). It is important to note that shorter recall periods provide less precision in estimating rates.

Longer recall periods are defined as ranging typically between seven to 12 months. Longer recall periods have the advantage of being less sensitive to “calendar error”, and they offer greater precision in estimating mortality rates. As disadvantages, longer recall periods do tend to capture more household change, thus making considerations of household membership (and, thus, denominator calculations) more complex, and they are subject to greater recall bias.

Table 15: Advantages and Disadvantages of Shorter and Longer Recall Periods.

Advantages	Disadvantages
<i>Shorter Recall Periods (Between 3-6 Months)</i>	<i>Shorter Recall Periods (Between 3-6 Months)</i>
<ul style="list-style-type: none"> ▪ Less sensitive to recall bias. ▪ Less sensitive to changes in household composition. ▪ More relevant for immediate programmatic purposes. 	<ul style="list-style-type: none"> ▪ Require a larger a sample size. ▪ Miss earlier mortality data that may have had an impact on the population of interest. ▪ More sensitive to “calendar error” (while respondents may have fewer problems recalling more recent mortality events, they may have trouble distinguishing a death that occurred 11 months ago from one that occurred 13 months ago). ▪ May include a relatively small number of deaths, even in a crisis, magnifying the effect of an erroneous count.
<i>Longer Recall Periods (Between 6-12 Months)</i>	<i>Longer Recall Periods (Between 6-12 Months)</i>
<ul style="list-style-type: none"> ▪ Smaller sample size needed for the same precision, potentially saving resources and time. ▪ If deaths are recorded for specific parts of the recall period, it is possible to look at sub-intervals of time (e.g. before and after major violence). 	<ul style="list-style-type: none"> ▪ Mortality rate may be less relevant for immediate programmatic purposes. ▪ Recall bias (accurate recall of events in terms of specific details — dates, ages, locations, circumstances, etc. — is more unreliable).

While it is critical for an implementing organisation to make a decision about the proper recall period based on the objectives of the mortality and demography survey, the local context and calendar also matter. Local context including complex displacement patterns and multiple risk factors need to be taken into account in the design and analysis of mortality. The local calendar is important to consider when determining a recall period both because mortality risk can have high seasonal variation (malaria, for example), and because the recall period should start at a common, well-known date in the past in order to reduce error.

The number of days in the recall period is needed for the mortality sample size calculations and the mortality analysis. **It is important to note that the number of days in the recall period is the number of days from the start date of the recall period to the mid-point date of the data collection.**

If a recall period of approximately 90 days is established for the survey, then in the household interview, the respondent will be asked to identify all persons who were household members between the day of the interview (or the night before) and a date approximately 90 days in the past. As noted previously, this could be a prominent date on the calendar, a religious festival

day, a seasonal marker, or some other date of high relevance and common reference for the community. What matters most is that this start point is clearly defined and understood by all interviewers and respondents. This should be established through qualitative research and pilot-testing of the survey instrument.

The start (prominent date) and end point (survey date) of the recall period are fixed points in time. As a survey question example, consider: “Has anyone in the household died between the beginning of Ramadan and today?” For purposes of calculating a sample recall period, the sample period end point would be mid-point of the data collection. In other words, if the beginning of the recall period is January 1 and the interviews take place during the period of March 24 to April 7, then the mid-point of data collection used to calculate the number of days in the recall period for sample calculation would be March 31. It is important to note that the mortality interviews continue throughout data collection. When the total of days of data collection is odd (for example, 31 days), you can divide by two (15.5) and then round up (so 16) to have the number of days added to the start of the recall period and the start of data collection.

Sample Size Calculation — CDR

With a mortality survey, the calculation of the sample size depends on six elements:

1. Estimated Crude Death Rate;
2. Desired precision;
3. Likely design effect (if the survey is to use cluster sampling);
4. Selected recall period;
5. Estimated household size;
6. Estimated non-response rate.

Because the mortality rate measures the frequency of an event over an interval of time, a sample size calculation must include exposure time as one of the parameters. Calculating the sample size to estimate a mortality rate is similar to the formulas used to estimate prevalence. The difference is that the mortality rate incorporates exposure time into the equation.

Calculating Sample Size: Simple and Systematic Random Samples.

To estimate a rate using a simple or systematic random sampling design, the formula is:

$$n = t^2 \times \frac{CDR}{RP \times d^2}$$

Where:

n = sample size

t = t-score associated with a desired confidence limit (for example, 1.96 = 95%)

CDR = estimated Crude Death Rate

d = desired precision

RP= recall period

Estimated CDR

When estimating the mortality rate for calculating the sample size, it is important to take into account the mortality benchmarks for defining crisis situations. Both CMR and U5DR are commonly-used indicators to set the thresholds or benchmarks used to define a humanitarian emergency and its levels of severity⁵⁶. The most widely-used approach (though certainly under challenge as well) has been to assume a baseline CMR of 0.5 deaths/10,000/day and a

⁵⁶ This section draws significantly from Checchi and Roberts, 2005.

baseline U5DR of 1 death/10,000/day. Double or more those rates (CDR ≥ 1 death/10,000/day and U5DR ≥ 2 deaths/10,000/day) is considered an emergency (Table 16). Mortality estimates must always be presented with confidence intervals and plausibility scores to provide sufficient evidence for interpretation.

The United Nations High Commissioner for Refugees (UNHCR) *Handbook for Emergencies* uses the same baseline indicators, but sets gradations of higher levels of mortality to define emergencies that are “very serious,” “out of control,” and “major catastrophe” (also Table 16). The Sphere Project applies region-specific baseline and emergency rates for crude mortality and under-five mortality (also Table 16).

Table 16: Mortality Benchmarks for Defining Crisis Situations.

Method	Assumed Baseline	Emergency Thresholds
Standard Thresholds	Fixed at: CDR: 0.5/10,000/day U5DR: 1/10,000/day	Emergency if: CDR: $\geq 1/10,000/day$ U5DR: $\geq 2/10,000/day$
Standard Thresholds (with levels of severity)	Fixed at: CDR: 0.5/10,000/day U5DR: 1/10,000/day	Levels of severity: CDR: $> 1/10,000/day$: “very serious” CDR: $> 2/10,000/day$: “out of control” CDR: $> 5/10,000/day$: “major catastrophe” Double for U5DR thresholds.
Context-Specific Thresholds	Fixed at – CDR (U5DR): Sub-Saharan Africa: 0.44 (1.14) Latin America: 0.16 (0.19) South Asia: 0.25 (0.59) Eastern Europe, Former Soviet Union: 0.30 (0.20)	Emergency if - CDR (U5DR): Sub-Saharan Africa: 0.9 (2.3) Latin America: 0.3 (0.4) South Asia: 0.5 (1.2) Eastern Europe, former Soviet Union: 0.6 (0.4)

Table adapted from Checchi and Roberts, 2005.
Context-specific thresholds from SPHERE, 2011.

Precision

It is not always clear what precision is needed in a mortality and demography survey though “at the very least, surveys should have a sample size that is sufficient to clearly identify seriously elevated mortality.”⁵⁷ In other words, if an organisation wanted to be reasonably sure that mortality was around a benchmark of 1 per 10,000 per day, it would be appropriate to establish a precision of ± 0.5 deaths per 10,000 per day. If there was some clear indication that mortality was well above 1 per 10,000 per day, a precision of ± 1 death per 10,000 per day might be appropriate. Where mortality rates are unknown, it is recommended to err on the side of caution and seek a precision of at least ± 0.3 deaths per 10,000 per day.

Design effect

If a cluster sample design is used, then a design effect would need to be added to the sample size calculations. As noted previously, using a cluster sample design requires a larger sample size to obtain the same precision as a simple or systematic random sample design. Because subjects within the same cluster are generally more similar to each other than to members of different clusters, there is a net loss of independent data, which results in decreased precision. The imprecision of cluster sampling is compensated for by multiplying the sample size calculated for a simple random sample by an inflation factor called the “design effect” (DEFF),

⁵⁷ Checchi and Roberts, 2005.

defined as “the ratio of the variance obtained with cluster sampling to that obtained with simple random sampling.”⁵⁸

Calculating Sample Size: Cluster Sampling.

To estimate a rate using a cluster sampling design, the formula is:

$$n = t^2 \times \frac{CDR \times DEFF}{RP \times d^2}$$

Where:

n = sample size

t = t-score associated with a desired confidence limit (for example, 2.045 = 95%)

CDR = estimated Crude Death Rate

d = desired precision

RP= recall period

DEFF = design effect

Design effects in crisis settings can vary from 1.1 (if the population is relatively homogeneous between and among clusters) to 4 or higher where some clusters are severely affected while others are not. A study of nine cluster sampling surveys conducted in emergency settings between 1993 and 2004 found that for estimates of crude mortality, observed design effects ranged from 1.1 to 4.0, while for estimates of under-five mortality, observed design effects ranged from 0.9 to 2.1⁵⁹.

Probably the best sources of estimates for the design effect are prior surveys done in the same population or data from comparable populations and/or situations. Design effects from prior surveys, however, should be used with caution if the sampling designs are different or if there is reason to think that the patterns and causes of mortality are not comparable. Generally speaking, if a survey is seeking to measure crude mortality (all causes and all ages), and particularly if there is expected to be some level of violence-related mortality affecting differently the areas to be surveyed, using a design effect of 2 or higher might be appropriate. In measuring mortality among children under five, however — perhaps because it is due more to malnutrition and disease than to violence, even during conflict — using a design effect of 1.5 to calculate a sample size may be sufficient.

It is recommended that organisations seeking to carry out a mortality and demography survey should first explore what local and/or comparable data may exist to select a design effect used for estimating sample size. Choosing a design effect of 1.5 as a “default” setting may be appropriate for estimating mortality in many crisis settings, but it could prove too high or too low, depending on what sort of clustering around mortality events is observed.

Recall Period

The recall period is a time interval, providing the denominator for mortality as “person-time” lived in the interval and being at risk of dying, whether that time is measured in days, weeks, months, or years. In a survey, it obviously matters a great deal whether 10,000 household members are included in a survey about deaths in the last 24 hours, or 1,000 household members in a survey about deaths in the last ten days, or 100 household members in a survey about deaths in the last 100 days. In all three cases, the denominator is roughly 10,000 person-days (a person-day being the number of people in the population multiplied by the number of days in the recall period, or interval of interest).

⁵⁸ Katz and Zeger, 1994.

⁵⁹ Kaiser R, Woodruff BA, Bilukha O, Spiegel PB, Salama P. Using design effects from previous cluster surveys to guide sample size calculation in emergency settings. *Disasters* 2006;30(2):199-211.

As a matter of straight mathematics, it is true that extending the recall period can reduce the sample size (Table 17). But, as was previously noted, the recall period should be established based on the objectives of the survey, the mortality experiences of interest, and with regard to local calendars and timelines. The decision to interview 100 people about deaths in the past 100 days versus 1000 people about deaths in the past ten days, in other words, should not be based on concerns for sample size efficiency. That said, adding two or three weeks to a recall period (considering survey objectives and local context) in order to achieve a usefully smaller sample size (also in consideration of local context and resources) may be quite appropriate.

Table 17: The Effect of Changing Recall Period in Mortality & Demography Survey Sample Size Calculations.

CDR /10,000/day	Desired Precision /10,000/day	Recall Period (days)	Household size	Households to sample		
				Design = 1.0	Design = 1.5	Design = 2.0
1	0.4	60	5	796	1,193	1,591
1	0.4	90	5	529	793	1,058
1	0.4	120	5	395	593	791
3	1.0	45	5	505	758	1,011
3	1.0	60	5	377	566	754
3	1.0	90	5	249	374	498
5	2.0	30	5	315	473	631
5	2.0	45	5	209	313	417
5	2.0	60	5	155	233	311

Household size

As noted earlier, the basic sampling unit for a mortality and demography survey is **always** the household, including anyone who was a member of the household for all or part of the recall period. While cluster sampling will likely rely on estimates of the numbers of households at a given stage in the sampling in order to apply the probability proportional to size (PPS) methodology, and the sample size generally will be referenced in terms of the number of households to be interviewed, the denominator used for calculating mortality rates comprises individual people and the “person-time” each of them contributes to exposure to mortality risk in the household. Sample size calculations, therefore, must estimate the average size of the households to be included in the survey.

The concept of a household is both complex and malleable, and definitions are diverse. Demographic and Health Surveys in sub-Saharan Africa define a household as “a person or a group of persons, related or unrelated, who live together and share a common source of food.”⁶⁰ Based on this definition, the Tanzanian DHS for 2004-2005 estimated the average household size to be 4.9 persons. Though this is not the specific derivation, the Tanzanian average household size approximates that of other sub-Saharan countries, thus it is not surprising that the recommended “default” for estimated household size in calculating sample size in mortality and demography surveys is five persons per household.

It should be emphasised that while the definition of household provided here is broadly accepted, it refers to a “statistical household”⁶¹. It is recommended to research local terms and concepts to be sure that interviewers and respondents alike are referring to the same concept when discussing household members and events. In crisis and displacement contexts, it is also likely that households will break up, move, and reconstitute themselves in a dynamic way. This, too, needs to be explored in a local context, as some of the issues of household membership that make enumeration complex even in static situations, such how to account

⁶⁰ Bureau of Statistics [Tanzania] and Macro International Inc. *Tanzania Demographic and Health Survey 2004-2005*. December 2005. Calverton, Maryland: Bureau of Statistics and Macro International Inc.

⁶¹ Van de Walle E, ed. *African Households: Censuses and Surveys*. Armonk, NY: M.E. Sharpe Inc; 2006.

for migrants, lodgers/boarders, visitors, multiple wives, children away at school, or adults away at work, for example, could pose even greater challenges in emergencies.

Non-response rate

A final element to be considered in estimating sample size is the estimated non-response rate, including households visited where no one is willing to participate in the survey, or where no one is home⁶². While there is no clear evidence of what level of non-response rate might be expected for surveys in crisis settings, it is generally accepted that for personal interview surveys, a response rate of 80% is considered good and 90% is exceptional. It is recommended, therefore, that unless there is good reason to expect otherwise (better or worse), a non-response rate of around 10% should be factored into sample size calculations. More than 10% would be problematic and could require either re-sampling or post-sampling adjustments (including weighting of different sized clusters); less than 10% seems unrealistic in crisis settings.

Table 18 below gives examples of estimated sample sizes (calculated using the ENA for SMART software), assuming different levels of estimated CDR, desired precision, recall period and design effects (all assuming an average household size of five and an estimated non-response rate of 10%). Looking at Table 18, which shows calculations for estimating CRD, several patterns can be observed:

- **All other inputs remaining constant, the higher the desired precision, the larger the sample size needed.** For a given design effect and recall period, greater desired precision (that is, a narrower confidence interval) will require a larger sample size. Seeking the same level of precision with a higher level of mortality (thus, greater population variability) will require a larger sample size.
- **All other inputs remaining constant, the longer the recall period, the smaller the sample size needed.** Because the denominator in mortality rates is a measurement of “person-time”, and not the number of people, the longer the period of recall, the larger the sample size of person-days of exposure.
- **All other inputs remaining constant, the larger the design effect, the larger the sample size needed.** A larger design effect requires a larger inflation factor to maintain an effective sample size equivalent to that for simple or systematic random sampling.

Table 18: Calculating Sample Sizes for Mortality & Demography Surveys (CDR).

Estimated CDR	Desired Precision	Recall Period	Household Size	Number of Households	Number of Households
10,000/day	10,000/day	Days	Persons	DEFF = 1.5	DEFF = 2
0.5	0.3	30	5	2581	3442
0.5	0.3	93	5	833	1110
0.5	0.3	180	5	430	574
0.5	0.3	275	5	282	375
0.5	0.3	360	5	215	287
1	0.5	30	5	1859	2478
1	0.5	93	5	600	799
1	0.5	180	5	310	413
1	0.5	275	5	203	270
1	0.5	360	5	155	207
2	1	30	5	929	1239
2	1	93	5	300	400
2	1	180	5	155	207
2	1	275	5	101	135
2	1	360	5	77	103

Note: A non-response rate of 10% is factored into the sample size calculations.

⁶² Working Group for Mortality Estimation in Emergencies, 2007.

Sample Size Calculation — U5DR

Assuming that the objective of the survey is not simply to measure crude mortality but under-five mortality as well, an additional sample size calculation is necessary. Table 19 below gives examples of estimated sample sizes (calculated using the ENA for SMART software), assuming different levels of estimated under-five mortality, desired precision, recall period, and design effects, all assuming an estimated non-response rate of 10%.

If children under the age of five comprise 20% of the population in higher fertility settings, then it can be assumed that every fifth person in the population will be under the age of five. On average, a household of five people will include one child under five years old.

If one objective of the mortality and demography survey is to establish a benchmark of the severity of the crisis, it is probable that sufficient precision will be needed to establish whether under-five mortality is above or below the level of 2 deaths/10,000/day \pm 1 death/10,000/day. And assuming that a recall period is selected of between 93 and 180 days and that the design effect will be between 1.5 and 2, sample size for estimating under-five mortality would need to include between 711 and 2,371 households (Table 19). While this lower number might be plausible for a survey in a crisis setting, the upper figure is likely to exceed the costs and capacities for an average survey. One option (not recommended unless the study objectives and context call for it) would be to increase the recall period to closer to 270 days, or nine months. Another approach would be to relax the desired precision from \pm 1 death/10,000/day to \pm 1.5 deaths/10,000/day.

As can be seen in Table 19, if U5DR is estimated to be 2/10,000/day and desired precision is set at \pm 1.5 deaths/10,000/day, then, with a recall period between 90 and 180 days and a design effect between 1.5 and 2.5, a sample size of under-five children would need to include from 316 to 1,054 households (factoring a non-response rate of 10%) in order to include the 285 to 949 children needed. Because children under the age of five are only a percentage of the total population, measurements of under-five mortality in a sample of between 300 and 900 households (a plausible range for household mortality and demography surveys in crisis settings) will be less precise than measures of crude mortality: as seen from the example, confidence intervals would be three times as large, all other inputs remaining constant. If that level of imprecision is not acceptable, it will be necessary to consider a larger overall sample size for the mortality and demography survey in order to measure under-five children (or other population sub-groups) with the desired precision.

Table 19: Calculating Sample Sizes for Mortality & Demography Surveys (U5DR).

Estimated U5DR	Desired Precision	Recall Period	Average Number of U5 Children per Household	Number of Households	Number of Households
10,000/day	10,000/day	Days	Children	DEFF = 1.5	DEFF = 2
1	0.5	30	1	2134	2846
1	0.5	93	1	711	949
1	0.5	180	1	356	474
1	0.5	275	1	237	316
1	0.5	360	1	178	237
2	1	30	1	1897	2529
2	1	93	1	632	843
2	1	180	1	316	422
2	1	275	1	211	281
2	1	360	1	158	211

Note: A non-response rate of 10% is factored into the sample size calculations.

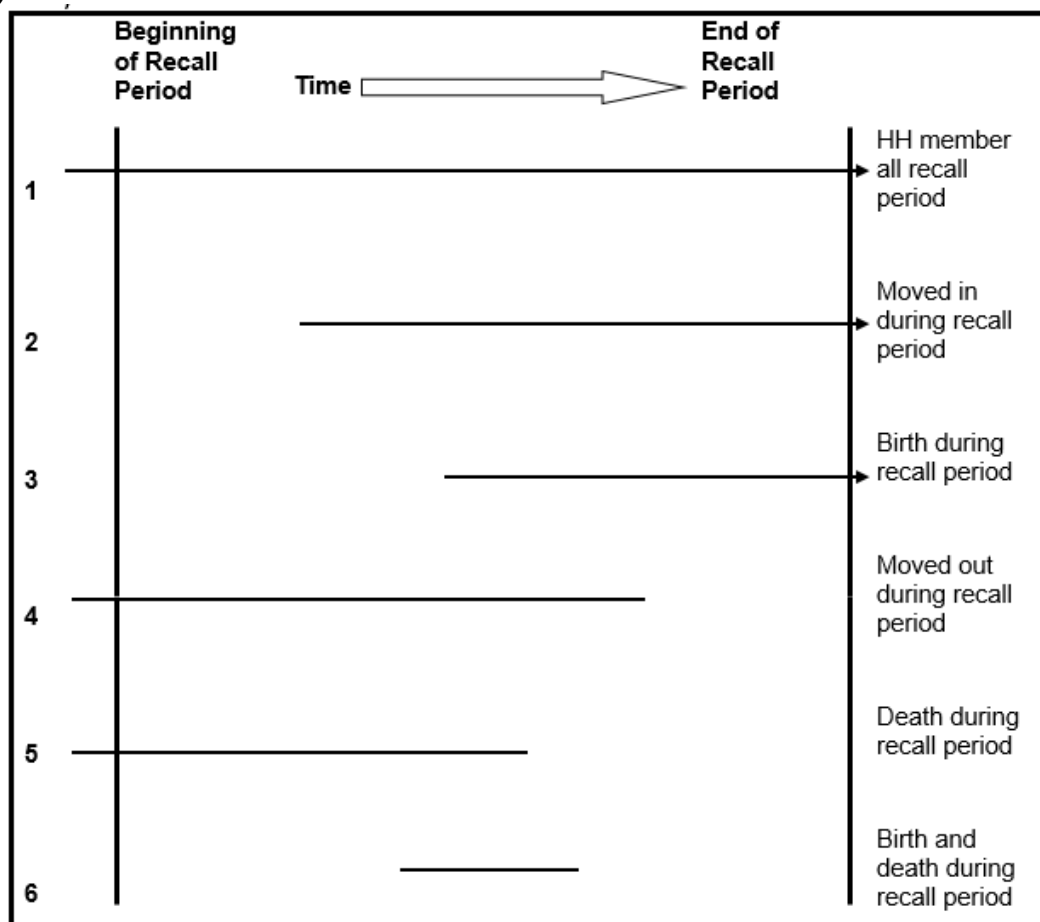
Enumerating Household Members

To estimate a mortality rate using a retrospective household survey, a denominator is needed, preferably one that most closely approximates “person-time” lived during the interval of interest. The household events (births, deaths, in-migration and out-migration) that are measured within the recall period both changes the number of people within the household at different times within the interval and affects the exposure time that each of them contributes to risk of mortality within the household.

Before exploring how household changes during a recall period might best be captured, it is important that a clear and contextually appropriate definition of household be established for use in the mortality and demography survey. As noted previously, one broadly accepted definition of household is “a person or a group of persons, related or unrelated, who live together and share a common source of food”⁶³. This definition might be an appropriate starting point, but it should be culturally-normed and explored in the context of the crisis. Of particular importance will be identifying how household membership is affected by patterns of short-term migration in and out of the household, by multiple marriage, or by work and schooling patterns. Also important would be clarifying definitions and concepts of live birth and neonatal deaths, as these could affect counts of household events as well.

Looking at Figure 37 below might help clarify the types of events that households experience during a given interval of interest, how they contribute to “person-time” lived in the interval, and how they might best be captured in a retrospective household survey.

Figure 37: Household Members During the Recall Period (Adapted from Woodruff, 2002).



⁶³ Tanzania Demographic and Health Survey 2004-2005. December 2005.

As the diagram illustrates, there are six individuals who are present in the household for all or part of the recall period (six months, for instance).

- Person 1 is a member of the household for the full recall period, and is in the household at time of interview.
- Person 2 moves into the household during the recall, and is in the household at time of interview.
- Person 3 is born to a household member during the recall period, and is in the household at time of interview.
- Person 4 is a member of the household at the beginning of the recall period, but moves out at some point during the recall period.
- Person 5 is a member of the household at the beginning of the recall period, but dies during the recall period.
- Person 6 is born to a household member during the recall period, and dies during the recall period.

There are several ways that an enumeration of household members can be carried out in the context of a retrospective household survey. The SMART Methodology recommends using the “full household census,” which is the approach used in the ENA for SMART software⁶⁴. Using the “full household census” method, it is possible to obtain information not only about births and deaths in the household during the recall period, but also about movements into and out of the household. Given that migration is likely to contribute more significantly to household change during times of crisis, it would seem prudent to account for these patterns not only for purposes of calculating a “person-time” denominator, but also because migration and/or displacement may be associated with greater household risk for mortality.

The “full household census” approach provides a more accurate denominator (and possibly numerator as well) in estimating household exposure time to risk of mortality. It also captures information about migration patterns and provides potentially important insights into household dynamics. The disadvantages of the approach are that it is likely to be more time-consuming during the planning and implementation phases.

During the planning phase, it is recommended that the survey team engage in exploratory field work to identify patterns of migration and/or displacement during the interval of interest, whether that is a few months, a year, or more. Training of interviewers should focus not only on definitions of terms like “household” and “family”, but also on how recent patterns of mobility may have affected households and families. It will be particularly important to explore how migration has been and is being used as a coping strategy and what relationship it may have to other household dynamics like employment, education, marriage, as well as births and deaths.

Mortality and Demography Questionnaires

Household survey questionnaires come in many forms and varieties, and mortality and demography surveys are no exception. Specialised modules may focus on specific age groups (neonatal, infant, under-five, youth and adolescent, women of childbearing age, older adults, etc.), specific causes of death (infectious disease, chronic conditions, injury, violence, etc.), specific settings (rural, urban, developed country, developing country, etc.), or the particular factors and conditions affecting mortality. In crisis settings — and particularly in acute-phase emergencies — mortality and demography surveys generally focus on answering a core set of basic questions about household demographics (including household size) and the total number of deaths, as well as deaths of children under five. In some cases, cause of death is

⁶⁴ In previous discussions and presentations of the SMART Methodology, the “full household census” was sometimes referred to as the “hybrid” approach.

investigated, often to distinguish violent and non-violent deaths in conflict settings, but also sometimes to conduct somewhat more in-depth verbal autopsies.

While mortality is the ultimate end of every human life and therefore a universal condition, it is in fact a complex phenomenon to measure, occurring as it does at any time, from one second to more than one century after birth, and for a myriad of proximate reasons, ranging from infection to gunshot wound to more distal determinants like genetics, environment, behaviour, household dynamics, social inequalities, and the like. The motives for asking about mortality are also diverse, ranging from a wish to identify current or recent risk factors in a local area to a desire for a “big picture” of mortality conditions in a country across a span of a year or more.

After consultation with numerous field practitioners and experts in mortality, SMART recommends using the following mortality questionnaire (Figure 38), adapted from a survey developed by the World Food Programme, the United Nations Children’s Fund, and the Centers for Disease Control and Prevention to estimate mortality in Darfur in 2005⁶⁵. This questionnaire includes a core set of questions that should be asked in any retrospective household survey. It is recommended that this form be used as a starting point for developing a mortality questionnaire in a given setting. It may be added to and possibly simplified (though there are caveats to a more simplified approach). Even organisations deciding to use it more or less “as is” should review it carefully to consider its application in the specific context of implementation and for the identified purposes.

As presented previously, the following information is needed to estimate the mortality rate:

- Total number of people in the household at the start of the recall period;
- Total number of births in the household during the recall period;
- Total number of people who moved into the household during the recall period;
- Total number of people who moved out or went missing from the household;
- Total number of deaths in the household during the recall period.

Before beginning the survey in a household, the respondent needs to be identified according to the survey protocol and their consent (verbal or written) should be obtained before administering the questionnaire. It is also important to identify who should be counted as a household member, depending on the established definition and the local context.

The questionnaire has three main sections:

- A. Current household members;
- B. Household members that have left since the start of the recall period, and;
- C. Household members that have died since the start of the recall period.

Each section should be filled out in its entirety before moving on to subsequent sections.

⁶⁵Talley L, Richardson L, Bilukha O, Brennan M, Blanton C. *A validation of mortality estimation methodologies at the household level: summary and household census methodologies*. A collaboration between the Centers for Disease Control and Prevention and the United Nations World Food Programme. May 2006.

Figure 38: Demography and Mortality Questionnaire.

DEMOGRAPHY & MORTALITY QUESTIONNAIRE

DATE OF INTERVIEW: [D][D]/[M][M]/[Y][Y]

COUNTRY:	REGION:	NAME OF INTERVIEWER:
CLUSTER NO. [][]	TEAM NO. [][]	HOUSEHOLD⁶⁶ NO. [][]

01	02	03	04	05	06	07	08	09	10
No.	Name	Sex (M/F)	Age (years)	Joined after:	Left on or after:	Born on or after:	Died on or after:	Cause of death	Location of death

				(Start date of recall period (ex. Jan 1 2017))					
WRITE 'Y' for YES. Leave BLANK if NO.									
a) List all the household members that are currently living in this household.									
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
b) List all the household members that have left this household (out migrants) since the start of the recall period.									
1					Y				
2					Y				
3					Y				
4					Y				
5					Y				
c) List all the household members who died since the start of the recall period.									
1							Y		
2							Y		
3							Y		
4							Y		
5							Y		

Was anyone in the household pregnant at the start of the recall period? No [] Yes []
If yes, how many? _____

⁶⁶ Insert the household definition used in the survey in footnote here.

Filling out the questionnaire

For each section, the interviewer begins by listing information about the respondent, followed by spouse(s), parents, siblings, children and then other current household members. **Children less than one year of age are noted as “0” under the column for age (Column 04).** The shaded columns should remain blank.

Section A: Household members that are currently living in the household

Ask the head of the household to list the names, sex (M/F), and age (in years) of all household members that are currently living in the household.

Once the listing is completed, ask the respondent which of the individuals listed were present at the beginning of the recall period. Indicate “Y” in Column 05 for anyone who was not present at the beginning of the recall period. Marking a “Y” indicates that they joined after the start date of the recall period. If the individual was present at the beginning of the recall period, leave Column 05 blank.

For every child whose age is indicated as “0”, ask the head of the household if the child was born on or after the beginning of the recall period. Indicate “Y” under Column 07 if the child was born on or after the beginning of the recall period.

Once section A is completed, read the information back to the respondent. Ensure that no household member that is currently living the household was forgotten. Move to section B.

Section B: Household members that have left since the beginning of the recall period

Ask the head of the household to list the names, sex (M/F), and age (in years) of all individuals that were part of the household but have left since the beginning of the recall period. List them under section B.

Ask the respondent if anyone listed in section B was present at the beginning of the recall period. Indicate “Y” in Column 05 for anyone who was not present at the beginning of the recall period. Marking a “Y” indicates that they joined after the start date of the recall period. If the individual was present at the beginning of the recall period, leave Column 05 blank.

For every child whose age is indicated as “0”, ask the head of the household if the child was born on or after the beginning of the recall period. Indicate “Y” under Column 07 if the child was born on or after the beginning of the recall period.

The column for “Left on or after” (Column 06) has been prefilled with the letter “Y” in section B. Do not write anything in Column 06.

Once section B is completed, read the information back to the respondent. Ensure that no household member that has left since the beginning of the recall period was forgotten. Move to section C.

Section C: Household members that have died since the beginning of the recall period

Ask the head of the household to list the names, sex (M/F), and age (in years) of all individuals that were part of the household but have died since the beginning of the recall period. List them under section C.

Ask the respondent if anyone listed in section C was present at the beginning of the recall period. Indicate “Y” in Column 05 for anyone who was not present at the beginning of the recall period. Marking a “Y” indicates that they joined after the start date of the recall period. If the individual was present at the beginning of the recall period, leave Column 05 blank.

For every child whose age is indicated as “0”, ask the head of the household if the child was born on or after the beginning of the recall period. Indicate “Y” under Column 07 if the child was born on or after the beginning of the recall period.

If the survey is collecting cause and/or location of death, ask respondents about the circumstances of the death. Indicate the cause of death (Column 09) and location of death

(column 10) using the codes provided in the questionnaire for everyone listed as died in section C. It is recommended to separate causes of death into the following categories that were set in the ENA for SMART software:

Cause of death	
1] Unknown	5]
2] Injury/Traumatic	6]
3] Illness	7]
4]	

Once section C is completed, read the information back to the respondent. Ensure that no household member that has died since the beginning of the recall period was forgotten.

Pregnancies

Ask the respondent if anyone in the household was pregnant at the start of the recall period. If the respondent answers yes, record the number of women who were pregnant at the beginning of the recall period and ascertain the outcome of each pregnancy. Make sure that births and possibly deaths that resulted from these pregnancies are reflected in sections A, B or C.

Ensure that any neonatal death is recorded on the form. A neonatal death is defined as any child who was born (live birth: **the child took at least one breath**) and died within the recall period. A neonatal death should be recorded as BOTH a birth and a death. Pregnancies that do not result in a live birth (abortions, miscarriage, and stillbirths) are not counted as births or deaths.

Adapting the Questionnaire

The questionnaire above is a template. It may need to be adapted to account for different contexts and survey needs. The following should be considered before finalising the questionnaire:

- **Start date of the recall period.** Type in the start date of the recall period before printing the questionnaire to ensure everyone is using the same start date. The date should be written clearly (e.g. January 2, 1990, NOT 01/02/1990).
- **Cause of death.** Determine whether the survey will collect information about cause of death. If cause of death is not going to be collected, remove Column 09 from the template. If cause of death is desired, SMART recommends to only differentiate between traumatic and non-traumatic deaths. Without advanced training, taking a verbal autopsy is complicated and therefore data may not be reliable. If cause of death is to be collected, add a key with a clear code for each cause at the bottom of the questionnaire. Instruct surveyors to fill in column 09 with the appropriate code.
- **Place of death.** Determine whether the survey will collect information about place of death. If place of death is not going to be collected, remove Column 10 from the template. If place of death is desired, SMART recommends only a few categories (e.g. in the refugee camp, in transit, in host country, other). If place is to be collected, add a key with a clear code for each option at the bottom of the questionnaire. Instruct surveyors to fill in column 10 with the appropriate code.
- **Household definition.** Write the household definition surveyors should use for the survey at the bottom of the questionnaire (footnote 1).
- **Number of lines.** The template has 20 lines for section A, 5 lines for section B, and 5 lines for section C. Depending on the context, the number of lines in each section may need to be adjusted.

Preliminary issues

To improve the likelihood of obtaining a reliable enumeration of household members during the recall period, several steps should be reviewed during the training of the surveyors:

- **Define “household.”** As noted previously, the concept of a household is complex and definitions are diverse. Demographic and Health Surveys in Sub-Saharan Africa define a household as “a person or a group of persons, related or unrelated, who live together and share a common source of food”⁶⁷. While this definition of household may be widely accepted, it refers to a “statistical household”⁶⁸. It is highly recommended that local terms and concepts be explored before beginning the survey for cultural norming in order to be sure that interviewers and respondents alike are using the same definition for a household and that interviewers are able to clearly explain the definition to the respondent. In crisis contexts, it is also possible that households will break up, move, and reconstitute themselves in a highly dynamic way, particularly after a death⁶⁹. The survey team should conduct formative research to explore household composition, including shorter-term and longer-term movements into and out of the household during ordinary circumstances — including lodgers/boarders, visitors, multiple wives, children away at school, adults away at work, etc. — as during periods of crisis —when, for example, orphaned children may join another household following the deaths of their parents, or whole households may disintegrate and reconstitute in different patterns during displacement. In situations involving significant displacement, it will be important to clarify where the respondent and his/her household were living at the start of the recall period.
- **Establish survey protocols for identifying a household respondent.** During the planning phase of the mortality and demography survey, the team should establish common procedures to identify the household’s most appropriate respondent and provide the individual an opportunity to give informed consent. The most appropriate respondent will depend on context. The person identified as the respondent should be able to provide the most accurate demographic information about the family; this may or may not be the head of household. For its Multiple Indicator Cluster Surveys (MICS), UNICEF recommends that the respondent be a “knowledgeable adult or mother/primary caretaker of children in the household”⁷⁰. For a more accurate record of births and deaths of children, it is preferred that an adult female serve as the respondent⁷¹. In some cases (especially in child-headed households), an older child (15 to 17 years old) may serve as a respondent, though this may raise concerns about accuracy of recall and ethics (interviews with children may also require consent of a parent or guardian). It may also be useful also to establish protocols as to when illness, frailty, or other conditions may exclude an individual from eligibility.
- **Clarify the start date of the recall period.** While many sample sizes are calculated with a Western calendar interval in mind (approximately 90 days, six months, etc.), the best start date for a recall period may not necessarily be 30, 60, or 120 days prior to the interview, but a day that is likely to be most memorable to the population. This could be a religious festival, a national holiday, a date marking a seasonal transition, etc. The salience of this date should be validated with community-based field research. Accurate

⁶⁷ Tanzania Demographic and Health Survey 2004-2005. December 2005.

⁶⁸ Van de Walle E, ed. *African Households: Censuses and Surveys*. Armonk, NY: M.E. Sharpe Inc; 2006.

⁶⁹ Lopman BA, Barnabas R, Hallett TB, Nyamukapa C, Mundandi C, Mushati P, et al. Assessing adult mortality in HIV-1-afflicted Zimbabwe (1998-2003). *Bulletin of the World Health Organization* 2006;84:189-197

⁷⁰ UNICEF, Multiple Indicator Cluster Surveys Regional Training Workshop I - Survey Design, Powerpoint presentation.

⁷¹ Kalter HD, Salgado R, Gittelsohn J, Parades P. *A Guide to Conducting Mortality Surveys and Surveillance*. Published by Basic Support for Institutionalizing Child Survival Project (BASICS II) for the United States Agency for International Development. Arlington, Virginia, June 2004.

recall of this date during the interview can be improved through development and use of a timeline and/or seasonal calendar⁷².

- **Clarify the end date of the recall period.** In conducting an interview in the household, interviewers will be asking the total number of people currently living in the household. It is recommended that “currently” should refer to anyone the respondent would consider a household member (“a person or a group of persons, related or unrelated, who live together and share a common source of food”) at the time of interview. For purposes of calculation and estimation, as was noted previously, the end of the recall period will be the mid-point between the date of the first interview and the date of the last interview in the survey.
- **Migration into and out of the household.** In many societies, even under ordinary circumstances, movements in and out of the household are routine occurrences — a parent may leave to take a seasonal or temporary job, a child may go away to school, a family member may come for a visit. In emergencies, there may be elevated levels of movement into and out of households, both short-term and longer-term, as families seek safety, shelter, food and water, or to meet other basic needs. Generally, the longer the recall period of the survey, the more in-migration and out-migration may occur, thus potentially having a significant effect on the person-days of exposure that household members contribute to household mortality risk. While it is important to measure migration into and out of the household, however, it is also reasonable, under most circumstances, to assume that short-term movements in and out of the household will not significantly affect the mortality estimates. SMART recommends that in-migration should only measure those who joined the household during the recall period and stayed (either up to the current time or until time of death) and that out-migration should only measure those who left the household and stayed away (although if they died while away from the household, that would not be counted as a household death). It is possible that this approach could lead to over- or under-estimation of person-time of exposure if there were significant movements into or out of households just prior to the survey, due to a substantial influx of new arrivals into an area, major displacements, special holiday movements, seasonal labor, etc. Formative research should assess current conditions, recent events, and develop an understanding of ordinary and extraordinary migration patterns at household- and community-levels to identify common situations and determine clear procedures.
- **Persons joining the household during the recall period.** Any persons who joined the household during the recall period should be enumerated. In addition to age and sex of the in-migrant, note that a column (05) in the questionnaire asks about whether this individual joined the household. As stated above, it is recommended that, for purposes of simplification, in-migrants to be counted for mortality estimation should only include those who joined the household during the recall period and stayed (either up to the current time or until time of death). It is essential to account for all persons who joined the household and stayed to the end of the recall period, as well as for people who joined the household and died during the recall period. Persons who joined the household and then left again or go missing should be dropped from the analysis (cancel each other out in the formula of CDR) on the simplifying assumption that they are contributing little in the way of the exposure time in the denominator.
- **Births in the household during the recall period.** From a methodological standpoint, births during the recall period are not household members at the start of the recall period, so they must be accounted for. Young infants not only are most likely to face the highest mortality risk, but their death are most likely to go unreported, whether due to recall error,

⁷² UNICEF/CDC. *Nutrition and Health Survey, Badghis Province, Afghanistan, February – March 2002*. A collaborative survey by: UNICEF and U.S. Centers for Disease Control and Prevention (CDC), 2002.

methodological issues, or misunderstanding due to cultural or other factors of what constitutes a live birth⁷³.

IMPORTANT: According to the WHO definition, a live birth “refers to the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of the pregnancy, which, after such separation, breathes or shows any other evidence of life — e.g. beating of the heart, pulsation of the umbilical cord or definite movement of voluntary muscles — whether or not the umbilical cord has been cut or the placenta is attached. Each product of such a birth is considered live born.”⁷⁴ It is recommended that interviewers receive adequate training regarding the definition of live births as well as neonatal deaths (deaths occurring during the first four weeks after birth), and be able to distinguish these from fetal deaths (sometimes also referred to as stillbirths, spontaneous abortions, or miscarriages), which should not be reported. Consequently, SMART’s approach to capturing infant deaths recommends asking the respondent about any pregnancies that occurred during the recall period and each of their outcome. In addition to having a clear idea of what constitutes a live birth, survey teams need to have an understanding of what local cultural and social factors may lead to under-reporting of infant deaths. These should be established prior to survey implementation through formative research with local key informants.

- **Incorporating a summary table of household members into the questionnaire.** Previously, a summary table of questions that allowed both the interviewer and the respondent to review and summarize the household information recorded in the previous sections was added at the bottom of the questionnaire. This may assist in the identification of people or events that were either previously missed or perhaps double-counted. However, numerous discrepancies have been witnessed in the field with the survey team improperly counting the number of people for each section in the household. Using summary tables are therefore **not recommended**.

⁷³ Alarcón D, Robles M. *The challenges of measuring child mortality when birth registration is incomplete*. Global Forum on Gender Statistics. Rome, Italy. December 2007.

⁷⁴World Health Organization. *Health Status Statistics: Mortality*. (WHO; <http://www.who.int/healthinfo/statistics/indneonatalmortality/en>, accessed on 1 June 2009).

7. Interpretation of Results

Before doing the interpretation of survey results, any errors in the data need to be identified and, if possible, corrected with the cross-checking of the data collection forms. This is done partly during data entry and double data entry. As seen in the plausibility check section of this manual, the ENA for SMART software automatically examines the data to see if there are values outside the usual or expected range and lists them. These values should be reviewed and checked against the original written data collection sheets, if applicable. Any error in data entry should be corrected immediately. A data line entry should **NEVER** be deleted even if it falls outside the acceptable range highlighted in the plausibility check and will not be part of the final analysis.

If certain critical pieces of information are missing from a child's survey record, it will not be possible to include the child in some of the anthropometric data analyses:

- **Age:** If information on age is missing, the child can still be included in the assessment of wasting and oedema, because these do not require age.
- **Sex:** If information on sex is missing, the child can still be included in the assessment of MUAC and oedema. However, the reference population information on height and weight (wasting or stunting) is sex-specific and therefore will be not included in the assessment.
- **Height:** If information on height is missing, the child cannot be included in the assessment of wasting. However, the child can still be included in an analysis of oedema, because any child with oedema is severely malnourished.
- **Weight:** If information on weight is missing, the child cannot be included in the assessment of wasting. However, the child can still be included in an analysis of oedema.

The survey report should always be presented in a standard format and contain all the information that allow the reader to understand why the survey was conducted, the methods used, the population to which the results apply, the results themselves, any additional relevant information, and a summary of limitations and problems encountered. The report can also contain recommendations. These recommendations should not be simply general recommendations, but should be directly supported and justified by the data in the report. Working with nutrition programme managers intervening in the survey area, for instance, would help contextualise the proposed recommendations.

The presentation of the information in a standardised manner ensures that no important information is omitted. It also allows a reader who is familiar with the format to quickly find the particular information he or she is looking for. The ENA for SMART software takes the data that have been entered during the survey, does the analysis, and presents the data in a standard format. It provides headings for all paragraphs that need to be completed by the person responsible for the report. The graphs of the distribution of the combined variables, by gender, should be examined. The quantitative data are also presented in tables. To put the illustrations into the report, the graphs are transferred to the clipboard within the ENA for SMART software and pasted into the report in the appropriate place.

ANNEX 1: Possible Biases for SMART Surveys and Recommendations on How to Avoid Them

Type of Bias	Type of Error and Characteristics	Recommendations
<p>Selection Bias Error due to the data sources chosen to design the sample. When data are inaccurate or imprecise, this generates a systematic error that jeopardizes the sample's representativeness (external validity).</p>	<p>Sampling: When a sample is drawn from a population, the fundamental question should be: Is this sample representative of the larger population of interest? In mortality and demography surveys in particular, estimates can be biased by the fact that mortality could have destroyed whole households - creating a "survivor bias" among the remaining households. Mortality events can also lead to patterns of household movement and dissolution that can affect later selection. Survey teams should critically examine the population lists and estimates used at all stages of sampling to assess their completeness and accuracy. Formative research in the targeted areas should explore patterns of mortality and migration to assess the extent to which these may affect household selection.</p> <p>Non-Response Bias: Non-response refers to situations in which, for various reasons, data are not collected from selected households or respondents.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Check data sources. <input type="checkbox"/> Insist on exploratory research to understand the context properly. <ul style="list-style-type: none"> <input type="checkbox"/> Inflate the final sample by a non-response rate (NRR) when designing the survey. This NRR would depend on the context of the survey (setting, season, etc.) <input type="checkbox"/> Implement a protocol stating how many times a household should be revisited when it is absent the first time.
	<p>Timing Bias: Time bias often occurs when the recall period has not been properly chosen, and does not capture the mortality effects for which it was set. Timing bias may also be due to the time of day selected to conduct the survey. Mortality can be strongly affected by seasonality or time-sensitive factors, whether due to agricultural seasons, disease patterns, or conflict. Failure to account for these can bias mortality upwards or downwards. Survey teams should use (or develop) a seasonal calendar for at least the 12-month period preceding the survey to assess how survey timing (both when the survey is conducted and the length of the recall period used) may affect mortality estimates. Timing bias can also occur depending on the time of day selected for surveys. While early morning or early evening are likely to find more household members at home (rather than out working, going to</p>	<ul style="list-style-type: none"> <input type="checkbox"/> During the exploratory work, ask key-informants about the community's daily and seasonal schedules for occupations. <input type="checkbox"/> Communicate the survey dates sufficiently in advance.

Type of Bias	Type of Error and Characteristics	Recommendations
<p>Measurement Bias</p> <p>Error due to data collection, entry and analysis techniques. When data are inaccurate or imprecise, this generates a systematic error that jeopardizes the accuracy of results (validity).</p>	<p>the market, visiting neighbours, etc.), morning and evening surveys may also be more intrusive on meals, rest, and “family time.”</p>	
	<p>Event Recall Bias: Respondents often do not remember the exact number of deaths during a recall period.</p> <p>Deaths that occur a short time after birth are particularly under-reported. When families are separated because of instability or insecurity, family members often do not know if their relatives are dead or alive. In addition, violent deaths may be recalled more vividly than other types of deaths. Formative research can help identify the main types of mortality present in the population, as well as the type of recall bias that may affect estimates.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Refine the event calendar so that it reflects as closely as possible the events that occurred in the community. <input type="checkbox"/> Avoid recall periods that are too long: limit recall periods to 12 months or less, except if the survey objectives warrant longer periods.
	<p>Response Bias:</p> <p><i>Sensitivity and/or taboos about death:</i> In general, it is not easy to talk about death with a stranger. In some cultures, taboos regarding death might prevent gathering reliable information (in this case, the death rate will be underestimated, since the numerator will be smaller than in reality).</p> <p><i>Information deliberately falsified:</i> In some populations used to humanitarian interventions, some respondents may give false information in order to have aid continue or even increase (in this case, the death rate will be overestimated, since the numerator will be bigger than in reality).</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Formulate specific questions. <input type="checkbox"/> Explain the objectives of the survey to community leaders and get their collaboration to sensitise households about the importance of sharing accurate information during interviews.
	<p>Age Heaping: Respondents tend to round numbers up or down to 0 or 5.</p> <p>In reporting the ages of household members (whether current age or age at death), respondents may round ages. This can be particularly problematic when measuring age-specific mortality among children. In the first year of life (age 0 in completed years), children may be reported as one year old rather than a certain number of months. In other instances, six-year olds may be reported as being aged five.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Ask respondents to report dates of birth (and/or death) rather than age whenever possible, and authenticate them with an official document. <input type="checkbox"/> Make the event calendar as precise as possible. Calendars and timelines may also be of help in determining dates of events.
<p>Instrument Bias: Improper translation or wording of questions may lead to this kind of error.</p> <p>Written survey instruments can contain mistranslations of key terms and concepts (“family” vs. “household”, for example). The content and wording of the questionnaire may also be misleading, and/or the layout and formatting of the questionnaire may make it difficult to record responses.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Translate and back translate the questionnaire. 	

Type of Bias	Type of Error and Characteristics	Recommendations
	<p>Individual questions may be leading, misleading, double-barreled (asking two questions at the same time), ambiguous, or not relevant to the objectives of the survey. Mortality questionnaires should be discussed with community representatives and pre-tested with a small sample of respondents before they are finalised.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Carefully validate the questionnaire with key informants. <input type="checkbox"/> Use role play (for example) during the training of interviewers. <input type="checkbox"/> Use the field test to test the questionnaire.
	<p>Interviewer Bias: Poorly trained or inexperienced interviewers may ask questions incorrectly and/or record the answers incorrectly, skip questions, or rush respondents in an effort to finish quickly. Intonations or different words employed by the interviewer for different households (e.g.: using the word "family" instead of "household") might lead to different understandings of the same concept by the respondents. Similarly, asking questions too rapidly or skipping some of them might increase bias. In situations where interpreters are used, they may misunderstand questions or mistranslate answers. Practice may not make perfect, but proper training and preparation of interviewers (and interpreters, if applicable) is an essential element of any successful survey and adequate time should be built in the preparation phase for interviewer training.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Do a lot of practice during training, and focus on adjusting behaviours. <input type="checkbox"/> Use role play whenever possible. <input type="checkbox"/> Provide a feedback after the field test to adjust behaviours.
	<p>Data Processing Error: This could occur when data is recorded in the field, if data is not correctly entered into the software, or excessively or inappropriately cleaned. In the process of converting written responses on a questionnaire into data for analysis, errors can occur in the stages of data checking (where mistakes and inconsistencies in the questionnaires are not caught before being entered), data entry (answers can be miscoded or otherwise entered incorrectly), and data cleaning (changes are not made or recorded consistently from one case to the next). Proper training and preparation for field surveys also requires support for data entry personnel and systems support.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Organise practice sessions. <input type="checkbox"/> Provide constructive feedback throughout the whole survey, but especially during training. <input type="checkbox"/> Double check the data entered into the software.
	<p>Analysis and Interpretation Error: This type of error occurs when the statistical tests to verify data are not carried out, or are done in an inadequate manner. In the analysis of any data, especially quantitative data requiring statistical analysis, mathematical and conceptual errors can generate flawed results or interpretations. Survey teams need to decide how much data analysis they are prepared to do themselves, how much may need to be tasked to outside experts, and what the trade-offs are.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Assess statistical competencies available in the team in order to objectively determine in which cases expert support is necessary.

ANNEX 2: SMART Methodology – Chronic Malnutrition as the Main Indicator

Sample Size Calculation

General guidelines for precision when estimating a higher prevalence can be found below in Table 1.

Table 1: Rule of Thumb for Precision with High Expected Prevalence.

Expected prevalence	Precision
>20%	±5%
>25%	±6-8%
>40% and ≤50%	±8-10%

Higher expected prevalence allows for lower precision in sample size calculation. Consequently, teams generally spend fewer days in the field when using chronic malnutrition (vs. GAM) as the main indicator for sample size calculation.

Plausibility Check

The **overall score of the summary table** should be used as an indication for further scrutiny of data and highlight the areas that may need more detailed analysis, and should not be used by itself as the primary criterion to validate or invalidate survey results.

Even though the plausibility check was designed for GAM as the main indicator, it can still be used to evaluate the quality of data for chronic malnutrition. In the **summary table** of the plausibility check, the penalty points for overall sex ratio, age ratio and digit preference score for weight, height and MUAC can be used to evaluate **all nutrition data, including chronic malnutrition**. However, the penalty points for flagged data, standard deviation, skewness, kurtosis and Poisson distribution are based on Weight-for-Height (WHZ).

The **body of the report** contains detailed analyses of **all three anthropometric indices** and their distributions (WHZ, Height-for-Age (HAZ) and Weight-for-Age (WAZ)). Table 2 summarizes where to find information for HAZ in the plausibility check.

The body of the report also includes important information on the **percentage of children with no exact birthdays**, a key indicator of the overall quality of age data. This should not be overlooked when interpreting the observed standard deviation for HAZ.

Table 2: Summary of Information in the Plausibility Check on Height-for-Age.

Criteria	Height-for-Age (Chronic malnutrition)	
	Summary Table (Penalty Points)	Report
Flagged Data (SMART flags, ± 3 Z-score from observed mean)	No	Yes
Overall Sex Ratio	Yes	Yes
Age Ratio	Yes	Yes
Digit Preference Score Weight	Yes	Yes
Digit Preference Score Height	Yes	Yes
Digit Preference Score MUAC	Yes	Yes
Standard Deviation	No	Yes
Skewness	No	Yes
Kurtosis	No	Yes
Poisson distribution	No	Yes

ANNEX 3: Mortality & Demography Survey Training Schedule Example

Mortality and Demography Survey

Training on Mortality and Demography focuses on task-specific training for field supervisors, interviewers and logisticians, and data entry personnel. Given an adequate number of trainers, these sessions could be run simultaneously in one day or consecutively (over two or more days). Below is an example of training schedule for a mortality and demography survey that has been adapted from a training schedule developed by the Centers for Disease Control and Prevention and the United Nations Children's Fund for a combination malnutrition and mortality and demography survey in Afghanistan⁷⁵. The schedule presented below presents only the mortality-related elements of the training.

Field supervisors

Role of supervisors

- Quality control procedures:
 - Monitoring interviews;
 - Checking sheets for completeness before leaving site;
 - Taking field notes.
- Stress key points:
 - Importance of standardising questionnaire administration;
 - Importance of consent procedure.

Interviewers and logisticians

Interview training

- Identification of household and household members;
- Recording household and household member numbers;
- Recording non-response (refusal, not home, moved away);
- Introduction to survey (script);
 - Review of each question and its data recording procedure.

Demonstration/practice interviews and role playing

Discussion on question clarity, cultural appropriateness, other problems

Data entry personnel

Role of data entry personnel

- Quality control procedures:
 - Checking questionnaires for completeness before entering data;
 - Coding responses;
 - Data entry systems (dual entry, random check).
- Stress key points:
 - Importance of data checking, coding, cleaning and entry;
 - Importance of standardised procedures for all questionnaires.

⁷⁵ UNICEF/CDC. *Nutrition and Health Survey, Badghis Province, Afghanistan, February – March 2002.*

Field Test

Field test allows practical field practice of final-stage sampling and household interviews, as well as practice data entry.

- Review logistics, field procedures, and data quality control.
- Field practice:
 - All procedures should be practiced at selected households.
- Data entry practice:
 - All procedures should be practiced with selected questionnaires
- Training wrap-up:
 - Discuss problems encountered during field practice and data entry ;
 - Final questions and clarifications.

ANNEX 4: Epi Info/ENA Manual

1. Overview

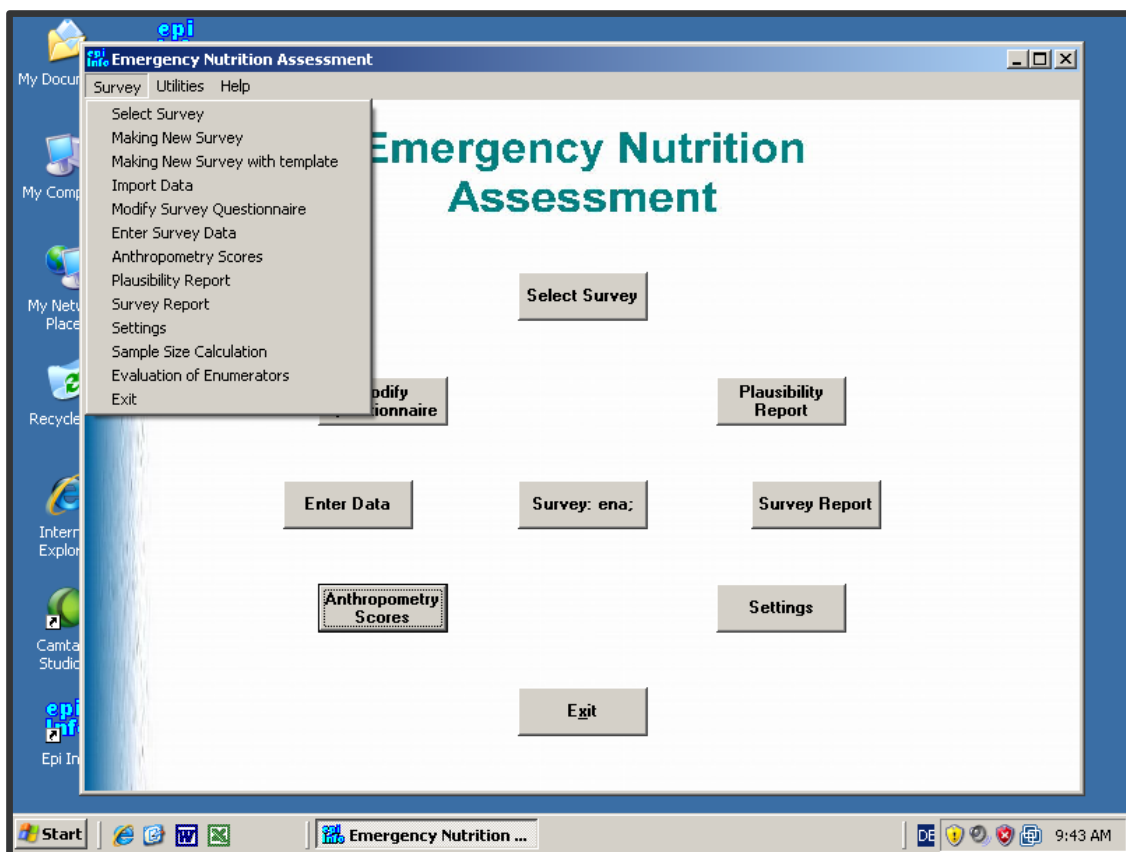
The purpose of Epi Info/ENA is to combine the advantages of the two software programmes, ENA for SMART and Epi Info.

The ENA for SMART software is a user-friendly software to analyse data from mortality and nutrition surveys: to generate anthropometric scores, perform automated analyses of anthropometry and mortality data, perform data quality checks, and generate report templates with populated data tables.

Epi Info provides sophisticated functions for data entry, management, and advanced statistical analysis.

The combined software provides an opportunity to enter data with Epi Info, to evaluate the results with ENA for SMART, and to do further statistical analyses with Epi Info.

The following screen shows the menu of Epi Info/ENA:



The buttons in this menu enable the user to:

- select an existing survey (the current, active one is shown in the central button);
- create or modify the survey questionnaire;
- enter data;
- generate the anthropometry scores (z-scores, % median, flags, etc.) of the survey;
- get the automatic reports (data quality check and survey report);
- change the settings of the software, and;
- generate plausibility check report.

Some more special functions, like sample size calculation or training of the teams for anthropometric measurements, can be accessed via “*Survey*” in the pull down menu. In Epi Info/ENA, a survey usually consists of a questionnaire and a data table. The questionnaire is necessary for data entry, but for survey analysis only a data table is necessary to get the automatic reports.

The sequence of steps to use the software depends on how data is entered or imported into the software.

There are three possibilities:

1. Creating a survey questionnaire from scratch and entering the survey data;
2. Creating a survey questionnaire using the template from the ENA for SMART software and entering the survey data;
3. Importing already entered survey data from another software.

All three procedures can be accessed via “*Survey*” from the pull down menu. After the survey data is entered or imported into the Epi Info/ENA software, the steps for analysing survey data are the same for each of these options.

In the following sections, the installation and functions of the software are described. For the details on handling Epi Info and the interpretation of the automatic reports from the ENA for SMART software, the reader should refer to the Epi Info manual and training tools (<https://www.cdc.gov/epiinfo/index.html>) and the ENA for SMART and Epi Info/ENA manuals (www.smartmethodology.org). This short manual mainly aims to explain features which are specific to the Epi Info/ENA combined software.

2. Software Installation

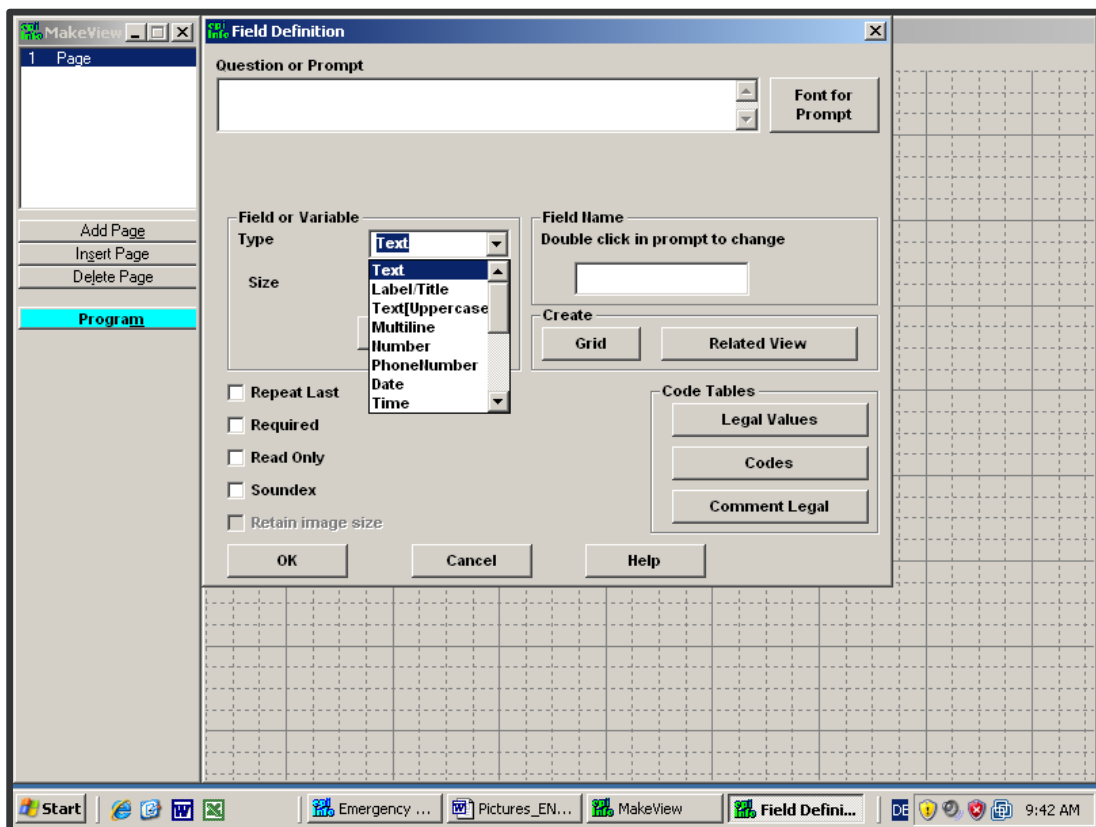
To install Epi Info/ENA:

1. Install **Epi Info** from the Centers for Disease Control and Prevention (CDC) at <https://www.cdc.gov/epiinfo/support/downloads.html> or from the SMART Initiative at <http://smartmethodology.org/survey-planning-tools/smart-emergency-nutrition-assessment/>. During installation, make sure that the installation directory is “C:\Epi_info”.
2. After the installation of Epi Info is finished, download the **installation file for ENA (“enaepi.exe”)** at <http://smartmethodology.org/survey-planning-tools/smart-emergency-nutrition-assessment/> and start the “enaepi.exe” file. The installation programme automatically creates the “ena” folder in Epi Info (c:\epi_info\ena) and copies the programme files there. Additionally, a shortcut is put on the desktop to enable direct access to the software.

To uninstall ENA for Epi Info, delete the “ena” folder (“c:\Epi_info\ena”) and the shortcut icon on the desktop.

Please note that all the survey data which are entered and imported into Epi Info/ENA are saved in the file “**ena.mdb**” in the directory “C:\Epi_info\ena”. Therefore, this is the only file for which a backup is necessary. Everything else can be reinstalled and restored from the websites.

3. Creating a Survey Questionnaire



To create a questionnaire for data entry from scratch:

- Go to menu option “*Survey/Making New Survey*”. Epi Info opens the *Make View* function, which allows to create a data entry form with a variety of possibilities to enter data and to check the validity of the entered data. For detail on how to create a questionnaire in Epi Info, the reader should refer to the online help of *Make View* or the manuals and tools from the Epi Info website mentioned above.
- For convenience of data entry, the electronic form should look as similar as possible to the paper questionnaire (same page layout, sequence of questions, etc.).
- **Very important:** after the questionnaire is finished, it should be saved under a new name in the ena.mdb file (c:\Epi_info\ena\ena.mdb). If it is not saved in this location, Epi Info/ENA will not be able to analyse the data it contains.
- This survey questionnaire can be accessed via the “*Select Survey*” button from the main menu of Epi Info/ENA. The central button then shows the name of the current active survey, and survey data can be entered or analysed.

Below are more detailed instructions to create a few types of fields (questions) commonly used in surveys. For more detail, the reader should refer to the manuals mentioned above.

Open the blank Make View page

1. Select “*Survey/Making New Survey*” in the top menu.
2. The “*Make/Edit View*” window will open.
3. Select “*File/New*” in the top menu.
4. The “*Create or Open Project*” window will open.
5. Browse to “C:\Epi_info\ena\ena.mdb” and click “*Open*”.
6. The “*Name the View*” window will open.
7. Type the name of the survey data file (for example, Kenya1), then click “*OK*”.

The blank view (questionnaire) is now created and saved under the name Kenya1 in ena.mdb file.

Create basic fields

a. title/label

1. Select “*Insert/Field*” in the top menu, or right-click on the empty questionnaire.
2. The “*Field Definition*” window will open.
3. Type Kenya1 Survey (example) in the “*Question or Prompt*” field.
4. Select the font size and style using the “*Font for Prompt*” button on the right.
5. Select “*Label/Title*” in the “*Field or Variable Type*” field.
6. “*Field Name*” will be filled automatically. To edit the field name, double click on the text.
7. Click “*OK*”. The label will be created.

To move the label field (or another field), click and drag the field.

b. text field

1. Select “*Insert/Field*” in the top menu, or right-click on the empty questionnaire.
2. The “*Field Definition*” window will open.
3. Type the Village Name (example) in the “*Question or Prompt*” field.
4. Select the font size and style using the “*Font for Prompt*” button on the right.
5. Select “*Label/Title*” in the “*Field or Variable Type*” field.
6. “*Field Name*” will be filled automatically. To edit the field name, double click on the text.
Note that the field name will be the name of this variable in the data table, so make sure to edit as needed. Once data is entered into this view, the field name cannot be changed.
7. Click “*OK*”. The question (field) will be created.

To move the label field (or another field), click and drag the field.

c. text field with legal values

Text field with legal values is needed when there are pre-defined options for the text field (for instance, Sex can only be Male or Female).

1. Select “*Insert/Field*” in the top menu, or right-click on the empty questionnaire.
2. The “*Field Definition*” window will open.
3. Type Sex in the “*Question or Prompt*” field.
4. Select the font size and style using the “*Font for Prompt*” button on the right.
5. Select “*Text*” in the “*Field or Variable Type*” field.
6. “*Field Name*” will be filled automatically. To edit the field name, double click on the text.
Note that the field name will be the name of this variable in the data table, so make sure to edit as needed. Once data is entered into this view, the field name cannot be changed.
7. Click on “*Code Tables/Legal Values*” in the right bottom corner.
8. The “*Set Up Code/Legal Links*” window will open.
9. Click “*Create New*” button.
10. In the new text field on the left type Male, click “*Enter*” and type Female.
11. Check the “*Do not sort*” box.
12. Click “*OK*” and again “*OK*”. The question (field) will be created.
13. Click on the arrow at the right end of the field to see the legal values for this field.

To move the label field (or another field), click and drag the field.

d. numeric field, specify number of digits before and after comma, put range

This type of field is for entering numeric data, such as height, weight or cluster number.

1. Select “*Insert/Field*” in the top menu, or right-click on the empty questionnaire.
2. The “*Field Definition*” window will open.
3. Type Height in the “*Question or Prompt*” field.
4. Select the font size and style using the “*Font for Prompt*” button on the right.
5. Select “*Number*” in the “*Field or Variable Type*” field.
6. In the “*Pattern*” field, enter the number of digits and decimal required using hashtag signs: for example, for height in centimetres, we need to allow for three digits and one decimal - therefore, enter “###.#”.
7. “*Field Name*” will be filled automatically. To edit the field name, double click on the text.
Note that the field name will be the name of this variable in the data table, so make sure to edit as needed. Once data is entered into this view, the field name cannot be changed.

In this field, it is also possible to specify a range. For example, if children six to 59 months are to be surveyed, values for height would not be expected to be lower than 50cm and more than 120cm.

8. Check the “*Range*” box. Range fields will appear. Enter lower and higher range values, for example, enter 50 in “*Lower*” and 120 in “*Higher*”.
9. Click “*OK*”. The question (field) will be created.

To move the label field (or another field), click and drag the field.

e. date field with different formats

This type of field is for entering a date, such as date of birth or date of survey.

1. Select “*Insert/Field*” in the top menu, or right-click on the empty questionnaire.
2. The “*Field Definition*” window will open.
3. Type Date of Birth in the “*Question or Prompt*” field.
4. Select the font size and style using the “*Font for Prompt*” button on the right.
5. Select “*Number*” in the “*Field or Variable Type*” field.
6. In the “*Pattern*” field, select the date format.
7. “*Field Name*” will be filled automatically. To edit the field name, double click on the text.
Note that the field name will be the name of this variable in the data table, so make sure to edit as needed. Once data is entered into this view, the field name cannot be changed.

In this field, it is also possible to specify an acceptable range of dates. For example, if children six to 59 months are to be surveyed and data collection takes place on January 1, 2008, only children whose birthdate is after January 1, 2003 and before July 1, 2007 would be included.

8. Check the “*Range*” box. Range fields will appear. Enter lower and higher range values, for example enter 01-01-2003 in “*Lower*” and 01-07-2007 in “*Higher*”.
9. Click “*OK*”. You will see the question (field) created

To move the label field (or another field), click and drag the field.

Create and name new page

If there is no more space on the first page, more pages can be created.

1. To add a page, go to the narrow window on the left of the screen and click “*Add Page*”.
2. A new blank page is created after the existing page, and the line 2 *Page* is added in the small window in the upper left corner. Note, “*Insert Page*” will create the new page before (not after) the existing page.

It is now possible to start entering fields on this page. Note that if no question is entered before moving to another page or closing the questionnaire, this page will be deleted. To rename the page, right-click on *2 Page* in the small window on the left side and enter new name of the page.

The button “*Delete Page*” on the left side will delete the current page.

Change background color

With Epi Info, the appearance of the pages can be changed using the “*Format*” option in the top menu. This option allows for a wide array of formatting actions.

For example, to change the background colour of the page:

1. Select “*Format/Background*”.
2. The “*Background*” window will open.
3. Click on “*Change colour*”, select the colour in the window, and click “*OK*”.
4. At the bottom of the “*Background*” window, select whether this colour should be applied only to the current page or to all pages in the questionnaire. Click “*OK*”.

The background colour is now changed.

Edit or delete an existing field

1. Double click on the field. The “*Field Definition*” window will open.
2. Make the edits and click “*OK*”.

To delete the field, double click on the field and click “*Delete*” in the “*Field Definition*” window.

Save, close, open and edit view (questionnaire) in ena.mdb file

To save and close the finished questionnaire:

1. Select “*File/Save*” and then “*File/Exit*”.
2. The “*New Data Table*” window will appear asking to confirm the creation of a data table for this questionnaire. Click “*OK*” to create a data table, otherwise click “*Cancel*”.
Note, once a data table is created, the field name can no longer be changed in the “Field Definition” window.

To open the view (questionnaire) and add new fields (questions) later on:

1. Select “*Survey/Making New Survey*” in the top menu of the main Epi Info/ENA screen.
2. The “*Make/Edit view*” window will open.
3. Select “*File/Open*” in the top menu.
4. The “*Create or Open Project*” window will open.
5. Browse to “*C:\Epi_info\ena\ena.mdb*” and click “*Open*”.
6. Select the view to open (for example, Kenya1), click “*OK*”.

4. Making a Survey with the Questionnaire Template of EpiInfo/ENA

Creating the data entry screen

The screenshot displays the 'MakeView' application window. The title bar reads 'Make/Edit View: C:\EPI_INFO\ENA\ENA.MDB:ENA' and 'Page:1'. The menu bar includes 'File', 'Edit', 'View', 'Insert', 'Format', 'Tools', and 'Help'. The main area is a grid-based form titled 'Data Entry for Emergency Nutrition Assessment'. The form fields are: Survey Date, Cluster, Team, ID, Household, Sex (with a dropdown arrow), Birthdate, Months, Weight (Kg), Height (cm), Edema, and MUAC. On the left side, there is a sidebar with buttons for 'Add Page', 'Insert Page', and 'Delete Page', and a 'Program' section. The Windows taskbar at the bottom shows the Start button, several icons, and the system tray with the time '8:16 AM'.

Since Epi Info/ENA requires a standard set of variables, it may be useful to start a new survey with the questionnaire template from Epi Info/ENA and to add to this existing questionnaire the necessary fields (questions). To do this, go to “*Survey/Making New Survey with template*”: the *Make View* programme from Epi Info will open the ENA questionnaire template. To avoid changing the master ENA template, copy and save this template under a new name before modifying it. Go first to “*File/Copy View*” to copy the questionnaire, give the questionnaire a new name, and then open this new questionnaire with “*File/New*”. It is then possible to add and modify the questions in the questionnaire.

Below are more detailed instructions on how to save and modify an existing template. For more detail on how to create, copy, and modify the view (questionnaire) in Epi Info, the reader should refer to the manuals mentioned above.

Open the ENA questionnaire template

1. Select “*Survey/Making New Survey with template*” in the top menu.

The ENA template view will open in the “*Make/Edit view*” window.

Copy the questionnaire under a new name

1. Select “*File/Copy View*” in the top menu.
2. The “*Copy view as*” dialogue box will open.
3. Select the “*Make New View Only*” option. Select “*Copy Code Tables*” option in the “*Code Options*” section. Click “*OK*”.
4. The “*Destination file-MDB*” window will open.

5. Browse to “C:\Epi_info\ena\ena.mdb” and click “Open”.
6. The “Copy Rename Tables” window will open.
7. In the “New Table Name” column, type the name of the questionnaire preceded by the word “view” (for example, to name the questionnaire Kenya2, type viewKenya2).
8. Click “OK”.

The template questionnaire is now copied under a new name in the same file (“ena.mdb”). It can now be opened using the new name assigned, make changes, add new questions, and more to match the questionnaire that will be used in the survey.

The reader should refer to the previous section **Creating a survey questionnaire** to learn how to:

- open the questionnaire;
- add new questions or pages to the questionnaire;
- save and close the questionnaire.

Entering data into the questionnaire

This section will briefly describe how to open the questionnaire in data entering mode (as opposed to the editing mode described above), and to enter data. For more detail on how to enter data into the view (questionnaire) in Epi Info, the reader should refer to the manuals mentioned above.

a. Select (activate) the view (questionnaire)

To select the questionnaire called Kenya1 (as an example):

1. Click the “Select Survey” button on the main screen.
2. The “Select Survey” window will open.
3. Type the name of the questionnaire, for example Kenya1. Click “OK”.

The name of the selected questionnaire should now be shown on the button in the centre of the main screen.

b. Open your questionnaire (view) for data entry

When the questionnaire is selected and its name is showing on the button in the centre of the main screen, this questionnaire can be opened in two ways:

1. Click the “Enter Data” button.
2. Select “Survey/Enter Survey Data” from the top menu.

The questionnaire will open for data entry.

Note: When the view to enter data is opened for the first time and the data table for this view has not been created yet, a message will appear informing the user that a data table will be created for this view. Click “OK” to proceed with data entry.

c. Entering data

Enter data by typing data into the fields.

Select legal values, check boxes and more using the pointer.

d. Moving from field to field and from page to page

Move from field to field using the pointer, the “Tab” key or the “Enter” key.

To go to a different page of the view (questionnaire), select the page in the small window at the top on the left side of the screen.

e. Navigating between different records

On the left-hand side, under the “*Record*” section, click on the arrows to navigate the entered records.

The “<<” sign brings the data-entry screen to the first record, while the “<” sign brings the data-entry screen to the previous record.

Conversely, the “>” brings the data-entry screen to the next record, and the “>>” brings the data-entry screen to the last record.

To go directly to a specific record number, click into the white box, type in the record number, and press “*Enter*” on the keyboard.

f. Save entered data and exit

Select “*File/Save*” and then “*File/Exit*”.

5. Importing Survey Data from another Software

For the analysis of data in Epi Info/ENA, the programme does not need a questionnaire. A simple data table is sufficient. To import data, select “*Survey/Import Data*”. The analysis function of Epi Info opens, and it is possible to import data with the “*Read*” (Import) function on the left side. A window then opens, which asks for the data format of the file which has to be imported and the data source. After the file is imported, the survey has to be saved into the “*ena.mdb*” file with the “*Write*” (Export) function.

Below are more detailed instructions on how to import data table into Epi Info. For more detail, the reader should refer to the manuals mentioned above.

Open in Epi Info survey data file which is in Excel or other format

To analyse, for example survey data which is now in Excel format (file name “*Kenya3.xls*”) in the location “*C:\Epi_info\ena\Kenya3.xls*”.

1. Select “*Survey/Import Data*” from the top menu.
2. The analysis programme of Epi Info will open. Commands are listed in the left window.
3. Click on “*Read*” (Import) command in “*Data*” section.
4. The “*Read*” dialogue box will open.
5. In the “*Current Project*” line, make sure that “*c:\Epi_info\ena\ena.mdb*” file is showing.
6. If not, change it to “*c:\Epi_info\ena\ena.mdb*” using “*Change Project*” button on the bottom of the window.
7. In the “*Data Formats*” line, select the data format that corresponds to the file you are importing (in this example, select Excel 8.0 format).
8. In the “*Data Source*” line, click on “*...*” button on the right and browse to the file to be imported (in this case browse, to “*C:\Epi_info\ena\Kenya3.xls*”). Click “*Open*”.
9. In the “*Worksheets*” window, highlight the worksheet with the survey data. Click “*OK*”.

Your data file is now open in Epi Info.

The script in the upper right window, “*Analysis Output*”, shows the path to the file and the number of data records found in this data table.

Save the file in Epi Info format in “ena.mdb” database under a new name

1. Click on the “*Write*” (Export) command in the “*Data*” section.
2. The “*Write*” dialogue box will open.
3. If all the variables from the data set are to be imported, make sure that the “*All (*)*” box in the upper left corner is checked.
4. Check “*Replace*” in “*Output Mode*”.

5. In the “*Output Formats*” window, select “*Epi 2000*”.
6. In the “*File Name*” line, click on “...” button on the right and browse to the file in which to save the table – “C:\Epi_info\ena\ena.mdb”.
7. In the “*Data Table*” window, type the name of the survey data set to be imported (for example, Kenya3)
8. Click “OK”.

Your data is now saved as a new table named Kenya3 in the “ena.mdb” file.

Open the new survey table from ena.mdb file

To check whether the data table is imported by opening this imported table:

1. Click on the “*Read*” (Import) command in the “*Data*” section.
2. The “*Read*” dialogue box will open.
3. In the “*Current Project*” line, make sure that “C:\Epi_info\ena\ena.mdb” file is showing.
4. If not, change it to “C:\Epi_info\ena\ena.mdb” using the “*Change Project*” button on the bottom of the window.
5. In the “*Data Formats*” line, select “*Epi 2000*”.
6. In the “*Data Source*” line, click on “...” button on the right and browse to “C:\Epi_info\ena\ena.mdb”. Click “*Open*”.
7. In the “*Show*” section, select “*All*”.
8. In “*All*” window, highlight the table you just imported (in this example, Kenya3). Click “OK”.

Your imported table located in ena.mdb file is now open in Epi Info.

The script in the upper right window, “*Analysis Output*”, shows the path to the file and the number of data records found in this data table.

You can now close the analysis programme windows.

Select new survey using “Select Survey” button on main menu screen

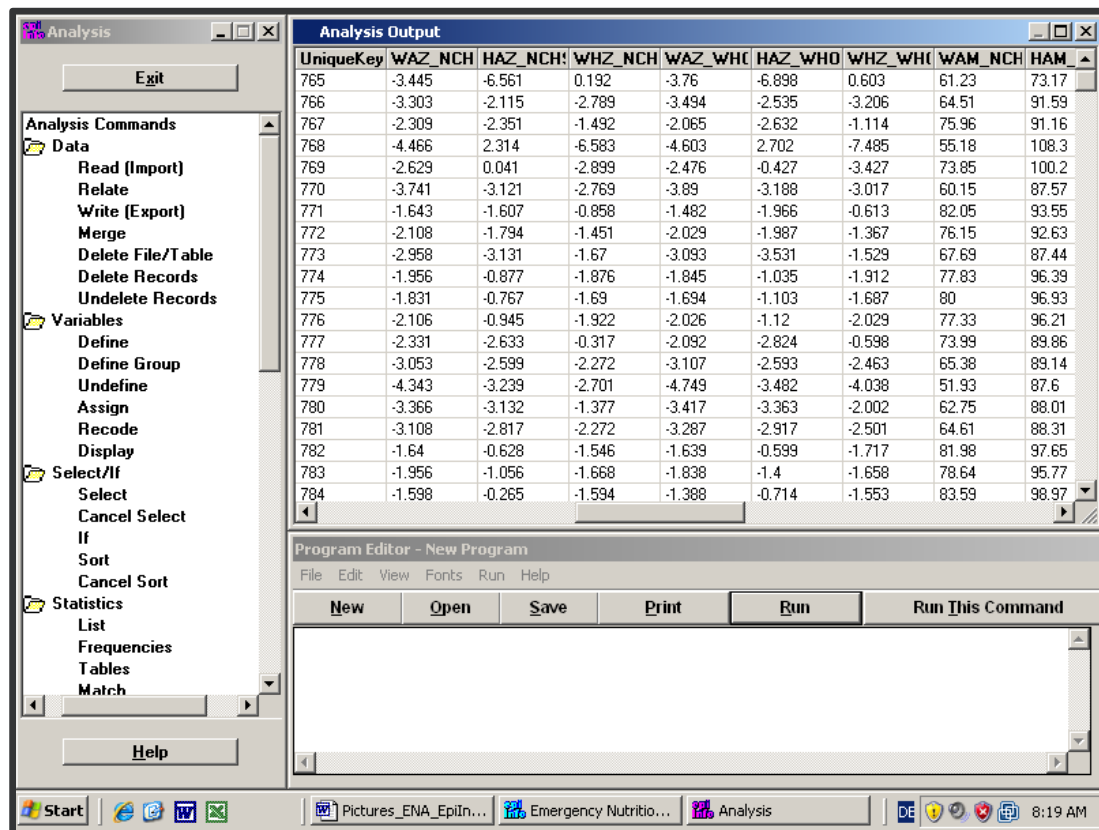
To perform automated analysis on the imported data, select (“activate”) the data table in the Epi Info/ENA main screen:

1. Click the “*Select Survey*” button on the main screen.
2. The “*Select Survey*” dialogue will open.
3. Type the name of the data table, for example Kenya3. Click “OK”.

The name of the data table should now be shown on the button in the centre of the main screen.

The data table is now ready for automated analysis.

6. Generating the Anthropometric Indicators of a Survey



To generate the z-scores of a survey, select the “*Anthropometry Scores*” button from the main screen. Before anthropometry scores are generated, the variables necessary to the z-score calculation have to be matched with the names that Epi Info/ENA is using. After anthropometry scores are generated, Epi Info includes them as new variables into the original survey data table and opens it in the “*Analysis*” function of Epi Info. This table can now be saved under a new name, and further analyses performed using the Epi Info analysis programme.

Instructions below provide more detail on how to generate anthropometry scores and save the newly generated data table with anthropometry variables added to the original dataset.

Select survey button

To generate anthropometry scores from existing data, you need to select (“activate”) your data table as described in the previous section:

1. Click “*Select Survey*” button on the main screen.
2. The “*Select Survey*” dialogue will open.
3. Type the name of the data table, for example Kenya3. Click “*OK*”.

The name of the data table should now be shown on the button in the centre of the main screen.

The data table is now ready for automated analysis.

Anthropometry scores button

1. Click the “*Anthropometry Scores*” button.
2. The “*Select Survey*” window will appear, in which the user must match variables.

Matching variables screen

The names of variables in the left column are fixed. The fields in the right column have drop-down lists with all the variables in the dataset: select in each field the variable from the dataset that matches the variable on the left.

For example, the HEIGHT variable on the left needs to be matched with the variable in the dataset that has height data on the children, and so on.

Note that:

- HEIGHT variable must be expressed in centimetres;
- WEIGHT variable must be expressed in kilograms;
- OEDEMA variable must be expressed as y/n, 1/0 or yes/no;
- SEX variable must be expressed as 1/2, M/F or Male/Female;
- MONTHS variable (age of the children) must be expressed in months.

Note:

As a default, “Epi Info/ENA” will automatically calculate a child’s age from the date of birth and the survey date if both dates are available. In this case, age provided in MONTHS variable will not be used.

If either of the dates is missing, age from the MONTHS variable will be used.

Note:

Please check the match of each variable carefully. Epi Info/ENA suggests the matching variable, but this may not be a correct match in the dataset.

The screenshot shows a window titled "Select Survey" with a subtitle "If necessary please change the variable assignment". The window is divided into two columns: "ENA for SMART" and "EPI Info". Each column contains a list of variables in text boxes. The "EPI Info" column has corresponding dropdown menus for each variable. To the right of the dropdowns is a "Date format" section with three radio buttons: "ddmmyy", "mmddyy" (which is selected), and "yymmdd". At the bottom right of the window is an "OK" button with a green checkmark icon.

ENA for SMART	EPI Info
SURVDATE	Survdate
CLUSTER	Cluster
TEAM	Team
ID	Id
SEX	Sex
BIRTHDAT	Birthdat
MONTHS	Months
WEIGHT	Weight
HEIGHT	Height
EDEMA	Edema

Selecting date format

In the window on the right, select the date format which used in the dataset. If the date format is selected incorrectly, the programme will not be able to generate anthropometry scores.

After the variables are matched and date format selected, click “OK”.

New table with anthropometry scores, flags and categorical variables

The new table opens in Epi Info “Analysis” window.

This data table contains all the variables from the survey data set, plus the following new variables:

- **WAZ_NCHS, HAZ_NCHS, WHZ_NCHS** – weight-for-age, height-for-age, and weight-for-height z-scores based on NCHS growth reference;
- **WAZ_WHO, HAZ_WHO, WHZ_WHO** – weight-for-age, height-for-age, and weight-for-height z-scores based on WHO growth reference;
- **WAM_NCHS, HAM_NCHS, WHM_NCHS** – weight-for-age, height-for-age, and weight-for-height percentage of median scores based on NCHS growth reference;
- **WAZ_NCHS_c, HAZ_NCHS_c, WHZ_NCHS_c** – categorical variables (normal/moderate/severe) based on weight-for-age, height-for-age, and weight-for-height z-scores based on NCHS growth reference (normal for ≥ -2 z-scores, moderate for ≥ -3 and < -2 and severe for < -3).
- **WAZ_WHO_c, HAZ_WHO_c, WHZ_WHO_c** – categorical variables (normal/moderate/severe) based on weight-for-age, height-for-age, and weight-for-height z-scores based on WHO growth reference;
- **WHM_NCHS_c** -- categorical variable (normal/moderate/severe) based on weight-for-height percentage of median based on NCHS growth reference;
Note: all children with oedema (irrespective of their z scores or percentage of median) are classified as severe for weight-for-height and weight-for-age, but not height-for-age
- **Flag_WAZ_NCHS, Flag_HAZ_NCHS, Flag_WHZ_NCHS** – flags for weight-for-age, height-for-age, and weight-for-height z-scores based on NCHS growth reference (flags are in yes/no format);
- **Flag_WAZ_WHO, Flag_HAZ_WHO, Flag_WHZ_WHO** – flags for weight-for-age, height-for-age, and weight-for-height z-scores based on WHO growth reference (flags are in yes/no format).

Note: Flagged are records where the corresponding anthropometry score is either missing, or out of range as defined on the “Option” page of “Settings” (see Defining the range for exclusion on the “Options” page described below). The default is no exclusion, and therefore all values in these columns will not be flagged except those with missing data. If one of the flagging procedures is active, the values which are outside the plausible range or missing are flagged with “yes”. The flag variable is very useful – flagged records can now be easily excluded from the statistical analysis in Epi Info.

Save the table in Epi Info format or in other format under a new name

This new data table must be saved under a new name, otherwise it will disappear after it is closed. Epi Info allows the user to save this table in a variety of formats.

1. Click on the “Write” (Export) command in the “Data” section on the left panel.
2. The “Write” dialogue box will open.
3. Make sure that “All (*)” box in the upper left corner is checked.
4. In the “Output Formats” window, select the format in which to save this data table.
5. In the “File Name” line, click on the “...” button on the right and browse to the folder in which to save this table. Type the name of the file, or select an existing file. Click “Save”.

6. In the “Data Table” window, type the name of the survey data set to save.
7. Click “OK”.

Your data is now saved as a new file or a new table or spreadsheet in existing file.

7. Automatic Generation of the Data Quality Report

Select survey button

Make sure the data table is “active” on the main Epi Info/ENA screen (the name of data table is shown on the central button). If it is not, select it as described in previous section.

Plausibility check button

1. Click the “Plausibility Check” button.
2. The “Select Survey” window will appear, in which the user must match variables.

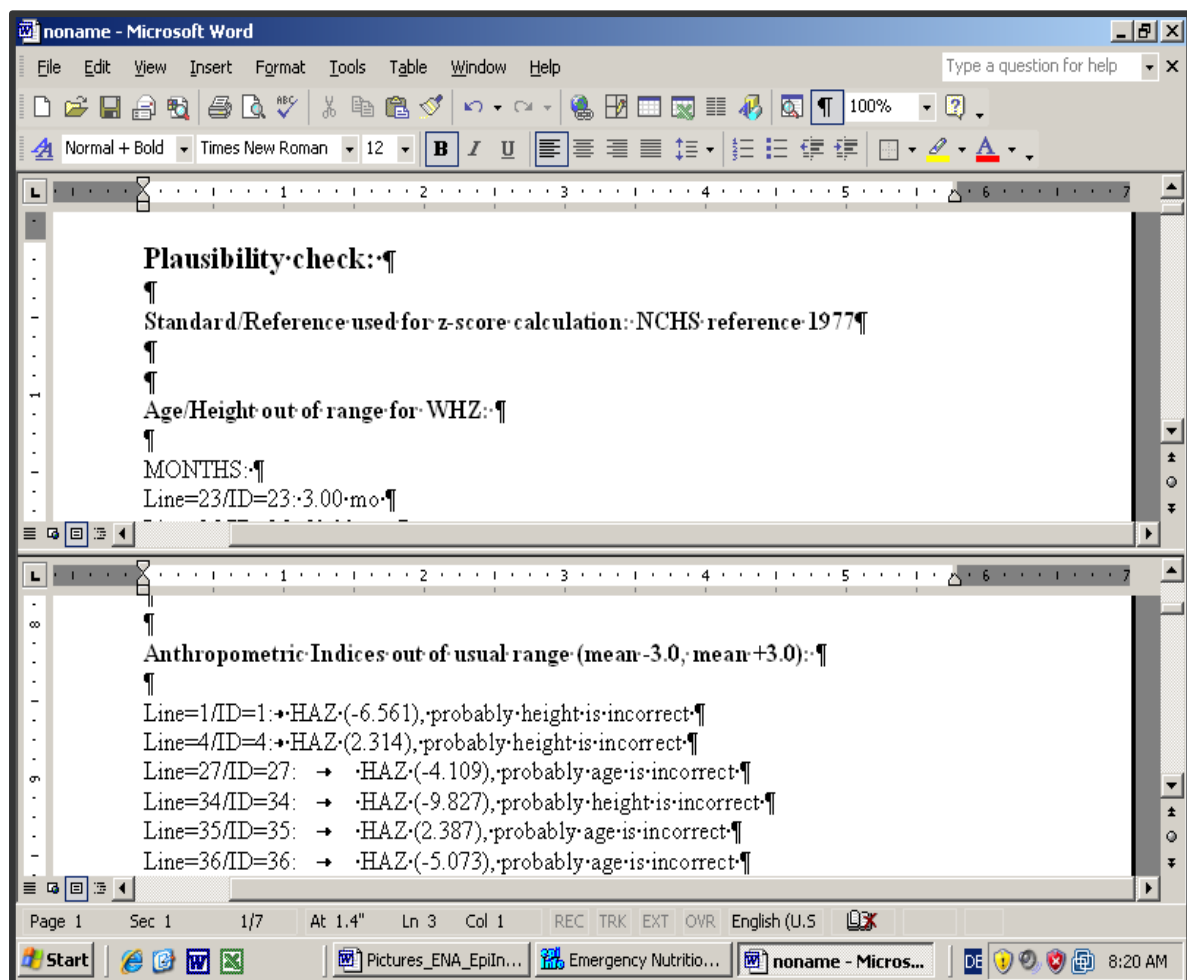
Matching variables screen

1. Match variables and select the date format as described in previous section. Click “OK”.

The plausibility check report is now opened in MS Word format.

2. It must be saved under a new name, otherwise it will disappear after it is closed.

The reader should refer to the plausibility check chapter for a detailed explanation of plausibility check outputs and interpretation.



8. Automatic Generation of the Anthropometry Graphs and Survey Report

Select Survey button

Make sure your data table is “active” on the main Epi Info/ENA screen (the name of the data table should be shown on the central button). If it is not, select it as described in previous section.

Survey Results button

1. Click “Survey Report” button.
2. “Select Survey” window will appear, in which the user must match variables.

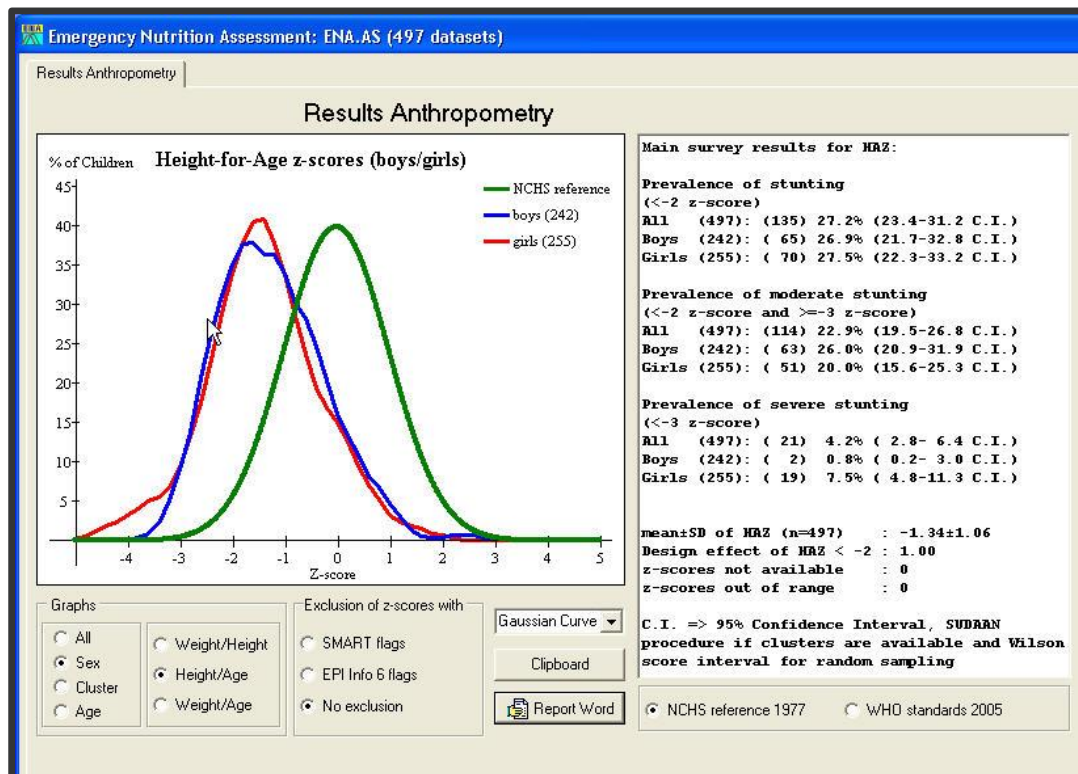
Matching variables screen and Results Anthropometry page, generating and copying anthropometry graphs

1. Match variables and select the date format as described in previous section. Click “OK”.
2. The “Results Anthropometry” page will open.

From the “Results Anthropometry” page, it is possible to generate various graphs of anthropometric scores by age group, sex, or cluster (first column of radio buttons from the left).

1. Choose the anthropometric indicator – W/H, H/A or W/A (second column of buttons from the left).
2. Select the criteria for exclusion of outliers (third column from the left): “SMART flags” mean exclusion from observed mean, “EPI Info 6 flags” mean exclusion from the reference mean (zero).
3. Select the type of graph from the drop-down list.

To copy the graph into the clipboard and then paste it into the report, click on the “Clipboard” button. The current graph is copied into the clipboard. It can now be pasted from the clipboard into a MS Word document (or another programme).



Generating a survey report

To generate a survey report, click “Report Word” button.

The survey report template is now opened in MS Word format.

It must be saved under a new name, otherwise it will disappear after it is closed.

Note: Some tables in the survey report template are populated with calculated survey results. Survey results based on NCHS growth reference are in the main body of the report, while survey results based on WHO growth reference are in the appendix.

Report of Nutrition and Mortality in (region, camp etc) of (country) during (date to date).

Summary of the report (one to two pages only)

- area surveyed
- date of survey
- methodology employed
- main anthropometric results (prevalence of global and severe acute malnutrition in terms of z-scores and/or oedema and 95 per cent confidence intervals)
 - GAM: 30.1% (24.1-36.9 95% C.I.)
 - SAM: 11.4% (7.3-17.4 95% C.I.)

Table 3.2: Prevalence of acute malnutrition based on weight-for-height z-scores (and/or oedema) and by sex

	All n=342	Boys n=188	Girls n=154
Prevalence of global malnutrition (≤ -2 z-score and/or oedema)	(103) 30.1% (24.1-36.9 95% C.I.)	(54) 28.7% (21.7-36.9 95% C.I.)	(49) 31.8% (24.1-40.6 95% C.I.)
Prevalence of moderate malnutrition (≤ -2 z-score and ≥ -3 z-score, no oedema)	(64) 18.7% (15.5-22.5 95% C.I.)	(31) 16.5% (12.4-21.6 95% C.I.)	(33) 21.4% (15.9-28.2 95% C.I.)

9. Software Settings

Age groups (months)		Expected proportion of children	
		male	female
6	- 17	1000	1000
18	- 29	975	975
30	- 41	945	945
42	- 53	930	930
54	- 59	920	920

Analyzing children only for
6.00 to 59.99 months

Exclusion of z-scores from

Observed mean

SMART flags

WHZ -3 to 3

HAZ -3 to 3

WAZ -3 to 3

Zero (reference mean)

EPI Info 6 flags

WHZ -5 to 5

HAZ -6 to 6

WAZ -6 to 5

No exclusion

Reset

Save

The settings for flags generation, plausibility and survey report can be changed in this part of the software.

The first section on the right is to define the age range of children for inclusion in the automated analysis. Children that are outside the defined age range (default is six to 59.99 months) will be automatically excluded from the survey report analysis.

The second section (“Age groups”) is relevant for parts of the plausibility and survey report. It determines the age groups which will be used for data analysis. It is possible to change assumptions about the proportion of males to females, and of different age groups. This is relevant mostly for the plausibility check.

The third section (“Exclusion of implausible z-scores”) has three options available:

- The first option uses a range from zero (mean of the reference). Cut-off values currently recommended by WHO are used as defaults for this option.
- The second option uses a range which is defined from the observed mean of the population. A value of ± 3 is used as a default for this option.
- The third option is “No exclusion” – no values are excluded no matter how implausible they are.

Note: Values more extreme than the ones defined in this section will be excluded from automated analysis in the survey report. They also will be flagged (along with the missing values) when anthropometry scores are generated.

10. Survey Planning and Sample Size Calculation

Emergency Nutrition Assessment: ENA_TEMPLATE.AS (0 datasets)

Planning | Training | Options

Planning Nutrition Survey

Name of Survey

Sampling
 Random Cluster
 Correction small population size

Sample size calculation for a cross sectional anthropometric survey*

Estimated prevalence %: 20
 ± desired precision %: 5
 Design effect: 2
 Children to be included: 535

Average household size: 5
 % children under 5: 15
 % of non-response households: 3
 Households to be included: 818

Sample size calculation for a death rate survey*

Estimated death rate per 10000/day: 2
 ± desired precision per 10000/day: 1
 Design effect: 2
 Recall period in days: 30
 Population to be included: 5576

Average household size: 5
 % of non-response households: 3
 Households to be included: 1150

*Please change the default values to get more correct estimations for the sample size

Table for Cluster sampling

Population Size: 7000

Geographical unit	Population size	Cluster
A	100	
B	2000	1,2,3,4,5,6...
C	3000	10,11,12,1...
D	1000	22,23,24,2...
E	100	
F	500	27,28
G	200	29
H	100	30

Number of Cluster: 30

Random Number Table

Range from: to Numbers: Generate Table

EN 6:38 AM 09/06/2012

11. Anthropometry Exercise for Training in Epi Info/ENA

Emergency Nutrition Assessment: ENA_TEMPLATE.AS (0 datasets)

Planning | Training | Options

Evaluation of Enumerators

Numeric continuous data | Categorical data

Please enter for the supervisor and all enumerators the training measurements (children measured twice for weight, height or MUAC)
 After entering the data press the report button to get an evaluation for each enumerator.

Subject	Supervisor					Enumerator 1					
	Weight 1	Weight 2	Height 1	Height 2	MUAC1	MUAC2	Weight 1	Weight 2	Height 1	Height 2	MUAC
1	12.3	12.2	82.8	82.2			12.3	12.2	81.9	82.6	
2	11.2	11.3	83.8	84.6			11.3	11.2	84.6	84.6	
3	8.4	8.5	86	85.6			8.5	8.3	86.3	86.1	
4	4.7	4.4	86.2	86			4.6	4.4	86.2	85	
5	13.2	13.2	82	82			13.1	13.2	82.5	82.3	
6	15	15.1	85.6	85.4			15.1	15.3	85.7	86.2	
7	16	15.9	82.3	82.4			16	16.1	82.4	82.5	
8	17	16.9	87.6	87.6			16.9	17.1	88	87.5	
9	18.3	18.4	80.1	80.6			18.2	18.2	81	80.4	
10	12.4	12.5	85.3	86.5			12.5	12.6	85.8	85.2	
11											
12											
13											
14											
15											
16											
17											

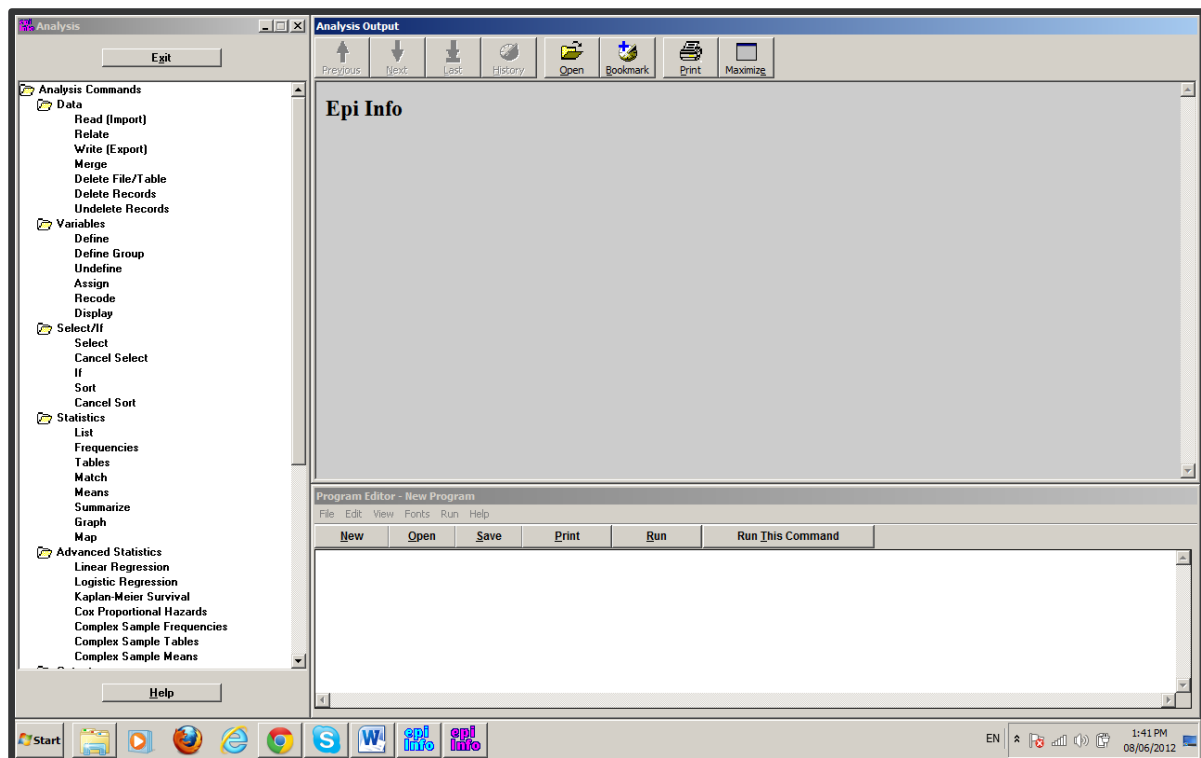
Replace supervisor values with mean from enumerators

Report from previous ENA versions

New Report

EN 6:37 AM 09/06/2012

12. Data Analysis



Data analysis in Epi Info/ENA is carried out in the “*Analysis*” screen (the screen used to import data, as presented above).

Select “*Analyse Data*” from the “*Utilities*” menu. The “*Analysis*” screen will open (Note: the “*Analysis*” screen can also be opened by clicking on the “*Import Data*” command under the “*Survey*” menu).

Once the “*Analysis*” screen is open, the advanced statistics option first needs to be set before conducting any analysis.

1. Select the “*Set*” command under the “*Options*” section.
2. The “*Set*” dialog box will open. Find “*Advanced*” under “*Statistics*”, and select it. Click on “*OK*”.

The software is now ready to carry out the analysis.

Depending on the type of sampling employed (simple/systematic sampling, cluster sampling) in the survey, different commands will be required to analyse survey data. Both of these are described below.

Simple/systematic random surveys

For data from simple and systematic random surveys, the commands under the “*Statistics*” section should be used.

Note that although there are several commands under the “*Statistics*” section, only three commands (“*Frequency*”, “*Tables*”, and “*Means*”) are described here with examples. The reader should refer to the Epi Info manual mentioned above for more information on the other commands.

Let us assume that the data in Kenya3 Excel file comes from a simple random sample (the user should ignore the variable, clusters, in the dataset for this exercise).

To begin, “*Read*” (Import) the Kenya3 file from the “*c:\Epi_info\ena\ena.mdb*” folder using the instructions given above under *Importing data from another software*.

To look at the **frequency** of (e.g.) girls and boys in the dataset:

1. Click on “*Frequency*” command under the “*Statistics*” section.
2. The “*FREQ*” dialog box will appear.
3. Select “*Sex*” from the dropdown menu under “*Frequency of*” and select click “*OK*”.
4. The frequency of girls and boys will be displayed in a table.

Note that you could select multiple variables here and the frequencies of each variable will be displayed in the order the variable is selected. For example, to look at the frequencies of (e.g.) Sex, GAM_WHO_c, Oedema, and MIAC_115:

1. Click on the “*Frequency*” command under “*Statistics*”.
2. Select “*Sex*”, “*GAM_WHO_c*”, “*Oedema*”, and “*MUAC_115*” from the dropdown menu under “*Frequency of*” and select “*OK*”.
3. The frequencies of each variable will be displayed in tables one after the other.

To look at the (e.g) number of girls and boys who are with GAM. It will be necessary to the “**Tables**” command to generate this information.

1. Select the “*Tables*” command under the “*Statistics*” section.
2. Select “*Sex*” from the dropdown menu under “*Exposure Variable*” and “*GAM_WHO_c*” under the “*Outcome Variable*”. Click on “*OK*”.

The “**Means**” command, under the “**Statistics**” section can be used to analyse a single variable or to compare two means. When the mean of a single variable is analysed, the following results are obtained:

- total number of observations;
- the sum of all observations;
- the mean;
- the variance;
- standard deviation of the observations;
- minimum and maximum values;
- 25th, 50th, and 75th percentiles, and;
- the mode.

For example, to get the mean of MUAC:

1. Click on “*Means*” under the “*Statistics*” section.
2. Select “*MUAC*” from the dropdown menu under the “*Means of*”. Click “*OK*”.

At the end of a large table, the mean and standards deviation of MUAC data will be displayed.

Cluster surveys

To analyse data from cluster surveys, the commands under the “*Advanced Statistics*” should be used to account for the survey design and to calculate the variance accordingly. Although the “*Advanced Statistics*” section has several commands, for the purpose of this manual, only the following will be exemplified: “*Complex Sample Frequencies*”, “*Complex Sample Tables*”, and “*Complex Sample Means*”.

Let us now assume that the data in Kenya3 Excel file comes from a cluster survey design.

Read (Import) Kenya3 from the c:\Epi_info\ena\ena.mdb folder.

To begin, “*Read*” (Import) Kenya3 from the “c:\Epi_info\ena\ena.mdb folder”.

If you want to find out (e.g.) about the frequency of GAM (i.e. GAM prevalence):

1. Click on “*Complex Sample Frequencies*” under the “*Advanced Statistics*” section.
2. The “*Complex Sample Frequencies*” dialog box will open.
3. Select “*GAM_WHO_c*” from the dropdown menu under “*Frequencies of*”.
Note: When analysing data from cluster surveys, in addition to the variable to be analysed, the Primary Sampling Unit (PSU) should always be specified⁷⁶. In this example, the “PSU” variable is “Cluster”. So, in addition to “GAM_WHO_c”, you will also need to select “Cluster” from the dropdown menu under the “PSU”.
4. The frequency of GAM will be displayed in a table.

To generate **tables** for cluster survey variables, the “Complex Sample Survey” will need to be used.

To find out about the prevalence of GAM among girls and boys:

1. Click on the “*Complex Sample Tables*” command.
2. The “*Complex Sample Tables*” dialog box will appear.
3. Select “*Sex*” from the dropdown menu under “*Exposure Variable*” and “*GAM_WHO_c*” from the dropdown menu under “*Outcome Variable*”.
4. Select “*Cluster*” under “*PSU*” and click “*OK*”.
5. A table with GAM prevalence disaggregated by sex will be displayed.

The “**Complex Sample Means**” command can be used to analyse the numeric continuous data in the survey dataset.

To get the mean of (e.g.) MUAC in the survey dataset:

1. Click on the “*Complex Sample Means*” command.
2. The “*Complex Sample Means*” dialog box will appear.
3. Select the “*MUAC*” from the dropdown menu under “*Means of*”.
4. Select “*Cluster*” under “*PSU*” and click “*OK*”.
5. A table with the following information will be displayed: “*Count*”, “*Mean*”, “*Standard Error*”, “*Confidence Limits*”, “*Minimum*”, and “*Maximum*”.

⁷⁶ Other designs, such as stratification and weighing may also be specified, but these are beyond the scope of this manual. The reader should refer to the Epi Info manual for more detail.

ANNEX 5: References

- Alarcón D, Robles M. *The challenges of measuring child mortality when birth registration is incomplete*. Global Forum on Gender Statistics. Rome, Italy. December 2007.
- Anker M, Black RE, Coldham C, et al. *A standard verbal autopsy method for investigating causes of death in infants and children*. Geneva, World Health Organization, 1999 (WHO/CDS/CSR/ISR/99.4; http://whqlibdoc.who.int/hq/1999/WHO_CDS_CSR_ISR_99.4.pdf, accessed 1 June 2009).
- Barton T, Borrini-Feyerabend G, de Sherbinin A, Warren P. *Our People, Our Resources: Supporting rural communities in participatory action research on population dynamics and the local environment*. IUCN-The World Conservation Union, 1997.
- Bennett S, Radalowicz A, Vella V, Tomkins A. A computer simulation of household sampling schemes for health surveys in developing countries. *Int J Epidemiol* 1994; 23(6):1282 – 91.
- Bennett S, Woods T, Liyanage WM, Smith DL. A simplified general method for cluster-sample surveys of health in developing countries. *World Health Stat Q* 1991; 44:98 - 106.
- Bilukha O. Old and new cluster designs in emergency field surveys: in search of a one-fits-all solution. *Emerging Themes in Epidemiology* 2008; 5(7).
- Binkin S, Sullivan K, Staehling N, Nieburg P, Rapid nutrition surveys: how many clusters are enough? *Disasters* 1995, 16(2): 97-103.
- Boss L, Toole M, Yip R. Assessments of mortality, morbidity and nutritional status in Somalia during the 1991-1992 famine: recommendations for Standardisation of methods, *JAMA*, 1994; 272 (5).
- Bureau of Statistics [Tanzania] and Macro International Inc. *Tanzania Demographic and Health Survey 2004-2005*. December 2005. Calverton, Maryland: Bureau of Statistics and Macro International Inc.
- CERTI. *Crisis and Transition Tool Kit: Strategic Framework for Crisis Prevention, Mitigation and Transition*, 2000.
- Checchi F, Roberts L. Documenting mortality in crises: what keeps us from doing better? *PLoS Med* 2008; 5(7): e146.
- Checchi F, Roberts L. Interpreting and using mortality data in humanitarian emergencies; A primer for non-epidemiologists. *Humanitarian Practice Network Paper* 2005.
- Doocy S, Rofi A, Moodie C, Spring, CE, Bradley S, Burnham G, Robinson C. Tsunami mortality in Aceh Province, Indonesia. *Bulletin of the World Health Organization* 2007; 85(2): 273-278.
- Golden MH. The Clinical Assessment of Oedema: Implications for Feeding the Malnourished Child. *European Journal of Clinical Nutrition* 1989; 43: 581–82.
- Grais R, Rose A, Guthmann JP. Don't spin the pen: two alternative methods for second-stage sampling in urban cluster surveys. *Emerging Themes in Epidemiology* 2007;4(1):8.
- Grellety Y. *Management of Severe Malnutrition in Africa [PhD Thesis]*. Scotland: University of Aberdeen. 2000.
- Haupt A, Kane T. *Population Handbook, Fourth Edition*, Washington, DC: Population Reference Bureau, 1998.
- Henry G. *Practical Sampling*, Applied Social Research Methods Series, Volume 21. Newbury Park: SAGE Publications, 1990, p. 11.
- Herold J, Peavy JV. Surveys and Sampling, in Gregg MB (ed.), *Field Epidemiology*, Oxford: Oxford University Press, 2002.

Kaiser R, Spiegel PB, Henderson AK, Gerber ML. The application of geographic information systems and global positioning systems in humanitarian emergencies: lessons learned, programme implications and future research. *Disasters* 2003; 27(2):127 - 40.

Kaiser R, Woodruff BA, Bilukha O, Spiegel PB, Salama P. Using design effects from previous cluster surveys to guide sample size calculation in emergency settings. *Disasters* 2006; 30(2):199-211.

Kalter HD, Salgado R, Gittelsohn J, Parades P. *A Guide to Conducting Mortality Surveys and Surveillance*. Published by Basic Support for Institutionalizing Child Survival Project (BASICS II) for the United States Agency for International Development. Arlington, Virginia, June 2004.

Kalton G, *Introduction to Survey Sampling*, Sage Publications, 1983.

Katz J, Zeger S, Estimation of design effects in cluster surveys, *AEP*, 1994; 4(4) 295-301.

Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Annals of Family Medicine* 2004; 2(3): 204-08.

Kish L, *Survey Sampling*, New York: John Wiley, 1965.

Lopman BA, Barnabas R, Hallett TB, Nyamukapa C, Mundandi C, Mushati P, et al. Assessing adult mortality in HIV-1-afflicted Zimbabwe (1998 -2003). *Bulletin of the World Health Organization* 2006;84: 189-197

Lulu K, Berhane Y, The use of simplified verbal autopsy in identifying causes of adult death in a predominantly rural population in Ethiopia, *BMC Public Health*, 2005: 5(58).

Milligan P, Njie A, Bennett S. Comparison of two cluster sampling methods for health surveys in developing countries. *Int J Epidemiol* 2004; 33(3):469 – 76.

Noji E. *The Public Health Consequences of Disaster*, New York and Oxford: Oxford University Press, 1997.

National Statistical Service, *Basic Survey Design: Errors in Statistical Data*. (NSS; <http://www.nss.gov.au/nss/home.nsf/SurveyDesignDoc>/ accessed on 1 June 2009).

Palmore, JA, Gardner RW. *Measuring Mortality, Fertility and Natural Increase: A Self-Teaching Guide to Elementary Measures*, Honolulu: East-West Center, 1994.

Prudhon C, Spiegel PB. A review of methodology and analysis of nutrition and mortality surveys conducted in humanitarian emergencies from October 1993 to April 2004. *Emerging Themes in Epidemiology* 2007, 4:10.

Robinson C, Lee MK, Hill K, Burnham G. Mortality in North Korean migrant households: A retrospective study. *Lancet* 1999; 354.

Setel PW, Macfarlane SB, Szreter S, Mikkelsen L, Jha P, Stout S, AbouZahr C et al. A scandal of invisibility: making everyone count by counting everyone. *Lancet* 2007; 370: 1569-1577

Shorr I. *How to weigh and to measure children: Assessing the nutritional status of young children in household surveys*, New York: United Nations, 1986.

SMART Methodology Website. Recommended survey equipment. 2017. <http://smartmethodology.org/resources/recommend-survey-equipment/>

Tailhades M, Toole M. Disasters: What are the needs? How can they be assessed? *Tropical Doctor* 1991; 21: 18-23.

Talley L, Richardson L, Bilukha O, Brennan M, Blanton C. A validation of mortality estimation methodologies at the household level: summary and household census methodologies. A collaboration between the Centers for Disease Control and Prevention and the United Nations World Food Programme. May 2006.

UNICEF/CDC. *Nutrition and Health Survey, Badghis Province, Afghanistan, February – March 2002*. A collaborative survey by: UNICEF and U.S. Centers for Disease Control and Prevention (CDC), 2002.

United Nations Refugee Agency (UNHCR). *Health Information System, Quarterly Bulletin* 2008; 3. (UNHCR; <http://www.unhcr.org/protect/4614bce84.html> accessed on 1 June 2009).

Van de Walle E, ed. *African Households: Censuses and Surveys*. Armonk, NY: M.E. Sharpe Inc; 2006.

Working Group for Mortality Estimation in Emergencies. Wanted: studies on mortality estimation methods for humanitarian emergencies, suggestions for future research. *Emerging Themes in Epidemiology* 2007, 4:9.

World Health Organization, *Communicable Disease Surveillance and Response*. "What is Disease Surveillance?" Available at www.who.int/emc/surveill/index.html.

World Health Organization. *Health Status Statistics: Mortality*. (WHO; <http://www.who.int/healthinfo/statistics/indneonatalmortality/en>, accessed on 1 June 2009).

World Health Organization. *The WHO Child Growth Standards*. 2006. <http://www.who.int/childgrowth/en/>

World Health Organization. *The WHO Multicentre Growth Reference Study (MGRS)*. 2006. <http://www.who.int/childgrowth/mgrs/en/>