

Comprehensive Food Security & Vulnerability Analysis Guidelines

January 2009 • first edition

WFP

World Food Programme

Comprehensive Food Security & Vulnerability Analysis Guidelines

January 2009 • first edition

WFP
World Food Programme

vam
wfp food security analysis

Comprehensive Food Security & Vulnerability Analysis Guidelines

United Nations World Food Programme
Via Cesare Giulio Viola 68/70, Parco de' Medici, 00148 Rome, Italy

Food Security Analysis Service
Chief: Joyce Luma
Tel: +39-06-6513-2168
E-mail: joyce.luma@wfp.org or wfp.vaminfo@wfp.org

CFSVA Guidelines
Project Manager: Jan Delbaere, Deputy Chief
Tel: +39-06-6513-3265
E-mail: jan.delbaere@wfp.org

first edition

citigroup foundation

EUROPEAN COMMISSION
Humanitarian Aid

Citigroup Foundation and ECHO provided financial support for the development and production of these guidelines.

Foreword

Fighting hunger in a changing world demands that we stay vigilant in our efforts to collect, analyze and disseminate information that is so very critical for designing and implementing hunger solutions which can save lives in emergencies, as well as putting the hungry poor on the path to food security.

Understanding food security and vulnerability has always been challenging. Yet the emergence of relatively new phenomena such as the recent high food and fuel prices, the global financial crisis, and climate change, all highlight the need to better understand the lives and livelihoods of vulnerable populations so that effective policies and actions can be implemented to save lives and address the root causes of hunger.

To tackle hunger, we first need to understand three key factors: how food is made available to people; how they economically and physically access food; and how they utilize the food. Understanding the constraints underlying each of these factors is a necessary condition for designing and implementing appropriate and effective hunger reduction strategies.

The Comprehensive Food Security and Vulnerability Analysis (CFSVA) is a unique tool designed to understand these factors. It describes the profile of the food-insecure and vulnerable households, identifies the root causes of hunger, and analyzes risks and emerging vulnerabilities among populations. It provides crucial information on the type of interventions that would be the most effective in reducing hunger, targeting the neediest, informing preparedness and developing contingencies. The range and depth of information provided by CFSVAs are invaluable, not only for WFP, but for the entire humanitarian and development community.

Over the last four years WFP, along with partners, has completed 27 CFSVAs worldwide. This was in large part made possible due to generous support from the European Commission's Humanitarian Aid department (ECHO), the Citigroup Foundation and the Gates Foundation.

This document, built on this experience, will guide WFP food security analysts, programme officers and partner's staff as they undertake Comprehensive Food Security and Vulnerability Analyses.

Ramiro Lopes da Silva
Deputy Chief Operating Officer and Director of Operations

Acknowledgements

The Comprehensive Food Security and Vulnerability Analysis Guidelines were made possible thanks to the generous support of the European Commission's Humanitarian Aid department (ECHO), which provided funding also for consultation meetings. The Citigroup Foundation supported one preparatory technical meeting.

The guidelines were written mainly by WFP staff and by Tango International, under the umbrella of the Strengthening Emergency Needs Assessment Capacity Plan. The purpose of the plan (2004–2007) was to reinforce WFP's capacity to assess humanitarian needs in the food sector during emergencies and their aftermath through accurate and impartial needs assessments. The effort was funded from internal and external sources, among them, ECHO, the Canadian International Development Agency (CIDA), the Citigroup Foundation, the UK Department for International Development (DFID), and the French, Danish, German, and Belgian Governments.

These guidelines benefited from several peer-review and consultation mechanisms. An initial workshop was held in April 2007, in Rome; a second one, in March 2008, in Cairo. The participants were mainly CFSVA practitioners from within WFP.

A draft of these guidelines was thoroughly reviewed by John Hoddinott (International Food Policy Research Institute), Nancy Mock (Tulane University), and Dan Maxwell (Tufts University), and most of their important comments and suggestions were taken into account in the final version.

WFP expresses special thanks to the many VAM staff, in headquarters and in the field, who in writing these guidelines provided their invaluable knowledge of food security analysis, based on many years of experience in a large variety of countries where food security is an issue.

All the contributions are gratefully acknowledged, although responsibility for the present text, including errors, remains with WFP.

Contents

! The CFSVA Annexes referred to throughout this document are available on the accompanying DVD and at www.wfp.org/food-security.

List of Tables

List of Figures

List of Boxes

Untitled boxes

Definitions

Additional information

Examples

Acronyms

ADRA	Adventist Development and Relief Agency
AET	Actual Evapo-Transpiration
AIDS	Acquired Immunodeficiency Syndrome
AIS	AIDS Indicator Survey
AMS	AgroMetShell
ANOVA	Analysis of Variance
AoE	Area of Estimation
ART	Anti-Retroviral Treatment
BSS	Behavioural Surveillance Survey
BMI	Body Mass Index
CBO	Community-Based Organization
CCA	Common Country Assessment
CDC	United States Centers for Disease Control and Prevention
CED	Chronic Energy Deficiency
CFSAM	Crop and Food Supply Assessment Mission
CFSNS	Comprehensive Food Security and Nutrition Survey
CFSVA	Comprehensive Food Security and Vulnerability Analysis
CGIAR	Consultative Group on International Agricultural Research
CFW	cash for work
CHS	Community and Household Surveillance
CI	Confidence Interval
CIDA	Canadian International Development Agency
CO	Country Office
CRED	Centre for Research on Epidemiology of Disasters
CRS	Catholic Relief Services
CSB	Corn-Soya Blend
CSI	Coping Strategies Index
CSPro	Census and Survey Processing System
CTC	Community-Based Therapeutic Care
DEM	Digital Elevation Model
DFID	Department for International Development (United Kingdom)
DHS	Demographic and Health Survey
DOT	Direct Observation Therapy
DOTS	directly observed treatment, shortcourse
DTP	Desktop Publishing Applications
EFSA	Emergency Food Security Assessment
EMOP	Emergency Operation
EM DAT	Emergency Events Database
EMF	Enhanced Metafile
EOS	End of Season
EPI	Expanded Programme on Immunization
FANTA	Food and Nutrition Technical Assistance
FAO	Food and Agriculture Organization of the United Nations
FCG	Food Consumption Group

FCS	Food Consumption Score
FEWS NET	Famine Early Warning Systems Network
FFW	food for work
FGD	Focus Group Discussion
FSMS	Food Security Monitoring System
GAM	Global Acute Malnutrition
GDDS	General Data Dissemination System
GDI	Gender-related Development Index
GDP	Gross Domestic Product
GEM	Gender Empowerment Measure
GER	Gross Enrolment Rate
GIS	Geographic Information Systems
GLM	General Linear Model
GMR	Global Malnutrition Rate
GMR	Global Mortality Rate
GPS	Global Positioning Systems
GTZ	German Agency for Technical Cooperation
HAZ	height for age
HBC	Home-Based Care
HDDS	Household Dietary Diversity Score
HDI	Human Development Index
HDR	Human Development Report
HH	Household
HIC	Habitat International Coalition
HIC	Humanitarian Information Centre
HIV/AIDS	Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome
HLS	Household Livelihood Security
IASC	Inter-Agency Standing Committee
IDP	Internally Displaced Person
IFPRI	International Food Policy Research Institute
ILO	International Labour Organization
ILWIS	Integrated Land and Water Information System
IMF	International Monetary Fund
IOM	International Organization for Migration
IUGR	Intrauterine Growth Retardation
JMP	Joint Monitoring Programme
LCA	Logistics Capacity Assessment
LGP	length of growing period
LSMS	Living Standards Measurement Survey
MCDA	Military and Civil Defence Assets
MCH	Mother-and-Child Health
MICS	Multiple-Indicator Cluster Survey
MTCT	Mother-to-Child Transmission
MUAC	mid-upper arm circumference

NCHS	National Center for Health Statistics
NDVI	Normalized-Difference-Vegetation-Index
NER	Net Enrolment Rate
NGO	non-governmental organization
NHDR	National Human Development Report
NOAA	National Oceanic and Atmospheric Administration
OCHA	Office for the Coordination of Humanitarian Affairs
ODAN	WFP Emergency Needs Assessment Service
OEN	WFP Needs Assessment Unit
OLS	Ordinary Least Square
OMXF	WFP Food Security Analysis Service
OVC	orphans and other vulnerable children
PCA	Principal Components Analysis
PDA	Personal Digital Assistant
PET	Potential Evapo-Transpiration
PLHIV	people living with HIV
PMTCT	prevention of mother-to-child transmission
PPS	Probability Proportional to Size
PRA	Participatory Rural Appraisal
PRRO	Protracted Relief and Recovery Operation
PRSP	Poverty Reduction Strategy Paper
PSU	Primary Sampling Unit
RB	Regional Bureau
RFE	Rainfall Estimate
SADC	Southern African Development Community
SAF	Standard Analytical Framework
SAM	Severe Acute Malnutrition
SD	Standard Deviation
SENAC	Strengthening Emergency Needs Assessment Capacity (project)
SO	Special Operation
SOP	Standard Operating Procedure
SOS	Start of Season
SRS	Simple Random Sampling
SSE	sum of squares for error
SSU	Secondary Sampling Unit
SWOT	Strength, Weakness, Opportunity and Threat (analysis)
TANGO	Technical Assistance to NGOs
TB	tuberculosis
TFC	Therapeutic Feeding Centre
TOR	Terms of Reference
UNCCS	United Nations Common Coding System
UNCT	United Nations Country Team
UNDAF	United Nations Development Assistance Framework
UNDP	United Nations Development Programme

UNHCR	Office of the United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
UN-ISDR	United Nations International Strategy for Disaster Reduction
UNU	United Nations University
USAID	United States Agency for International Development
USDA	United States Department of Agriculture
USGS	United States Geological Survey
USU	Ultimate Sampling Unit
VAM	Vulnerability Analysis and Mapping
VAM-SIE	VAM Spatial Information Environment
VBA	Visual Basic for Applications
WFP	World Food Programme
WHO	World Health Organization
WAZ	weight for age
WHZ	weight for height
WR	Water Requirement
WRSI	Water Requirement Satisfaction Index

Glossary

Analysis plan – A plan, based on the key hypotheses to be tested, detailing how the collected data will be analysed. It may also guide which data need to be collected from primary sources and which from secondary.

Asset – Anything considered valuable or useful, such as a skill, a quality, or a person.[1] In the Sustainable Livelihoods Framework, the following six categories of assets are defined:

- human: health and nutrition status, physical capacity, skills, level of education, etc.;
- social: household, gender, kinship and other networks; community groups; values and attitudes; etc.;
- financial: income; credit and loans; savings; liquid assets; etc.;
- physical: productive items such as tools and equipment, stores, housing, livestock, and infrastructure;
- natural: land, water, forests, etc.;
- political: power relationships, access to – and influence over – local and higher-level government processes.

Chronic food insecurity – A long-term or persistent inability to meet minimum food requirements.

Cluster sampling – A sampling technique in which the sample is defined in two or more stages. The population of interest is first divided into groups (clusters), usually according to geographical area or location (e.g. villages are used as clusters). Second, a random sample of clusters is selected. Third, households or individuals from the selected clusters are then randomly sampled. There can be additional stages. Sampling units in the second and subsequent stages are selected from within the selected clusters from the previous stage – e.g. households from within a sampled village, or individuals from within a household.

Community group discussion – A discussion with a mixed group of community members that includes men, women, and young people from all subgroups within the community – village, camp, urban neighbourhood, etc.

Comprehensive Food Security and Vulnerability Analysis (CFSVA) – A study, typically conducted in a crisis-prone food-insecure country, that describes the food security status of various segments of the population over various parts of a country or region, with the purpose of indicating the broader underlying causes of vulnerability, and recommending appropriate interventions to deal with problems identified.

Coping strategies – Activities to which people resort to obtain food, income, and/or services when their normal means of livelihood have been disrupted.

1. Chambers Compact Dictionary, Edinburgh, UK: Chambers Harrap Publishers Ltd, 2005.

Coping strategies index (CSI) – A quick and simple indicator of household food insecurity behaviour that reveals how households manage or cope with shortfalls in food consumption. Two CSI have been proposed in the literature: a context-specific CSI and a reduced CSI. While the first is based on a series of context-specific strategies and context-specific severity scores, the second always relies on the same short list of (five) coping strategies and the same severity weights.

Data – Information collected from primary or secondary sources.

Design effect – In the context of sampling, the loss or (exceptionally) the gain in precision of statistical estimates when sampling design differs from simple random sampling.

Focus group discussion – A structured discussion to obtain qualitative information on a particular topic (the focus) with people who are knowledgeable and who have experience in that topic.

Food access – A household's ability to acquire food regularly through one or a combination of home production and stocks, purchases, barter, gifts, borrowing, and food aid.

Food availability – The food that is physically present in the area of concern, through all forms of domestic production, commercial imports, reserves and food aid. This might be aggregated at the regional, national, district, or community level.

Food consumption score (FCS) – A composite score based on the dietary diversity, food frequency, and relative nutritional importance of the various food groups consumed. The higher the FCS, the higher is the dietary diversity and frequency. High food consumption increases the possibility that a household achieves nutrient adequacy.

Food security – The state at which all people, at all times, have physical and economic access to sufficient, safe, and nutritious food to meet their dietary needs and food preferences for an active and healthy life (World Food Summit, 1996).

Food utilization – (i) A household's use of the food to which they have access; and (ii) individuals' ability to absorb and metabolize nutrients (i.e., the efficiency of food conversion by the body).

Gender perspective – An approach that includes a comparative analysis of the roles and relations between men and women, and boys and girls, with respect to division of labour, productive and reproductive activities, access to and control over resources and benefits. The perspective includes systematic investigation of socio-economic and environmental factors that influence roles and relations as well as the differential impacts of humanitarian or development intervention on women and men, girls and boys.

Gender-sensitive indicators – Indicators used to measure the extent of gender inequality (e.g. female share of total, ratio between females and males, gender gap).

Geospatial data set – A set of data that includes detailed location information. The data may be organized by geographic area or with geographic features such as coordinates, and line and polygon attributes.

Hazard – Something that has the potential to cause harm; also, the probability of occurrence of a potentially damaging phenomenon within a given time period and area.

Hazard analysis – The identification, study, and monitoring of any hazard to determine its potential, origin, characteristics, and behaviour.

HIV perspective – An approach that takes into consideration the immediate, medium-term, and long-term effect of HIV and AIDS on food security.

Indicator – A variable or combination of variables that give insight into a particular aspect of a situation.

Information needs – The data that must be collected and processed from primary and secondary sources in order to fulfil assessment objectives.

Key informant interviews – Interviews with individuals who have good knowledge about particular aspects of a community or a given emergency.

Literature review – The collection and assessment of findings in existing documents relevant to a food security analysis being conducted.

Livelihood group – A group of people who share the same basic means of livelihood and lifestyle – the same main subsistence and income-generating activities, and social and cultural practices – and who face the same risks of food and nutrition insecurity.

Livelihoods – The capabilities, assets (both material and social), and activities required for a means of living linked to survival and future well-being.[2]

Livelihood strategies – The means by which households use resources, household assets, and skills to obtain the income necessary for welfare goals such as enjoying food security, living a healthy life, having sufficient shelter, and educating their children.

Primary data – The data collected during the assessment, (e.g. interviews with key informants, focus groups, households and individuals). Primary data analysis is the process of analysing primary data.

Proxy indicator – An indicator that is used to indirectly measure a variable that is difficult to measure or cannot be measured directly.

Purposive sampling (non-probability sampling) – A method by which groups are selected for interview according to the researcher's choice. Purposive sampling does not involve random selection, so extrapolation of results to wider populations is not

2. Humanitarian Charter and Minimum Standards in Disaster Response, The Sphere Project, Geneva, Switzerland, 2004.

possible; the method's value lies in selecting information-rich cases for in-depth analysis related to the issue being studied.

Qualitative data – Observations that are categorical rather than numerical; qualitative data often includes attitudes, perceptions, and intentions.

Quantitative data – Measurements of quantities, amounts, or ranges, expressed as numbers, that can be analysed using statistical methods and models.

Questionnaire – A series of questions that have been carefully formulated and ordered to provide information from individuals, households, and communities. In a selected sample, the same individual, household, or community questionnaire is addressed to each individual, household, or community, respectively.

Random sampling (probability sampling) – A sampling method in which all members of the sampled population have a known, non-zero chance of being selected. Results can be extrapolated to the entire population with a degree of accuracy that depends on the sample size and the variability of the indicator. Based on formal statistical theory, random sampling allows reliable estimates to be calculated and minimizes bias.

Resilience – The ability to recover after being affected by a shock.

Response analysis – Analysis to determine the need, or otherwise, for an intervention and, when appropriate, to identify the most suitable types of interventions, an intervention's timing, and its targeting criteria.

Risk to food insecurity – The probability of food insecurity resulting from interactions between a natural or human-induced hazard and vulnerable conditions.

Sample – A subset of households or individuals extracted from the total population under study. Samples can be probability or non-probability samples.

Sampling frame – A complete list of potential sampling units. If households are the primary sampling units (PSU), the sampling frame is the list of all the households living in the area under study. Most of the time, food security assessments use villages/clusters as primary PSUs. In such cases, the sampling frame is the comprehensive list of villages of the study area, and for each selected village/cluster, there is a related household sampling frame consisting of all households from the village.

Secondary data – Data collected from outside the current assessment. Examples include data collected by other agencies. Secondary data analysis is the act of re-analysing existing data so that the findings inform the conclusions of the CFSVA.

Semi-structured interview – An interview based on a prepared series of questions and a checklist, the phrasing, order, and form of which are not fixed.

Shock – An event that has a negative impact on food and nutrition security. Shocks can be natural or caused by human action.

Simple random sampling – A technique in which the primary sampling units (PSUs) are selected directly from the sampling frame. Each unit has the same probability of being selected. No intermediate steps are undertaken.

Stage sampling – See cluster sampling.

Stratified sampling – A sampling method by which the population of interest is split up into subgroups (i.e. strata) that have something in common. In the context of food security analysis, administrative boundaries or food security zones can be strata.

Thematic map – A map that displays the spatial pattern of a single theme or series of attributes related to a singe subject matter.

Transitory food insecurity – A short-term or temporary inability to meet minimum food requirements, indicating a capacity to recover.

Triangulation – A process for comparing information from different sources to determine if evidence converges.

Vulnerability to food insecurity – Conditions that increase the susceptibility of a household to the impact on food security in case of a shock. Vulnerability is a function of how a household's livelihood would be affected by a specific hazard and how it would manage to cope with this impact.

CHAPTER 1

Introduction

WFP's mandate to address hunger and food insecurity demands a comprehensive understanding of household food security situations, particularly in fragile nations, to enable an effective response. WFP food security/vulnerability assessment and analysis is a key tool for programme formulation, and at country level, the Vulnerability Analysis and Mapping (VAM) units provide regular and comprehensive analysis of the prevailing food security situation in a given country to guide WFP and its partners' operational responses. Comprehensively addressing various aspects of food security ensures that WFP plans relevant and efficient interventions based on timely identification of the problem and thorough analysis of its impact on the affected population. This in-depth understanding ensures that WFP targets only those most in need. It also ensures that WFP's response strategies (general food distributions, food for work, food for education, and food for health) are appropriate for saving lives and strengthening livelihoods, thereby reducing future vulnerability.

As of 2008, WFP is operational in 80 countries, and its large technical field network of more than 100 VAM staff members gives it a comparative advantage to undertake food security analysis. Since 2005, WFP, through the multi-donor-funded Strengthening Emergency Needs Assessment Capacity (SENAC) project, has invested millions of dollars to strengthen its food security analysis methodology.
Throughout its existence, the Food Security Analysis Service (OMXF, which incorporates VAM) has worked to provide normative technical guidance in the analysis of food security and vulnerability. Although some of this guidance was published, much of the methodological support given to the field remained without a formal written record.

Under the SENAC project, a refinement and consolidation of food security analysis methodology, primarily as it is used in Comprehensive Food Security and Vulnerability Assessments (CFSVAs), is taking place. The existing draft guidelines for this analysis methodology, along with guidelines on the core principles and themes surrounding CFSVAs, needed to be consolidated, edited, and harmonized. They aim to guide VAM officers and partners with a food security analysis background, in the conduct of the CFSVA.

What are CFSVAs?

The CFSVA process generates a document that describes the food security status of various segments of a population over various parts of a country or region, analyses the underlying causes of vulnerability, and recommends appropriate interventions to deal with the problems. CFSVAs are undertaken in all crisis-prone food-insecure countries. Given their relevance, analytical rigour, and comprehensiveness, CFSVAs have become a key reference for decision makers involved in planning and implementing food security–related programmes. The location-, population-, and livelihood group–specific information and analysis provided through these studies are used to design and implement broader sectoral interventions to address the supply-and demand-side constraints to food security.

CFVSAs involve activities ranging from reviewing existing literature and data to undertaking surveys to collect and analyse primary data. A CFSVA can be an extensive exercise, usually taking around four to eight months from initiation to dissemination of results. The timeframe depends on the size of the study, the political environment, and other characteristics of the country.

The overall objective of a CFSVA is to analyse the food security and vulnerability condition of population groups and communities, and to provide baseline information on the population in a “normal” situation. CFSVAs provide to WFP decision-makers and partners in government, civil society, and the donor community timely and relevant information on household food insecurity and vulnerability, who and how many people are affected, and where they are located, allowing for recommendations on (food) interventions to improve the situation.

CFSVA principles

- CFSVAs should be undertaken, where possible, in partnership with other United Nations system agencies, government counterparts, and key civil society organizations.
- All CFSVAs should include a thorough literature review and secondary data analysis to identify data gaps and, when necessary, to justify primary data collection.
- Analytical methods found in CFSVAs should be clearly defined so as to ensure transparency.
- The shelf life of CFSVAs is determined by the indicators being collected and reported. In most situations, CFSVA findings are valid for three to five years, unless there are drastic food security changes in the meantime.
- CFSVAs can be completed using secondary data in countries and regions where such data are recent and of good quality.

How are CFSVAs used?

Needs assessments:

- CFSVAs can inform the design of WFP operations - especially in the context of protracted relief and recovery operations (PRROs), country programmes, special development activities, and, in some situations, emergency operations (EMOPs).
- In complex emergencies and post-conflict situations, CFSVAs can provide an important form of updated information for all sectors and partners until standard surveys (e.g. income/expenditure, demographic and health surveys [DHS]) can be conducted.

Baseline vulnerability analysis:

- Information found in CFSVAs can be used to design and implement food security monitoring systems, which track key trends and regions within a country.
- CFSVAs can act as a benchmark for emergency food security assessments in the event of a crisis or shock (pre-crisis baseline). Since comparability is a key element, definition of indicators should be standardized and the sampling approach made compatible with eventual subsequent emergency food security assessments (EFSAs).
- In some situations, CFSVAs can be used as a “global food security baseline” against which the impact of specific WFP projects/programmes can be compared (project/results monitoring).

Government policy and rural development initiatives:

- CFSVAs are a strategic entry point for partnership and collaboration with other United Nations system agencies - especially in the context of Poverty Reduction Strategy Papers (PRSPs), UNDAF Common Country Assessments (CCAs), and the United Nations Cluster Approach.
- CFSVAs can be the umbrella under which new WFP research efforts - such as market analyses, cross-border trade, safety nets - are launched.

- CFSVAs can be used as the basis for planning capacity-building initiatives in collaboration with government partners.

CFSV analysis[3] is based on a particular understanding of food security and vulnerability. The Food and Nutrition Security Conceptual Framework presented in Figure 1.1 informs not only the selection of indicators for analysis and use in geographic targeting, but also the design of field assessment instruments and the organization of standardized reporting formats.

1.1 PURPOSE AND STRUCTURE OF THIS DOCUMENT

The purpose of these guidelines is to provide the CFSVA analysis team with currently recommended procedures and protocols for undertaking a CFSVA. It is not a manual or protocol, but rather a collection of guiding frameworks, tools, and approaches to CFSVA planning, implementation, analysis, and dissemination. It consolidates existing CFSVA guidelines into one unique, comprehensive document. It should be used as a reference to ensure that most aspects of a CFSVA exercise are adequately covered. The reader should already possess basic knowledge of food security and social research techniques, which should be applied according to the guidance in this document. These guidelines are organized in the following way. First there is a section on the key planning steps for implementing a CFSVA. This is followed by sections organized by the different types of data used: desk review, household and community data collection, and how such data are typically analysed. The next section covers food security analysis, and how information from all sources is combined to answer the key questions of the CFSVA. The document ends with sections on conclusions and response options and report preparation and dissemination. Gender and HIV/AIDS are cross-cutting elements of the document.

1.2 A LIVELIHOODS APPROACH FOR CFSVA[4]

1.2.1 Key terms, concepts, and issues

In the last decade, international and national agencies have used the concept of sustainable livelihoods and the application of livelihoods analysis as a means to better understand and respond to the multidimensionality of poverty and food insecurity. Given that the causes of poverty are complex, it is essential in a CFSVA to understand the web of poverty and people's mechanisms for dealing with it (CARE, 2002).

A DEFINITION OF LIVELIHOOD
A livelihood comprises the capabilities, assets (stores, resources, claims, and access) and activities required for a means of living: a livelihood is sustainable which can cope with and recover from stress and shocks, maintain or enhance its capabilities and assets, and provide sustainable livelihood opportunities for the next generation: and which contributes net benefits to other livelihoods at the local and global levels in the long and short term.
Source: Chambers and Conway, 1992

3. Part of this section is taken from VAM, *VAM Standard Analytical Framework: Role and Objectives of VAM Activities to Support WFP Food-Oriented Intervention*, WFP, Rome, 2002.
4. Information is partly taken from *Integrating "Livelihoods" into Food Security and Vulnerability Analysis: Some Initial Guidance*, WFP draft, 2005.

Livelihood systems are made up of several components:

- The activities households engage in to earn income and make a living. These include a range of on-farm and off-farm activities that together provide a variety of procurement strategies for food and cash.
- The assets and other resources a household possesses.
- Social networks and safety nets, the human and social capital that a household possesses or can call on in times of need.

Thus, livelihood systems are quite diverse. Each household can have many possible sources of entitlement (i.e. the rights, privileges, and assets a household has, and its position in the legal, political, and social fabric of society). (CARE, 2002).

Livelihood strategies and outcomes

Generally speaking, the goal of household livelihood strategies is to improve welfare levels in some way, ensuring that: (a) the household has enough to eat; (b) their fluctuating income is stabilized; (c) children are able to go to school; (d) the household can afford or access health services; or (d) natural resources are better managed.
Households often use their skills and know-how to diversify income sources and offset risks.

Households are able to meet their needs through six main tangible and intangible assets.

1. **Human capital:** skills, knowledge, ability to labour, nutritional status of adults and children;
2. **Financial capital:** financial resources, savings, credit, liquid assets;
3. **Natural capital:** types and quantities of crops grown and harvested;
4. **Physical capital:** assets and land available to households;
5. **Social capital:** informal community support networks, extended family structures, or community labour-sharing systems; and
6. **Political capital:** participation in community decisions and power relations.

Levels of security derived from these assets are generally termed **livelihood outcomes**, a set of factors that govern household welfare. It is important to take into account: which resources must be combined or transformed to ensure sustainable livelihoods; the tradeoffs that exist between resources; which resources are prerequisites to others; and the trends in long-term use (adapted from Scoones, 1998).
Household livelihood security is defined as adequate and sustainable access to income and resources to meet basic needs. Basic needs include food, proper nutrition, clean water, health and health facilities, economic and educational opportunities, housing, physical safety, and time for community participation and social integration.
Having enough to eat is one of the livelihood outcomes. A household's attempts to secure sufficient amounts of food are a central component of its livelihood strategies. Therefore, food security is a sub-component of household livelihood security. However, food is only one important basic need among several, and adequate food consumption is sometimes sacrificed for other important needs.

1.2.2 Policies, institutions, and organizations

Policies and institutions affect the livelihood choices of poor and food-insecure households and are the last piece of the conceptual framework. Policies are generally split into the following three broad categories:

- **Macroeconomic policies** place the focus on medium- and long-term measures that aim to stabilize a given economy (currency devaluations, labour markets, interest rates on borrowing capital, privatization, financial liberalization, public investments, and trade liberalization).
- **Social policies** place the focus on measures that can improve health and nutrition, education, safety nets, and social protection schemes for the disadvantaged.
- **Sectoral policies** place the focus on specific areas within an overall economy and society (e.g. agriculture, water supply, management and sanitation, energy infrastructure, and the environment).

Institutions and organizations, the structures through which policies are formulated and implemented, represent the interface between households and policymakers.

- The **state**, in addition to services, may provide safety nets, change policies, or limit freedoms, all of which can have positive or adverse affects on livelihood systems.
- **Formal civil society** may offer support of conditions that enable households; or may confine household opportunities.
- **Informal civil society** may negatively or positively influence the livelihood strategies pursued by households.
- The **private sector** may augment or constrict opportunities for households.

The political, institutional, and economic environment has a profound effect on household livelihoods (assets, strategies, and outcomes). For example, the local agricultural policy governing input and output markets has an effect on whether households whose main activity is farming can effectively use their land, labour, water, and livestock (livelihood assets). Input and output markets facilitate the production, movement, and exchange of agricultural commodities (e.g. seeds, fertilizer, storage, marketing, farm-gate purchases). If such systems are inefficient, then farming-based livelihoods are rendered ineffective, leading to losses in income and contributing to a broader decrease in household welfare. This shows clearly how policies and institutions can affect **availability** of assets, **access** to those assets, and ability to utilize assets productively.

1.3 THE FOOD AND NUTRITION SECURITY CONCEPTUAL FRAMEWORK

CFSVA is based on a particular understanding of food security and vulnerability. The Food and Nutrition Security Conceptual Framework informs not only the selection of indicators for analysis and use in geographic targeting, but also the design of field assessment instruments and the organization of standardized reporting formats. The Food and Nutrition Security Conceptual Framework adopted by CFSVAs considers food availability, food access, and food utilization as core elements of food security, and links them to households' asset endowments, livelihood strategies, and political, social, institutional, and economic environment. The strength of the household livelihoods approach lies in its ability to obtain a holistic and multidimensional profile of a micro-level context

- food, nutrition, livelihood, and rights-realization - with strong regional and national contextualization, allowing for the scaling-up of interventions (CARE, 2002).

Food security was broadly defined in the 1996 World Food Summit Plan of Action with the following text:

Food security exists when all people, at all times, have physical and economic access to sufficient, safe, and nutritious food to meet their dietary needs and food preferences for an active and healthy life.

While this is a goal-level notion, it is important to define operational measures of food insecurity, vulnerability to food insecurity, and its determinants. The CFSVA focuses on identifying specific metrics for food insecurity and vulnerability and it adopts a risk analysis framework for understanding the distribution and causes of vulnerability and resiliency of countries, regions, communities, and households.

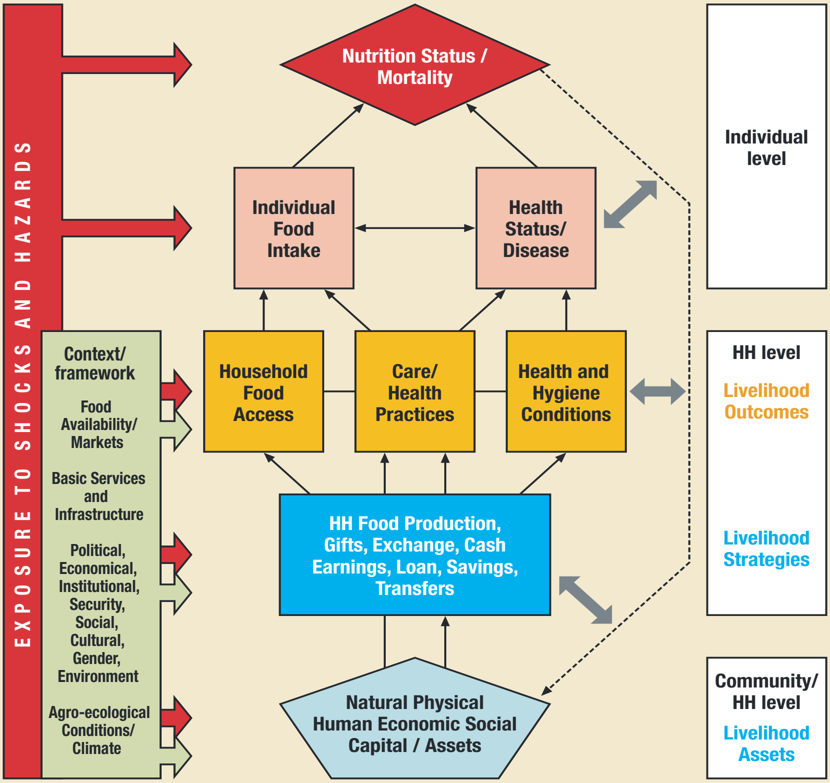
During a CFSVA, this framework serves two purposes by providing:
- a basis for developing initial hypotheses on the level of vulnerability and food insecurity, and the causes and effects of both; and
- a succinct way of visualizing the relationships among factors that affect food and nutrition security, which is helpful during data collection and analysis.

The Food and Nutrition Security Conceptual Framework is based on UNICEF's Nutrition Framework and the (DFID) Sustainable Livelihoods Framework.
The analysis of food security begins with an examination of livelihood assets; the agro-ecological, political and institutional context of the area; and the resulting livelihood strategies adopted by the people that may lead to food security. Various hazards and more gradual changes affect the macro context and household-level assets and strategies, and hence household food security.

The food security status of any household or individual is typically determined by the interaction among a broad range of agro-environmental, socio-economic, and biological factors. As with the concepts of health or social welfare, there is no single, direct measure of food security. However, the complexity of the food security problem can be simplified by focusing on three distinct, but interrelated, dimensions: aggregate food availability, household food access, and individual food utilization.
Achieving food security requires addressing all three of these separate dimensions, ensuring that:
- the aggregate availability of physical supplies of food from domestic production, commercial imports, food aid, and national stocks is sufficient;
- household livelihoods provide adequate access for all members of the household to those food supplies through home production, market purchases, or transfers from other sources; and
- the utilization of those food supplies is appropriate to meet the specific dietary and health needs of all individuals within a household.

Vulnerability is a forward-looking concept aimed at assessing community and household exposure and sensitivity to future shocks. Ultimately, the vulnerability of a household or community is determined by their ability to cope with their exposure to

Figure 1.1: Food and Nutrition Security Conceptual Framework



the risk posed by shocks such as droughts, floods, crop blight or infestation, economic fluctuations, and conflict. This ability is determined largely by household and community characteristics, most notably a household's or community's asset base and the livelihood and food security strategies it pursues.

The framework shows that exposure to risk is determined by the frequency and severity of natural and man-made hazards, and their socio-economic and geographic scope. The determinants of coping capacity include household levels of natural, physical, economic, human, social, and political assets; levels of household production; levels of income and consumption, and, most important, the ability of households to diversify their income and consumption sources to mitigate the effects of any risks they face.

Coping behaviour involves activities such as the sale of land or other productive assets, the cutting of trees for sale as firewood, and, in an extreme example, the sale of girls into prostitution. These practices undermine not only the long-term productive potential of vulnerable households, but also important social institutions and relationships. The extent of reliance on these destructive practices is an indicator of vulnerability levels during a crisis.

While an understanding of how households cope is important to analysis, knowing how well households cope, or the resilience of household livelihoods, is more important.

How well the local economy can absorb the additional labour or products, such as livestock or firewood, that appear on the market as the result of coping behaviour during a disaster, and the stability of wages and prices for those products, are critical factors in understanding vulnerability.

Food security analysis is primarily a static view of food access and household constraints to that access, from either a short- or long-term perspective. In contrast, risk and vulnerability analysis, because it includes the element of risk that households face in their day-to-day decision-making and their capacity to respond effectively over time, views food access from a more dynamic, forward-looking perspective.

In the end, there is a significant overlap between households that are currently food insecure and those at risk to the severe fluctuations in food access that threaten well-being. While, in concept, all households may be considered vulnerable to a certain degree, from an operational perspective, the primary emphasis of vulnerability analysis should be on households that are nearly, or already food insecure.

CFSVAs should provide stakeholders with an analysis of food insecurity and livelihoods at the sub-national level by addressing the five VAM questions:

1. Who are the food-insecure or vulnerable people?

VAM surveys are conducted at the household level. The collected information is regrouped and analysed in order to create livelihood groups. Looking at household expenditure and income, the analyst is able to determine which are the most vulnerable households and what risks (drought, flood, pest, insecurity) will affect them the most. In Liberia, it was noted that households that had recently returned were particularly vulnerable to food insecurity, as they had to restore their livelihoods in an environment that had been destroyed by the war. These households are now a priority for WFP.

2. How many people are food insecure or vulnerable?

During the design phase of the survey, a sample of households is drawn using probability sampling methods. The prevalence of food insecurity and vulnerability found in the sample is applied to the entire population from which the sample was drawn in order to estimate the total number of food-insecure and vulnerable people. (For example, in Mali, VAM estimated in 2005 that 6.2 million people were food insecure and vulnerable.) These numbers are then used to target WFP PRROs and EMOPs.

3. Where do the food-insecure and vulnerable people live?

CFSVAs provide an essential package of maps showing the areas most affected by food insecurity and vulnerability. These are crucial tools for decision-makers and for targeting aid. The maps are produced by VAM staff with considerable experience in geographical information systems.

4. What are the underlying causes and threats of food security and malnutrition?

CFSVAs collect a wide range of information that allows VAM and WFP to explore the determinants of food insecurity/vulnerability. Using qualitative and quantitative techniques, together with local expert judgment, the CFSVA analysis team is able to identify the local contextual causes of food insecurity and vulnerability.

5. What are the implications for food security interventions?
VAM gives recommendations for interventions in a country based on the conclusions of the CFSVA and input from WFP programme officers and partners in development involved in the CFSVA and in the field of food security.

1.4 BASIC GENDER CONCEPTS, FRAMEWORKS, AND INDICATORS[5]

1.4.1 Gender analysis and CFSVA studies

The purpose of gender analysis is to determine gender disparity. This knowledge can then be incorporated into gender-responsive programming with positive measures taken to level the playing field. The effective integration of gender analysis into CFSVA studies entails exploring how gender roles relate to all aspects of food security (availability, access, and utilization) and food aid interventions. Specific issues include:

- Understanding how gendered division of labour and decision-making power are related to food availability and access;
- Exploring variability of food consumption (i.e. utilization), health, and nutrition by gender and how these factors affect food utilization for both genders;
- Analysing how the benefits of food aid interventions can be effectively targeted to both men and women and used to promote gender equality; and
- Anticipating any negative impacts interventions may have on women or men, girls or boys, or on gender relationships.

Applying a gender perspective to CFSVA studies demands that a gender-sensitive approach be taken during research design, data collection, data analysis, reporting, and, ultimately, programme planning. This requires an explicit sensitivity to the varying needs of men and women. It is therefore crucial to involve men and women in all stages of the research, and to sensitize enumerators and other research team members to gender issues relevant to the context in which a study is being conducted.

Avoiding assumptions
Although the term gender has often been misinterpreted as focusing on women, a gender perspective requires a comparative analysis of men and women, as well as the relations between them.

Assumptions concerning the relationship between gender and vulnerable groups are inappropriate prior to the analysis of the particular context under study, and run the risk of introducing bias into the research design.

1.4.2 Gender analysis frameworks

Gender relates to all three aspects of food security:
- **food availability** - productive, reproductive and community roles;
- **food access** - differentiated access to and control over resources, power, and decision-making at the household and community level; and
- **food utilization** - caring practices, reproductive health, gender-specific diseases.

5. Information in this section draws heavily on the *Thematic Guidelines: Integrating a Gender Perspective into Vulnerability Analysis*, WFP, 2005.

Box 1.1: Challenges of incorporating a gender perspective in different settings

In some of the settings in which CFSVA operates, gender disparities are obvious, and inform and shape everyday life. In these settings, the challenge is to bring a gender perspective to the analysis of food security and vulnerability, while respecting local culture.

In other settings, gender disparities are more subtle. The challenge in these settings is to design perceptive studies that are able to capture less obvious gender variance, inequity in gender relations, and the relationship of these factors to food security and vulnerability.

The following case provides an example of how gender inequities can be both pervasive and obvious, and more subtle and nuanced.

In Malawi, gender inequalities exist because of discrimination within families and institutions, and because of the social and cultural norms that perpetuate the beliefs and practices detrimental to women. The greatest challenge to achieving gender equality is overcoming socially accepted cultural beliefs and ideologies that emphasize male dominance. In Malawi, marriage customs that establish men as dominant heads of household are an important form of gender discrimination. These customs underlie property rights and inheritance practices that disadvantage women. The socialization of boys and girls to assume different roles and the norms limiting women's mobility are also important factors in gender inequality in the country. The universal acceptance of gender inequality helps bring about unequal access to opportunities, resources, and assets for women and men. Results of the Malawi Integrated Household Survey show that the sex of the head of household is a statistically significant variable for poverty. Forty-one percent of rural households are food insecure, and 40 percent of these are female-headed. Studies have found that even female-headed households without significantly lower average incomes fare poorly in indicators of human capabilities including health, education, and employment.

Source: TANGO International, Gender Exploitation in Malawi, prepared for CARE Malawi, 2004.

Several conceptual frameworks provide examples of how a gender perspective can be applied to studies of food security and vulnerability.

DFID[6] has developed a Gender Analysis Framework that offers key issues to consider in four areas of enquiry: gender roles, assets and livelihoods, power and decision-making, and needs analysis.

Roles and responsibilities

- What do men and women do?
- Where (location/patterns of mobility)?

6. DFID Infrastructure Department

- When (daily and seasonal patterns)?
- What are their productive roles (paid work, self-employment, and subsistence production)?
- What are their reproductive roles (domestic work, child care, and care of the sick and elderly)?
- What is their community participation/self-help (voluntary work for the benefit of the community as a whole)?
- What is their community politics (decision-making/representation on behalf of the community as a whole)?

Assets

- What livelihood assets/opportunities do men and women have access to and control over?
- What constraints do they face?
- What are their human assets (e.g. health services, education)?
- What are their natural assets (e.g. land, natural resources)?
- What are their social assets (e.g. social networks)?
- What are their physical assets (e.g. infrastructure)?
- What are their economic assets (e.g. capital/income, credit)?

Power and decision-making

- What decision-making do men/women participate in?
  - Household level (e.g. expenditure decisions, use of savings)?
  - Community level (e.g. decisions on the management of community water supplies)?
- What decision-making do men/women usually control?
  - Household level (e.g. expenditure decisions, use of savings)?
  - Community level (e.g. decisions on the management of community water supplies)?
- What constraints do they face?

Needs and priorities

- What are the needs and priorities of both men and women?
  - “Practical” gender needs - inadequacies in immediate necessities such as water access, food, and employment (e.g. a more convenient water point to save women time and energy)?
  - “Strategic” gender needs - structural changes that challenge subordinate roles and create greater equality (e.g. legal rights, equal wages, reproductive choice)?
- What perspectives do they have on appropriate and sustainable ways of addressing their needs?

Gender, while not explicitly illustrated in the Food and Nutrition Security Conceptual Framework pictured in Figure 1.2, is a critical dimension of food security and should always be taken into consideration. The questions listed in section 1.4.2 provide good guidance on how a gender lens can be applied to this framework and guide analysis and project design.

1.4.3 Gender-sensitive indicators[7]

Various indicators can be used to measure the extent of gender inequality, based on the number of females and males in a given context (female share of total, ratio between females and males, gender gap) or by comparing the variable outcomes among the two subgroups.

An illustrative list of generic, quantitative indicators that can be used to incorporate gender analysis into food security and vulnerability studies is provided in Table 1.1. Examples of their application in developing indicators are also provided (in the right-hand column). These indicators are a useful tool for ensuring that sex-disaggregated quantitative data are generated during primary data collection and allow for gender analysis to be incorporated into the overall food security and vulnerability analysis.

Table 1.1: Gender-sensitive indicators

Indicator	Formula	Interpretation	Example
Female share of a total	# (females) *100 / # (females + males)	• 50% = gender equality • <50% = females are underrepresented • >50% = males are underrepresented	Share of women participating in political meetings at the community level
Ratio between females and males	# (females) / # (males)	• 1 = gender equality • The closer to 0, the more females are underrepresented • >1 = males are underrepresented	The ratio between girls' and boys' school enrolment rates (no. of girls per 1 boy)
Female characteristic as percentage of male characteristic	mean female characteristic *100 / mean male characteristic	• 100% = gender equality • The closer to 0%, the more females are disadvantaged • Values >100% = males are disadvantaged	Average earnings of women as percentage of average earnings of men
Gender gap (% difference between no. of females and males vs. the no. of males in the same population)	(# males - # females) *100 / # males	• 0% = gender equality • The closer to 100%, the more females are disadvantaged • Values <0% = females are advantaged	Differences in school enrolment between boys and girls; differences in access to (or control over) productive assets between men and women

1.4.4 Gender dynamics

Understanding gender relations and dynamics is critical to our understanding of livelihood systems and intra-household issues. Often gender plays a large role in the division of labour, access to goods and services, control over resources, and power relations and rights.

Women and men often allocate resources differently, which has a differential impact on household welfare. Women frequently allocate more resources to meet a household's

7. See also CIDA: *Guide to Gender-Sensitive Indicators*, Quebec 1997, 9-13.

basic needs than men do. However, much of women's work is unpaid and/or taken for granted. As a result, it is often not counted, and their contribution to household livelihood security is thus undervalued. Additionally, women often have limited power in household decision-making, and in choosing how they will contribute to the household livelihood system.

1.5 UNDERSTANDING THE IMPACT OF HIV/AIDS ON LIVELIHOODS[8]

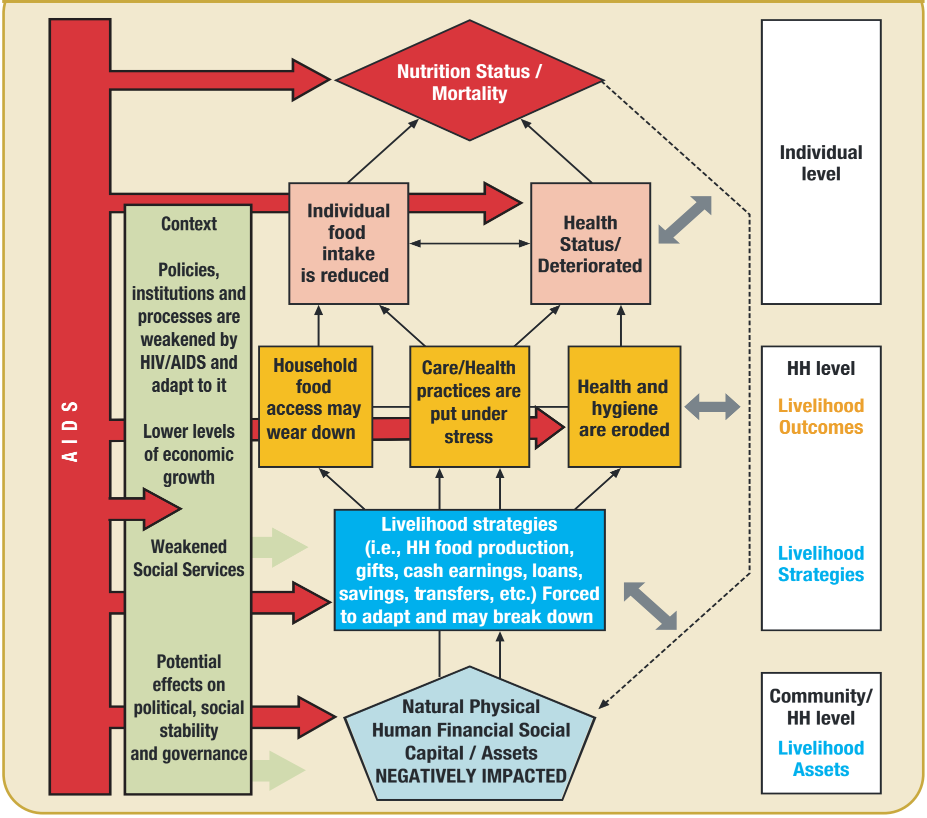
The Food and Nutrition Security Conceptual Framework helps to demonstrate the impact of HIV/AIDS on food security (see Figure 1.2). The framework underlines that illnesses and deaths due to AIDS have both an immediate and a long-term impact on households' and communities' vulnerability to food insecurity. It suggests considering both the direct impact of AIDS at all livelihood levels (human, financial, social, natural, and physical) and the indirect impact of policies, institutions, and processes on livelihoods.

Finally, it draws attention to the feedback loop generated by the epidemic: livelihood assets are often negatively impacted by AIDS; livelihood strategies are usually adapted in response to HIV/AIDS, but the strategies can hardly prevent the increase in poverty and food insecurity. This increases susceptibility to HIV/AIDS. Most of the studies that adopted the Sustainable Livelihoods Framework found that HIV/AIDS had significant impacts on all capital assets, including human, financial, social, natural, and physical.

In countries highly affected by HIV/AIDS, consideration for HIV/AIDS should be mainstreamed in each component of the CFSVA, including:
- Secondary data/literature review;
- Collection of household-level data; and
- Collection of community-level data (e.g., infrastructures, perception of the community).

8. Detailed guidance on how to mainstream HIV/AIDS into CFSVA is provided by the technical guidelines *"HIV/AIDS Analysis: Integrating HIV/AIDS in Food Security and Vulnerability Analyses"* developed by WFP VAM in 2007. Guidelines are available online on the Food Security Analysis/VAM website, www.wfp.org/food-security.

Figure 1.2: HIV/AIDS in the Food and Nutrition Security Conceptual Framework



1.6 KEY REFERENCES

- ACDI/VOCA. 2006. *Food Security Needs Assessment Toolkit.* Prepared for ACDI/VOCA by TANGO International Inc., Tucson, Arizona, 2006.
- CARE. 1999. *Household Livelihood Security Training and Facilitation Manual.* CARE USA, Atlanta, PHLS Unit.
- ibid. 2002. *Household Livelihood Security Assessments: A Toolkit for Practitioners.* CARE USA, Atlanta.
- DFID. 2001. *Sustainable Livelihoods Guidance Sheets*. London, UK.
- Scoones, I. 1998. *Sustainable Rural Livelihoods: A Framework for Analysis.* Discussion Paper. Brighton, UK: Institute of Development Studies.
- UNICEF. 1990. *Strategy for Improved Nutrition of Children and Women in Developing Countries.* New York.
- VAM. 2002. *VAM Standard Analytical Framework: Role and Objectives of VAM Activities to Support WFP Food-Oriented Intervention*, WFP Rome.
- WFP 2006. *Qualitative Analysis Guidelines for Comprehensive Food Security and Vulnerability Analysis*. Rome.

CHAPTER 2

Managing the implementation of a CFSVA

2.1 INTRODUCTION

The CFSVA process must be carefully managed. It has a number of moving parts that, when they work together, result in high-quality information useful for programming and accountability. Therefore, when adequate time and resources are invested in managing the process, the end result is a product that is useful and of high value.

2.2 INITIAL STEPS FOR SMOOTH IMPLEMENTATION

Adhering to the following sequence of steps encourages the smooth implementation of a CFSVA.

Box 2.1: Steps for implementing a CFSVA

- Country offices, in agreement with their local partners and government agencies (e.g. through an established committee on food security), inform the regional bureau and Headquarters about the necessity of undertaking a CFSVA.
- Headquarters, regional bureaux, and country offices determine the scope of the study and finalize funding and potential sources.
- All partners (including country offices, regional bureaux, and Headquarters) agree on their respective roles during the entire CFSVA process and include this in the study's terms of reference.
- Communication and management lines are established at the country office, regional bureau, and Headquarters.
- The country office requests permission from the national authorities to undertake a country-wide survey.
- The country office, regional bureau, and Headquarters identify and deploy qualified staff/consultants to lead the management and technical aspects of the study.
- Results are shared with local partners/government agencies for discussion.
- The report is finalized.
- The report is officially launched.
- The results are incorporated into country office planning and government food security initiatives.
- The relevance of the CFSVA is periodically reviewed to determine when a new CFSVA is required.

2.3 SCOPE OF A CFSVA

The level of resources required to conduct a CFSVA depends on a number of factors, including:
- the number of partners involved and their total contribution (financial, logistical, and/or human resources);
- the size of the sample for the quantitative survey;
- the amount of qualitative information employed to complement the quantitative survey;
- the amount of external assistance used to manage and conduct the literature review, survey, and analysis of primary and secondary data; and
- the costs associated with the survey logistics.

Costs can vary greatly due to the variable contexts within which CFSVAs are conducted. CFSVAs are not necessarily more expensive than large rapid assessments, but they can be costly if a large household survey is included. Large surveys on the order of 2,000 to 3,000 households can easily cost in the range of US$75,000–US$100,000[9] and consume considerable staff time and effort. CFSVAs with larger samples covering many different locations cost well over US$200,000.

Beyond the financial cost of conducting a CFSVA is the required level of effort. From initial planning to report writing, an assessment can typically last from four to nine months. It requires day-to-day management from WFP staff or contracted personnel (often an assessment coordinator, for example, a WFP staff member or consultant, is assigned full time to the effort). Some of the skills needed to plan and conduct a CFSVA (such as sampling design and data analysis) may not reside in the WFP country office conducting the CFSVA. When the required skills are not present, then the WFP project officer is more focused on managing the process, ensuring that appropriate procedures are followed and that the quality of the information is high. This role will also include administrative, communication, logistical, and other duties. This may mean that functions such as assessment design, sampling, data analysis, and reporting are contracted out to individuals or firms with the necessary skills for these technical areas.

2.4 MOBILIZING PARTNERS AND FUNDING

CFSVAs require specific skills, including quantitative methods with some grounding in basic statistics and experimental design, food security, livelihood security, qualitative data collection, and data analysis and interpretation. Aspects to be considered may include: production and marketing of food or cash crops; livestock, hunting, fishing, and gathering activities; wage employment; provision of skilled/ professional services; home production of items for sale; access to micro-finance services; access to markets; and power relationships.

Relevant expertise and practical experience must be mobilized to:

- help draw up the terms of reference for the assessment and elaborate the work plan;
- design the sampling strategy;
- design the logistic plan associated with data collection; and
- participate either in the assessment or in the selection of an appropriate entity to be commissioned to undertake it.

Such expertise may be mobilized within the country through:

- the UN country team — especially FAO for agriculture, forestry, livestock, and fisheries aspects; UNICEF for health, nutrition, and education aspects; and on occasion ILO, for employment, micro-finance, skills training, and income-generating aspects;
- national/governmental institutions, universities, specialized NGOs, and donor missions; and
- national and international consultants, accessed through either reputable firms or individuals.
- regional offices or WFP Headquarters may be able to assist in identifying appropriate assistance.

9. Direct costs for a CFSVA do not include the time and cost of core staff at Headquarters, regional bureaux, or country offices.

2.5 INITIAL PLANNING WITH PARTNERS

Once the need for a CFSVA has been identified and relevant national and international institutions have been invited to participate, the analysis is launched with an exploratory meeting. In many countries, the main food security actors have established committees to coordinate their activities, and such committees should be included in initial planning. At this meeting, all relevant partners confer upon managerial, financial, and technical aspects of the study. Meeting participants should:

- Gauge partner interest in participating;
- Determine if major partners are missing;
- Identify topics of particular interest to partners;
- Get an overview of existing reports and data and determine the need for additional data collection;
- Highlight envisioned constraints, particularly the survey timing;
- Discuss proposed methodology;
- Agree on lead responsibility for literature review, secondary data analysis, and the survey modules;
- Propose methods of data collection, e.g. personal digital assistants (PDAs), paper forms;
- Review survey data use and ownership issues; and
- Identify support that partners could provide, including:
  - *Technical staff in their respective competencies;*
  - *Enumerators, field supervisors, team leaders;*
  - *Logistics;*
  - *Administrative; and*
  - *Financial.*

2.5.1 CFSVA planning and preparation

CFSVAs require careful planning and preparation. Following the initial meeting, an action plan with details of the overall approach must be developed. The plan will, at a minimum, include:

- Establishment of a CFSVA or Food Security Committee;
- A review of secondary data or literature as an initial source of information for planning;
- A definition of scope and objectives; and
- The terms of reference:
  - *assessment objectives;*
  - *a timetable for the entire process that specifies when each set of tasks must be started and completed;*
  - *roles, activities, and responsibilities for each participating member; and*
  - *geographic scope of the assessment.*
- Development of a sampling strategy in collaboration with partners;
- A methodological overview of data collection and the analytical approach to data analysis (production of an analysis plan);
- Site selection and rationale;
- Logistical planning;

- An estimate of total resources needed to implement the survey:
  - personnel;
  - survey equipment;
  - logistics; and
  - duration of fieldwork.
- Reporting processes; and
- Use of the information in food security interventions.

2.5.2 Establishing a CFSVA or food security committee

In many countries where WFP is a participant, and at times the lead agency, food security committees already exist. It is in the interest of WFP to use existing structures when possible. If a structure does not exist, it may be useful to establish a committee to:

- agree on the terms of reference and the selection of the CFSVA team (staff and/or consultants);
- create the opportunity for technical discussion within and among sectors;
- secure the collaboration of all parties in the field to facilitate the survey;
- get buy-in and support from various agencies;
- resolve any problems that arise;
- review the report and collaborate in following up on the findings and recommendations; and
- agree on the procedure for validating and disseminating the report.

Such a committee would normally include representatives of the same entities that could be involved in a joint assessment or review of the operation as a whole, namely: WFP (in particular, focal points from VAM, nutrition, and food security), interagency representative(s), one or more representatives from national and local government, one or two representatives of the major donors, and representatives of the major NGOs involved in food security and emergency response.

2.5.3 Defining scope and objectives

The scope of the assessment will vary according to the types of shocks experienced by areas or populations and by their relative impacts on the food security of particular households and communities. For example, sudden-onset disasters (floods, earthquakes, cyclones, etc.) may call for a different approach to assessing vulnerability than a slow-onset crisis such as drought, or a complex emergency involving violent conflict. Each of these scenarios might call for a different approach to CFSVA assessments in order to adequately address the food insecurity and vulnerability experienced by the given population.

Box 2.2: CFSVA objectives

a) Answer the five key CFSVA questions.
b) Train and capacity-build to enhance WFP staff and partner capabilities, particularly in-country, to manage and/or conduct such assessments.
c) Answer additional questions introduced by WFP or key partners.

CFSVAs make use of systematic quantitative methods to estimate the prevalence of food insecurity vulnerability and its determinants, and to collect qualitative information and secondary data related to trends and risks. Triangulation using qualitative and quantitative methods permits identification of food-insecure and particularly vulnerable population groups. It also permits a more detailed understanding of livelihoods, and of the modalities and targeting for food and non-food aid that will best meet the needs of the population. Assumptions that describe causal relationships between two or more variables are developed based on the determined assessment objectives.

2.5.4 Develop the terms of reference

A CFSVA always requires clear terms of reference (TOR).[10] The TOR for a CFSVA mission should include a rationale, the objectives and timing, roles, activities and responsibilities for each participating member, the phases, and the geographic scope of the assessment. The objectives are typically the five VAM questions (see section 1.3). The TOR should provide a timeline with a list of expected outputs. They also specify which partners are responsible for which outputs. For instance, who will analyse the data, and who is responsible for reporting, writing, approving, and disseminating the final document? The WFP country office is responsible for coordinating with local representatives of partner organizations to ensure logistical arrangements and coordinate secondary data reviews in preparation for the CFSVA.

The WFP regional bureau and Headquarters should be involved in the initial consultations between the country office and partners, to determine the timing of the assessment and the wording of the TOR. Deciding whether WFP staff from the regional bureaux and Headquarters, and/or consultants will be part of the team will normally be decided by the country director; in some cases, though, others, such as the regional director, may be consulted.

The team leader is critical to the success of the CFSVA as she/he should have extensive prior experience in CFSVA/assessments and possess training and writing skills. The team leader's role is normally included in the study's overall TOR, but a specialized TOR can also be developed.

2.6 ADMINISTRATIVE DUTIES FOR THE CSFVA

CFSVA administrative duties are numerous and time-consuming. The CFSVA team will have to make sure WFP staff (and especially administrative and finance officers) can dedicate crucial time to the following tasks:

10. See example of CFSVA TOR in annexes.

- drafting a budget for the CFSVA (after agreement on contribution, sampling, etc. - see section 2.3 and Annex 14 for a sample budget) and managing incoming expenses;
- hiring enumerators, field supervisors, and data entry personnel;
- preparing training for supervisors/enumerators;
- arranging for external assistance (travel, DSA, contract, etc. for consultant, Headquarters, or regional bureau staff);
- logistical planning (see section 2.6.1); and
- organizing any additional meetings (restitution of findings, press conference, etc.).

2.6.1 Logistical planning

Implementing a successful assessment requires logistical planning and preparation. Logistics is a big part of the survey, and in some countries can be a cumbersome exercise, and so should be considered early on in the process. Logistics considerations include selection of field sites and advance notification of sample communities or sites, coordination of transport and communications operations, and distribution and collection of data-collection instruments.

Logistical and administrative steps for survey implementation include, at a minimum, the following:
- Agree on the number of survey teams to be deployed.
- Agree on the number of individuals in each survey team.
- Agree on the number of vehicles per team.
- Ensure that each team has appropriate equipment (radio, satellite phone, PDAs or questionnaires, measurement boards, weighing equipment, money, maps, introduction letters, etc.).
- Ensure that each team has received appropriate training in health, ethics, security, and data collection.
- Ensure that each team has the required documents (ID cards, local travel authorizations, security clearances, insurance, etc.).
- Ensure that communities and key informants are informed of survey dates and requirements. Supply local authorities with a list of villages to be surveyed. They can help in informing communities.

Several challenges that increase logistics requirements need to be considered in advance:
- **Team size and composition:** Large assessment teams, the participation of multiple agencies and United Nations representatives or donors;
- **Administrative issues:** Administrative regulations that are not flexible enough to allow for rapid planning and schedule changes; per diem and expenses for government and NGO officials;
- **Physical environment or emergency context:** Lack of access due to poor or limited infrastructure, natural barriers, rainy season, or security limitations; and
- **Security constraints:** The designated official of the United Nations Security Management Team in all countries establishes limits for safety and security and requires all staff to operate within these limits. Security concerns in conjunction with the conduct of a CFSVA may influence logistics, for example in requiring that data-

collection teams return to their base by a certain hour or proclaiming certain areas as off-limits. Such security matters can add extra planning requirements.

For multi-agency assessments, it is recommended that agencies pool resources for logistics support, drawing on the specific needs of the assessment and relative capacities of the partnering agencies.

2.6.2 Development and use of questionnaires

The probability household survey requires a fairly extensive household questionnaire. It is important, however, not to reinvent the wheel. WFP now has considerable experience in conducting surveys. For the CFSVA, the Programme now recommends a combination of standard core modules and optional modules, elements of which are used by other international survey programmes such as the Demographic and Health Surveys or the Living Standards Measurement Surveys. Formulation of questions is a time-consuming task and should be done only as necessary and as described here. The questionnaire should be the result of a consensus among all partners, technicians, programme specialists, and the government. The CFSVA or food security committee is the best forum for developing and agreeing on questionnaires. The questionnaires should also be shared with relevant specialists at the regional bureau and Headquarters.

Components of a CFSVA questionnaire should address the following data needs:[11]
- Demographic and life cycle information;
- Education;
- Water and sanitation (WatSan);
- Food and livelihood security (including availability, access, and production; consumption and utilization; stability and creation of assets; coping strategies);
- Household incomes and expenditures; marketing systems; and
- Health (including HIV/AIDS) and nutrition.

The following outline delineates an appropriate path for the development and implementation of questionnaires and enumerator training.[12] Once questionnaires have been shared among partners and agreed upon, the following steps must be taken:
1. Translate questionnaires systematically into the local language when needed, including back-translation. An interviewer who has to translate on the spot with the respondent has a lot of room for interpretation. After the back-translation, the expert has to reconcile any translation problem.
2. Prepare for enumerator training and pilot-testing by printing questionnaires or uploading data needs to hand-held computers (PDAs).
3. Assign and train enumerators, measurers, supervisors, team leaders, and data entry personnel on their respective roles (food security enumerators, nutrition data collectors and measurers, field supervisors, team leaders).
4. Pilot-test (at least 30 interviews) and finalize questionnaires.

11. See Chapter 5 for a complete description of questionnaire design guidelines.
12. See section 4.2.5.1 for more on enumerator training.

5. Once tested and corrected, paper questionnaires should be printed in sufficient numbers and distributed to supervisors. If PDAs are used, supervisors should make sure that batteries are fully charged and that there are enough PDAs, memory cards, stylus, chargers, and so on.
6. Launch survey:
   a. Design a complete survey implementation plan for each team:
      I. Date, day, and location to be surveyed;
      II. Estimated arrival and departure times;
      III. Overnight locations; and
      IV. Tasks to be achieved.
   b. Ensure that each completed questionnaire is checked by the team leader;
   c. Ensure that all questionnaires are completed and accounted for before leaving a particular location;
   d. Ensure that data is downloaded and secured at the end of each day;
   e. Ensure that the PDAs are charged for the next day; and
   f. Provide enumerators with sufficient rest days.

2.6.3 Quality control

The manager of the CFSVA should ensure that the highest quality data is collected. Data quality is influenced by many factors, such as whether PDAs or paper questionnaires are used; the experience of enumerators (knowledge of local languages, area surveyed, food security, data collection); the quality of the training; how motivated staff are to collect accurate data (this can be seen during the test); and how well the data-collection process is supervised. If PDAs are used, enumerators should have a basic knowledge of computers.

A daily and random quality control of the collected data is highly recommended, whether paper questionnaires or PDAs are used. Having back up help available to replace enumerators (or supervisors) who get sick, leave for personal reasons, or who are not performing well is essential to the quality of the data. Supervisors might consider performing unannounced spot checks on data-collection teams in order to verify the quality of the work in progress, and make any corrections necessary.

CHAPTER 3

Desk study: literature review and secondary data

A desk study is a review of the food security–related information available before the in-depth food security study takes place, including documents and historic data. Its purpose is to help build a good understanding of the food security situation in a country and that situation's evolution, and to uncover data gaps. More specifically, this analysis aids in uncovering long-term socio-economic trends; designing the sampling frame; defining the original scope, geographic area, and target populations; determining sampling methodology; analysing the risks posed by various hazards; and informing recommendations for future interventions.

When starting the CFSVA, as much relevant information as possible should be collected on the region and population of interest. Secondary data consist of existing data sets, reports, and documents, usually compiled by other persons or organizations, and often for purposes other than those of the present analysis. A literature review is often the main source of information on the political and economic environment of a given area. When properly incorporated into the process of analysis, these data provide an essential complement to the primary data collected.

The information and findings of the desk study should be integrated with the data and findings from the primary data collection and analysis, and not placed in a separate section. In all cases, the source of secondary data should be adequately cited in the report.

3.1 PURPOSE OF A DESK STUDY

A desk study will identify key factors affecting vulnerability to food insecurity for a particular population or within a specific geographic area, helping to determine the overall scope and objectives of upcoming CFSVAs. It will also often help identify important information needs that have not been addressed through previous research. Finally, a desk study can save time and resources directed toward primary data collection in the field. The specific purposes of secondary data analysis are to:

- Clarify the context of primary data research, define the depth and breadth required, and formulate appropriate research questions and instruments;
- Identify the relevant socio-economic groupings, livelihood groups, and livelihood strategies that determine food security/insecurity in the present context;
- Discover the macro-level socio-economic and agro-climatological environments and how they may impact food security and vulnerability;
- Identify long-term trends in poverty, aggregate food availability, consumption, undernutrition, and food security;
- Study hazards and their historical impact on food security;
- Exclude information from the primary data collection process that is already available and does not require verification (however, the collection of previously existing data may be warranted in certain cases for the purpose of trend analysis);
- Verify information that may no longer be accurate or that can serve as a baseline for understanding changes resulting from a given shock or intervention;
- Include indicators and information in the primary data process not found in previous studies; and

- Include the same indicators in the current data collection that were analysed in previous studies, so that deviations from normal periods can be assessed.

For assessments conducted after a crisis, pre-crisis background information can be of particular value. Secondary information can provide crucial insights into the most effective means for providing relief and rehabilitation assistance to an affected population, and prior research can reveal key indicators of “slow-onset” emergencies such as droughts and crop failures. Background information facilitates a rapid assessment of the scope and scale of emergency food requirements and the capacity of different populations to cope with a given shock. Other significant pre-crisis information that can be gleaned from secondary data includes management capacity, infrastructure, and historical experiences related to emergency interventions (Seaman and Leather 2003; WFP Thematic Report 2003).

3.2 ISSUES IN COLLECTING AND ANALYSING SECONDARY DATA

The quality of secondary data dictates the scope of primary data collection: in countries where secondary data is good and regularly collected, primary data may be largely limited to verifying secondary data, or to merely filling in the gaps left from the literature review. In complex humanitarian situations characterized by population displacement, it is unlikely that current secondary data will be available. Given the time and resources required for conducting CFSVAs, it is essential that every effort be made to collect secondary data beforehand to streamline the process and provide the essential contextual information that can orient primary data collection (Riely 2002).

3.2.1 Evaluating the quality of secondary information

Secondary data, documents, and reports are generally not prepared by the same people or institutions that do the primary data collection; hence the goals and purpose of primary and secondary data may not be the same. It is therefore essential that data quality be carefully assessed, particularly where part of the analysis involves a direct comparison of secondary data with the data being collected for the CFSVA. Key questions to help assess data quality include:
- What is the original purpose of the data or publication? How do the goals of the original study differ from the measurement objectives of the current research?
- What is the information source? What are the source’s credentials? What is the potential level of bias? Is the material well referenced?
- Is the information relevant to the current context or is it out-of-date? Five-year-old data may be too old to be useful for analysing the current food security situation; however, such information may be very relevant to identify trends.
- Is the intended audience researchers or the general public? Is the source too elementary?
- What is the coverage of the data source?
- To what level is the data disaggregated?
- For which population is the data representative and with what degree of precision (confidence intervals)? Has the document been officially validated? Is there consensus in the community on the document’s quality?

The level of disaggregation is of particular importance for comparing data from different sources. Generally, the level of data disaggregation varies across or between political and geographic units. Secondary data should be gathered at the lowest possible level of disaggregation to ensure comparability after primary data collection.

Remember: the more aggregated the data, the more invisible the people (McCaston 1998).[13]

3.3 MAIN TOPICS IN A DESK STUDY

Topics for consideration in a desk study are determined by the objectives of the CFSVA. In this section, likely topics for literature review and secondary data analysis are divided into broad, overlapping categories: socio-economic and demographic data, livelihood information, institutional processes and structures, factors affecting food security, sources of risk and vulnerability, health and nutrition, and market conditions.

Table 3.1: Common areas covered by a desk study

	Types of secondary information
Institutional/ stakeholder information	• Existing institutions (e.g., public, NGOs, community-based organizations [CBOs], religious, trade, and labour associations; industry) • Nature of institutional programming and strategic plans • Interest in collaboration • Access to political decision-making at village, regional, and national levels • Relations with governments and communities • Mechanisms normally available to target food assistance to the most vulnerable/food insecure • Effectiveness of government, participating agencies, or NGOs in facilitating recovery and minimizing acute food insecurity or acute malnutrition
Socio-economic and demographic data	• Social and political structures affecting food security – e.g. government policies affecting production, marketing and trade, rationing, fiscal policies, taxation and subsidies • Integrity of infrastructure: water, health services, schools, roads and railways • Census population data at various levels; age and gender distribution • Map data (digital) with political/administrative boundaries, roads, and localities
Market conditions	• Data and trends on national/regional/local food production • Description of existing market systems, trends in market prices and flows • Import/export data production, consumption, and food balance sheets • Time series of market prices of major food and cash crops, by region and season, seasonality • Markets (e.g., locations, access, integration, and functioning) • Previous experiences/analyses of problems with market access
Livelihood Information	• Livelihood strategies, access to resources and all asset categories, desired livelihood outcomes and levels of achievements disaggregated by ethnicity (if appropriate) or mother tongue, wealth, livelihood groups, or gender or other groups, depending on the context (religion) • A socio-anthropological outline of different ethnic groups • A description of exposure to social exploitation and discrimination • A description of livelihood zones or food economy zones • Information on rural and urban poverty • Prevalent diseases and seasonality, by region • School attendance and literacy rates • Seasonal migration patterns, by region and reasons • Levels of debt

13. For additional guidance on secondary data collection, see *Guidelines for Undertaking a Secondary Data Analysis (WFP and VAM 2001) and Tips for Collecting, Reviewing and Analyzing Secondary Data* (McCaston 1998).

	Types of secondary information
Livelihood Information	• Traditional economic roles/control of resources • Income-generating activities, remittances, and expenditures • Labour supply and dependency ratio • Productive assets, landholdings and utilization, livestock ownership and sales, and relative importance in food economy by group or zone • Major food and cash crops, by region; cultivation calendars • Agricultural inputs, utilization and access, food stocks, alternative food/income sources, and food storage
Food security	• Existing safety-net programmes • Emergency response capacities (e.g., United Nations country team [UNCT], international NGOs/civil society/local NGOs/donor presence/interest and flows) • Early warning systems and contingency plans • Number of months of normal self-provisioning • Food consumption, preparation patterns, and diet diversity • Restrictions or taboos on food consumption, preparation, or usage • Malnutrition rates and trends • Normal sources of food, by region and different social groups • Existing food aid and food security programmes • Experience of disaster food assistance, including food basket, duration, and different types of assistance
Risks to food security	• History of natural or man-made disasters, shocks and stresses • Geographic and historic information on the occurrence of hazards • Current exposure to sudden-onset natural disasters • Understanding/awareness of long-term threats to food and livelihood security (soil erosion, depletion of forest resources, water scarcity, etc.) • Issues involving migration, conflict, or large-scale exodus of refugees • Traditional coping strategies • Coping mechanisms, distress coping mechanisms, and grades of livelihood groups

3.3.1 Institutional/stakeholder information

Information on larger socio-political structures and their direct (and indirect) influence on individual and household food security must also be collected. When gathering institutional information it is important to capture the fabric of the society - "the rules of the game" and the factors imposed on the choices people make.

At the macroeconomic level, data on national policies and priorities can provide valuable information for the analysis of local processes and structures. National priorities can be researched by reviewing the latest Poverty Reduction Strategy Paper (PRSP), United Nations Development Assistance Framework/Common Country Assessment (UNDAF/CCA) strategies, or National Action Plans. Relevant national policies and regulations must be understood in context in order to assess their potential direct and indirect impact on food security. Depending on the context, this institutional information may provide insight into the causes and nature of conflict, the capacity for disaster management at the national and local levels, or the effectiveness of existing national food assistance programmes.

3.3.2 Socio-economic and demographic data

From existing publications and through secondary data sources, information detailing basic demographic data is obtained. Various zoning systems (such as administrative jurisdictions, socio-economic areas, agro-ecological zones, livelihoods zones, or purposely constructed food security zones based on multivariate statistical analysis of secondary data) will help determine the sample frame and stratifications in which primary data collection will be conducted. Table 3.1 offers examples of basic socio-economic and demographic information to obtain.

Where possible, secondary data analysis should provide the essential social context on the region most relevant to issues of food security and vulnerability: the forms of local leadership and authority; the relevant ethnic groups; formal and informal social networks; existing social programmes; political systems; historical trends and policies; and issues of personal security for individuals and households. Female-headed poor households, orphans and other vulnerable children (OVCs), and other groups might be of particular interest. Where good background information exists, secondary data analysis can provide the necessary information for CFSVA planners to identify the regions of greatest vulnerability (Frankenberger 1992).

3.3.3 Market conditions

An understanding of market prices, and of regional and seasonal fluctuations for major food and cash crops and livestock (as appropriate), provides essential comparative data in analysing markets under duress. A CFSVA review of the secondary information on markets should focus on trends in national, regional, and local food production, prices (real prices and import parity prices of main food staples; cash crop and livestock prices, as appropriate; terms of trade), flows and market integration; import and export data; the location, access, market chains, and structure; and historical experiences with access by the target populations. Secondary data on prices, volumes, and trade patterns is often available, and is particularly valuable when explicitly linked to household experiences of food security and vulnerability.

3.3.4 Livelihood information

Once the primary groups of interest within the target region or population have been identified (livelihood groups, gender-based groups, age-based groups, or other possible relevant socio-economic groups), preliminary profiles of each group should be prepared from the secondary data. Where such information is available, key statistics on the relative contribution of various employment sectors (farmers, pastoralists, daily wage workers, etc.) to national gross domestic product (GDP) can help in identifying sectors significant to livelihood analysis.

In the best-case scenario, available secondary data will provide sufficient information to create preliminary livelihood profiles in the region or area of interest. These livelihood profiles should indicate how different areas derive their incomes, food, and access to social services. Table 3.1 provides an outline of livelihood areas that a review of secondary data can provide insight into.

The livelihood profiles of these various groups should provide sufficient information to allow a preliminary identification of the poorest and most vulnerable groups, with a particular focus on describing their level of food insecurity and nutritional status. Ideally, livelihood profiles will inform the assessment process by providing information on the causes and levels of vulnerability, the capacity to cope, factors that determine or influence intra-household distribution of food, and the potential roles of targeted food aid.

Additionally, the information contained in livelihood profiles can be used as a benchmark against which the impact of the shock to food insecurity can be measured (WFP 2002).

3.3.5 Food security

Secondary data for a CFSVA should focus on food security – particularly food availability, access, and utilization – and hazards and threats to it. Information on food security is linked to knowledge of institutions, and should include available data on existing food aid and food security programmes and local experiences with disaster food assistance (WFP 2003). In CFSVAs conducted after crises, this information provides essential insights into baseline food security and can be used as a benchmark for a given shock's effects on the food security of vulnerable populations.

Food Availability
Availability is defined as the physical presence of food in a geographic area, not its affordability. Measures of food availability may be derived from production statistics, seasonality information, data on market and food supply infrastructure, import and export statistics, and national policy information (government trade policies, exchange rates, balance of payment constraints, etc.). FAO/WFP Crop and Food Supply Assessment Missions (CFSAMs) and Bellmon determinations also provide valuable secondary information for a preliminary analysis of food availability.

The CFSVA is also interested in looking at surplus food production areas and identifying net food sellers. This information is interesting for programmes aimed at developing local agriculture through local purchases.

Food Access
Indicators of food access are typically focused on social and economic characteristics at the household level and the relationship with markets: the distribution of income within a household; the (seasonal) prices of foodstuffs, cash crops and livestock, as appropriate; the access of individuals to entitlements such as private transfers and gifts or public distributions of resources; household purchasing power as determined by household income and food prices; and household assets and savings in times of duress. Relevant data can thus be collected from information on socio-political structures influencing the household, sources of food and income, the history of shocks and their impact on food access within the household, the number of months of self-provisioning in a normal year, and land distribution and use. Mobility and migration trends, and the effects of coping mechanisms on food accessibility should also be considered if such data is available.

Food Utilization
Food utilization incorporates water, sanitation, food safety, and nutrition data. Bio-cultural information may also be relevant: data may consider medical statistics concerning individual intake and the conversion efficiency of food by the body, but should also include cultural factors influencing food use (e.g. restrictions or taboos on food consumption, preparation, or usage). Food utilization may also be affected by endemic disease, unsafe drinking water, poor sanitation, or lack of appropriate nutritional knowledge. Typical secondary data on food utilization includes: anthropometric measurements of children

under 5 (wasting, stunting underweight); mid-upper arm circumference (MUAC) of children under 5; body mass index (BMI), especially for women; MUAC for women; measures of micronutrient malnutrition; and other measures of nutritional status.[14] General information on health should also be included, such as health status; prevalence or incidence of major diseases; dietary habits and typical consumption patterns, to serve as a benchmark for primary data; and access to water and sanitation.

Seasonality and trends

The causes and effects of food security are often cyclical, not simply because of annual climate variation and the corresponding agricultural cycle, but also in terms of food availability, food prices, prices of agricultural produce and other means of living, household purchasing power, household stocks, household food security status, health and hygiene conditions, and individual nutrition status. Typical seasonal patterns, crop calendars, and time series of rainfall data and prices can often be obtained from secondary sources. This data contextualizes the findings of a CFSVA, and indicates to what extent annual variation may be responsible for identified trends.

Long-term trends (secular change) are also an important element of CFSVA. Whether quality of life indicators are generally going up or down in a country or province reflects the vulnerability and resilience of communities. Secular change in climate and environmental conditions signals changing risk exposure.

3.3.6 Risks to food security

Secondary data analysis of vulnerability should first seek to understand the nature of hazards and shocks and their differential impacts by collecting data on emergency and disaster types (including the rate of onset of disasters and the complexity and scale of emergencies). This data should include a history of natural disasters, price and income shocks and stresses experienced by the population of interest, their spatial and temporal patterns of food availability and access in times of vulnerability, traditional coping strategies, and the mechanisms available to target food assistance to the most vulnerable under normal circumstances.

A broad range of causes may render households more vulnerable, including: physical vulnerability (e.g. people living with HIV/AIDS, disabled populations); socio-economic vulnerability (e.g. children, women, and other potentially disenfranchised groups); political vulnerability (e.g. ethnic or gender discrimination); vulnerability due to physical insecurity or limited governance (e.g. internally displaced populations). Populations affected by a shock experience varying levels of vulnerability depending on the mix of these factors.

3.3.7 HIV/AIDS in the secondary data review

As mentioned in section 1.5, HIV/AIDS issues should be incorporated in all the food security analyses undertaken in countries with high prevalence of HIV and AIDS. This

14. See WFP nutrition website
http://www.wfp.org/policies/introduction/policy

should be done at all the levels of analysis, including secondary data analysis and literature review.

National and sub-national data related to HIV/AIDS can inform a situation analysis in several ways:

- **Estimates on HIV/AIDS prevalence** can be used to establish whether it is relevant to include HIV/AIDS in a food security and vulnerability analysis. These estimates can also identify areas with both high HIV/AIDS prevalence and high food insecurity.
- **Data on knowledge and awareness of HIV/AIDS**, presence of risky behaviours, identification of main roads, and analysis of migration flows can suggest areas and groups where food support is most needed for prevention, mitigation, and care activities.
- **Data on health centres** (e.g. location, type of centre, and kind of services offered) can identify areas and institutions where WFP assistance can complement care and treatment services to people living with HIV/AIDS.
- Understanding the **prevalence of other chronic illnesses**[15] and serious diseases (e.g. malaria and tuberculosis [TB]) is useful for determining whether, and to what extent, the validity of chronic illness as a proxy indicator of HIV/AIDS may be jeopardized by such confounding factors.

A number of secondary sources of data are available for conducting initial situation analysis on HIV/AIDS. Major sources include:

- AIDS Indicator Surveys (AIS)
  http://www.measuredhs.com/aboutsurveys/ais.cfm
- Multiple-Indicator Cluster Surveys (MICS)
  http://www.childinfo.org/MICS2/Gj99306m.htm
- Demographic and Health Surveys (DHS)
  http://www.measuredhs.com/aboutsurveys/dhs/start.cfm
- Behavioural Surveillance Surveys (BSS)
  http://www.fhi.org/en/topics/bss.htm
- UNAIDS
  http://www.unaids.org
- WHO
  http://www.who.int/statistics/en/
- HIV/AIDS Survey Indicator Database
  http://www.measuredhs.com/hivdata/

This list of sources does not aim to be comprehensive. At regional and country levels, it is possible to identify others. Especially in countries with high HIV prevalence, official bodies frequently collect (or have access to) updated information. It is therefore good practice to conduct a desk review of information available in the country.

Data on HIV/AIDS and OVC from standardized surveys or other official bodies should always be reviewed. However, these surveys usually do not include data on food security or information on household response to AIDS. Therefore, they cannot be

15. WHO's definition of *chronic disease* is "a disease of long duration and generally slow progression." Chronic diseases include heart disease, stroke, cancer, chronic respiratory diseases, and diabetes. For each country, WHO is able to provide a picture of major illnesses and causes of mortality. Such information can be retrieved at the WHO website (www.who.org).

used to explore the relationship between HIV/AIDS, food security, and livelihoods. In order to study such relationships, it is crucial to include proxy indicators of HIV/AIDS in food security and vulnerability analyses.

3.3.8 Gender analysis using secondary data sources

Annual Human Development Reports (HDRs) produced by the United Nations Development Programme (UNDP) provide a general overview on the status of gender equality across countries and regions at the national level. Since 1995 the Human Development Index (HDI), also produced by UNDP, has been complemented by the gender-related development index (GDI) and Gender Empowerment Measure (GEM). The GDI compares women's and men's life expectancy, educational attainment, and income, while the GEM concentrates on gender differences in income, access to jobs classified as professional and technical and administrative and managerial, and the percentage of parliamentary seats held by women and men.[16]

Some National Human Development Reports (NHDRs) produced by UNDP country offices also provide GDI and GEM information at the sub-national level.[17] The secondary data provided by these composite indices are often useful as a complement to primary data collection exercises.

Secondary data on HIV/AIDS should be consulted (and reported) taking into account the gender perspective. It is therefore suggested to report:
- Gender-disaggregated figures on HIV prevalence (to see if women are actually more exposed than men, and at which ages);
- Women's access (and utilization) of antenatal clinics;
- Women's enrolment in mother-to-child transmission (MTCT) prevention programmes;
- Women's awareness of transmission and prevention; and
- The consequences of inheritance practices on women's productive role.

3.4 SPATIAL INFORMATION MANAGEMENT

3.4.1 Use of standardized codes in analysis and reporting

All analysis should be performed using standardized location coding, in order to ensure that the correct codes appear in the results and in the final reporting. Standard coding ensures compatibility with past and future analysis, and with data from other projects and agencies. Standard location codes consist of:

- **Pcodes** for populated places (points). Pcodes are generally decided upon by the main mapping agency in the country or region (HIC, WFP, OCHA, UNHCR). The standard Pcodes should be copied to the regional bureaux and Headquarters for the sake of disseminating the standard.

16. For more information, see http://hdr.undp.org/docs/statistics/indices/technote_1.pdf and http://hdr.undp.org/reports/global/1995/en/pdf/hdr_1995_ch3.pdf.
17. National Human Development Reports can be accessed on the following website: http://hdr.undp.org/nhdr/.

- **GAUL codes** for administrative unit subdivision. FAO GAUL codes ensure compatibility with the UNCCS boundaries and standards, and ensure traceability of administrative changes over time (merging and splitting of administrative units). Currently the second administrative level is present for all countries, and the third and fourth levels for some countries. The FAO GAUL team ensures a quick reaction and that changes sent to them will be incorporated in the next annual release of GAUL. If a higher level of administrative boundaries is used at the country level, this should be sent to the regional bureaux and Headquarters, which will copy FAO for inclusion in the GAUL. In this case it will be possible to easily map the administrative codes used to the new GAUL codes when they are released.

If other spatial divisions are used for sampling or reporting (e.g., livelihood zones or agro-ecological zones), then codes should be supplied and the shape file containing the delineation of these zones and the coding should be published on the VAM-SIE with appropriate metadata, which will allow a future reader to properly identify these zones as those used in the analysis.

3.4.2 Minimum data sets

A set of minimum data sets for each country is maintained at the regional bureau and Headquarters level. These data sets should remain synchronized with those used at the country office, which is based on global layers and secondary data available from international institutions. If the country office or regional bureau uses a data set of higher resolution or better quality, this should be copied to the other offices to maintain standardization.

3.4.3 Metadata and sources

Documentation regarding the data sets used in an analysis is a fundamentally necessary step that is often overlooked because its significance is not immediately apparent.

Standards in metadata specifications exist (ISO 19115 and ISO 19139) but are far too detailed for practical use. VAM SIE (see section 8.1.3.8, item 7) implements ISO 19119, and the upcoming release will include ISO 19139, but the fields are not compulsory, so it is up to the user to ensure that the minimum essential information is present.

In the case of layers and data sets used in spatial analysis, it is fundamental to document the source of the data, the methodology used in deriving it (or a link to an external web page of documentation), the reference dates, and the accuracy of the data.

3.4.4 Specific tool set

Integrating the various data sets at different times and spatial scales to provide food security analysis is challenging. Appropriate methodologies and analytical tools will be developed to facilitate the production of analysis. The integration of global- and country-level data sets in a structured way through a “Dynamic Food Security Country Atlas” will facilitate the management of food security information.

3.5 GEOSPATIAL DATA ANALYSIS

With recent advances in geographic information systems (GIS), satellite technologies, and information and communications technology (ICT), geospatial data is becoming widely available for civilian use. Satellite data, computing capacity, and GIS software are becoming more and more affordable and accessible. Geospatial applications in the context of CFSVAs encompass spatial data collection, geospatial analytical tools, and storage and dissemination of geo-referenced information.

GIS tools and geospatial techniques improve CFSVA by increasing data availability and improving the integration of various data sets such as environmental data and socio-economic information, thus improving food and livelihood security.

In undertaking a CFSVA, relevant work done by other institutions or organizations, and potential partners (both national and international), should be identified. Local capacity and institutions in the area of remote sensing, GIS, and risk analysis should be assessed to identify any gaps or need for capacity-building.

Databases are disaggregated at the sub-national level. Databases may include relevant documentation (e.g. text documents, photographs, tabular data, maps, statistics, earth observation data). Here is a sample list of secondary data information needs:

- Patterns of migration
- Climate and weather
- Irrigated areas
- Land use and cover
- Areas of crop production and timeliness of planting
- Cropping pattern and stages
- Occurrence of crop pests and diseases
- Crop diversity
- Market price
- Major reasons for poor performance
- Farming practices
- Farming inputs
- Livestock size, movement, productivity
- Food production, accessibility, and availability
- Crop and livestock production
- Seed access, availability, and quality
- Household livelihood characteristics
- Epidemic diseases
- Consumption patterns
- Size and number of meals
- Type of food items consumed
- Evidence of malnutrition
- School attendance

Low-resolution (e.g. Terra-Modis, Spot Vegetation, Meteosat, NOAA AVHHR) and high-resolution satellite data (e.g. LANDSAT TM, SPOT, IKONOS, Radarsat 1 and 2) should also be employed when possible.

3.5.1 Compilation and organization of secondary geospatial data sets

The various secondary geospatial data sets should now be reorganized according to:
- scale (global, national, sub-national);
- periodicity (for those sets that evolve regularly);
- type (vector or raster); or
- theme.

Data sets available at the global level as secondary data (for all or most countries worldwide and most available for download) and those available locally (especially at the sub-national level) are listed here (note that this is not an exhaustive or mutually exclusive list):

Globally available data sets (mainly but not exclusively at the national level)
- Administrative boundaries
- Population and population density (and populated places)
- Urbanization estimates
  - Migration estimates
  - Gender and age composition
- Infrastructure
  - Roads and railroads
  - Markets
  - Public buildings (town halls, schools, etc.)
- Geophysical characteristics
  - Land cover: forest, grassland, desert, urban, etc.
  - Soils and geology
  - Water networks: rivers, streams, lakes, etc.
  - Elevation
  - Land use, farming systems, irrigation
  - Climatic variables (from global climate models): mean precipitation, temperature, etc.
- Meteorological/climatic information that is periodically available (dekadal) is compiled by Headquarters on a regular basis to include:
  - NDVI: Normalized Difference Vegetation Index
  - RFE: Rainfall Estimation
  - WRSI: Water Resources Satisfaction Index
  - Country Crop Calendar
- Disaster/hazard exposure: occurrence of hazard (date/location), number killed, affected and homeless, etc.

National-level data sets (sub-national level)
- Most of the above
- Health-related coverage
- Malnutrition (underweight, stunting, wasting)
- HIV prevalence
- Socio-economic variables: income, education, etc.
- Hazard coverage: drought, flooding, etc.
- Livelihoods maps

A careful inventory of all data acquired and the respective metadata needs to be prepared before continuing with the next step. It is useful here to map each individual variable/indicator in order to start to explore the varying spatial trends. It may be useful at this step to integrate the hazard and vulnerability data; this would depend on the range of data identified as "missing."

3.5.2 Identifying data gaps

Before proceeding with the CFSVA spatial analysis, it is essential to review the data sets compiled in light of identifying the information gaps. Information deemed ideal and important to the CFSVA and not currently available will be prioritized in the subsequent primary data collection efforts. To identify gaps, it is useful to refer to the list of ideal data established in section 3.5 and compare it to the data sets compiled.

3.5.3 Integration: Data management and analysis

Mapping of determinants of food security, hazard mapping, and vulnerability and food security mapping

The spatial data sets can be organized into two main categories according to the well-accepted components of risk analysis: (1) hazard, and (2) vulnerability. These can be assessed separately, as well as overlaid geographically, to produce levels of risk. This is described in section 6.2.4.

3.6 CONDUCTING THE LITERATURE REVIEW

If resources allow, assign one person or a small team fully conversant with the assessment objectives and terms of reference to go through secondary data and identify useful material. Table 3.2 shows how a template for secondary information review helps identify the gaps that determine the type of primary data to be collected during the CFSVA.

Table 3.2: Identification of information gaps, with examples

Question	Information requirement	Secondary source 1	Secondary source 2	Secondary source 3	Ways to fill information gap (with primary information)
Is there a food security or nutrition problem?	**Is there evidence of excess mortality?**	Survey data from 6 months ago show mortality rate of 1/10 000 people per day.	Doctor interviewed in local press says that mortality rate "seems to be declining."	Recent NGO rapid assessment indicates increased burial ceremonies in last 3 months.	Collect data on number of deaths over last 6 months in sampled villages.
	Is there evidence of acute malnutrition?	An NGO specialized in nutrition carried out an anthropometric survey in 3 of the 5 affected districts 3 weeks ago.	Livelihoods and the emergency impact in the other 2 districts seem to differ from those in the 3 surveyed districts.	No further information could be identified.	Undertake an anthropometric survey in the remaining 2 districts, using the same methodology that the NGO used.
	Have people lost land or access to land?	Government economic data show increased land sales and decreased land prices in drought-affected areas.	A market survey by the local university shows no increase in the number of people looking for casual labour opportunities.	A report by the International Organization for Migration (IOM) indicates no unusual migration within or out of the affected area.	Identify the people selling land. Can these be considered crisis sales? If so, how are people compensating for their lost land assets?

Question	Information requirement	Secondary source 1	Secondary source 2	Secondary source 3	Ways to fill information gap (with primary information)
Is there a food security or nutrition problem?	**How do people currently obtain food?**	A market survey by a local NGO indicates that demand for expensive food, such as meat, has declined.	A WFP rapid emergency food security assessment (EFSA) undertaken 3 months ago showed most households' food consumption – measured through the food consumption score – was acceptable.	The local clinic reports increased micronutrient deficiency among under-5 children.	Evidence suggests that food access has deteriorated over the last 3 months. The CFSVA will check the current situation and look for the causes of this.

3.7 SOURCES OF SECONDARY DATA

Table 3.3 provides examples of common secondary data sources that include national governments, the United Nations, academic publications, Internet resources, and literature from NGOs.

Table 3.3: Common sources of secondary data

Government Documents	Partner NGOs	Professional and Academic Institutions	Internet Websites
• Municipal development plans • Official statistics • Technical reports • Departments of agriculture, rural development, environment, nutrition, social welfare, roads and transport, disaster management, etc.	• Project reports • Baseline studies • Project evaluations • Technical reports	• Journals/articles • Reference books • Public and private research organizations • Public and private universities • Public and private libraries, including the WFP library • Computerized databases	• Eldis Food Security Resource Guide http://www.eldis.org/food/index.htm • Famine Early Warning Systems Network (FEWS NET) http://www.fews.net/ • Food and Nutrition Technical Assistance http://www.fantaproject.org/ • United Nations Food and Agriculture Organization (FAO) http://www.fao.org/ • UNDP Human Development Report http://hdr.undp.org/ • World Bank Global and National Development Reports http://www.worldbank.org/ • United Nations Children's Fund (UNICEF) http://www.unicef.org/ • WFP http://www.wfp.org/ • WHO http://www.who.int/en/

3.7.1 Sources of food security and logistical information

Information on food security can often be found in-country, or from organizational reviews within the WFP country office. Food availability data may be obtained from the Ministries of Agriculture, and of Finance and Commerce, or from the National Statistics Office.

Reports from USAID/FEWS or the World Bank, data on market information systems, and European Union Food Security Units often provide further information on food availability. Food access data may be available from local government or NGO reports. Food utilization and nutrition data are gathered from secondary sources such as demographic and health surveys (DHS), or other national surveys carried out by the government's Ministry of Health. UNICEF nutrition surveys and WHO health surveys also frequently provide health and nutritional data, which may also be available from local health centres.

3.8 KEY REFERENCES: SECONDARY DATA ANALYSIS

- Frankenberger, T. 1992. Indicators and data collection methods for assessing household food security. In S. Maxwell and T. Frankenberger, eds. *Household food security: concepts, indicators, and measurements: a technical review*. New York and Rome: UNICEF and IFAD.
- McCaston, K. 1998. *Tips for Collecting, Reviewing, and Analyzing Secondary Data.* Atlanta: CARE.
- Novak, Thomas P. 1996. *Secondary Data Analysis Lecture Notes.* Marketing Research, Vanderbilt University.
- Riely, F. 2002. *Pilot Food Aids Needs Assessment (FANA) Methodology: A Step-by-Step Guide.* Rome: World Food Programme.
- Seaman, J. and C. Leather. 2003. Emergency Needs Assessment Methodologies. Non-food aid response to food insecurity: How do we identify the most appropriate types of intervention through emergency food security assessments? Annex A in Report of the Technical Meeting: Key Terms in Emergency Needs Assessment. October 28–30. Rome: Emergency Needs Assessment Unit (OEN), WFP.
- WFP. 2009. *Emergency Food Security Assessment Handbook*, second edition. Rome.
- WFP. 2002. Emergency Needs Assessments: Report on the Proceeding of the Expert Consultation. November 12–14. Rome.
- ibid. 2003. Expert Consultation on Four Thematic Areas for Emergency Needs Assessment: Sampling Group. November. Rome.
- WFP and VAM. 2001. *Guidelines for Undertaking a Secondary Data Analysis*: Within the Framework of a SAF Comprehensive Vulnerability Analysis. Draft.

CHAPTER 4

Household-level data in a CFSVA

Much of the information used in the CFSVA comes from data collected at household level for the purpose of the CFSVA. Usually secondary data is not sufficient because some key indicators are missing from it is. It is rare to find recent information on household dietary diversity, coping strategies, income and expenditures, and nutrition status all in the same secondary data set, hence primary data are collected.

4.1 SAMPLING IN A CFSVA

Sampling is the methodology by which specific individuals, households, and communities are selected to be surveyed as part of the CFSVA. Sampling is a highly technical activity, and it is critical that the sample design be carefully undertaken. The most common mistake found in CFSVAs and many other data collection exercises is to make errors when designing the sample. Most field offices will require special support from Headquarters or a specialized consultant with sampling expertise. This section focuses on approaches typically used in the primary data collection of a CFSVA.

Why draw a sample? The alternative would be to obtain information on all households, as in a population census. This would provide a very accurate "snapshot" of the population at a particular moment in time. Even groups that were numerically small (and hence possibly missed in a survey) would be counted. However, a census is usually much more expensive than a survey, and processing and cleaning census data would be enormously time-consuming.

When we draw a **random sample** of that same population, we can infer the findings of this sample to the entire population with a known degree of precision. Hence, a smaller sample allows the researcher to devote extra effort to ensure the information obtained is accurate, and it allows for more detail: a CFSVA requires an intensity of interview or observation that cannot be carried out in a census.

Therefore, issues of cost, time, precision, and quantity of data all suggest that a survey is preferred to a census. CFSVAs typically require the conducting of surveys, which have several steps:

- Decide on the sample unit (n), for example, a village or household;
- Determine the "universe" (e.g., rural part of country A);
- Construct a sampling frame (list of all villages in country A; list of all households in a village);
- Decide on the sample size (N); and
- Choose the sample (sampling).

4.1.1 Key terms and concepts[18]

4.1.1.1 Sampling

The term sampling refers to the selection of a limited number of individual units of analysis (denoted as n) from a population of interest (denoted as N) with the purpose

18. WFP, ODAV (VAM), December 2004. *Thematic Guidelines: Sampling Guidelines for Vulnerability Analysis.*

of inferring something about that population from the individual units selected in the sample. Almost all CFSVAs have a primary data collection component. Households and individuals (e.g. mothers and children) are the most common units of analysis in a CFSVA survey.[19]
There are two broad categories of sampling relevant to CFSVAs: probability sampling (also called formal sampling) and non-probability sampling (also called informal sampling).

4.1.1.2 Probability sampling

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

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

Example From an exhaustive list of all households in the peri-urban area of Port au Prince, Haiti, 200 households were randomly selected, by assigning numbers to each household and then randomly drawing 200 numbers. An assessment employing probability sampling methods estimates that 28 percent (95 percent confidence interval of +/- 6 percentage points) of households in the peri-urban areas outside Port au Prince consume fewer than two meals per day. In other words, based on a sample survey, we are 95 percent sure that the estimated percentage of households in the peri-urban area outside Port au Prince consuming fewer than two meals per day is between 22 percent and 34 percent.[20]

The types of probability sampling discussed in this chapter include:

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

Probability sampling is strongly recommended for all CFSVAs, even in purely qualitative studies.

4.1.1.3 Non-probability sampling

Non-probability sampling relies on a more subjective means of inferring something about the population of interest from a sample. It is not based on statistical theory, and without a statistical basis it is impossible to assess precision and reliability (accuracy)

19. By contrast, nutritional surveys that collect anthropometric data normally treat individuals within households as the unit of analysis. Combined food security and nutritional surveys may use a combination of household- and individual-level analyses.
20. This range estimate is known as a *confidence interval* and is discussed in detail in section 4.1.3.

of estimates. Sample households or individuals are selected because there is reason to believe that they "represent" the population well or that they are well positioned to provide information about the population (as with key informants). The inherent subjectivity and bias associated with non-probability methods are both its strengths and its weaknesses.

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

Non-probability sampling methods are appropriate for meeting many of WFP's information needs. Specifically, they are widely used for selecting communities/villages for qualitative studies.

Non-probability sampling is rarely used in CFSVAs for household data collection, and therefore is not covered in this chapter. However, it is frequently used in qualitative data collection, such as focus group and community discussions. Refer to Chapter 5, on community data collection for guidance on sampling in these circumstances.

4.1.1.4 Sampling frames

A sampling frame is an exhaustive list of all sampling units[21] and their physical locations within the population of interest (N) from which the units that will be sampled are selected. The purpose of constructing a sampling frame is to ensure that each household within the population of interest has an equal or known probability of being randomly selected for inclusion in the sample. Random selection of sampling units from a sampling frame allows for estimates from the sample population (n) to be generalized to the larger population of interest (N) defined by the sampling frame.

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

In practice, sampling frames that are 100 percent complete and accurate do not exist. However, the sampling frames constructed for CFSVAs should be as accurate and complete as possible, but should rely primarily on existing data sources rather than primary data collection.[22] Government census data or demographic data from other surveys are among the most useful sources for constructing sampling frames.

It is important to be transparent about groups or areas that will be intentionally left out of the sampling frame because population (N) level estimates generated by the

21. See section 4.1.1.6 for a detailed explanation.

22. In this instance, primary data collection refers to population data collected in the field by WFP for the purpose of constructing a sampling frame. By contrast, secondary data refers to existing data collected for another purpose that can be used to construct a sampling frame.

sample population (n) do not apply to these groups. Security is perhaps the most common reason for intentionally excluding groups or areas. Some studies may exclude urban areas purposely. However, some individual households or villages are often omitted from the sampling frame unintentionally. Although estimates derived from the sample population (n) cannot be used to generalize about these households, a limited number of chance omissions will not undermine the validity of findings.

4.1.1.5 Sources of information on sampling frames and research design

Secondary data may also be helpful in developing a sampling frame for primary data collection. Four common sources of data of this type are the National Population and Housing Census, the Demographic and Health Survey, the Multiple-Indicator Cluster Survey (MICS), and the Living Standards Measurement Survey (LSMS). These can inform questionnaire design and validate the results of the CFSVA. In cases where data collection for such surveys coincides with the CFSVA, it may be useful to coordinate activities so as to avoid duplication of effort, provided that the sampling and coverage issues can be satisfactorily resolved.

National Population and Housing Census information should be used to develop the sampling framework, since the demographic information often includes rural/urban classification, gender, age, disability, shelter, education level, and migration status by the smallest administrative unit. However, care should be taken to ensure that the census results are still valid (no more than 10 years old) and that no extraordinary events have occurred (wars, conflicts, environmental disasters) that could have significantly changed the census findings.

A review of the latest household consumption and expenditure surveys and of the agricultural census is also very important. These data and analyses assist in defining the critical regional baselines, including long-term average production and consumption patterns. The income, price, supply, and demand elasticity generated from these surveys are extremely useful in determining the expected income, price, and substitution effect for food commodities.

The IMF Generalized Data Dissemination System (GDDS) is also a good source of information about the type of economic, financial, and socio-demographic data available in the country, along with its characteristics, quality, access, and integrity.

4.1.1.6 Primary and ultimate sampling units

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

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

However, in most cases a complete list of households for a population of interest is unavailable and would be costly and time consuming to construct. Even if a complete list for the population of interest were available, the cost of visiting households dispersed all over the region of interest would be excessive. In these cases, the sampling frame is constructed at the lowest aggregation of households for which accurate information on the existence, location, and relative size[23] of aggregates is available. In rural settings, this aggregation is often villages, such that an exhaustive list of villages (primary sampling units) within the population of interest can be constructed. In urban settings, neighbourhoods or blocks often provide a suitable aggregation of households and can be used when constructing a sampling frame. Households (the most common unit of analysis in CFSVA surveys) remain the ultimate sampling units.[24]
Several options exist for choosing households for inclusion in the sample when the primary sampling units are an aggregation of households such as a village or neighbourhood/block. Choice of a method of household selection is driven by the information available and time/cost constraints. Guidance on choosing an appropriate household selection method is described in detail under each of the sampling methods described in section 4.1.2.

4.1.2 Choosing an appropriate sampling method

A variety of probability and non-probability sampling methods exist to suit different situations encountered in the field. The most commonly used sampling methods for CFSVAs are one or more of the following: simple random sampling, systematic sampling, cluster (or area) sampling, two-stage cluster sampling, and/or stratification. The household survey of a CFSVA typically uses a stratified two-stage cluster sample.

4.1.2.1 Simple random sampling

As the name implies, simple random sampling (SRS) is the most straightforward of the probability sampling methods. A simple random sample involves the random selection of households from a complete list of all households[25] within the entire population of interest (e.g. sampling frame). Households are therefore both the primary and ultimate sampling units. Simple random sampling has a statistical advantage over other sampling methods[26] and requires a smaller sample size (approximately half the size required for cluster or two-stage cluster sampling). However, an exhaustive population list is required, and the cost of visiting geographically dispersed households may be high.

When to apply simple random sampling

In practice, household-level sampling frames are rarely available. However, assessments conducted in long-term refugee camps or areas in which a census has

23. The utility of size estimates is discussed in detail under sections 4.1.2.4 and 4.1.2.5 Cluster Sampling, Two-Stage Sampling, and Multi-Stage Sampling.
24. In rare cases it may be necessary to have multiple levels of sampling units. For example, if no information on villages and their location is available, a higher aggregate, such as a district, may be used. In this example, district is the primary sampling unit (PSU), villages are the secondary sampling unit (SSU), and households (the desired unit of analysis) remain the ultimate sampling unit (USU). A more detailed discussion of this issue is provided in section 4.1.2.4.
25. It is rare to find a complete list of all households to construct a sampling frame for simple random sampling.
26. Systematic sampling shares this advantage.

recently been conducted may provide enough information at the household level to construct one.

For CFSVAs, simple random sampling is almost never applied (except sometimes within a cluster; see section 4.1.2.3 on cluster sampling and selecting households within a cluster).

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

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

How to apply simple random sampling

Step 1. Each household in the sampling frame is assigned a unique number between 1 and the total number of households in the sampling frame.

Step 2. A randomization method is then used to select households for inclusion in the sample.[27] Microsoft Excel can also be used to generate random numbers, and even the serial numbers on currency can be used.

Step 3. Next, selected households are mapped to facilitate data collection. The data collection team must also have a household replacement strategy for when (a) a household cannot be located (due to inaccurate information in the sampling frame) or (b) an appropriate respondent is not available.

Step 4. Replacement households can be preselected prior to data collection by identifying the next household in the sampling frame. Alternatively, a protocol[28] for replacing households in the field can be agreed upon prior to data collection. Examples include choosing the next closest household or spinning a pencil in front of the absentee household to select a transect line and choosing the first house encountered in that line as the replacement household. The means of household replacement is less important than the uniform application of the procedure chosen.

Example of applications of simple random sampling

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

(cont...)

27. The total number of households to be randomly selected from the sampling frame is determined by the sample size requirements (see section 4.1.3).
28. The protocol should be written and provided to each enumerator for reference during data collection.

(...cont)

Each household is assigned a number between 1 and 5,050 (the total number of households in the camp). A random numbers generator (www.randomizer.org, or the RAND function of Excel can be used) is used to select four hundred households. The selected households are then mapped. The workload is divided among four data collection teams, with each team given a mapped area containing approximately 100 households.
Given the proximity of households to one another within the camp, data collection teams are able to walk between selected households. Households that are non-existent or that do not have a suitable respondent available at the time of data collection are replaced by the closest household to the mapped location of the original household.

4.1.2.2 Systematic sampling[29]

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

When to apply systematic sampling

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

For CFSVAs, systematic sampling is almost never applied (except sometimes within a cluster; see section 4.1.2.3, on cluster sampling and selecting households within a cluster). Care must be taken to assess what patterns, if any, exist in the sampling frame. If the ordered pattern has any relation at all to food security, simple random sampling must be applied.

4.1.2.3 Two-stage cluster sampling

In practice, (stratified) two-stage cluster sampling is used in almost all CFSVAs. The combination of minimal information requirements and logistical ease make this method particularly well suited to many of the scenarios encountered during CFSVA surveys.

As the name implies, two-stage cluster sampling is a variant of cluster sampling. A cluster is simply an aggregation of households that can be clearly and unambiguously defined (Magnani 1997). For CFSVA surveys in rural areas, villages are the most common cluster used in sampling. For urban studies, blocks or neighbourhoods may be more appropriate. Two-stage cluster sampling involves selection of a limited

29. Scientific researchers will insist that the only correct way is through random sampling (instead of systematic sampling) because there is always a possibility of “hidden patterns” in the list of households, which may lead to a bias when applying systematic sampling.

number of villages (usually between 25 and 30 in CFSVAs) in each stratum (non-stratified samples have only one stratum). Two-stage cluster sampling uses a second step to select a limited and fixed number of households within each selected cluster. The number of households per cluster varies, but is usually between 8 and 20 for CFSVAs. A 30-by-30 cluster sample is a common form of two-stage cluster sampling, often used in nutrition surveys, where 30 households are selected in each of 30 villages.

Cluster sampling in the CFSVA always uses a “probability proportional to size” selection of clusters. This means that a village with 500 households is 5 times more likely to be selected than a village of 100 households. This ensures that all households, whether from a small or a big village, always have an equal probability of being selected.

When to apply two-stage cluster sampling

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

Even when a household-level sampling frame does exist, using a random or systematic sampling method is likely to produce a geographically dispersed sample (see sections 4.1.2.1 and 4.1.2.2). Therefore, a large number of villages may need to be visited to select a relatively small number of households.

To reduce the financial costs and time needed to conduct an assessment, particularly one covering a large physical area, a decision may be made to use a two-stage cluster sampling. Two-stage cluster sampling reduces costs and time because it limits the number of villages/neighbourhoods/blocks to be visited and the number of households to be interviewed in each selected village/neighbourhood/block. However, the precision of the results obtained may suffer. For most assessments, the sample size required for a two-stage cluster sampling approach is approximately twice that required for a simple random or systematic sample.[30]

Two-stage cluster sampling is used in nearly all CFSVA sampling approaches.

Example It is determined that the minimum sample size[31] for each rural stratum in a rural CFSVA in Ethiopia is around 180 households. Since we use cluster sampling, assuming a design effect of 2 (180 x 2 =), 360 households are required. Although there has not been a recent census, a reasonably accurate list of all villages in each stratum exists. Looking at one particular stratum, there are 150 villages in total, and their approximate size is available through the government’s statistics department. Villages range in size from 20 to 300 households, and on average contain 150 households. At the first stage of selection, 30 villages are randomly selected (with probability proportional to size) for inclusion in the assessment. At the second stage of selection, 12 households are selected within each of the 30 villages, a total sample size of n = 360 (e.g. 30 x 12 = 360).

30. This is due to the design effect of using a cluster sampling methodology. This issue is discussed in detail in this chapter. A design effect of around 2 is common for many indicators if the cluster size is around 10 to 15 households.

31. Assuming simple random sampling.

How to apply two-stage cluster sampling

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

Selecting clusters to include in the sample

Step 1a. The first step in (two stage) cluster sampling is defining the aggregation of households that will be used as "clusters." The following criteria are helpful for defining appropriate clusters[32]:

- Aggregations should be pre-existing and recognized. Villages, blocks, neighbourhoods, and census blocks are good examples.
- Aggregations used for clusters should be as unrelated to food security as possible. Unlike stratification – in which households are categorized into sub-groups on the basis of criteria related to food security such as livelihoods and land-use zones (e.g. homogeneity) – the aim of clustering is just the opposite (e.g. heterogeneity). Ideally, each cluster should contain households that reflect the diversity found in the entire population of interest (in terms of food security–related factors such as livelihoods and land use). For the majority of CFSVA surveys, the use of administrative aggregations (such as villages) as clusters will most closely approximate this ideal.
- Information on the size of the cluster (number of households or population size) is available.

Where population estimates are unavailable, key informants can be used to provide rough/relative estimates for all villages in the sampling frame (e.g. very large, large, medium, small, very small).

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

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

Step 1c. Use the third column to list the cumulative size values for all clusters. The cumulative size value for cluster B is the sum of clusters A and B. The cumulative size value for cluster C is the sum of clusters A, B, and C... and so on.

32. The first, third, and fourth criterion were adapted from the FANTA Sampling Guide (Magnani 1997).

Example Sampling Frame with Cluster Population Estimates

Cluster	Size (pop.)	Cumulative Size
A	50	50
B	125	175
C	35	210
D	20	230
E	80	310
F	20	330
G	25	355
H	40	395
I	25	420

Example Sampling Frame with Key Informant-Generated Cluster Size Estimates

Cluster	Size (category)	Cumulative Size
A	3	3
B	1	4
C	5	9
D	2	11
E	1	12
F	1	13
G	4	17
H	5	22
I	3	25

Step 2a. The next step is to decide how many clusters will be included in the sample. As indicated above, 25 to 30 clusters per stratum are typical for most settings (non-stratified samples have only one stratum). The recommended size of the clusters is between 8 and 20 households. The recommendation of 30 clusters per stratum is somewhat arbitrary, but provides a commonly used and technically sound standard. Choosing the most appropriate number of clusters requires striking a balance between technical and logistical considerations. A bare minimum of 20 clusters (preferably 25) per stratum provides a lower limit for surveys where cost and time considerations are major constraints.[33] Typically, CFSVAs have around 25 or 30 clusters per stratum, and if increasing the number of clusters (and decreasing the sample size per cluster) does not affect the survey logistics or cost, this is a preferable option, as it decreases the design effect with a constant sample size.

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

Example A CFSVA survey in rural India required a sample size of 300 households in each of 5 strata (sub-groups defined by land-use zones) for a total sample size of n = 1,500. Information from the government allows for the use of villages as clusters. The following options were considered for each of the 5 strata:

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

Since there are 5 strata, a decision is made to take the minimum acceptable number of clusters to reduce the number of vehicles and other costs associated with the assessment. The total number of clusters/villages to be visited is 100 (25 clusters in each of 5 strata), for a total sample size of n = 1,500 (12 households in each cluster). This worked well because 1 team of enumerators, with 1 vehicle, was able to interview 12 households in a village in one day, with enough time left to travel to the next location. Fifteen households was too many to interview in a day, and 10 households would have left extra time but not enough to start on a different village in the same day. Additionally, the limited impact on the survey's design effect, by increasing the cluster size from 10 to 12, and the number of clusters from 30 to 25, was considered acceptable in this case.

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

Example Continuing from the Indian example (with 25 clusters in each of 5 strata, with 12 households taken per cluster for a total sample size of n = 1500), it is estimated that each enumerator can complete 4 interviews per day. Therefore a team of 3 enumerators (with each doing 4 households) and 1 team leader (working on the community questionnaire and supervising household data collection) can complete 1 cluster per day. Twenty days have been allotted for data collection. Since there are 125 clusters (25 x 5), 8 such teams could complete the work in about 16 days, with 4 extra travel days (or more, depending on the distance between clusters), it is estimated that 8 teams will be needed (24 enumerators plus 8 team leaders).

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

Box 4.1: Probability proportional to size (PPS)

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

Example The required sample size for a survey in rural northern Uganda is 250 households. Information on the location and approximate size of villages is available through the government. A total of 75 villages is listed in the cluster-level sampling frame. Twenty-five villages will be chosen for the sample and ten households will be taken in each of the selected villages for a total sample size of n = 250.

Random selection – Generate 25 random numbers between 1 and the total cumulative population (or household or size code values). The clusters containing each of the cumulative numbers selected are included in the sample. Statistically, if a cluster is selected twice in this example, 20 households should be taken in that cluster. However, in practice in CFSVAs, this is not always applied. To avoid this problem, one should be cautious when creating clusters. If clusters sizes are often large, they should be subdivided so that they are not double-selected.

Systematic Selection – To determine the sampling interval (SI), divide the total cumulative size indicated in the last cluster listed in the sampling frame by the number of clusters to be selected (25). Generate one random starting number between 1 and the sampling interval. The cluster containing the cumulative number selected is the random starting household.

Example
111 is the randomly selected "first household" selected from the range 1–200.04 (range defined by the sampling interval). This cumulative size corresponds with cluster B in the example here:

Cluster	Size (pop.)	Cumulative Size
A	50	50
B	125	175
C	35	210
D	20	230
E	80	310
F	20	330
G	25	355
H	40	395
I	25	420
J	100	520
etc.	etc.	etc.

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

Example Second household (200.04 + 111 = 311.04) located in cluster F. Third household (200.04 + 311.04 = 511.08) located in cluster J, and so on.

A common trick when selecting clusters systematically is to order the villages (or other cluster unit) by their location within the strata. For example, if in a survey where the main stratification is by province, and a two-stage cluster sample is being drawn in each province, the list of villages can be ordered by geographic area, such as district and livelihood zone, before taking the systematic PPS sample of villages. This can reduce the chance (even if small) of having all villages located within one district (or livelihood zone, or other geographic region) in a province, even if that province has three districts.

Selecting households within selected clusters[34]

Several options exist for selecting households within selected clusters. Each option can be applied regardless of whether the clusters were selected randomly or systematically (Step 2c). Two options are listed here in order of preference; however, the second option is cheaper and faster than the first. Choosing the right method for household selection will vary by assessment. Assessments should strive to use the preferred method (Option 1), choosing Option 2 or an alternative method only when absolutely required due to logistical, time, and resource constraints.

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

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

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

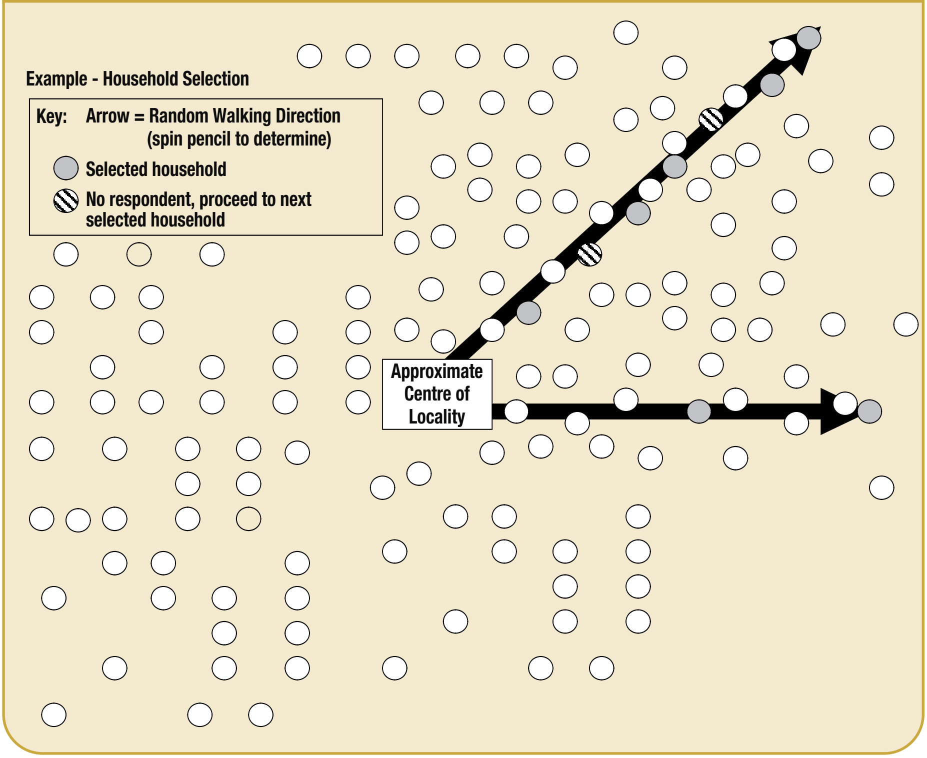
34. This section borrows heavily from the procedures outlined in the FANTA Sampling Guide (Magnani 1997).

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

Option 1b. A second option is to select households systematically. A sampling interval of 3.5 is calculated (35 HHs in the village divided by 10 HHs needed for the sample) in this example. Household 2 is selected as the random starting household (chosen in the range of 1 to 3, since 3.5 contains a decimal). The sampling interval of 3.5 is added to the random start to select the second household (5.5, rounded up to household 6). Add the sampling interval again to get the third household (5.5 + 3.5 = 9), and so on.

Option 2. The second option for selecting households is the most rapid, but also the less preferred method. This method is sometimes used in the Expanded Programme on Immunization (EPI) surveys and in UNICEF anthropometric surveys. Once the data collection team arrives in the cluster, the approximate middle of the cluster is identified. A pencil or bottle is spun to select a random walking direction (also called a transect line). The data collection team then counts the number of households encountered along the transect line between the centre and the perimeter of the cluster. This number is then divided to determine the interval at which households along the transect line will be selected.

Figure 4.1: Household Selection



When the transect line contains fewer than the number of households required, all households in the line are included in the sample and the data collection team returns to the centre of the cluster to pick a second random walking direction, and the process is repeated. If a household without an appropriate respondent is encountered, skip it and proceed to the next selected household. This may require returning to the centre and repeating the process for transects with fewer than the number of required households. This method usually results in a bias, because households from the centre of the village can be overrepresented. Additionally, enumerator teams tend to bias themselves toward transects along main roads or paths.

Example An assessment was carried out in Tambura District in Southern Sudan. Villages served as clusters. Thirty villages were selected in each of 2 livelihood zones, with each representing a stratum. Ten households were selected in each village for a per-stratum sample size of n = 300 and an overall sample size of n = 600. Upon arrival in each selected village, the data collection team asked two key informants to help locate the centre of the village. A pencil was spun to pick a random walking direction (transect). The number of households encountered when walking from the centre of the village to the perimeter was 20. Therefore, every other household was selected for inclusion in the sample.

In two households, an appropriate respondent was unavailable. Therefore, the data collection team was required to repeat the process by returning to the centre, picking a transect line, dividing the number of households in that line by 2 (the number of replacement households needed). With 8 households in that transect, this resulted in every fourth household in the second transect line being sampled.

4.1.2.4 Multi-stage sampling

In the majority of CFSVAs, a two-stage cluster sampling methodology is used. However, on rare occasions a multi-stage method may be required.

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

Stage 1. Random selection of villages
Stage 2. Random selection of households within selected villages
Stage 3. Random or systematic selection of household members within selected households

The design effect, and therefore sample size requirements, increase considerably with each additional sampling stage. Therefore, multi-stage sampling (where districts are sampled, and then, within them, villages, and then, within those, households) is not recommended.

A common mistake when designing a survey is, for logistical reasons, sampling a limited number of districts (one or two) for each province first (stage 1), and in these districts sampling a number of villages (stage 2), and from those villages, sampling 10 households for interview. Such a design will have a huge design effect, and hence very imprecise population estimates. If only one district is sampled from a province, no generalized statements about that province can be made.

4.1.2.5 Stratification or stratified sampling

Stratification, or stratified sampling, involves dividing the population of interest into sub-groups (e.g. strata) that share something in common based on criteria related to the assessment objectives.[35] Stratification is used when separate food security estimates are desired at a predefined, minimum level of precision for each of these sub-groups. When used appropriately, stratification can increase the precision of overall food security estimates for the population of interest.

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

However, stratification is most effective when it is used to define sub-groups within the population that share characteristics related to vulnerability or food security. Livelihoods and land-use zones are examples. If there are distinct livelihood zones in the area where the CFSVA is to be conducted (e.g. agricultural, pastoral, agro-pastoral groups), they can be used to stratify the population. Defining groups in this way serves two functions. First, administrative boundaries rarely correspond with household characteristics related to food insecurity and estimates for administrative aggregations are likely to mask meaningful differences between sub-groups. Second, defining sub-groups for stratification using criteria related to vulnerability or food insecurity improves the precision of both sub-group and overall food security estimates.[36]

Stratified sampling is a key component of all CFSVA sample designs, and is used for comparing sub-groups within the population of interest, an important objective of any CFSVA.

Example The estimated percentage of food-insecure households for Garissa, Kenya, a rural district containing both an area with primarily nomadic pastoralists and one with primarily sedentary farmers (livelihood zones), is 35 percent (+/- 5 percentage points). However, this average at the district level masks the fact that 70 percent of pastoralists are food insecure, while only 10 percent of sedentary farmers are food insecure.
Stratified sampling requires that each sub-group (stratum) be mutually exclusive, meaning that every household in the population of interest must be assigned to only one sub-group. The strata should also be collectively exhaustive, meaning that every household in the population of interest must belong to a sub-group.

35. The purpose of stratification is to define homogenous sub-groups within a heterogeneous population for comparison and, to a lesser extent in CFSVAs, to increase the overall precision of estimates derived from the sample.
36. Stratification by sub-groups defined by criteria related to food security results in more homogenous groupings in terms of food security outcomes. The result is an increase in the precision/accuracy of estimates for each sub-group and of the combined overall estimate for the population due to reduced sampling error. By contrast, stratification by administrative boundary is likely to result in heterogeneous groupings similar to the heterogeneity found in the overall population under study.

In many CFSVAs, two separate geographic stratification systems are used simultaneously. For example, both administrative boundaries and livelihood zones could be used to define the strata. It is important to match what is commonly used in the country, to allow for comparability.

Since information related to food security and vulnerability is most often found for administrative aggregations (districts, divisions, provinces, departments, etc.) or agro-ecological zones, stratification in a CFSVA is always done on a geographic basis. We may prefer to stratify by population group (livelihood groups, gender, wealth groups).[37] However, lack of data almost always makes this impossible.

If it is the intention to report on every cross-section of the two stratification systems, which entails the inclusion of additional villages and households in the sample, each additional sub-group (e.g., stratum) represents an increase in cost and time required to conduct the assessment. If the reporting domain is each stratification system separately (and not the cross-sections), cost increase is limited. Therefore, cost and time constraints will figure heavily into if and how a sample can be stratified. If, for example, the sample size required for a province level of estimate at a reasonable level of precision is 200 households, stratifying the province into two livelihood zone sub-groups would require applying the same sample size to each of the two livelihood zones if the same level of precision were desired for each sub-group (200 x 2 = 400).

Example A food security assessment in Country ABC was originally designed to yield district-level estimates for four districts (four strata). The estimated sample size required was 400 households per district for a total of 1,600 households.
Upon further reflection, the country office wanted results reported by major land-use zones within each district (requiring stratifying by two criteria). Land-use maps suggested that two of the districts had four land-use zones and the other two districts had three land-use zones, for a total of 14 land-use zones/district strata. To keep the same precision in each of these 14 combinations, a sample size of 400 is needed in each zone, increasing the sample size to 5,600 households.
However, another option was to increase the sample with just 450 HHs to 2,050 HH. As a result, each of the four land-use zones also had a sample size of at least 400 HH. Data could now be reported either by district or by land-use zone.

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

All CFSVAs use some sort of stratified sampling, usually based on geographic areas (administrative boundaries and/or livelihood or food security zones). A best practice is

37. Of course, this does not prevent reporting according to these individual household characteristics, such as livelihood.

to design the sample so that it equally satisfies both an administrative stratification and a food security zone stratification. This means that the analysis can provide aggregates for both stratifications. In the report, one can highlight the stratification that best explains the observed differences in a particular indicator.

4.1.2.6 Implications of a complex sample design

The sampling strategy must be taken into account in data analysis. The software SPSS assumes a simple random sample, which assumes that each subject (household) has the same probability of being selected, and that the selection of each subject is independent of the selection of any other subject. This is rarely the case with CFSVAs, which typically use a complex sample design using stratification and multi-stage (usually two-) cluster sampling.

4.1.2.7 Error of estimates/design effect

Cluster sampling produces a less precise estimate than a simple random sample. This is referred to as the design effect, which is the number by which the sample size is multiplied to get the same margin of error as a simple random sample. The design effect will be different for each variable, and can be calculated only *post hoc*. In calculating sample size, an educated guess is made at the design effect, often assuming a design effect of 2. However, this is just an assumption; the design effect can be smaller or much larger, depending on the cluster size and the circumstances in the country.

The reason for the design effect is that households in the same village are often similar to each other (i.e. there is an intra-cluster correlation). Twenty households from 2 villages will not reveal as much about the entire population as 20 households from different villages. The higher the intra-cluster correlation, and the more households coming from the same cluster, the higher the design effect.

To calculate all measurements of errors (including confidence intervals, standard deviation, standard error, and variance), SPSS assumes a simple random sample. Only by using the complex samples module of SPSS (which is not included in the standard version) can complex sampling designs be taken into account.

4.1.2.8 Weights

The most important effect of a complex sampling design is the need for weighting. If each household in the sampling frame (and therefore, the resulting sample) has an equal probability of being selected, then no weighting system is needed. If this is not the case, then a weighting system needs to be used. Weights are needed to compensate for the unequal probabilities of a household being included in the sample.

As the goal of a CFSVA is to make estimates to the larger population, the use of weighting needs to be taken into account at each level of a survey, from design to reporting. Statements such as “40 percent of the sample responded that . . .” are a subtle way to get around data that is not representative of the greater population, but also reveal little or nothing about the greater population. For example, the research team is not interested in the food security situation of the 2,000 sampled households; rather, they want to infer conclusions to larger groups and regions and therefore need to make statements such as “40 percent of all rural households are . . .”

For example, in country A, estimates are desired for each of the three provinces, X, Y, and Z. An equal number of households are selected from these three provinces, and all households within each province have the same probability of being selected. However, the populations within the provinces are unequal. This means that when analyses are run on unweighted data, the estimates will be accurate by province, but when estimated by any other stratifying variable or together for a national average, the results will be biased – that is, the numbers will reflect the sample but not the population from which the sample is taken.

Design weights can be described as "the inverse of the probability that a household could be selected." Another definition for the weight is "the number of HHs represented by one sampled household in a substratum." For example, if in a province with 100,000 households 200 HHs are sampled, the weight of each household equals 500 (every household represents 500 HHs from the population); if in another province with 50,000 HH, again 200 HHs were sampled, the weight for that province would be 250 (every household represents 250 HHs of the sampling universe).

Box 4.2: Calculating design weight

$$Ws = \frac{N_s}{n_s}$$

W_S: Design weight in sampling stratum s
n_S: Sample size of sampling stratum s
N_S: Number of households in sampling stratum s

To correct for this unequal sampling probability using weights, a weighting factor must be added in SPSS. Although the weights as calculated in Box 4.2 are correct, in practice the "normalized weights" are used as correction factors.

Box 4.3: Calculating normalized weight

"Normalized weight" can be arrived at by the following formula:

$$w_s = \frac{N_s / N}{n_s / n} = \frac{N_s / n_s}{N / n} = \frac{N_s}{n_s} \cdot \frac{n}{N}$$

w_s:	Normalized weight for sampling stratum s
N_s:	Number of households in sampling stratum s
N:	Total number of households in the entire sampling universe
n_s:	Sample size of sampling stratum s
n:	Total sample size of all sampling stratums
N_s/N:	Proportion of all households living in sampling stratum s
n_s/n:	Proportion of sampled households coming from sampling stratum s
N_s/n_s:	The design weight in stratum s
n/N:	The sampling fraction of the survey

Table 4.1 illustrates an example of normalized weight calculation. The example illustrates that 15 percent of all households of a country live in province A; however 20 percent of the households are sampled from province A. The normalized weight to be applied to households from province A in SPSS is: 15%/20%= 0.75.

Table 4.1: Example of normalized weights

Province	Total households (N_S)	(Ns/N) (%)	Sampled households (n_S)	(ns/n) (%)	Design weight (N_S/n_S)	**Normalized Weight (ws) [(Ns/N)/(ns/n)]**	Weighted number of households
A	30 000	15	300	20	100.0	**0.75**	225
B	80 000	40	450	30	177.8	**1.33**	600
C	50 000	25	400	27	125.0	**0.94**	375
D	40 000	20	350	23	114.3	**0.86**	300
All	200 000	100	1 500	100	133.3	**1.00**	1 500

Once these weights are calculated, new variables in SPSS must be created, and each case will have its design weight and its normalized weight recorded. Using the example in Table 4.1, all the households in Province A will have a normalized weight of 0.75 and a design weight of 100. All households in Province B will have a normalized weight of 1.33 and design weight of 177.8, and so on.

The weights can then be applied in SPSS by activating them. Under "Data" select weight cases, and then select the variable that contains the appropriate normalized weight. In the complex sample procedure of SPSS, the design weight is used.

Box 4.4: How to double-check weights

The weighted total number of all households should be equal to the original unweighted number of households, if normalized weight is to be used.

The unweighted mean of the normalized weights in the data set should equal 1. Think about normalized weights logically: in areas that are very over-sampled (where the sampling fraction is larger – e.g., a province with a small population where the same number of households were sampled as in other, more populous provinces), the weight should be smaller. In areas that are very under-sampled (where the sampling fraction is smaller), the weight should be larger.

In an efficient sample design, the normalized weights should not be very different from 1 (ranging from 0.75 to 1.5 or so). Very large weights (above 2 or so) and very small weights (below 0.5) can decrease accuracy.

4.1.2.9 Considerations in complex samples

A CFSVA typically uses a two-stage sampling design. First, clusters (villages) are sampled, and during the second stage, households are sampled in each cluster. The clusters must be taken into account during analysis, particularly when calculating tests of significance, standard deviations/variations, and confidence intervals. Compensating for clusters in analysis will not alter point estimates, but will change (widen) confidence intervals and variations. Clusters can be dealt with in a variety of ways. The design effect is influenced primarily by two things (assuming constant sample size): (1) the number of households in each cluster, and (2) the intra-cluster correlation.[38] Another way to think about it is that, keeping constant sample size, more clusters (and therefore fewer households per cluster) means a smaller design effect.

- With a constant sample size, if there is a large number of clusters in the area of estimation (greater than 50 or so) and the number of households in each cluster is small (less than 10 or 15), then the design effect will be small (typically around 2 for many indicators used in CFSVAs). In this case, analysis can be run without considering the design effect (this should be clearly stated in the report). Additionally, statistical tests are more likely to find significant differences when there are none, and confidence intervals will be larger than calculated. Although not ideal, this is also the most common approach.
- If there are a small number of larger clusters in the area of estimation (not a typical situation), complex sample analysis (a procedure in SPSS) must be used.

A typical CFSVA generally uses many clusters (e.g. 250 clusters of 10 households). Typically small cluster size results in a higher number of clusters, which gives a smaller design effect (although it depends on the sample and the indicator). Other surveys, such as nutrition surveys, use 30 clusters of 30 households, which tend to give a design effect of around 2 for nutritional indicators. Compared to other socio-economic or livelihood indicators, nutritional indicators have lower intra-cluster correlation. WFP's experience from Niger shows that 10 households per cluster gives a design effect of 2.0.

When analysing nutritional indicators (stunting, wasting, underweight, etc.) and other methodologically bound and important indicators, the effect of clusters should always be accounted for in the analysis.

Another positive but little exploited effect of a complex sample is stratification. If a set of strata in the survey explain a lot of the variation in the indicator (the variation between strata is large, but within a stratum, small), then using SPSS's complex samples, this can be taken into account, and a greater level of precision can be achieved in the statistical tests.[39] However, the benefits rarely add much value to the results, so this is rarely used in CFSVA analysis.

4.1.2.10 Typical survey design: sampling, weights, and analysis

CFSVAs usually use a complex sample design, using both stratification and a two-stage cluster approach. Usually, areas of estimation have a minimum of 20 or 25 clusters, where each cluster has a minimum of 10 households. Most of the time, these are not self-weighting samples, and need probability weights in analysis.

Considering this, it is absolutely necessary to use the weights in analysis, and their use is strongly recommended to account for the cluster design.[40] However, if the clusters

38. Intra-cluster correlation is a measurement of the relative homogeneity or heterogeneity within clusters as compared to the homogeneity or heterogeneity between clusters.
39. This means a design effect of 0.9 is possible in this case.
40. To account for cluster design, use the module "complex samples" in SPSS; the alternative: simply assuming a design effect equal to 2 is not especially accurate.

are not accounted for, it should be reported in the methodology that any confidence intervals (CI) reported are likely to be wider, and that the tests of significance may indicate significance when there is in fact none.

4.1.3 Determining the appropriate sample size

The aim of this section is twofold: to provide a basic understanding of the factors to be considered in the calculation of sample size, and, more important, to offer easy-to-use[41] sample size guidance for common scenarios found in CFSVA surveys. Only the most basic sample size calculation guidance is given here, which is even more than what is needed in CFSVAs, since almost always the "rule of thumb" sizes of 200–300 HHs per reporting domain is applied.

The choice of sample size formulas depends on whether the key food security indicator (or indicators) of interest for the assessment is mean or proportionate.[42] A primary objective of most CFSVA surveys is to estimate the percentage of food-insecure households within the population. However, some CFSVA surveys will use other indicators expressed as means.

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

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

4.1.3.1 Sampling when key indicators are expressed as percentages

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

$n = (D)(z^2 * p *(1-p))/d^2$

Where:

n = Required minimum sample size

D = Design effect (often assumed to be 2, but varies by type of sampling and by indicator)

z = z-score corresponding to the degree of confidence (1.96 if degree of confidence is 95 percent)

p = Estimated proportion of key indicator expressed as a decimal (e.g. 20 percent = .20)

d = Minimum desired precision or maximum tolerable error expressed in decimal form (e.g. +/- 10 percentage points = .10)

41. Guidance is provided that does not require users to make the calculation themselves.

42. The term *proportion* includes percentages and prevalence.

Taken as a whole, the formula can be intimidating, particularly for those unfamiliar with mathematical notation. However, taken separately, each parameter in the formula is relatively easy to define, and automated sample size calculators are available to perform the necessary computation. In addition, recommended sample sizes (not requiring computation) are provided for common scenarios encountered in CFSVA studies.

D The design effect for simple random sampling is equal to 1 (meaning there is no design effect). The design effect for cluster or two-stage cluster sampling is the factor by which the sample size must be increased in order to produce survey estimates with the same precision as with a simple random sample.[43] A typical value for cluster and two-stage cluster sampling is 2, resulting in a doubling of the sample size requirement. However, it may be possible to reduce this value by increasing the number of clusters, and hence having a lesser number of households in each cluster, or when design effect estimates for the same indicator are available from previous surveys.[44] For a given number of clusters, the gains from adding a few additional households are usually minimal. In a CFSVA, typically 10 to 15 households are selected and interviewed in each cluster (a village), with 17 households, instead of 15, interviewed resulting in a minimal gain in precision.

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

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

Since CFSVAs report a variety of indicators (not just percentage of food insecure), it is generally recommended to use the default of 50 percent, knowing that certain indicators with a higher or lower prevalence than 50 percent will have tighter confidence intervals (i.e. more precision).

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

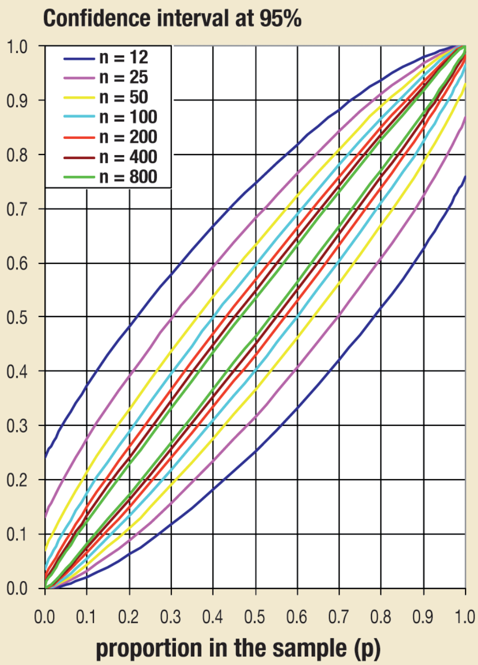
43. See FANTA Sampling Guide for a more in-depth discussion (Magnani 1987).
44. Demographic and Health Surveys (DHS) often have estimates of the design effect of two-stage cluster sampling for food security indicators.

Example It is estimated that 28 percent (+/- 5 percentage points) of households in a rural district in Bolivia consume meat less than once a week. The "+/- 5 percentage points" is the degree of error around the estimate; it defines the confidence interval. The point estimate 28 percent reflects the percentage actually found in the sample population. The range, or confidence interval, of 23–33 percent better reflects the larger population from which the sample was taken.[45] The larger the sample, the narrower the confidence interval.

4.1.3.2 Sample size guidance

Figure 4.2 depicts the intervals that will be obtained with a 95 percent level of confidence, for different simple random sample sizes and different values of an indicator expressed as a proportion. Examples of such indicators are the proportion of food-insecure households, proportion of illiterate household heads, and the global acute malnutrition rate. Figure 4.2 clearly shows that the precision of the estimate is less for proportions around 0.5, and that the higher the sample size, the higher the precision. It also shows that in order to increase the precision appreciably, if there are already 200 units, one has to add a large number of additional units. This graph can be used to gain an idea of what sample size is needed to obtain the desired precision.

Figure 4.2: Confidence intervals for proportion estimates using simple random sampling[46]



Example: The assessment will employ a two-stage cluster sampling method (design effect = 2). The estimated population proportion is 0.4. The graph shows that for a simple random sample of 100 HH, the interval would be 0.31 to 0.50. Hence a cluster design sample with a design effect of 2 would need (100 x 2 =) 200 HH to obtain the same confidence interval. A cluster design sample of 400 households would have a confidence interval of 0.34 to 0.46 by following the lines for n = (400/2 =) 200.

Example: An assessment in West Bank/Gaza will employ a two-stage, cluster sampling method. An estimate for the key food security indicator is 60 percent for the target population. The assessment team decides that the estimate for the population should have a degree of error no larger than 5 percentage points (pp in column) in either direction (+/- 5 pp). This corresponds with the n = 400 line; assuming a design effect of 2, the sample needs (400 x 2 =) 800 HH. (The exact calculation results in a required sample size of n = 762.)

45. As discussed under z, in section 4.1.3.1, the convention for confidence intervals is 95 percent. A comprehensive statement about the estimate given in the example would be "we are 95 percent confident that the true proportion of households in X District, Bolivia, consuming meat less than once a week falls between 23 percent and 33 percent" or "it is estimated that 28 percent (95 percent CI 23–33 percent) of households in X District, Bolivia, consume meat less than once a week."
46. Based on the Wilson score interval. See the formula at: http://en.wikipedia.org/wiki/Binomial_proportion_confidence_interval.

Box 4.5: Nutrition surveys vs. food security sampling requirements

Food security analysis has less stringent demands on the precision of a single indicator, since convergence of evidence is used to make conclusions about food insecurity, whereas nutrition indicators (stunting, wasting, and underweight) demand a greater level of precision. A few percentage points difference in prevalence of "food insecurity" is acceptable as a margin of error. However, a difference of a few percentage points in the wasting prevalence can have considerable implications for programmes and responses.

Nutrition survey

Most surveys measure more than one nutrition outcome (wasting, stunting, etc.). In such cases, the sample size for each outcome should be calculated. The largest sample size is used in order to ensure that all outcomes have an adequate sample size.

Many nutrition survey managers select a sampling consisting of 30 clusters, with 30 children or households in each cluster (often called 30-by-30). Such sampling has also been recommended by many organizations in order to ensure sufficient precision when the survey is completed. However, in some situations, the 30-by-30 cluster survey provides more precision than is actually needed. The 30-by-30 sample size assumes that:

- the prevalence is 50 percent;
- the desired precision is -+ 5 percentage points;
- the design effect is 2; and
- 15 percent of the households or children will refuse or be unavailable.

However, in most emergency-effected populations, the prevalence of acute malnutrition is much lower than 50 percent, far less than 15 percent of households refuse or are unavailable, and the design effect may be less than 2. Therefore, for most nutrition surveys where wasting is the primary indicator of interest, the desired precision can be obtained with a sample size substantially less than the 900 included in the 30-by-30 survey. The solution should be to calculate the sample size that provides the desired statistical precision, but not more.[47]

Household food security survey

Food security surveys estimate several key proxies of food security, so it is difficult to determine one prevalence a priority to be used in the calculation of the sample size. It is best, therefore, to use 50 percent in the sample size calculations, which gives you the most conservative (largest) sample sizes. On the other hand, a larger margin of error is usually acceptable with food security indicators than with nutrition indicators. For example, a sample size of 200 households with a design effect of 2 gives a 95 percent confidence interval of between 40.5 and 59.5 percent. This accuracy is sufficient for most food security analysis.[48]

Reconciling food security survey and nutrition survey requirements

When reconciling the required sample sizes for nutrition surveys versus food security indicators, the first question to ask is the purpose of the nutrition data. As presented in section 6.2.2, the primary goal of collecting nutrition information in a CFSVA is to analyse the link between food security and nutrition. This can be accomplished even with the smaller sample sizes. However, if the secondary goal of collecting nutrition data is also included in the scope of the survey – to provide accurate and precise prevalences of key nutrition indicators – then the nutrition sample size has to be adequate for that goal.

47. Identified as "common mistake #6" by CDC and WFP, 2005, *A Manual: Measuring and Interpreting Malnutrition and Mortality.*
48. We could also state with 75 percent confidence that the interval is between 44.3 and 55.7 percent. Another argument in WFP's analysis is that we do not need these "scientific" levels of confidence to inform programmes.

In CFSVA practice
If the CFSVA aims only to study the link between food security and nutrition with the nutritional information gathered (using under-5 anthropometry as an example), then the smaller sample size as recommended for food security surveys is acceptable: All households are eligible to be included in the survey. All children under 5 from these households are weighed and measured. This is a representative sample resulting in unbiased, but less precise, population estimates of the anthropometric indicators that, when reported, must indicate the correct confidence intervals.

If the CFSVA needs to make accurate and precise estimates of malnutrition prevalence, larger sample sizes will be needed. Three options are generally considered:

1. Sampling proceeds as described above, but nutrition indicators are reported at a more aggregate level. For example, if a sample of 200 households per "province" is taken, nutrition indicators are reported for a higher aggregation level, a "region," which joins three or four provinces.

2. Sampling proceeds as described above, and in each cluster, additional households are selected, and only anthropometry (and possibly one or two other key indicators) are collected from these households, in order to satisfy the nutrition sampling requirements.

3. The sample size for the household survey is increased to take into consideration the nutritional indicator sample requirements. Then sampling proceeds as above.

4.1.3.3 Rules of thumb: A typical CFSVA sample procedure

Typical sampling methodology for CFSVAs.
Although there are many valid possibilities for taking representative samples in a CFSVA, most countries find similar demands for their sampling. Here is an example of a sampling methodology that follows this "typical" situation.

In Haiti, the WFP country office wanted estimates of food insecurity and other key indicators for the rural population at the department level. (Departments are the main administrative boundaries.) There are 10 departments in Haiti. Additionally, FEWS NET and the Government, both key partners, wanted estimates for the FEWS livelihood zones. There are eight livelihood zones in Haiti, although "urban" was excluded, and "sea salt" included as part of "Plains" under "Monoculture zone" due to its very small size and population. This reduced the zones to six strata. Livelihood zones and departments do not share common boundaries.

The following map shows the departments and livelihood zones in Haiti.

Figure 4.3: Livelihood zones in Haiti

To determine a sample size for any geographic stratum where an estimate will be made in the report, CFSVA sampling typically follows some rules of thumb. These include:

- A two-stage cluster sample is drawn. Usually a minimum of 25 clusters (typically villages) per stratum is desired.
- A minimum sample size of 250 households per stratum is desired, 300 is better, if time and logistics allow. A sample of over 300 is rarely necessary (see table here).
- Between 10 and 15 households are sampled in each cluster; in the case of Haiti, 12 HHs per cluster.
- Typically a design effect of 2 is assumed under these conditions.

Following these criteria, the 95 percent confidence intervals for the different cluster sample sizes are the following:

Sample Size (HHs)	Design Effect	Assumed Prevalence (%)	95% Confidence Interval (%)
200	2	50	40.4–59.6
	2	20	13.3–28.9
250	2	50	41.4–58.6
	2	20	13.9–27.9
300	2	50	42.1–57.9
	2	20	14.4–27.1
400	2	50	43.1–56.9
	2	20	15.0–26.1
500	2	50	43.8–56.2
	2	20	15.5–25.4

As seen in the table, only small decreases in confidence intervals occur when increasing sample size from 250 or 300 to 500.

There may be a need to make a statement about a sub-stratum. This is particularly common post-survey, where a particular group (geographic, livelihood, etc.) of households displays unexpected findings that might be worth highlighting in the report. In the Haiti example, the livelihood zone “Plateau agro-pastoral” was thought to have some possible geographic distributions of livelihoods that would better inform future use of the zones. This implies that, in analysis, this zone may be further split into separate strata. For example, if 30 percent of the zone falls into one area during the analysis, a sample of 250 households in the zone will yield information on only 75 households, making statements about that area imprecise. However, if 500 households were sampled in that zone, 150 households would be included in the survey in that area – which would allow for tighter confidence intervals and therefore more confidence around statements made about that area. In the case of Haiti, it was decided to draw a sample that could provide estimates either at the livelihood zone level, or at the department level.

Although estimates in each of the intersections would be interesting and useful, this would result in 35 strata. If the minimum sample of 250 per stratum were used, this would give (35 strata x 250 HH per stratum =) 8750 HH. This sample size would be far beyond the available resources for the survey.

The next step is to gather relative population estimates. The following population information was calculated. (These numbers reflect population, but households could equally be used.) This information shows the intersections of all departments and livelihood zones.

This data could come from a recent census, other surveys, or Landsat information using GIS techniques to estimate the populations by geographic area.

Table 4.2: Intersections of departments and livelihood zones for sampled rural populations

Rural Population in the Sampling Frame	**Livelihood zone**					
Department Name	Dry agro-pastoral (plus sea salt production)	Plains monoculture	Humid mountain agri.	Plateau agro-pastoral	Agri. semi-humid	Dry agri. and fishing
Artibonite	141 591	267 176	285 739	32 452	n/a	177 205
Centre	n/a	n/a	84 904	405 120	n/a	1 954
Grand'Anse	n/a	n/a	304 313	n/a	47 287	185 229
Nord-Est	69 058	23 306	108 253	n/a	n/a	n/a
Nord-Ouest	246 964	n/a	103 817	n/a	37 455	27 104
Nord	n/a	186 352	355 576	41 094	n/a	n/a
Ouest	41 160	380 199	231 616	n/a	13 319	412 895
Sud-Est	634	n/a	134 433	n/a	68 821	225 526
Sud	20 581	61 542	89 989	n/a	112 206	236 781

CHAPTER 4

If Haiti wanted estimates only for the 10 departments, using 250 HHs per stratum, this would give a total sample size of 2,500 HHs. If only the 6 livelihood zones were considered, a sample size of 1,500 HHs would be needed.

An initial total sample (all strata) was set at about 2,500. This would give about 250 HHs per department, and at least that per livelihood zone. A cap of about 3,000 HHs was set for the entire sample due to logistics and financial limitations. (Note that in this example, Grand'Anse and Nippes were combined, as they did not have separate population figures at the time of the sampling.)

In Excel, a dynamic spreadsheet (Table 4.3) was created to explore possible sample sizes that would provide an appropriate sample size at both department and livelihood zone.

- The population sizes (column C) were taken from Table 4.2.
- The percentage of total population in each line was calculated based on the population size of the line (value in column C cell) divided by total rural population (sum of column C).
- A proportionate sample size (column F) was estimated based on the percentage of total population (e.g. cell F1 = cell E1 x sum of column D).
- Column F was copied (values, not formulas) and placed in column D. (These values were later altered). Column G was calculated using the formula presented in section 4.1.2.8 on weighted data analysis.

Table 4.3: Example of CFSVA sampling procedure

	A	B	C	D	E	F	G
	Department	Livelihood zone	Total population size (rural)	Population of each combination as proportion of the population in entire sampling frame (%)	Planned sample of HHs	Sample size of each combination as proportion of entire sample (%)	Standardized weight (how to duplicate each HH to have a proportionate overall sample) G = D/F
1	Artibonite	Dry agro-pastoral	141 591	2.7	60	2.0	1.38
2	Artibonite	Plains monoculture	267 176	5.2	120	4.0	1.3
3	Artibonite	Humid mountain agri.	285 739	5.5	120	4.0	1.39
4	Artibonite	Plateau agro-pastoral	32 452	0.6	24	0.8	0.79
5	Artibonite	Dry agri. and fishing	177 205	3.4	72	2.4	1.44
6	Centre	Humid mountain agri.	86 858	1.7	48	1.6	1.06
7	Centre	Plateau agro-pastoral	405 120	7.8	300	9.9	0.79
8	Grand'Anse/Nippes	Humid mountain agri.	304 313	5.9	180	6.0	0.99
9	Grand'Anse Nippes	Agro-pastoral semi-humid	47 287	0.9	48	1.6	0.58
10	Grand'Anse/Nippes	Dry agri. and fishing	185 229	3.6	132	4.4	0.82
11	Nord-Est	Dry agro-pastoral	69 058	1.3	96	3.2	0.42
12	Nord-Est	Plains monoculture	23 306	0.5	36	1.2	0.38
13	Nord-Est	Humid mountain agri.	108 253	2.1	120	4.0	0.53
14	Nord-Ouest	Dry agro-pastoral	246 964	4.8	156	5.2	0.93

	A	B	C	D	E	F	G
	Department	Livelihood zone	Total population size (rural)	Population of each combination as proportion of the population in entire sampling frame (%)	Planned sample of HHs	Sample size of each combination as proportion of entire sample (%)	Standardized weight (how to duplicate each HH to have a proportionate overall sample) G = D/F
15	Nord-Ouest	Humid mountain agri.	103 817	2.0	60	2.0	1.01
16	Nord-Ouest	Agro-pastoral semi-humid	37 455	0.7	36	1.2	0.61
17	Nord-Ouest	Dry agri. and fishing	27 104	0.5	24	0.8	0.66
18	Nord	Plains monoculture	186 352	3.6	108	3.6	1.01
19	Nord	Humid mountain agri.	355 576	6.9	180	6.0	1.16
20	Nord	Plateau agro-pastoral	41 094	0.8	24	0.8	1.01
21	Ouest	Dry agro-pastoral	41 160	0.8	24	0.8	1.01
22	Ouest	Plains monoculture	380 199	7.4	180	6.0	1.24
23	Ouest	Humid mountain agri.	231 616	4.5	96	3.2	1.41
24	Ouest	Agro-pastoral semi-humid	13 319	0.3	24	0.8	0.33
25	Ouest	Dry agri. and fishing	413 529	8.0	192	6.3	1.26
26	Sud-Est	Humid mountain agri.	134 433	2.6	72	2.4	1.09
27	Sud-Est	Agro-pastoral semi-humid	68 821	1.3	60	2.0	0.67
28	Sud-Est	Dry ag. and fishing	225 526	4.4	132	4.4	1
29	Sud	Dry agro-pastoral	20 581	0.4	12	0.4	1.01
30	Sud	Plains monoculture	61 542	1.2	36	1.2	1.01
31	Sud	Humid mountain agri.	89 989	1.7	36	1.2	1.46
32	Sud	Agro-pastoral semi-humid	112 206	2.2	84	2.8	0.78
33	Sud	Dry agri. and fishing	236 781	4.6	132	4.4	1.05
		TOTAL	**5 161 651**	**100.0**	**3 024**		

In this example, it is assumed that household size does not change across the rural area. Hence for calculating weights, we use the number of people in a certain area, instead of the number of households, as the basis for the calculations.

Next, the totals from column D were calculated for each department and each livelihood zone using formulas to retain the dynamic nature of the spreadsheet.

	Total sample
Artibonite	396
Centre	348
Grand'Anse/Nippes	**360**
Nord-Est	252
Nord-Ouest	276
Nord	312
Ouest	516
Sud-Est	264
Sud	300

	Total sample
Dry agro-pastoral	348
Plains monoculture	480
Humid mountain agri.	912
Plateau agro-pastoral	**348**
Agro-pastoral semi-humid	252
Dry agri. and fishing	684

CHAPTER 4

Finally, the samples from each department-livelihood zone combination were adjusted following the following criteria:

- All strata (either department or livelihood zone) should have a minimum of 250 HH.
- All sample sizes should be a multiple of 12 (as it was planned to have 12 HH per cluster, so the sub-strata sample sizes would then all give a round number of clusters).
- The total sample size should remain at or below 3,000 households.
- The sampling weights should not become too extreme (it was attempted to keep them all between 0.75 and 1.5, where possible).
- All strata combinations should have a non-zero sample size, and preferably above 36 (if possible).
- The Grand'Anse/Nippes departments, as they are combined in drawing the sample, should be oversampled to the extent possible to allow for later disaggregation between them.
- The livelihood zone Plateau Agro-pastoral should be oversampled, as a secondary analysis of distribution of livelihoods within this livelihood zone was desired.

These adjustments to the sample sizes required some degree of "playing" with the numbers to achieve a sample that satisfied the requirements proposed, and to have a sample that appeared reasonable to the partners involved in the survey. Especially since each "combination" at the same time contributes to the total sample size of a livelihood zone and a department.

Rules of thumb for effective "playing":

- If the totals of some departments or livelihood zones are insufficient:
  - add households/clusters to the intersections where both the department and livelihood zone are lacking in numbers;
  - add households/clusters to the cells of strata with insufficient numbers, add them first where the sample weight is highest.

Once the sample sizes were finalized, the villages (clusters) were selected using PPS selection within each of the department/livelihood combinations, following the guidance for cluster selection.

Within each village, households were selected using simple random sampling, as described in the section on selecting households within clusters.

4.1.4 Key references: Sampling

- Magnani, R. 1997. *Sampling Guide*. Food and Nutrition Technical Assistance Project, Academy for Educational Development, Washington, D.C.
- TANGO International. 2002. *Household Livelihood Security Assessments: A Toolkit for Practitioners.* Prepared for CARE International.
- ibid. 2007. *Monitoring and Evaluation Manual. Prepared for ADRA International.*
- WFP. 2004. *Sampling Guidelines for Vulnerability Analysis: Thematic Guidelines*. Rome.

4.2 HOUSEHOLD DATA COLLECTION

4.2.1 Introduction

Developing questionnaires is a central component of the survey design and implementation process. The quality of any analysis depends on asking the right types of questions and getting reliable answers. At the same time, questionnaires are directly linked to data management and storage. The form a questionnaire takes has direct implications on how the resulting data will be organized for use by an analyst.

While VAM staff may be responsible for different types of assessments with varying objectives, food security remains the central theme. In this context, it is important to ensure that appropriate guidance is available for VAM staff on how best to develop the tools that capture the right information needed for food security analysis.

WFP/VAM surveys use both a quantitative and a qualitative approach in obtaining data. Hence the data collection tools necessarily incorporate components of both approaches.

These guidelines focus on the household questionnaire, as it is the most commonly used data collection tool in a CFSVA survey.

4.2.2 Objective of household data collection

The objective of a household survey is to gather quality indicators, in a standardized way, which after analysis will provide the useful statistics required to fulfil the objectives of a CFSVA, EFSA, or FSMS.

A good-quality questionnaire is a necessary, but not sufficient, tool for obtaining primary data reflecting the real situation. Other necessary tools are: enumerator training, an unbiased sample, and the collaboration of respondents.

The starting point for designing a questionnaire is the list of indicators we want to collect through the survey. The CFSVA guidelines give definitions of standardized indicators collected at the household level. The indicators should be, as much as possible, compatible with generally accepted indicators from other surveys (DHS, MICS, etc.) and other organizations.

4.2.3 Conducting household data collection[49]

Preparing the ground
Pre-survey publicity is essential. Enumerators should not show up unannounced to demand information, as that approach is unlikely to be successful. Letters of introduction should be sent to the appropriate officials, community leaders, etc. These should contain an explanation of the purpose of the research, the procedure for selection, the subject matter to be covered, and an assurance of confidentiality and anonymity.

49. This material is partly extracted from Devereux and Hoddinott, 1992, *Fieldwork in developing countries.*

The first step to a successful interview is properly introducing oneself to the respondent and obtaining informed consent.[50] Enumerators must briefly explain to those being interviewed the purpose of their study, who has funded and supported it, how the data will be collected, the expected duration of the interview, and how the results may be used. If a respondent does not understand the purpose of the interview or does not wish to participate, the survey may end up with inaccurate or misleading answers. The respondents should be aware that they will not receive payment or any other form of compensation, but that their participation is voluntary and that they may refuse to participate in the interview or stop it at any time. Consent must be derived from the actual people involved, not just officials or leaders.

Equally important is talking with an appropriate respondent. The respondent should be an adult member of the household – not a guest – and preferably the household head or her/his spouse. If nobody suitable is available, skip this household and move to the next on the list, returning later to interview the household, if possible.[51]

Interacting with respondents

An essential qualification for successful fieldworkers is a demonstrable and genuine interest in other people. The ideal interview is a friendly conversation between enumerator and respondent. The posing of questions and the noting of replies should have the flow and pattern of a dialogue.[52] This is not accomplished if enumerators are impolite or brusque; nor is it possible if poorly trained enumerators fumble their way through the questionnaire.

A common pitfall is the mechanical recitation of questions without thinking about the responses being given; often this leads to extensive work for those cleaning data during analysis. Time must be taken during the interview. Questions must be asked carefully, making sure that respondents have understood them correctly. It may be necessary to repeat questions and probe to be sure the answer recorded is the one intended.

The interview setting

Interviews are usually one-on-one encounters. When the research deals with personally sensitive matters, the presence of outsiders, or even other family members, may inhibit respondents, embarrassing them into evasion or silence. Onlookers may encourage respondents to answer untruthfully. For example, working during food crises is problematic because respondents have an incentive to understate their stocks of grain and their general wealth, with the expectation that food aid will be brought into the community. This tendency will be exacerbated during a public interview, since a respondent who admits to being wealthy may face demands for help from poorer neighbours or relatives. In general, the more sensitive the topic, the stronger the case for conducting the interview in private.

50. Wilson, 1992.
51. WFP. EFSA Handbook 2009.
52. Casley and Lury (1987:111)

It is critical to be mindful of when the interview is conducted, as the process can be an imposition on respondents' time. Day of the week and time of day are important in both rural and urban settings. Attempt to meet on days and at times convenient to respondents. For example, in urban settings, interviews outside of regular work week hours may be necessary. In rural settings, women may be busy on market days and when preparing meals; men may be busy working in the field at a particular time of day. One way of ensuring that interviews are not an imposition is to make appointments to see people. An interview should never last more than 90 minutes and should be held in a place convenient for the respondent. Urban teams and respondents should be protected from violence and crime by interviewing only at "safe times."

4.2.4 Modules of questions

The questions used to construct single indicators, or different indicators related to similar topics, are usually organized in modules, which should be ordered logically in a questionnaire. Within a module, questions follow a logical flow, and should not be redundant.

Types of modules in a food security household questionnaire

Based on an extensive review of household questionnaires, commonly used modules have been divided into four broad categories:

1. ***Core modules with standard, non-changeable, questions.*** These modules have been tested and used in several CFSVAs, EFSAs, and FSMS and must be used in all food security assessment questionnaires. They contain standard questions (formulated in a standardized way but with country-specific adaptations, e.g. food items, expenditure items) that have proven to be useful for data analysis and answering the five VAM questions. The modules were identified as:

   - Food consumption patterns (including sources of foods consumed)
   - Expenditures
   - Household assets
   - Sources of water
   - Access to sanitation

2. ***Core modules with questions that are flexible (i.e. changeable) depending on specific contexts and survey objectives.*** Some standards have been developed for these modules – but there are slight variations from country to country. These modules are central to WFP food security assessments and essential to answering the five VAM questions. Specifically, the modules containing flexible questions are:

   - Household composition/demography and education
   - Housing materials (walls, floors)
   - Access to credit, indebtedness
   - Livelihoods/sources of income
   - Agriculture
   - Livestock
   - External assistance (food and non-food)
   - Shocks, coping, and/or coping strategies index

3. ***Non-core modules with non-changeable questions.*** These modules are sometimes, but not always included in CFSVAs, EFSAs and FSMS. Global standards have been established for these modules, and therefore the questions should not be adapted or changed. The data are very relevant in terms of answering the five VAM questions. Specifically, this type of module includes:
   - Maternal health and nutrition
   - Child health and nutrition

4. ***Non-core modules with changeable/flexible questions.*** This type of module is sometimes, but not always included in WFP food security assessments. These modules are quite flexible in terms of structure and the types of questions posed. The main modules that fall within this category are:
   - Migration/movement, displacement status
   - Remittances

5. For each one of the four categories and modules contained therein, a series of guidance notes have been written that include the following information:
   - Main purpose of the module
   - Current limitations of the module
   - Creation of the module (i.e., how-to)
   - Modifications that can be made to the module
   - Links to other modules

Each category and associated modules are detailed in section 4.2.4.1.

4.2.4.1 Core modules with standard, non-changeable questions

Module Title: Food Consumption Patterns (including sources of food)

Main purpose of the module: This module allows the analyst to calculate the food consumption score (FCS) for each household and to investigate current food consumption patterns.

Limitations of the module
- The module cannot provide the caloric value or nutritional adequacy of the household diet.
- The module cannot measure the quantity of the food consumed.
- The module does not look at the intra-household differences in consumption.

Creation of the Food Consumption Patterns module
Before recording the dietary diversity and frequency of the household diet, ask the following questions as to the number of meals consumed:
i) Yesterday, how many times did adults eat?
ii) Yesterday, how many times did children 6 to 12 years of age eat (should link with demographic section)?

List the food items belonging to food groups typically eaten in a specific context. The list (which is country specific) should contain: (a) staples and food eaten commonly throughout the study area; and (b) preferred items (e.g. maize versus millet, cassava versus Irish potatoes). The list of food items is country specific and should reflect what is typically consumed in the country. However, food items should be listed in such a way that allows their aggregation into the food groups used for computation of the FCS.

- The list should include between 15 and 20 food items. The list is not meant to be comprehensive of all food items found in the country. Instead, it should reflect the basic items found in a general diet (e.g. oil, salt, meat, dairy products). Use of the word "other" can make the list comprehensive enough.

- Corn-soya blend (CSB) should be considered a separate food item.

- If particular condiments are consumed with staples, it is important to identify them as such and not group them with that food item (e.g. fish powder with fresh fish, milk in tea and glass of milk). Training should be given on whether to include condiments in the analysis.[53]

- It is important to ask about combination food items. For example, when a household indicates that maize and sauce were eaten, and the sauce is prepared with oil, vegetables, salt, and chicken, these items should be indicated as consumed, with the amounts of oil, salt, vegetables, and chicken regarded as more than mere condiments.

- If one member of a household consumes food away from the household, the items eaten should not be recorded. If the entire household ate outside of the household, then the items consumed should be recorded.[54]

- For all food items, the recall period is set at the previous seven days. The purpose is to capture the number of days out of seven that a particular food item was consumed.

- Aside from the food items it is important to identify the primary sources from which the food was acquired (this can be either the primary source or the two primary sources). Generally these sources are: own production; hunting, fishing and gathering;[55] exchange; borrowed; purchased; gift; food aid; and credit.

- All items should have a numeric value. There should be no empty cells. If no consumption is reported, then the source and number of days is recorded as zero.

53. This issue needs further discussion and consensus.
54. This is particularly acute in urban and peri-urban areas.
55. If any of these food sources is a specific and important activity, then hunting, fishing, and gathering can be split up.

Table 4.4: Example of a Food Consumption Patterns module

Yesterday, how many meals did the ______ in this house eat?

Adults Children (< 6 yrs)

Food Item	No. of days eaten over last 7 days	Food source (main and secondary)	Food Source Codes
Maize	I__I	I__I,I__I	1 = Own production (crops, animals)
Rice/paddy	I__I	I__I,I__I	2 = Hunting, fishing
Millets	I__I	I__I,I__I	3 = Gathering
Wheat/Barley and other cereal products	I__I	I__I,I__I	4 = Borrowed
Roots and tubers (potatoes, yam)	I__I	I__I,I__I	5 = Purchased with wages
Pulses/lentils	I__I	I__I,I__I	6 = Exchanged labour for food
Fish	I__I	I__I,I__I	7 = Exchanged items for food
White meat (poultry)	I__I	I__I,I__I	8 = Gift (food) from relatives
Pork	I__I	I__I,I__I	9 = Food aid (NGOs, etc.)
Red meat (goat, sheep)	I__I	I__I,I__I	10 = Other (Specify: ______________)
Red meat (buffalo)	I__I	I__I,I__I	
Eggs	I__I	I__I,I__I	
Milk/curd/other dairy products	I__I	I__I,I__I	
Vegetables	I__I	I__I,I__I	
Fresh fruits	I__I	I__I,I__I	
Oil/fats/ghee/butter	I__I	I__I,I__I	
Sugar/sweets	I__I	I__I,I__I	
Salt/spices/condiments	I__I	I__I,I__I	

- The main indicators emanating from the analysis of these data are: (a) number of days out of seven that items and food groups are consumed; (b) household FCS; and (c) percentage contribution of the sources to the household food basket over the previous seven days.
- This module has to be adapted to the context. The food items can be changed, but the exact same eight food groups (staple food, pulses, vegetables, fruits, meat and fish, milk, sugar, oil, condiments) should always be used.[56]

Modifications to the module

- A separate column can be added to the table if an item was consumed in the previous 24 hours. This would be a way to incorporate the Household Dietary Diversity Score (HDDS) indicator in the module. However, the HDDS uses a different data collection methodology and a different questionnaire.[57]
- The number of meals people consumed in the previous 24 hours by age cohort groups in the household (e.g. children under 6, children 6 to 12, and children 13 to 18) can be modified when specific information on child food consumption (or other age cohorts) is needed (e.g. in order to programme a school-age child or MCH food aid intervention).
- Contribution of sources in the previous year and quarters, which can provide information about seasonality for food security analysis, can be modified. Information for a few key staple food groups or for general food consumption may suffice. The source categories should be the same as those used for the food sources. See the following table for an example:

56. See WFP guidelines: *Food Consumption Analysis*, at http://vam.wfp.org/MATERIAL/FCS_Guidance.
57. See FAO, 2007, *Guidelines for measuring household and individual dietary diversity*, June.

In the last calendar year (2006) what was the contribution of (source) to your annual food consumption? How does this differ throughout the year?
(Use proportional piling, or divide the pie method, to estimate the relative contribution of each source to total food consumption)

Food source	Annual (%)	Jan.–March (%)	April–June (%)	July–Sept. (%)	Oct.–Dec. (%)
Own production					
Hunting, fishing, gathering					
Purchases					
Gifts/borrowing					
Food aid					
Total contribution	**100**	**100**	**100**	**100**	**100**

Food Consumption Score (FCS) and Household Dietary Diversity Score (HDDS)

WFP and FAO both use measurements of dietary diversity in their assessments and monitoring systems. WFP has adopted a methodology and tailored it to its own information needs in terms of data collection and analysis of food consumption. FAO uses a methodology based on the Demographic and Health Survey (DHS) procedures developed by FANTA. For both approaches, standard methodologies have been developed to calculate indicators of dietary diversity and consumption frequency.

The Food Consumption Score (FCS). Information is collected from a country-specific list of food items and food groups. The household is asked about the number of times (in days) a given food item was consumed over a recall period of the past seven days. Items are grouped into eight standard food groups (each group has a maximum value of seven days/week). The consumption frequency of each food group is multiplied by an assigned weight based on the nutrient content of a portion. Those values are then summed to obtain the FCS. The FCS has a theoretical range from 0 to 112; WFP has defined thresholds (WFP 2007) to convert the continuous FCS into categories creating three food consumption groups (FCGs): poor, borderline, and acceptable.

The Household Dietary Diversity Score (HDDS). A standard list of 16 food groups, the same for any country/context, is used to gather information on food consumed in the past 24 hours. Information for each group is of a bivariate type (yes/no). To calculate the HDDS, the 16 food groups are aggregated into 12 main groups. All food groups have the same importance (relative weights equal to 1), with each group consumed providing 1 point. The HDDS is the simple sum of the number of consumed food groups (it goes theoretically from 0 to 12). For analytical purposes, the HDDS is often ranked into thirds or quartiles.

Both the FCS and HDDS are used as proxy indicators of household access to food. Data collected for both indicators can also be used to consider dietary patterns and the consumption of specific foods. The FCS and HDDS are used for monitoring economic access to food and surveillance at decentralized levels; moreover, the FCS is used for classifying households who are food insecure, while the HDDS is used for monitoring dietary quality.

Link with other modules

- For internal consistency, it is important that the food items listed in the consumption module be reflected in the expenditures module.

Sources of inspiration

- The standard food consumption module currently adopted for the CFSVAs is unique in its format and methodology. FAO's food consumption module is based on a 24-hour recall. The list of food items in the FAO module is different from WFP's in that it focuses more on diversity and specific food groups (e.g., vitamin A–rich foods).

Module title: Expenditures

Main purpose of the module: This module allows the calculation of household expenditure (in cash)–related indicators. Expenditures are useful as a proxy for wider purchasing power, which is an important component of food access. Moreover, understanding expenditures on specific items allows the analyst to determine how households allocate scarce resources and give priority to competing needs.

Limitations of the module

- Cannot estimate the value of own production (section revolves around only cash expenditures). Collecting consumption and expenditure data is tricky because of the varying extent to which households consume out of their own production, which is not collected in this questionnaire and hence reduces its usefulness.
- Cannot estimate the quantity of food items purchased.
- Food expenditure is linked to the season (e.g. is lower after harvest).
- Non-food expenditures (especially education) are also seasonal.

Creation of the Expenditure module

- List of food and non-food items that are mutually exclusive and yet will encompass all essential expenditures (e.g. education and clothing – do not include school uniforms in both categories).
- List of food items should match those found in the Food Consumption Patterns module, with a few possible exceptions (e.g. collapse meat into one category; include expenditures on condiments).
- In addition to main food items, include "condiments" (e.g. salt, spices, beef-tea cubes, fish powder).
- For all food items, the recall period is set at one month.
- Some non-food items also have a one-month recall period. These are: soap, transport, firewood/charcoal, rent, paraffin, and alcohol/tobacco. (Tip: Collect information on likely daily expenditure for households, not infrequent bulk expenditures.)
- For the remaining non-food items, the recall period is set at six months (prior to the day of survey).
- If other CFSVA and/or EFSAs have been conducted, use the same lists.
- All items should have a numeric value. There should be no empty cells, as that would mean "missing data." If there are no expenditures, use a zero.
- Include debt expenditures (i.e. repayment of loans).

Table 4.5: Example of Expenditure module

In the past **MONTH**, how much money have you spent on each of the following items or services? (Write 0 if no expenditure.)		**a.** Estimated expenditure in cash	**b.** Estimated expenditure in credit (if applicable)
1	Maize		
2	Wheat/barley		
3	Millet		
4	Rice/paddy		
5	Roots and tubers (potatoes, yams)		
6	Pulses/lentils		
7	Vegetables		
8	Milk/yogurt/milk products		
9	Fresh fruits/nuts		
10	Fish		
11	White meat (poultry)		
12	Pork		
13	Red meat (goat, sheep)		
14	Red meat (buffalo)		
15	Eggs		
16	Oil/butter/ghee (fats)		
17	Sugar/salt		
18	Condiments		
19	Alcohol and tobacco		
20	Soap		
21	Transport		
22	Firewood/charcoal		
23	Kerosene		

In the past **6 MONTHS** (semester), how much money (in cash) have you spent on each of the following items or services? *Write 0 if no expenditure.*			
24	Equipment, tools, seeds	30	Celebrations, social events, funerals, weddings
25	Hiring labour	31	Fines/taxes
26	Medical expenses, health care	32	Debts
27	Education, school fees	33	Construction, house repair
28	Clothing, shoes	34	Other long-term expenditure (Specify: ___________)
29	Veterinary expenses		

The typical indicators emanating from this module include:

(a) total household expenditures (food and non-food);
(b) total per capita expenditures;
(c) percentage food expenditures;
(d) percentage non-food expenditures; and
(e) percentage individual food and non-food items.

Modifications to the module

- Recall period for non-food expenditures can be modified to one year.
- Recall period for education can be realigned with the calendar of payments; however, a conversion of the expenditures to "yearly expenditures" must be allowed.
- Recall period for food expenditures can be reduced (1 week) during an EFSA.
- FSMS may modify the non-food recall period to correspond to the last data collection round, as long as a pro-rated amount can be calculated.
- A credit or exchange/barter column could be added when appropriate, but this will require additional time for survey administering. Based on the experience with previous CFSVAs, such data add little value for food security analysis.
- Pastoralists: include information on veterinary fees, water costs for animals, livestock purchases (this could be placed in the Water/sanitation module, as long as both modules have the same recall period).
- Seasonality of expenditures — information on celebrations, school fees, agricultural inputs, seasonal disease outbreaks (e.g., malaria) and health —could be collected in a community/key informant questionnaire.
- Rent expenditures should be included in the monthly expenditures, unless specific questions on house ownership and rental are posed under the "Housing" section.

Links with other modules

- Questions on expenditures can be asked in other modules, as long as they are not duplicated.
- Link to income sources, food consumption modules (triangulation at field level as well as analysis stages).
- When deciding on a recall period, make sure that all relevant secondary data are reviewed to ensure that the proper period is chosen. Harmonize all recall periods in other parts of the questionnaire, otherwise enumerators will likely make mistakes with the period used.

Module Title: Household Assets

Main purpose of the module: This module allows the analyst to calculate proxy indicators of wealth and qualify the type of assets the household possesses.

Limitations of the module

- It does not allow for an exact measure of wealth.
- The list is a finite number of wealth and production assets. The selected assets must be typical for the context but allow for inter-household differences to be captured.
- It is advisable that the list contain between 10 and 15 non-perishable assets.

Creation of the Assets module

- Prepare a list of productive and non-productive assets. Guidance on the applicability for the specific country or region can be sought from previous WFP surveys or HEA, MICS, and DHS studies. In the development of this module, questionnaire designers should look at other household surveys done in the same country, especially large-scale government surveys or DHS, and make the asset module of the CFSVA survey compatible with these. A big advantage to this is the ability to compare levels and distributions of assets to determine, for example, if households in your survey are richer or poorer than households found in other data sets.

- Based on the context, assets should include agricultural and journeyman's tools (e.g., pesticide sprayer, plough, mason's tools, carpenter's tools), as these are examples of productive assets.
- "Luxury" assets are a reflection of the wealth and/or social status of a household. They can be used to generate income, though this is not their primary use. Examples are context specific, but can include a radio, television, satellite receiver, mobile phone, car, and microwave.
- Typical household assets should also be included and, though context-specific, can include: a mattress, a lantern, a mosquito net, or a manufactured cooking stove.
- Once the list has been created and verified by local experts as being applicable to the context, list the items by category (basic, productive, and wealth).
- To the right of the "Assets" column, place either 1 (yes) or 0 (no); the asset the household owns can also be circled.

Table 4.6: Example of Household Assets module

Indicator	Productive/transport assets			
Does your household own any of the following assets? ***Circle all that apply***	1	Shovel/spade	2	Plough
	3	Sickle	4	Weaving tool
	5	Fish net	6	Pounding mill (wood), foot or hand
	7	Rice mill (fuel)	8	Rice mill (electricity)
	9	Motorcycle	10	Hand tractor
	11	Bicycle	12	Boat/canoe
	Household assets			
	13	Sleeping mats	14	Bed
	15	Table	16	Radio
	17	Stove (gas/fuel)	18	Generator (run by water)
	19	Generator (run by fuel)	20	Mosquito net

- The main indicators emanating from this module are: (a) percentage of households owning an asset (e.g. radio), and (b) wealth index.
- Enumerators can be given guidance on excluding assets beyond repair.

Modification to the module

- The number of each asset can be recorded (e.g. 4 chairs, 2 beds, 6 radios), but this adds little value for food security analysis.

Link with other modules

- There should be an internal consistency between productive assets and household economic activities.

Source of inspiration

- Household questionnaires from the DHS always include comparable questions regarding the ownership of productive and non-productive assets. Similar to DHS surveys, CFSVAs look at assets ownership and adopt the same methodology to compute the wealth index (i.e. Principal Component Analysis).

Module Title: Sources of Water and Sanitation

Main purpose of the module: This module allows the analyst to estimate the percentage of the population using improved drinking water sources and improved sanitation, which is a commonly used indicator to assess hygiene at the household level.

Limitations of the module

- The module cannot tell us about the quality of the drinking water for each type of source, the quantity of water the household is drinking, or the household water storage practices.
- The module alone cannot tell us the impact of poor hygiene on food security and nutrition. The link between poor water access and hygienic conditions, malnutrition and food security has to be statistically explored and proven. Appropriate expertise is required for this.

Creation of the Water and Sanitation module

- Based on the accepted United Nations guidance, the primary household water sources must be grouped accordingly: water tap in household; water tap in compound; public stem pipe; borehole; protected dug well; protected spring; rainwater collection; unprotected well; unprotected spring, river, or pond; vendor provided; tanker truck; bottled water. The list is context-specific. An example of the question follows:

What is the main source of drinking water for your household? ***Circle one***	1	Piped water in-/outside	5	Mountain source (incl. gravity water feeder system)
	2	Well/borehole protected	6	Rainwater from tank
	3	Well/borehole unprotected		
	4	River, stream, or dam	7	Other

- Based on the accepted United Nations guidance, household sanitation sources must be grouped as follows: flush toilet, pail flush, simple pit latrine, ventilated improved pit latrine, public/shared latrine, open pit latrine, bucket latrine, bush. An example of the question follows here:

Where do members of your household normally go to the toilet? *(Do not read answers aloud.)* ***Circle one***	1	Flush latrine/toilet with water	4	Communal latrine
	2	Traditional pit latrine (no water)	5	None/bush (go into forest)
	3	(Partly) open pit (no roof or no wall)		

- The main indicators from these modules are: (a) percentage of households using improved drinking water sources; and (b) percentage using improved sanitation.

Modifications to the module

The following questions can be added to further the understanding of household access to water:

- How much time is required to collect water (round trip)? This can either be a specific time (e.g. 35 minutes) or a categorical response (e.g. 1 to 3 hours).
- What is the distance from the water source (one way; usually collected as continuous variable in kilometres)? This information can be used for spatial analysis of accessibility of resources. Appropriate skills and other relevant data must be available for this type of analysis.
- Who collects the water (e.g. girls, boys, women, men)? This information can be collected in the Key Informant or Community questionnaire.
- Do you pay for water? How much do you pay for the water per day/week/month (depending on the context)?
- What is the seasonal variation of your water source? Is there is a second source?

Links with other modules

- If questions are asked regarding expenditure on water, they should not be repeated in the Expenditure module, and the reference period should be the same (e.g. 1 month).
- The link between water source and sanitation and the nutritional status of the households can be explored.
- Components of a wealth index.

Sources of inspiration

- Questions on water sources are always included also in the MICS and DHS questionnaires. The main differences between CFSVA and DHS/MICS modules are:
  i) CFSVAs focus on the source of drinking water, whereas DHS and MICS consider sources of water used for cooking and hand-washing.
  ii) A more detailed list of options is typically included in the DHS/MICS questionnaires than in the CFSVA questionnaires.
- Questions on sanitation are always included in the MICS and DHS questionnaires. In DHS and MICS, the options for type of toilet are usually more detailed. CFSVA HH questionnaires do not include questions on the number of people sharing toilet facilities.

4.2.4.2 Core modules with flexible questions

Module Title: Demography and Education/Household Composition Roster

Main purpose of the module: This module records demographic characteristics of households.

Limitations of the module

- A household is usually defined as people living in the same compound and eating “from the same pot,” forming a clear socio-economic entity. The head of the household makes the major decisions. However, these definitions have to be adjusted to be in line with country-specific definitions. The definition of a household can bias results (e.g. polygamy can be incompatible with the standard definition of a household). Often the definition of the “head of household” is culturally defined and may not reflect “who is making the key decisions” or who is the bread winner.

- With a comprehensive roster (at the individual level), the module can become quite weighty and time-consuming.
- It will not provide rates of fertility or fecundity.
- Mortality and morbidity rates can be calculated provided that the module is organized in a suitable way (using roster type) and that appropriate questions are asked. However, the usual sampling design, while suitable for food security indicators, can give rather imprecise estimates of mortality and morbidity.
- This module is not meant to capture members of the household who have migrated.

Creation of the Demographics module

- There are two standardized methods for collecting household information. However, depending on the information desired, one approach will be recommended (e.g. mortality; non-age-bound population estimates will require a more detailed enumeration of the household, and a roster approach is the recommended tool). The type of tool will affect the amount of time required to administer and the level of skills required by the enumerators.

As with other modules, the simplest approach is presented as standard base for CFSVA surveys. It can provide a limited but sufficient amount of information. As more information is required, more questions can be added, but the module's structure then becomes more complex, as with the individual roster.

The basic standard: Household Summary module

The following questions are required to capture the household composition and child enrolment:

1. Sex of the head of the household (Male = 1 Female = 2)
2. Age of the head of the household (|__|__| years)
3. Marital status of the head of the household (1 = Married, 2 = Cohabitating, 3 = Divorced/Separated, 4 = Widow/Widower, 5 = Never married)[58]
4. Can the head of the household and the spouse of the head of the household read/write a simple message? (yes/no)
5. Age pyramid: Create a three-column table as here:

	Age	Male	Female
Please complete this household's demographics table on the right. This is to record the number of individuals in each age category. Make sure to differentiate between males and females.	**a.** 0–5 years **b.** 6–14 years **c.** 15–59 years **d.** 60 years or older	|__|__| |__|__| |__|__| |__|__|	|__|__| |__|__| |__|__| |__|__|

These are the basic categories used to calculate the dependency ratio (see "Modifications to the module" for further discussion on adaptation to the local context). There should be no blank cells, if there are no individuals in the age cohort and sex, then a zero must be used.

6. Attendance rate and absenteeism from school: the basic information is collected for primary school children only.

58. Issues of polygamy need to be considered.

Table 4.7: Example of household questionnaire for primary school attendance

		Male		Female	
1.10	Number of children attending primary school (6–11 years)?	|__|__|		|__|__|	
1.11	Did anyone miss school for at least 1 month in the last year?	1 Yes		2 No → go to 1.13	
1.12	If yes, why? **CIRCLE THE MOST IMPORTANT REASON**	**Male Children**		**Female Children**	
		1 2 3 4 5 6 7 8	Sickness Work for money or food Domestic work (gardening, fetching water) Taking care of siblings School is far away/located in insecure area No money for school fees/ school costs Refused to go Other (Specify:________________)	1 2 3 4 5 6 7 8	Sickness Work for money or food Domestic work (gardening, fetching water) Taking care of siblings School is far away/located in insecure area No money for school fees/ school costs Refused to go Other (Specify:____________)
1.13	If there are boys, girls, or both who do not attend school, what is the main reason? **CIRCLE THE MOST IMPORTANT REASON**	**Male Children**		**Female Children**	
		1 2 3 4 5 6 7 8	Sickness Work for money or food Domestic work (gardening, fetching water) Taking care of siblings School is far away/located in insecure area No money for school fees/ school costs Refused to go Other (Specify:________________)	1 2 3 4 5 6 7 8	Sickness Work for money or food Domestic work (gardening, fetching water) Taking care of siblings School is far away/located in insecure area No money for school fees/ school costs Refused to go Other (Specify:____________)

The main indicators created through the household roster/summary are:
- Average household size;
- Percentage of male- and female-headed households;
- Average age of head of household (aggregated by sex);
- Marital status;
- Age pyramid;
- Literacy rates of household heads and spouses;
- Attendance rates (check definition) and causes for not attending school;
- Absenteeism and causes; and
- Percentage of dependents/dependency ratio.

Modifications to the module

Household Summary
- Total years or completed level of education of household head and spouse.
- The generic child age cohort can vary depending on the primary school age category in the country. For example, in the 2006 Lao PDR CFSVA, the generic school age cohort was split in two (6 to 11 and 12 to 14 years) in order to reflect the school system of the country and the types of information WFP needed for its school feeding programmes.

- Number of days of absenteeism from school in a certain reference time (the two parameters can vary according to local agreement).
- Child labour and adult labour to household economic unit.
- Chronic illness.
- Disabilities.

The decision to move from household summary questions on demography and education to a household roster type of table, where each question is asked for each individual in the household, is usually driven by the need to obtain more articulated information about education, labour, chronic illness, and disabilities for both adults and children. The household roster provides information and statistics at the individual level. This adds valuable information but also difficulties: the administration of the questionnaire is time-consuming, and this method adds data management and analytical difficulties. Again, the choice of having household- or individual-level indicators must be driven by programme need, analytical capacity, and the intended use of this more detailed information.

Household Composition Roster (example in Table 4.8)
Create a table with the following columns, from left to right:

1. Household member code: unique number that must be consistently employed when the table spreads over several pages (e.g. the same individual in the households has the same household member code).
2. Name of the individual (this is usually not entered in the database but is used during the interview).
3. Gender of the individual.
4. Relationship of the individual to the head of the household (head of the household, spouse, child, orphan, uncle, other).
5. Age of the individual in years (never record the months of a child under the age of 1; use a zero). Where age is not known, an event calendar or other estimation tool should be provided.

Adult Cohort

6. Can the household member read/write a simple message (use DHS/MICS definition)?
7. What level of education does the individual have (this should follow the formal schooling system of the country)?

Child Cohort (age groups will depend on the country)

8. Does the child go to school? (0 = Does not go the school, 1 = primary, 2 = secondary, 3 = university; – this is only for school-age children).
9. If not, why?
10. Did the child miss school for at least five days or more in the past month?
11. If yes, why?

Table 4.8: Example of household composition roster

	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10
Household Member code	**First Name**	**Gender**	**Relationship to head**	**Age in years**	**Marital Status**	**Current level of education**	**Schooling status of children 6-14**	*If attending school*	
								Did [name] miss school for at least 1 month in the last year?	What was the reason for missing?
	Do not record full name, only a first name to refer to the household member	1 = Male 2 = Female	1 = Head 2 = Spouse 3 = Child 4 = Parent 5 = Sibling 6 = Grandchild 7 = Grandparent 8 = Orphan being taken care of 9 = Other relative 10 = No relation	For children < 12 months, write 0	1 = Married 2 = Cohabitating (not married) 3 = Divorced 4 = Living apart, not divorced 5 = Widow or widower 6 = Not married	1 = No schooling 2 = Some primary 3 = Completed primary 4 = Some secondary 5 = Completed secondary 6 = Vocational 7 = Some university 8 = Completed university 9 = N/A	1 = Attends primary 2 = Attends secondary 3 = Not attending school	1 = Yes 2 = No	1 = Sickness 2 = Work 3 = Household Work 4 = Taking care of siblings 5 = Long distance to school 6 = School fees not paid 7 = Insecurity 8 = Refused to go
01									
02									
03									
...									
...									
14									
15									

Additional modifications to the Household Composition Roster

- For mortality/morbidity refer to WFP nutrition guidelines. This addition should be made only if required and strongly supported and properly undertaken.
- Child labour and adult labour to household economic unit.
- Chronic illness.
- OVC issues (see HIV/AIDS guidelines).
- Disabilities.

Links with other modules

- Links with the maternal and child modules (same number of under-5s and women listed as in the modules; carry over the codes.
- The age cohorts need to be identical to those used for schooling of children.

Sources of inspiration

- Information on household demographics is typically collected by DHS, MICS, and LSMS questionnaires through a household roster (individual-level information). CFSVAs use either individual level rosters or household-level questions.
- When a roster is included in the CFSVAs, it usually covers the same areas addressed by MICS, DHS and LSMS, including sex, age, position of the HH members, health conditions, educational level, school enrolment/attendance (for children), questions related to OVC. In general, questions on health status and schooling are more detailed in MICS, DHS, and LSMS; in particular, MICS (used by UNICEF to assist countries in filling data gaps for monitoring the situation of children) systematically collects information on OVC and child labour.

Module Title: Housing Materials

Main purpose of the module: The construction materials used in a household are very basic indicators of living standards. They provide analysts with information on a household's standard of living that goes beyond consumption expenditures. Usually materials for floor, roof and walls are recorded, as per example in Table 4.9.

Limitations of the module

- Housing materials provide only an indirect measurement of wealth.
- Certain "luxury" materials might not have been available in the local context, thus preventing the use of these indicators to identify wealthy households.

Table 4.9: Creation of Housing Materials module

3.5	*What is the major construction material of the exterior walls?* **IF POSSIBLE, DO NOT ASK; ANSWER BASED ON YOUR OBSERVATIONS**	1 Concrete/burned bricks 2 Mud blocks 3 Mud and straw	4 Wood 5 Plastic shelter 6 Other (Specify: __________)
3.6	*What is the major material of the roof?* **IF POSSIBLE, DO NOT ASK; ANSWER BASED ON YOUR OBSERVATIONS**	1 Concrete 2 Tiles 3 Straw (grass, papyrus, banana fibres)	4 Wood 5 Plastic shelter 6 Galvanized iron 7 Other (Specify: ___________)
3.7	*What is the major material of the floor?* **IF POSSIBLE, DO NOT ASK; ANSWER BASED ON YOUR OBSERVATIONS**	1 Concrete 2 Mud 3 Straw 4 Wood	5 Plastic sheeting 6 Tiles 7 Other (Specify: ___________)

Modifications to the module

Additional elements can be included in the Housing Materials module other than construction materials:

- The type of dwelling (single family house, separate apartment, mud house, shelter, other).
- The number of rooms.
- Availability of electricity.

Links with other modules

- Water and sanitation facilities and assets to construct the wealth index.
- Expenditure to cross-check wealth status.

Sources of inspiration

- Information about the main material of the dwelling, floor, roof, and walls is collected in the MICS, DHS, and LSMS questionnaires. As with the DHS, the CFSVA takes into consideration this information while selecting the variables for the wealth index.

Module Title: Access to Credit

Type of questionnaire: The following presents the Access to Credit module as collected in a quantitative/household survey. However, access to credit can be deeply explored in focus group discussions or in community interviews. This would save time and would make the household-level module optional. More information on qualitative tools can be found in Section 5, "Qualitative and Community-level Data in CFSVA."

Main purpose of the module: The module provides for an estimation of household's access to sources of credit and their actual use of credit.

Limitations of the module

- The module does not aim to estimate the amount actually borrowed.
- There is no information collected with regard to the "interest rates" charged to borrowers or other credit conditions.

Table 4.10: Creation of Access to Credit module

3.5	Do you have **access** to a place to borrow money? **Circle all that apply**	1 Yes – relatives/friends 2 Yes – charities/NGOs 3 Yes – local lender 4 Yes – bank	5 Yes – cooperatives 6 Yes – village head 7 Yes – company/middle men 8 No access to credit
3.6	In the last 3 months, how **often** did you use credit or borrow money to purchase food? **Circle one**	1 = Never 2 = On one occasion 3 = On two occasions	4 = On three occasions 5 = On more than three occasions

Modifications to the module

- The first question can be broken into two: Do you have access to credit? If yes, where?
- Additional questions can be added to explore the average amount of debt in addition to the issue of access to credit sources. Experience from past surveys indicates that piece of information might not be fully reliable, perhaps because of people's reluctance to declare their financial status.

Links with other modules

- Expenditure module: if the household manages to pay back their debt.
- Information about access to credit facilities can be gathered in focus groups or community interviews.

Source of inspiration

- Questions on credit usually are not included in DHS and MICS questionnaires, but they are frequently inserted in the LSMS questionnaires, which collect this information at the individual level. Source, frequency, and time needed for reimbursement are also addressed by the LSMS module.

Module Title: Livelihoods and Economic Activities

Main purpose of the module: This module attempts to identify the activities and combinations of activities that sustain households, and their relative importance to a household's income strategy.

Limitations of the module

- This is not a comprehensive livelihood analysis, which includes but is not limited to economic activities. Its main goal is to identify and group households based on a common set of economic activities and their relative importance for risk analysis.
- If absolute values are collected from the economic activities, the sum of those values should not be considered as an income level for the household. This derived income is not intended for poverty analysis.

Creation of the Livelihoods and Economic Activities module

- Prepare a list of economic activities that households would undertake or the main income sources a household would exploit to earn cash or acquire food or services. The list of activities should be based on secondary data and local expert knowledge. It is important to include atypical sources that vulnerable households would exploit to sustain themselves.
- If another CFSVA and/or other EFSAs have been conducted previously, review the activities listed and include (1) those reported in the previous study, and (2) through review, those that might have been excluded. The aim is to differentiate households and minimize the reporting of undefined "other" activities, which are difficult to interpret and could confound results.
- Include a column where households are asked, using proportional piling, to estimate the relative importance of the activities to contributing to the household's income, food, and access to services.
- The module is not meant to be exhaustive in identifying all the activities undertaken by each household. Instead, it is critical to identify the three or four essential activities.
- It is likely that the three or four activities cover almost all income sources of the household. The sum of the three to four contributions should equal 100 percent.[59]

59. When there are more than three to four activities, it must be made clear to the enumerator that the proportions reported are valid only for the identified activities.

- The categories should not be duplicated. For example, if men undertake one type of agricultural activity and women undertake another type, the two activities should be grouped, as the level of analysis is the household.
- The main indicators emanating from this module are: (a) main economic livelihood activities; and (b) percentage contribution of main economic/livelihood activities to household income.

The *minimum* information required can be obtained through one of the following tables:

Table 4.11a: Recommended layout of economic livelihoods table

Activity(ies) undertaken to earn cash or acquire food or services	(√)	Using proportional piling or "divide the pie" methods, estimate the relative contribution to total income of each activity (%)
1 = Agriculture and sales of crops	I__I	I__II__II__I
2 = Livestock and sales of animals	I__I	I__II__II__I
3 = Brewing	I__I	I__II__II__I
4 = Fishing	I__I	I__II__II__I
5 = Unskilled wage labour	I__I	I__II__II__I
6 = Skilled labour	I__I	I__II__II__I
7 = Handicrafts/artisanal work	I__I	I__II__II__I
8 = Use of natural resources (firewood, charcoal, bricks, grass, wild foods, honey, etc.)	I__I	I__II__II__I
9 = Petty trading	I__I	I__II__II__I
10 = Seller, commercial activity	I__I	I__II__II__I
11 = Remittances	I__I	I__II__II__I
12 = Salaries, wages (employees)	I__I	I__II__II__I
13 = Begging, assistance	I__I	I__II__II__I
14 = Government allowance (pension, disability benefit)	I__I	I__II__II__I
15 = Others (Specify:____________)	I__I	I__II__II__I
TOTAL	**1 0 0**	

Table 4.11b: Alternative layout of economic livelihoods table

Activity	**a.** List, in order of importance, household income activities? (Use activity code from a list like the one in Table 4.11a)	**b.** Using proportional piling or "divide the pie" methods, estimate the relative contribution to total income of each activity (%)
Main	I__I	I__II__II__I
Second	I__I	I__II__II__I
Third	I__I	I__II__II__I
Fourth	I__I	I__II__II__I

CHAPTER 4

The advantage of Table 4.11a is that households can list as many activities as they want. Additionally, the output variables obtained from such a table during the data entry process are ready for analysis (see Annex 16). However, some data management has to be done to come up with the main activities at the household level (percentage of household undertaking agriculture, trading, etc.).

The second example presents the question in an easier way, both for the interviewer and interviewee. The interviewer asks what the main (or second, etc.) household's activity is and the interviewee is free to recall without having to reply yes/no to a long list of activities. If data are collected this way, it is easier to calculate percentages of households undertaking determinate activities as their main one. On the other hand, calculating "Contribution from the different livelihood activities" requires more data management skills.

The preference for one module option over the other should depend on the main indicator(s) needed and on the available analytical capacities.

Modifications to the module

- The table can also be expanded to include information on who undertakes or is the key actor in the activity (see above Tables 4.11a and 4.11b).
- The recall period for the combinations of the activities is typically one year. However, depending on the context (e.g., following a rapid EFSA where the period of time it takes household to adapt, and how they adapt, are relevant), the recall period can be reduced.
- The seasonality of activities can be included to identify when key economic activities are undertaken. This can also be done in a community questionnaire, community focus group, or key informant interview.
- Instead of getting relative contributions (percentage), the absolute cash value of the activity can be captured by recording either (1) the estimated value provided by the household; or (2) the provided value within a series of categorical variables. Even though value ranges are commonly more easily collected, categorical variables present more limitations during the analysis phase. On the other hand, the feasibility of collecting truly reliable absolute cash values has to be explored and tested.
- The respondent could also be asked to estimate the percentage of results/goods from each activity that is directly consumed by the household. This question is used to estimate the relative importance of self-production that is directly consumed and is not captured by expenditure indicators. However, this concept has been reported to be difficult to explain both to enumerators and to interviewees; and the analysis is quite complicated and is based on the assumption that a household's total income can be measured through total cash expenditures plus the value of own-produced and consumed goods.

An example of possible modification of Table 4.11b is shown in Table 4.12.

Table 4.12: Modified Livelihoods/Economic Activities module

Using the "Income activity" and "Participant" codes provided below, complete the following table, completely filling in the information for one activity at a time (to earn cash or acquire food or services)

		a. What is your household's [rank] income activity? (Use "income activity" codes)	**a.** What is your household's [rank] income activity? (Use "income activity" codes)	**c.** Using proportional piling or "divide the pie" methods, estimate the relative contribution to total income of each activity (%)
1.	**Main**	I__I__I	I__I	I__II__II__I %
2.	Second	I__I__I	I__I	I__II__II__I %
3.	Third	I__I__I	I__I	I__II__II__I %
4.	Fourth	I__I__I	I__I	I__II__II__I %

Income activity codes

1 = Agriculture and sales of crops
2 = Livestock and sales of animals
3 = Brewing
4 = Fishing
5 = Unskilled wage labour
6 = Skilled labour
7 = Handicrafts/artisanal work
8 = Use of natural resources (firewood, charcoal, bricks, grass, wild foods, honey, etc.)
9 = Petty trading
10 = Seller, commercial activity
11 = Remittances
12 = Salaries, wages (employees)
13 = Porter
14 = Begging, assistance
15 = Government allowance (pension, disability benefit)
16 = Others (Specify:________________)

Participant codes

1 = Head of the household only
2 = Spouse of the head of the household only
3 = Men only
4 = Women only
5 = Adults only
6 = Children only
7 = Women and children
8 = Men and children
9 = Everybody

CHAPTER 4

Links with other modules

- The income activities (e.g. agriculture and livestock) should agree with the households' responses regarding agriculture and pastoral activities.
- Link to expenditure and credit (the total of cash income and credit should correspond with total cash expenditures, which should be verified during data collection and analysis).

Module Title: Agriculture

Main purpose of the module: This module aims to gather information on the practice of agriculture at the household level. In most developing countries, not only is agriculture one of the main income-generating activities of a household, but the majority of the population also practice it. Furthermore, many households, especially in rural areas, produce at least part of the food they consume, through agriculture or home gardening.

The CFSVA should also identify net food sellers, especially in prime agricultural areas, to help programmes aimed at developing local agriculture through local purchases.

Limitations of the module

- Even though food security surveys almost always collect agriculture information, they are different from an agriculture extension survey or a crop and food supply assessment mission (CFSAM).

- In a food security survey, the aim is not to precisely measure the size of cultivated land and yields, but rather to cross-tabulate agriculture-related data with other socio-economic characteristics such as family size, data on income, expenditure, and consumption, for the purpose of presenting a better picture of the livelihood of rural households and to identify possible factors of food insecurity and inform response options later on.

Creation of the Agriculture module

Information obtained from a basic module on agriculture includes:

- Percentage of households having access to land
- Most common types/methods of land access
- Percentage of households cultivating land
- Common crops cultivated
- Prevalence of kitchen gardens

Modifications to the module

- The nature and scope of the study will determine the level of information sought from the module. Hence, the above list is not meant to be comprehensive.
- Additional questions for the key staples: duration of production for own consumption, share of the production sold, share of the production consumed.
- Additional questions for consideration: size of land, major crops grown, source of seeds, use of agrochemicals, source and extent of irrigation and average yields per harvest. Of course, the more information collected the more complex the module becomes for both the enumerator and for the interviewee. Carefully evaluate the needs of the survey in order to avoid overloading the questionnaire with questions that will not be analysed.
- Questions regarding agriculture can be very specific and detailed, or more general covering the agriculture sector as a whole.

Links with other modules

- Livelihood and economic activities
- Productive assets
- Livestock
- Sources of consumed food
- Expenditure

Module Title: Livestock

Main purpose of the module: to gather information on livestock ownership. Livestock can be seen as assets or as main livelihood activities for pastoralists and nomads, but also involves specific vulnerabilities.

Limitation of the module

- Although food security surveys collect livestock ownership information, they are different from livestock surveys.

Creation of the Livestock module

Usually a filter question opens the module. The list of commonly owned animals and wealth status livestock follows.

4.20 -	Does your household own any farm animal(s)?		1	Yes	2	No → next Section
4.21 -	If yes, how many of each of the following animals do you own? (write 00 if none)					

a	Chicken	|__|__|	g	Pigs	|__|__|
b	Ducks	|__|__|	h	Bulls	|__|__|
c	Other birds	|__|__|	k	Cows	|__|__|
d	Rabbits	|__|__|	l	Oxen	|__|__|
e	Goats	|__|__|	m	Donkey	|__|__|
f	Sheep	|__|__|	n	Camels	|__|__|

Modifications to the module

- The filter question can be removed. When the household does not own any livestock, it is important to enter a zero (0).
- The data collection can be simplified by recording single-species ownership only as a categorical bivariate (yes/no).
- However, in appropriate countries, collecting the number of animals allows for computing the synthetic indicator "Livestock Index" through the use of Livestock Tropical Unit values.
- Extra information on amount of fodder needed, average fodder price, and monthly expenditure on fodder can be useful.

Links with other modules

- Livelihood and economic activities
- Agriculture
- Sources of consumed food
- Expenditure

Module Title: External Assistance (food and non-food)

Main purpose of the module: Any external assistance going on in the surveyed area or in the country should be recorded and taken into account when evaluating households' self-reliance and food security.

Limitation of the module

- The module is not designed to record tonnage or quantities of aid delivered/received by each household.

Creation of the External Assistance module

Usually the module is introduced with a filter question. After that, there is the list of food aid programmes, the organizations that provide non-food assistance in the area, and the types of assistance received by the household. The two lists must be customized according to the local context.

Table 4.13: Example of External Assistance module

Has any member of your household received food aid in the last 6 months?	1	Yes	2	No
If yes, please specify the type of programme and the number of beneficiaries in your household? Circle all that apply and specify number of beneficiaries in the last column.	1 2 3 4	School feeding Food for work/food for assets Supplementary feeding Other (Specify: __________)		|__|__| |__|__| |__|__| |__|__|
Has any member of your household received any other type of external assistance besides food aid in the last 6 months?	1	Yes	2	No
If yes, from whom? Circle all that apply	1 2 3 4 5 6 7 8	World Food Programme SAPPROSC/DEPROSC Save the Children UNICEF GT2/SNV/DFID French Cooperation The Government Other (Specify: ____________)		
If yes, what type of assistance? Circle all that apply	1 2 3 4 5 6 7	Food products Money allowances/loans For education (fees, books, uniforms) For medical services Construction material, building Agricultural assistance (tools/seeds) Other (Specify:_______________)		

Modification to the module

- The list of assistance programmes and organizations has to be context-specific.

Link with other modules

- Livelihood activities – since assistance programmes sometimes focus on certain livelihood activities, there should be a correspondence; for instance, if the household benefits from agricultural assistance, we could expect crop production to be mentioned as a livelihood activity.

Source of inspiration

- Questions on external assistance are usually not included in DHS and MICS questionnaires, but can be found in LSMS questionnaires.

Module Title: Shocks and Coping

Main purpose of the module: This module aims to identify shocks that, in the recent past, have affected the household's ability to acquire food or cash perceived as important by each household, and the types of coping mechanisms used. This will determine which households are prone to be affected by shocks and which have poor coping capacity.

Limitations of the module

- Problem with the definition of shock: A shock is an abnormal event affecting a household's economic status and capacity to feed themselves. Sometimes events are reported as shocks that do not have real consequences on a household's status because they are (or should be considered) normal events (e.g. lack of rain in a desert area). Information on shocks can be more appropriately collected through

secondary data. The problem of shock definition should be carefully considered during questionnaire design and enumerator training.
- Shocks and coping strategies depend on household perception.
- Households may not be able to attribute their coping mechanisms or the consequences of a shock to a particular event.
- Neither intensity of shock nor coping mechanism is measured.
- This is not equivalent to the Coping Strategy Index. If this is a desired output, it should be included as a separate section.
- If an initial filter question is used (e.g. Have you experienced a shock in the last 12 months?) as a skip question, this may lead to unwarranted non-response by households, as some may not understand what is meant by the term shock (i.e. they might not consider specific shocks, such as fire, drought, or war).
- If no initial filter question is used, and each shock is enumerated (i.e. Have you experienced drought? Have you experienced floods?), this may increase the number of responses, as each may be a leading question.

Creation of the Shocks and Coping Strategies module
- The focus of this module is on **shocks that affect the economy of the household or the ability to acquire (produce/purchase) food.**
- The **minimum structure** of this module should be a list of shocks experienced and coping mechanisms used.
- The recommended standard recall period is one year. There are exceptions to this; see "Modifications to the module."
- The module should reflect:
  - Covariate shocks experienced during the period of recall (e.g. one year), including economic and environmental shocks;
  - Idiosyncratic shocks that likely affected the household;
  - Coping mechanisms commonly employed by households within the context, especially those used by known vulnerable households.
    - Coping mechanisms should reflect food and non-food responses (i.e. the list should come from the secondary data review or previous surveys/studies).

Modifications to the module
- Once the shock and coping strategy lists are made, the following options can be added (there is no current agreed-upon standard):
  - A filter question at the beginning of the section (Have you experienced a shock?) that will determine the enumeration of the rest of the section (if no shock, go to the next module);
  - When enumerating the shocks experienced, a question about each individual shock, or a general question, allowing the household to list all shocks;
  - Based on the suggested focus for this module, a filter question to ensure the HH understands they are responding with a list of shocks that have affected their economic status or their ability to acquire food;
  - Once the section on shock identification is complete, coping mechanisms used during the same period, ideally from a pre-coded list (i.e. the enumerator lists all the coping mechanisms used during the recall period).
- Rank the importance of a given shock for that household.
- Record the seasonality of the shock or coping mechanisms. When during the

previous recall period did the household experience this shock or use the coping mechanisms? (Note that this question has limited field experience but can have interesting results.)
- Link each specific shock (two to four main shocks) to a coping mechanism. However, feedback suggests that unless the problem is major – for example, an earthquake – households often struggle to link one unique coping strategy to one unique shock.
- Instead of 12 months, the recall period, if in an FSMS, may be back to the previous round of data collection. In an EFSA, it may be back to the period of the last main covariate shock.
- Relate specific shocks to their impact on assets (loss of, recovery of), usually in conjunction with linking specific shocks to coping strategies.
- Relate specific shocks to their impact on income (reduction of, return to the same level of).
- Relate specific shocks to the ability to feed the household.
- Relate specific shocks to current recovery status (not recovered, partially recovered, totally recovered).

Sources of inspiration

Questions on shocks and coping strategies can be found in the LSMS questionnaires. Such questions are quite similar to those included in the CFSVA module, even though they draw attention to the impact on household welfare, whereas the CFSVA module looks more specifically at the impact on food consumption.

Table 4.14 presents a typical set of questions and options for the Shocks and Coping Strategies module. However, note that analysing the number of indicators presented is difficult, as most of the data obtained from such a table are not significant due to the multiple combinations of answers that result in a small number of cases. Hence, it is suggested to group the indicators based on type. For example, “Reduced number of meals,” “Reduced proportions of the meals,” “Rely on less preferred, less expensive food,” “Reduced expenditures on health and education” can be grouped as “adjustment strategies,” while all borrowing can be combined as “borrowing strategies.” Similarly, all selling of assets could be combined as “divestment strategy,” and all instances of migration could be combined as “migration strategy.” The analyst can create other categories based on the importance of a particular coping strategy in a particular country, and on the number of households that adopted the strategy.

Links with other modules

- This section should not be combined with the Coping Strategies Index (CSI), nor is it a substitute for a CSI (a CSI is done in addition to this and has its own specific methodology).
- To the qualitative data, if collected.

Tips

- If using the initial skipping rule, be sure the enumerators do not lead respondents on this question in order to skip the entire section.
- It is imperative that a relevant list of shocks and food and non-food coping strategies is created.
- When linking shocks and coping strategies, enumerator training and clarification of questions is important.

Table 4.14: Example of Shocks and Coping Strategies module

Has your household experienced any of these shocks that have made it difficult to obtain sufficient means of livelihood in the last 12 months? *If yes, please rank the shocks and report the three most serious.* *If no shock affected the household, go to question 9.3.*		**1.** Rank three shocks 1 = main 2 = second 3 = third	**2.** Did [shock] decrease your household's ability to produce or purchase food? 1 = Yes 2 = No 3 = Don't know
a	Drought/irregular rains	|__|	|__|
b	Regular floods	|__|	|__|
c	Flash floods	|__|	|__|
d	Landslides, erosion	|__|	|__|
e	Severely high level of crop pests and disease	|__|	|__|
f	Severely high level of livestock diseases	|__|	|__|
g	Lack or loss of employment	|__|	|__|
h	Unusually high level of human disease	|__|	|__|
i	Fire	|__|	|__|
j	High costs of agricultural inputs (seed, fertilizer, etc.)	|__|	|__|
l	Earthquake	|__|	|__|
m	Reduced income of a household member	|__|	|__|
n	Serious illness or accident of household member	|__|	|__|
o	Death of a working household member	|__|	|__|
p	Death of other household member	|__|	|__|
q	Theft of money/valuables	|__|	|__|
r	Theft of animals	|__|	|__|
s	Conflict	|__|	|__|
t	Other (Specify:________________________)	|__|	|__|
3.	What did the household do to compensate for this loss of income and/or assets? (Report the three most important coping strategies. Choose from the strategy codes below).	|__|__| 1st |__|__| 2nd |__|__| 3rd	

Coping strategies codes:

01 Rely on less preferred, less expensive food
02 Borrowed food, helped by relatives
03 Purchased food on credit
04 Consumed seed stock held for next season
05 Reduced the proportions of the meals
06 Reduced number of meals per day
07 Skipped days without eating
08 Some HH members migrated
09 Sold durable household goods
10 Sent children to live with relatives
11 Reduced expenditures on health and education
12 Spent savings
13 Gathering
14 Sold or consumed livestock
15 Sold agricultural tools, seeds, or other inputs
16 Worked for food only
17 Sold crop before harvest
18 Rented out land
19 Sold land
20 Borrowed money
21 Other (Specify: _______________)

Module Title: Coping Strategies Index

Main purpose of the module: The CSI is a relatively simple and easy-to-use indicator of household food security; it is straightforward and correlates well with more complex measures of food security. A series of questions about how households manage to cope with a shortfall in food for consumption results in a simple numeric score. In its simplest form, monitoring changes in the CSI score indicates whether a household's food security status is declining or improving.

CHAPTER 4

Limitations of the module

See the limitations discussed under the Shocks and Coping Strategies module.

Creation of the Coping Strategies Index module

The question to ask is **"What do you do when you do not have enough food, and do not have enough money to buy food?"** The answers to this question are the basis for the CSI module.

- The minimum structure of this module should be a list of strategies used to cope with the food shortage or when households do not have enough money to buy food.
- The recall period for CSI is recommended to be the past seven days.

One category should be "daily" or "all the time," and one category should be "never." The intermediate categories can be changed around according to conditions and the amount of detail required. In general, at least five relative frequency categories are recommended, as shown in Table 4.15.

When using the CSI, the question at the top should be repeated for each of the strategies on the list, and the appropriate relative frequency box should be ticked.

The best way to assess the frequency of coping strategies is not to count the number of times a household has used them, but to ask a household respondent for a rough indication of the relative frequency of their use over the previous month. Precise recall is often difficult over a long period of time, and asking for the relative frequency provides adequate information. There are various ways a relative frequency count can work – this one asks roughly what proportion of the days of a week people had to rely on various strategies.

Modification to the module

- Although a generic list of strategies is presented in Table 4.15, list only those strategies applicable to the area. There is no point in asking about the strategies not adopted in the area (for instance, in a non-agricultural population, we do not need to ask about "consuming the seed stock").

Table 4.15: Example of questions for constructing a CSI

In the past 30 days, if there have been times when you did not have enough food or money to buy food, how often has your household had to:	**Never**	**Seldom (< 1 day a week)**	**Sometimes (1–2 days a week)**	**Often (3 or more days a week)**	**Daily**
a. Reduce number of meals eaten per day?	1	2	3	4	5
c. Borrow food or rely on help from friends or relatives?	1	2	3	4	5
d. Rely on less expensive or less preferred foods?	1	2	3	4	5
e. Purchase/borrow food on credit?	1	2	3	4	5
f. Gather unusual types or amounts of wild food/hunt?	1	2	3	4	5
g. Have household members eat at relatives' or neighbours'?	1	2	3	4	5
h. Reduce adult consumption so children can eat?	1	2	3	4	5
i. Rely on casual labour for food?	1	2	3	4	5
j. Feed working members of HH at the expense of non-working members?	1	2	3	4	5
k. Go entire day without eating	1	2	3	4	5
l. Consume seed stock to be saved for next season	1	2	3	4	5

The context-specific CSI "has been criticized for being relatively unhelpful in comparative analysis. However, fieldworkers have noted that several of the individual behaviours that the CSI measured recur across different contexts"[60]. Recognizing this, a reduced CSI was developed to compare food security across different contexts. It is based on the same short list of (5) coping strategies (see figure below) and the same severity weights.

The reduced CSI is less valuable in identifying the most vulnerable HHs in a given location, but it is very useful for comparison across countries as it focuses on the same set of behaviours. Furthermore, "extensive research demonstrated that the 'reduced' CSI reflects food insecurity nearly as well as the 'full' or context-specific CSI".[61]

The figure below describes how to calculate a reduced household CSI score:

In the past 7 days, if there have been times when you did not have enough food or money to buy food, how often has your household had to:	**Raw Score**	**Universal Severity Weight**	**Weighted Score = Frequency X weight**
Relative Frequency Score			
a. Rely on less preferred and less expensive foods?	5	1	5
b. Borrow food, or rely on help from a friend or relative?	2	2	4
c. Limit portion size at mealtimes?	7	1	7
d. Restrict consumption by adults in order for small children to eat?	2	3	6
e. Reduce number of meals eaten in a day?	5	1	5
TOTAL HOUSEHOLD SCORE – Reduced CSI	Sum down the totals for each individual strategy		**27**

In order to conduct the analysis of the CSI, however, you need a few more pieces of information. The first is a way to "weight" the individual strategies or behaviours.

The CSI tool relies on counting coping strategies that are not equal in severity. Different strategies are therefore "weighted" – multiplied by a weight that reflects their severity before being added together. The simplest procedure for doing this is to group the strategies according to similar levels of severity and assign a weight to each group.
The severity of coping strategies is, to some extent, a matter of perception.

Focus group discussions with different community groups are needed to determine the severity of the coping mechanisms. The first step is to try to group the strategies into categories of roughly the same level of severity. Since this task is carried out with different groups, it is useful to impose some structure from the outset. For example, one could divide them into four different categories: very severe, severe, moderate, and not severe.

- It is always easiest to determine the most severe coping strategies, so ask the group to select the most severe and least severe individual strategies first.
- Then ask if there are other individual strategies that are more or less the equivalent of these two, in terms of how severe they are perceived to be. Once the two extremes are established, it is easier to group the remaining behaviours into intermediate categories.

60. Source: "Coping Strategy Index: Field Methods Manual" II edition (2008)
61. ibid

- This must be done with enough groups to represent the diversity of the location or culture, to ensure that a reasonable consensus has emerged.
- Incorrectly weighting individual strategies will result in errors in the analysis.

Table 4.16: Example of coping strategies grouped and ranked by focus group

Strategy	FG1	FG2	FG3	FG4	FG5	FG6	FG7	FG8	Avg	Rank
a. Limit portion size	1	1	1	1	1	1	1	1	1	1
b. Reduce number of meals	1	2	1	1	3	1	1	1	1.4	1
c. Borrow food	2	2	3	3	2	2	2	3	2.4	2
d. Rely on less preferred/expensive foods	1	1	1	1	1	1	1	1	1	1
e. Purchase/borrow food	2	2	2	3	2	2	2	2	2.1	2
f. Gather unusual types	5	5	4	4	3	5	5	5	4.5	5
g. Eat at relatives' or neighbours'?	2	n/a	n/a	2	2	3	2	n/a	2.2	2
h. Reduce adult consumption	2	3	2	2	3	3	2	2	2.4	2
i. Rely on casual labour	1	1	2	2	1	-	2	2	1.6	2
j. Feed working members	3	3	3	4	3	3	5	3	3.4	3
k. Go entire day without eating	5	4	5	5	5	4	5	5	4.8	5
l. Consume seed stalk	4	3	4	3	3	3	4	4	3.5	4

Source of inspiration

- The CSI is consistent with the CARE/WFP methodology.

4.2.4.3 Non-core modules with non-changeable questions

Module Title: Maternal Health and Nutrition

Main purpose of the module: The aim of this module is to gather data about the health and nutrition status of mothers with children under 5 years of age. The information can be used to understand the relationship between malnutrition, diseases, and consumption.

Limitation of the module

- Health status is usually self-reported by women and not clinically demonstrated.

Creation of the Maternal Health and Nutrition module

Weight, height, and mid-upper arm circumference (MUAC) are collected to calculate the woman's nutrition status.

Measurements of mother		
11.22	**Mother's height** (in centimetres)	I__I__I__I.I__Icm
11.23	**Mother's weight** (in kilograms)	I__I__I__I.I__Ikg
11.24	**Mother's MUAC** (in centimetres)	I__I__I.I__Icm

Health is commonly assessed through questions about illness or diseases (fever, diarrhoea). However, many additional questions may be included.

Modification to the module

- Additional questions about a mother's health can be added: Is she currently pregnant or breastfeeding? The number of pregnancies, miscarriages, stillbirths? The number of living children? The age of the first delivery? Has she received iron-folate tablets, antenatal care, vitamin A capsules? What are her health and hygiene practices (sleeping under a mosquito net, boiling or treating drinking water, washing hands before preparing meals or after going to the toilet)? Her level of education, occupation, and control over food and income could be included as potentially relevant for the analysis of the results.

Links with other modules

- Child health and nutrition
- Water and sanitation
- Food consumption
- Community-level infrastructure (specifically, health facilities)

Source of inspiration

- MICS and DHS surveys have more extensive questionnaires for women, including information on child mortality, maternal and newborn health, and marriage. Likewise, the CFSVAs, MICS, and DHS collect anthropometric data on women.

Module Title: Child Health and Nutrition

Main purpose of the module: The aim of the module is to gather data about the nutrition status and health conditions of children under 5 years old (usually 6 to 59 months).

Limitations of the module

- Appropriate sampling approaches and sample sizes are required to calculate prevalence rates with a sufficient degree of precision.
- If that is not feasible, data can be collected and used to investigate relationships between nutritional outcome and other food security indicators, but not to provide prevalence rates.
- Collecting height and weight data requires some effort. Data are collected using special equipment that is bulky and troublesome to carry around to each household. Frequently, an individual appropriately trained in anthropometry must be added to each survey field team.

Creation of the Child Health and Nutrition module

The key data to assess child nutrition status are: sex, age, weight and height/length. Additionally, MUAC is often collected. Basic health information is related to illness, particularly to diarrhoea and fever. Table 4.17 offers an example of a simple Child Health and Nutrition module.

Table 4.17: Example of Child Health and Nutrition module

<table>
<tr><td colspan="13">ASK SELECTED RESPONDENT IF THERE ARE CHILDREN OF 6–59 MONTHS OF AGE IN THE HOUSEHOLD. IF NOT, TERMINATE INTERVIEW.
Read aloud: Now I would like to ask you some questions about the children in this household (Continue the interview with the same woman)
We would like you to come with all the children aged 6 to 59 months from your household. We would like to measure and weigh them as part of our assessment.
It is very important that children are measured, so be persuasive. Assist women in transportation, if need be.</td></tr>
<tr><td colspan="13">Starting with the youngest child, and focusing on one child at a time, enter each child's first name and ask for the following information:</td></tr>
<tr><td>1.a</td><td>1.b</td><td>2.</td><td>3.</td><td>4.</td><td>5.</td><td>6.</td><td>7.</td><td>8.</td><td>9.</td><td>10.</td></tr>
<tr><td>First name</td><td>Mother's/ caretaker's ID no.

(link with mother's section, if collected)

8 = missing at interview
9 = dead</td><td>If available, date of birth from the medical card

If no → 10.3

If yes → enter, then → 10.4

Use format dd/mm/yy</td><td>Child's age in months</td><td>Child's sex?

1= Male

2 = Female</td><td>Are you the mother/care-taker of [Name]

1 = yes

2 = no

If no → 10.8</td><td>Has [Name] been ill in the last 2 weeks?

1 = yes
2 = no → 10.8

3 = do not know → 10.8</td><td>Has [NAME] been ill with diarrhoea and/or fever at any time in the past 2 weeks? (Diarrhoea: perceived by mother as 3 or more loose stools per day for 3 days or one large watery stool or blood in stool)
1 = yes
2 = no
3 = do not know

Diarrhoea Fever</td><td>Child's height/ length

(in centimetres, with 1 decimal place)</td><td>Child's weight

Enter weight in kilograms, with one decimal place</td><td>Child's MUAC

(in centimetres, with 1 decimal place)</td></tr>
<tr><td></td><td>|_|</td><td>|_|_|/|_|_|/|_|_|</td><td>|_|_|</td><td>|_|</td><td>|_|</td><td>|_|</td><td>|_| |_|</td><td>|_|_|_|.|_|</td><td>|_|_|.|_|</td><td>|_|_|.|_|</td></tr>
</table>

Modification to the module

- Additional questions about the child's health can be added: size at birth, breastfeeding and weaning history, vaccination, de-worming or other treatment. Also useful would be questions on other diseases (e.g. malaria, measles, acute respiratory infections) and on the child's feeding patterns.

Links with other modules

- Mother health and nutrition
- Water and sanitation
- Food consumption
- Community-level infrastructure (specifically, health facilities)
- Sources of inspiration
- MICS and DHS surveys have more extensive questionnaires for children. Similar to the CFSVAs, MICS and DHS collect anthropometric data on children and information on health status. LSMS questionnaires sometimes include a children's module for anthropometric data.

4.2.4.4 Non-core modules with changeable/flexible questions

Gender-sensitive survey design and implementation

Study preparation

An in-depth literature review can be used to identify factors that shape gender relations, such as cultural beliefs, values and practices, religion, education, politics, legislation, economic situation and demographic factors. Generating this type of overview prior to primary data collection provides a context for tailoring generic data collection tools to ensure that they are gender-sensitive and appropriate for a particular setting.

Selection, composition, and training of survey teams

Although members of field teams do not need to have a technical background in gender analysis per se, it is crucial that enumerators are sensitized to the importance and rationale behind collecting sex-disaggregated data and phrasing questions in a way that allows for an analysis of the relationship between gender, food security, and vulnerability. This is even more important for facilitators applying qualitative tools such as focus group discussions and participatory rural appraisal techniques. A balanced mixture of male and female enumerators will minimize the extent to which bias is introduced due to enumerator gender. Where group discussions are to be held separately for men and women, same-sex discussion facilitators are likely to contribute to a more relaxed and open discussion.

Study design (household surveys)

In each questionnaire, sex, age, and relationship to the household head of the main respondent should be indicated to determine possible biases introduced during the data collection process. This will also assist with the identification of different perceptions of men, women, and age groups during analysis.

To the extent possible, all questions concerning food security and vulnerability included in household surveys should be designed in such a way as to differentiate between the experiences of women and men (girls and boys). Please refer to Box 4.6 for examples of key questions that should be disaggregated by gender. Gender-disaggregated data provides valuable information about those intra-household differences that can be masked by surveys that treat households as a single, homogenous unit. Quantitative indicators produced by household survey data can be used to measure the degree of gender inequalities related to food security and vulnerability.

Community discussions or interviews

The qualitative data generated through discussions or interviews with community members provides key insights for understanding the underlying causes and reasons for inequalities identified during household surveys, and allows for further elaboration of the causal mechanisms suggested by quantitative data.

Interviews with key informants (i.e. local authorities)

It is important to include knowledgeable women in the list of persons to be used as key informants. Women's organizations or women's affairs offices often provide suitable candidates. Discrepancies between authorities' perceptions and household- and community-level realities enable an assessment of whether key decision-makers are aware of gender-related differences and inequalities.

Timing of data collection

Appropriate timing is crucial for ensuring that women and men are able to participate in all data collection exercises. Although communities are busy throughout the year, there may be periods when their workload is slightly less burdensome. Similarly, the availability of community and household members is influenced by the daily pattern of agricultural work, and the income-generating and household activities of men and women. For example, women may not be able to attend meetings during evening hours due to domestic responsibilities.

Box 4.6: Key questions for use in conducting gender analysis (WFP 2006)

Household Roles/Social and Cultural Constraints
- What are the different needs, roles, and interests of women and men?
- What are the power dynamics between women and men?
- Which decisions are made by men and which by women?
- What are the social and cultural constraints and opportunities of women and men?
- What are the relations between women and men in society, the community, and the household?
- What different coping mechanisms are available to women and men to lessen the risk of food insecurity for their families?
- How do access to and control of resources, information, and services affect participation by women and men in the programme/project?
- How do gender roles (e.g. workload, time, mobility) influence the ability of women and men to participate in the programme/project?

Food and Livelihoods

Who manages food within the household?
- How is food distributed within the household?
- Who cultivates land and grows food?
- Who is the family's main income earner?
- What are the income-generating opportunities and needs of men and women?
- Where is it convenient for women and/or men to collect food assistance?
- Who collects food assistance?

Health Risks and Accessibility to Health Services
- What are the health risks for women and men? How and why are they different?
- What barriers (e.g., self-confidence, mobility, financial resources, role in decision-making) do women and men face in accessing health services and health information?
- Where do women and men go for health services and information?
- Which communication channels are most appropriate for women and men?
- Can women and men discuss their health problems/issues among themselves?
- Is this culturally accepted?
- Where can women and men learn more about how to address their health concerns?
- What social networks exist in the community for men and for women?
- Can these networks help address health concerns?

HIV-Affected Households
- For HIV-affected households, what are the different coping mechanisms of women and men? Of girls and boys?
- For HIV-affected households, what is the impact on girls' and boys' school attendance? Are girls withdrawn from school more often than boys?
- What are women's and men's responsibilities related to caring for PLHIV?

Collecting data on gender issues

As a general rule, gender-sensitive data is to be collected in each module of a household survey. In the context of a CFSVA, the gender-sensitive data usually incorporated (or that can be easily incorporated) into a household survey includes:

- **Demography:** sex and age of the head of the household; household composition;
- **Migration:** circumstances of migration (reasons, remittances, gender of the migrated members) in order to assess impact of migration on gender prevalence (at the household and community levels) and on food security;
- **Education:** primary and secondary school attendance of girls and boys, and literacy skills of the head and his/her spouse, informal training of men and women;
- **Income sources:** differentiated participation of household members in income-generating activities;
- **Food consumption:** intra-household distributions and sequence of family members eating food; and
- **Health and nutrition:** prevalence of malnutrition among women, prevalence of child malnutrition by gender, data on breastfeeding and reproductive health, and awareness of HIV/AIDS prevention and transmission.

Key questions in the household questionnaire for HIV/AIDS

Literature on HIV/AIDS identifies some key attributes of chronically ill or deceased adults that are crucial to studying the impact of and responses to HIV/AIDS. These attributes include age, gender, relationship to the household head, educational level, active role of the individual in the household, and decreased capability to work. Ideally, a survey on the impact of HIV/AIDS should collect information on all these attributes. Within the context of food security assessments, the minimum set of attributes to consider includes:

- age;
- relationship to the head of the household; and
- decreased capability to work.

The way we capture the presence and key attributes of chronically ill or deceased household members depends on how demographic data are collected during the household survey.

Option 1: Data are collected through a roster

If data are collected through a roster, ID, name, age, gender, and relationship to the household head are typically collected. Table 4.18 shows how a roster can be adapted to capture information related to chronically ill members. Yellow highlighted sections help identify chronically ill adult members and some key attributes.

Table 4.18: Demographic data collected through a roster

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
ID	**Name**	**Age** (in years, if <1yr. old, write 0)	**Gender**	**Relationship to the HH head**	**Has s/he been not fully functional for at least 3 months over the past 12 months?**	If yes, **with which kind of illness?**	**Is s/he engaged in paid work (cash/in-kind)?** If ill, consider the period before illness.	If chronically ill, **over the past 12 months, has s/he been able to work as before?**
01 02 03 04 05 … … N	____ ____ ____ ____ ____ ____ ____ ____	|__||__| |__||__| |__||__| |__||__| |__||__| |__||__| |__||__| |__||__|	|__| |__| |__| |__| |__| |__| |__| |__|	|__| |__| |__| |__| |__| |__| |__| |__|	|__| |__| |__| |__| |__| |__| |__| |__|	|__| |__| |__| |__| |__| |__| |__| |__|	|__| |__| |__| |__| |__| |__| |__| |__|	|__| |__| |__| |__| |__| |__| |__| |__|
		If 98 or more, write 98 99 = NK	0 = M 1 = F	1 = head 2 = spouse 3 = son/daughter 4 = father/mother 5 = brother/sister 6 = grandparent 7 = uncle/aunt/ cousin 8 = niece/nephew/ grandchild 9 = adopted/foster child 10 = stepchild 11 = no relation	0=no 1=yes	0=mentally/ physically disabled 1= chronic illness	0=no 1=yes	1=yes, able to work the same number of hours/days 2= no, working for fewer hours/days 3=completely unable to work

NK = not known

Data on deceased members require nesting a separate table within the questionnaire. Table 4.19 suggests how questions on recent deaths can be formulated. They are very similar to the questions on chronically ill members.

Table 4.19: Collecting data on deceased household members

Q1	Did any adult (ages 18–59) household member die during the 12 months before the survey after being sick for at least 3 months over the past 12?				|__| 0=no (skip the whole section) 1=yes
For each of the adult (ages 18–59) household members who died after being sick for at least 3 months over the past 12, report:					
	Q2 **Cause of death**	**Q3** **Gender**	**Q4** **Relationship to the HH head**	**Q5** **Was s/he engaged in paid work (cash/in kind)?** consider period before illness	**Q6** **In the period s/he was sick, was s/he able to work as before?**
1 2 … N	|__| |__| |__| |__|	|__| |__| |__| |__|	|__| |__| |__| |__|	|__| |__| |__| |__|	|__| |__| |__| |__|
	1=after chronic illness 2=after a period of physical disability 3=old age 4=problems caused by pregnancy 9=other (Spec. :________)	0=M 1=F	1=head 2=spouse 3=other member	0=no 1=yes	1=yes, able to work the same number of hours/days 2=no, working for fewer hours/days 3=completely unable to work

Option 2: Data are not collected through a roster

If the household questionnaire does not include a roster, questions on deceased members are the same as in Option 1; questions on chronically ill members need to be asked in a different format.

Table 4.20: Demographic data not collected through a roster

Q1	Has any adult (ages 18–59) household member been not fully functional for at least 3 months over the past 12 months?				|__| 0=no (skip the whole section) 1=yes
For each of the adult (ages 18–59) household members who have not been fully functional, report the following:					
	Q2 **Kind of illness**	**Q3** **Gender**	**Q4** **Relationship to the HH head**	**Q5** **Was engaged in paid work (cash/in kind)?** Consider period before illness	**Q6** **Over the past 12 months has s/he been able to work as before?**
1 2 … N	|__| |__| |__| |__|	|__| |__| |__| |__|	|__| |__| |__| |__|	|__| |__| |__| |__|	|__| |__| |__| |__|
	0=mentally/ physically disabled 1=chronic illness	0=M 1=F	1=head 2=spouse 3=other	0=no 1=yes	1=yes, able to work the same amount of hours/days 2= no, working for fewer hours/days 3=completely unable to work

4.2.5 Ensuring data quality

The manager of the CFSVA should ensure that the highest quality data are collected. Data quality is influenced by many factors, such as whether or not PDAs or paper questionnaires are used, the selection of previously experienced enumerators (knowledge of local languages and the language of the supervisors, country, food security, data collection), the quality of the training, how motivated staff are to collect accurate data (this can be seen during the test), and how well the data collection process is supervised. If PDAs are used, enumerators should possess basic computer knowledge.

4.2.5.1 Training of the enumerators

Quality training is essential to ensure that enumerators are able to ask questions in a non-leading way, probe during the interviews, and triangulate the information collected.

Although there are several ways to go about enumerator training, there are some rules of thumb for CFSVAs. A typical training schedule could include sections on:

- Administrative issues;
- Overview of survey purpose and objectives;
- The role of interviewers and interviewer comportment, standard operating procedure (SOP);
- Overview of questionnaire and discussion of individual items;
- Mock interview demonstrations;

- Role playing;
- Field exercise; and
- After action and review.

For the main survey instruments (household and key informant, not including child anthropometry), training usually takes about five to eight days.
  - three to four days of in-class training;
  - one to two days of field testing; and
  - one to two days to review issues after the field testing.

- If the survey includes child anthropometry data collection, an additional two to five days will be needed for the enumerators responsible for the child nutrition data collection, depending on the level of previous experience, including field testing.
- If PDAs are to be used, an additional one to three days of training will be necessary, depending on the level of computer experience the enumerators have. It is also recommended to have an ad hoc trainer for the PDA.
- If qualitative methods (focus groups) are part of the survey, an additional two to eight days will be needed, depending on the skill level of the enumerators.
- Team leaders will need a separate one to two days of training on issues such as key informant questionnaires (if they are the ones responsible), selecting households, and reviewing their responsibilities.
- It is best to have two or more trainers, preferably including a national staff member.
- Try to have the questionnaires as finalized as possible before the training begins.
- In training, try to mix formal presentations with discussions and group work, practice sessions, and other examples to keep the attention of the participants.
- The trainers are responsible for modifying the questions in the questionnaires and/or the PDA file accordingly with the recommendations coming from the field test.
- If PDAs are used, the trainers, in collaboration with the supervisors, are responsible for the logistics (charging the batteries and updating the latest version of the questionnaire file on the PDAs).

4.2.5.2 Field testing

The field testing is a crucial, and mandatory, step in ensuring data quality. It involves checking the questionnaire, raising problems/issues that arise with the questions, and answering the doubts the enumerators can face. Once the field test is done, it is essential to modify the questionnaire in accordance with the observations and discussions made by and with the enumerators.

The rules of thumb for field-testing during a CFVSA

- Generally, this is done over one to two days. Often there are two phases: the second field test uses the final version of the questionnaire and captures the latest mistakes and difficulties encountered by enumerators.
- Each enumerator should be responsible for a minimum of two household interviews, preferably more.
- It would be useful to send the enumerators to do the interviews in pairs, with one asking the questions and the other listening. After the first interview they can exchange roles. This process helps the enumerators better understand how to ask the questions and allows them to give their opinion.

- The team leaders should work with their assigned teams. They should also be field-testing their supervisory responsibilities.
- Pick a site for field-testing that is typical of real field conditions yet relatively easy to reach. The field test should be as similar to the real field conditions as possible.
- It is often preferred not to inform the community and informants they are only participating in a "test." Even if they are not included in the sample, they are filling an important role as the field test site, which will improve the quality of the survey, much as the participation of any community and informant selected in the sample will improve the quality of the survey.
- If child anthropometry data collection is expected, the field test is a good opportunity to check if the enumerators are able to perform the measurements under field conditions.
- If PDAs are used for data collection, it is advisable that the PDA trainers join the groups. This way they will immediately address issues that might come out during the test. This is particularly important if GPS units are used.

4.2.5.3 Field supervision

A daily quality control of the data collected is highly recommended whether paper questionnaires or PDAs are used. Having a replacement for any enumerator (or supervisor) who may become sick, who leaves for personal reason, or who is not adapted to the type of work required is also essential to the quality of the data. A supervisor oversees all or part of the enumerator teams; a team leader is the head of one team of enumerators. Supervisors should conduct spot visits of the data collection teams in the field in order to check on the work achieved, especially during the beginning of the survey. They should also revisit a random sample of households to ensure quality.

Team leader responsibilities

General

The team leader is in charge of managing all data collection, starting from the selection of the households, to the assignation of households to each enumerator, and the checking of each questionnaire filled out.

The team leader should support the enumerators during data collection and clarify and resolve issues or problems raised during data collection. He/she should also be in charge of the time, planning and managing data collection based on the number of hours the team will stay in the field. He checks the enumerators' work by sitting in on interviews, reading completed questionnaires for misunderstandings and recording errors, liaising and troubleshooting with respondents and local community leaders, and covering for enumerators in emergencies (e.g. arranging interviews if an enumerator falls sick).

Paper questionnaire

The supervisor should be able to check each questionnaire in the field before leaving the village, so any mistakes or missing information can be corrected/filled in by the enumerators. By marking any errors with a red pen, the supervisor makes them more visible and reminds enumerators not to repeat those mistakes.

The supervisor should also check that the numbering of the questionnaire questions is correct; he/she is also in charge of collecting all the questionnaires at the end of the day and storing them for the data entry process.

PDA

The supervisor should check the questionnaires for data quality. The PDA software allows supervisors to load the questionnaires saved on each SD (memory) card so that they can review entered data. Supervisors are also responsible for assigning the PDAs to the enumerators.

Data is automatically copied in two locations by the survey software: SD card and PDA internal memory. At the end of each day the files should be copied to a PC, if available, as a further backup.

Enumerator responsibilities

General

The enumerators are the first responsible for the quality of the data collected. A job description for an ideal enumerator would include: communication skills, good knowledge of the international language of the country (English or French, or Portuguese or Spanish) as well as the local language(s), a perceptive intelligence, inexhaustible patience, unfailing dependability, outstanding people skills, and a willingness to work long hours.

Documents

Each enumerator needs to bring enough copies of the questionnaire for the whole day, plus some spare copies. If pens are used to fill in the questionnaires, they should bring extras. If the enumerators prefer pencils, they need to bring erasers and a pencil sharpener.

Once the data collection is concluded, the questionnaire should be kept in a protective plastic folder. Each enumerator is responsible for bringing the questionnaires to the supervisor, checking for mistakes with him/her, and making the necessary changes.

Legible handwriting is important. When the enumerator checks an answer or writes it out, he/she should bear in mind that the data will be entered by a different person, who will need to be able to read the handwriting.

PDA management

If PDAs are used, the team leaders should be trained in how to manage them in order to provide support to enumerators during data collection. Necessary knowledge includes: how to maximize battery life, how to reset the PDAs, and how to reinstall the survey software, if needed.

Logistics for data collection

General

Implementing a successful assessment requires logistics planning and preparation. Logistics is an important part of the survey, and in some countries it can be a cumbersome exercise, and so should be considered early on in the process. It is linked to selection of field sites and advance notification of sample communities or sites, coordination of transport and communications operations, and distribution and collection of data collection instruments.

Paper questionnaire

To ensure proper data collection, the teams should arrive in the field with the adequate equipment, including:

- Enough copies of the questionnaire
- Pens
- Pencils (rubbers and pencil sharpener)
- Seasonal calendar

If the CFSVA includes anthropometric measurements, the enumerators taking the measurements are responsible for bringing along the necessary equipment:

- Scale (UNICEF standards)
- Height board (children and women)
- MUAC tape

PDA transport

It is important to inventory every item related to the PDA before and after travelling. The list of necessary items generally includes:

- PDAs and manuals
- Chargers (one per PDA)
- Batteries (one per PDA)
- Stylus
- Storage cards
- GPS units, if needed
- Car chargers, if needed (in zones with no power supply)
- SD card reader for installing the software (usually one is sufficient)

PDAs are delicate hardware. They should be stored in a durable bag for transportation, to avoid their getting crushed by other cargo. PDAs usually come with a cover for protection, which should be used to avoid damage to the PDA.

After each day of use, the PDA's battery should be fully charged. If electricity is not available, team leaders should be provided with car chargers. Some PDA models can be charged via a USB cable attached to a laptop or a desktop PC. If GPS units are used, these should be charged as well. Charging PDAs and GPS devices requires about two hours.

When using the car charger, it is important to follow these steps:

1. Before connecting the converter, start the car's engine.
2. Verify that the switch at the rear of the converter is off.
3. Plug the converter into the cigarette lighter.
4. Make sure the switch on the multiple connectors is off.
5. Connect the multiples connectors to the converter.
6. Switch the converter on.
7. Turn on the multiple connectors' switch.
8. Start connecting the processor for PDA and GPS.

9. The unit is charged when the light is green.
10. Extinguish the multiple connectors and then plug in the converters.
11. Unplug the converter from the cigarette lighter.
12. Run the engine for another 20 to 30 minutes before shutting it down.

Managing problems with PDAs

Battery problems
With normal continuous usage, the battery lasts about 5 hours. In order to maximize the battery life:
- Charge the PDA as often as possible.
- Keep the screen brightness low.
- Turn off the volume.
- Switch off the PDA when not used.
- Enable the automatic "power off" function in the PDA's battery settings.

Resetting
If the PDA freezes, it is possible to do a "soft reset" to restart it (Note: you will lose the data that is currently being entered.) In most models, there is a button on the back of the PDA that you can push using the stylus. If this button cannot be found, consult the manual that comes with the PDA.

Speed/memory
To increase performance of the application, quit unused applications using the PDAs memory manager tool.

4.2.6 Key references: Survey

- Devereux and Hoddinott. 1992. *Fieldwork in developing countries*.
- CARE, Second Edition. 2008. *The Coping Strategies Index: Field Methods Manual.*
- WFP Nutrition Service. 2005. *Measuring and Interpreting Malnutrition and Mortality.*

4.3 HOUSEHOLD DATA ENTRY

The analysis of household survey data is undertaken in Microsoft Excel, SPSS, or STATA. A key step in the process is to convert responses collected during the interview into a format that can be easily manipulated by the analyst. The two most popular choices currently used by WFP are direct data entry through PDAs and manual entry of responses into a data entry programme. This chapter focuses solely on manual data entry by data entry operators using desktop or laptop computers. Regardless of the method employed to digitize the responses, a series of standardized steps must be followed to ensure the quality of data. The process may vary depending on the context, availability of resources, and the circumstances.

Ideally the management of quantitative survey data, or closed question/semi-open question qualitative data survey, will take the following steps:

1. Prior to the study, review the questionnaire and ensure that the format is adapted for ease of data entry.

2. Set up the data management entry process before the paper questionnaire is used.
3. Development of a data entry method, usually using a data entry programme, if large amounts of data are to be organized into a database.
4. Conduct data entry and check the accuracy.

Different data collection methodologies require the use of different questionnaires. Likewise, the types of data that need to be captured will vary by type of survey. In close-ended (quantitative) questionnaires, coded categorical responses, yes/no answers, and figures are entered into a database. For qualitative studies, answers in the form of narratives need to be coded and then entered. For the purpose of a CFSVA, regardless of the data entry application employed, the output or captured data must be easily transferable into SPSS, STATA, or Microsoft Excel for analysis.

4.3.1 Review of the questionnaire

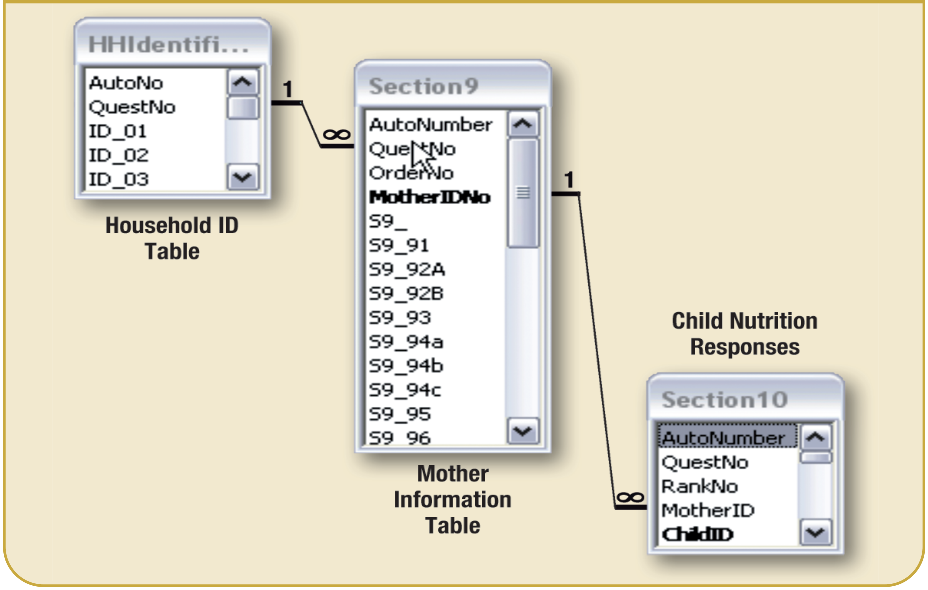
During questionnaire design, the person responsible for designing the data entry tool (who should be familiar with the software used) must review the format of the household and village questionnaire to ensure that all questions are properly numbered; that the layout of the questionnaire allows for easy construction of the relationships between the tables, and that unique ID numbers are logically constructed based on an agreed-upon coding system.

Each question and table should be clearly and logically numbered. In addition, the unique ID code assigned to each household should be clearly labelled.

4.3.1.1 Household and household member relationships

An important step in the design of the household questionnaire is creating relationship tables between the differing hierarchies within the households.

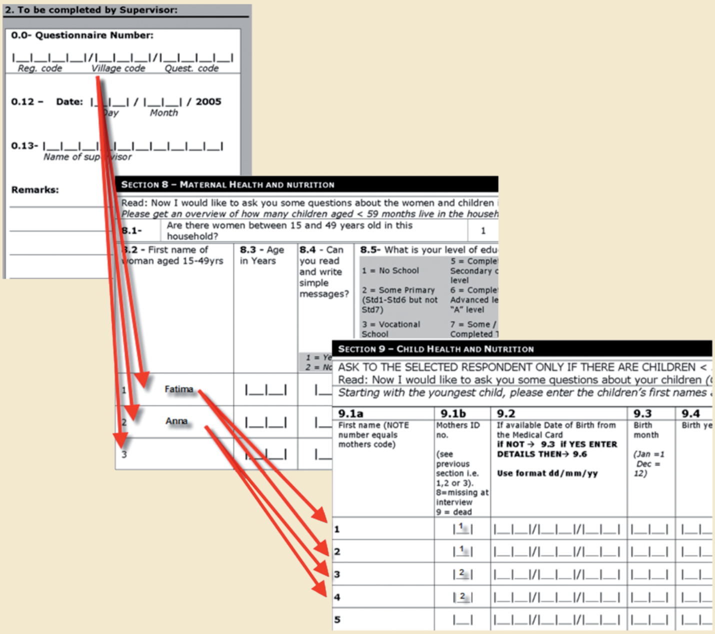
Figure 4.4: Household and member relationship tables



This is particularly important if the study is attempting to record the responses of the household, household members, mothers in the household, and information on each child. In terms of data relationships, the household is the highest unit; the household members and mothers have “one-to-many” relationships to the household (e.g. one unique household can have many members, each with his/her own characteristics), and the child has a “one-to-many” relationship to the mother. For example, in a single household (one unique HH ID), there are two mothers who have responses recorded in the survey (both have the same unique HH ID, but each has her own unique mother ID). Each of these mothers has two children under 5 whose age, weight, and height are measured (each child has the same unique HH ID, each set of siblings has the same mother ID, and each child has his/her own unique child ID). The household-to-mother and household-to-child relationship is one-to-many, the mother-to-child relationship is one-to-many, and the reverse relationships are many-to-one. Figure 4.4 is an example of the relationships from the East Timor CFSVA.

In order to facilitate the one-to-many relationships in a data entry application, it is much easier if the questionnaire uses a horizontal table, where each member of the sub-group (e.g. the children) is recorded on a horizontal line. The enumerator can add as many lines as required in the questionnaire based on the number of children in the household concerned; the programme can also insert the same number of lines using one-to-many relationship tables. The example from the Tanzania Household CFSVA questionnaire in Figure 4.5 shows the relationship.

Figure 4.5: One-to-many Relationships



4.3.2 Paper questionnaire management before data entry

Organizing filled-in questionnaires in a systematic fashion will prevent questionnaire loss. Good questionnaire management is a prerequisite for ensuring timely data entry. The steps necessary to develop an effective questionnaire management system at the data processing office are discussed here.

1. As envelopes/boxes of questionnaires return from the field, organize them into groups. The envelopes can then be further organized into larger groupings, usually by province or district. For example, all questionnaires from the same village/cluster can be put together in one envelope, with the name of the village, district, province, etc. written clearly on the envelope. Then all the envelopes from the same district can be put together in one box, with the names and codes of the province and district written on the box. Be sure that the system of organization uses information that is also included on the questionnaire and entered into the database, so that if a physical questionnaire needs to be located, the information leading to the correct envelope will be contained in the database.

Box 4.7: Correcting errors on paper

- When correcting errors or making changes to completed questionnaires, use a red pen to make it clear what is original data and what has been added or changed. Give red pens only to those responsible for making changes/corrections.
- Always cross out bad answers with one slash mark. Do not scratch out answers, or use liquid paper or erasers. This way, should a correction be made in error, the original answer will still be legible.
- Never fill in missing answers unless it is 100 percent clear what the answer should be. It is better to have missing data than incorrect data!

2. Each questionnaire should have a unique ID, i.e. a number that identifies a questionnaire from all other questionnaires in the survey. This number can be assigned at any time, but is best done before the teams are in the field. It is recommended that a questionnaire number be used. The questionnaire number is a sequential number that can be constructed with each team starting off with a number and then continuing in continuous order (e.g. 1001,1002, . . . 1999)

Example of Questionnaire Identification Number

QUESTIONNAIRE ID:
|_|_|/|_|/|_|/|_|_|/|_|_|//|_|_|
Prov. Dist. Terr Sec. Group. Men

3. As questionnaires are entered into the database, they should be filed in sequence according to their unique ID numbers.
4. Sometimes it is helpful to have a separate group of people (not the enumerator or data entry operators) responsible for carefully reviewing every

paper questionnaire, looking for problems, making corrections where possible, and ensuring that all answers and unique IDs are clearly recorded. The decision to add this step depends on the type of questionnaires and questions being used, the number of errors present, the recoding needed, and the experience and understanding of the team leaders and data entry operators. If the team leaders and data entry operators have the experience to identify and correct most of the mistakes, then this separate step might not add much value. Correction of mistakes should be done in a clear manner, as sometimes the corrections themselves can lead to confusion for the data entry operators.

When data is gathered at several levels in one survey, a decision must be made regarding how to organize the different levels of information. One option is to keep all the questionnaires together based only on the geographic or other filing system. Another is to separate the questionnaires based on level of data collection (household, focus group, key informant, etc.), and file them separately.

4.3.3 Development of a data entry application or programme

4.3.3.1 Data entry applications

Principles

Developing data entry applications can be complex and time-consuming. However, time spent developing an application will mean fewer errors at the data entry stage. Although each application requires some customization, the key principles in any data entry application are similar:

1. Maintain household records and the relationships between the sub-household units (household members, household mothers, and children).
2. Create data entry masks (forms), which allow the data entry person to enter records in a standardized and intuitive way to minimize key strokes and mouse clicks.
3. Use structured entries that unambiguously link every household to its administrative units and limit the ability of data entry persons to input improbable values and categories, and also use embedded logic commands (such as filters) that control for simple data entry errors.
4. Export entered data into statistical software to ensure minimal effort and no loss of integrity.

Keeping these principles in mind, this section guides the design of data entry templates that adhere to these four principles while allowing for the specific requirements of each study. This guideline recommends Microsoft Access, which has been used in this example, as the platform for data entry. However, many other software packages are available and are commonly used.

4.3.3.2 Software for data entry and management

A number of software packages are available to facilitate data entry into a computer system. Although this guideline uses Microsoft Access to show how to develop a data entry template, data mask, and how to enter data, there are other applications available which can be used for data entry. Brief descriptions of commonly used applications are provided below.

Microsoft Access

Access is a Microsoft database application that allows data to be stored in related tables based on unique IDs. Data entry in MS Access can be undertaken in designed forms. The forms can be customized with embedded Visual Basic for Application (VBA) code to allow filters and logic commands to control for clerical errors. Data can be exported into most data analysis applications retaining the hierarchical structure. Unlike with Excel, there is no limit to the number of records that can be stored in MS Access. However, the data table needs to be less than two gigabytes, which is largely sufficient for CFSVA surveys. The strength of the Access data entry programme is that it is user-friendly, as it shows the image of the questionnaires (graphic water mark) so that data entry operators can enter the information in the same manner in which it appears in the paper questionnaires. A data entry network can be built that centralizes all records entered by operators into a single computer that the data entry supervisor can use to randomly check questionnaires.

SPSS

SPSS Data Entry Builder allows the design of surveys and forms using the drag-and-drop interface and a library of sample questions that one can edit. One can also create new questions not in the library. In SPSS forms, you can embed rules and filters, and validation procedures. The SPSS data entry platform can also automatically create data files and dictionaries that can be immediately imported into SPSS for analysis.

Epi Info 3.x

Epi Info is a Windows-based programme developed by the U.S. Centers for Disease Control and Prevention. It is freely available on the web at www.cdc.gov/EpiInfo/, which also offers online support. It is one of the most widely-used applications for anthropometric data analysis. Epi Info allows the data entry operator to customize the data entry mask and control for categorical variables and filters. The earlier version of Epi Info (EpiInfo 6.0) was a DOS-based programme. In 2000, the CDC developed a Windows-based Epi Info. Although there were bugs in the earlier Windows-based version, the more recent releases function well.

Epi Info works closely with MS Access as it creates the latter's (.mdb) databases. However, to run Epi Info, it is not necessary to install MS Access, nor does the operator need to know how to work in MS Access.

CSPro

Census and Survey Processing System (CSPro) is a Windows-based data entry and editing programme developed by the U.S. Census Bureau. CSPro has the ability to support the development of effective data entry and the editing of programmes for complex national household surveys. The application is easy to acquire and to learn. Anybody with a basic understanding of databases can learn how to use CSPro in two weeks. It is a free data entry software (available at www.census.gov), and has the added benefit of exporting data easily to a wide variety of analysis software formats.

NutriSurvey

The main purpose of NutriSurvey is to integrate all steps of a nutrition baseline survey into a single programme. The strength of the software is in preparing a suitable questionnaire for entering data and evaluating the results. The programme's standard nutrition baseline questionnaire can easily be customized for a specific survey. For further statistical analysis, the data can be exported to SPSS or another statistical programme.

Microsoft Excel

The Excel application is usually included in a Microsoft Office software package. This easy-to-use spreadsheet programme works with two-dimensional data. However, with Excel it is not possible to develop one-to-many relationship links within household units in one spreadsheet. Likewise, the maximum number of records allowed in Excel is 65,536, and the programme can accommodate only 255 variables in one spreadsheet. Moreover, data masks cannot be developed to minimize entry-level errors. Hence, for complex surveys, such as the CFSVA, the use of Excel by enumerators will increase the probability of data entry errors and inconsistencies.

Table 4.21: Summary of data entry platforms

Software	SPSS	Epi Info	CSPro	Nutri-Survey	MS Excel	MS Access
Type	Commercial (if special option for data entry)	Freeware	Freeware	Freeware	Commercial (usually part of the Windows Office suite)	Commercial (usually part of the Windows Office suite)
Availability	Fair	Excellent	Excellent	Excellent	Excellent	Good
Ease of use (creation of data entry programme)	Fair	Fair	Fair	Fair	Good	Poor
Ease of use (data entry)	Good	Good	Good	Good	Fair	Excellent
Data output	All formats	All formats	All formats	All formats	All formats	All formats
Multi-user data entry	Yes (requires second licence)	No	No	No	No	Yes
Customized data entry forms	Yes	Yes	Yes	No	No	Yes (with graphic watermark)

4.3.3.3 Overview of MS Access data entry mask

The MS Access data entry tool contains a "front-end/back-end" configuration. Briefly, the front-end is the data entry template containing the forms and control code for data entry; it displays the image of the actual questionnaire (see: Image Watermark in Annex 9), with each question associated with a box where the data entry operators enter the answer. The front-end configuration is installed on each of the computers to be used for data entry.

Conversely, the back-end is where the data tables are saved and are linked to the front-end. The back-end can exist either in each data entry operator's computer or in a single computer when using a network; in the latter case, only the supervisor has access to the back-end. Microsoft provides a detailed explanation of the steps involved in creating a front-end and back-end. Visit the following URL for more assistance: **http://support.microsoft.com/?kbid=304932.**

A computer network is a setting in which the front-end is installed in each computer to be used by the data entry operators, and the back-end is installed only in the supervisor's computer. This back-end is meant to centralize all entries from the many data entry operators' computers. Setting up a computer network is a convenient way for the data to be automatically merged into the different tables in a single computer. Also, it allows the supervisor to have control over the entered data and limits data entry operators' ability to manipulate data once it is stored.

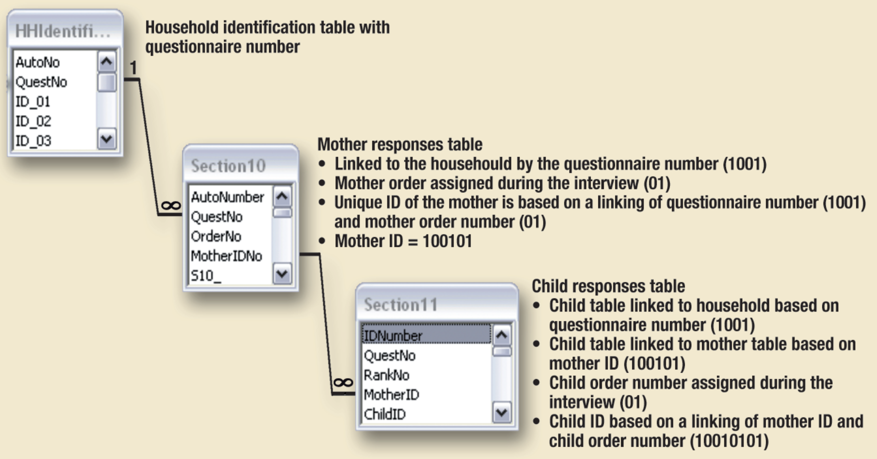
In building data entry masks using MS Access, it is critical to know the following key characteristics of the application in order to prevent entry errors:

1. Multiple Related Tables

Microsoft Access allows a maximum of 255 variables in one table. However, a standard CFSVA may have more than 500 variables. Therefore, household data will be stored in several tables linked together (in a one-to-one relationship) by their unique household ID. Since CFSVAs collect demographic data for all household members and anthropometric data for the members in the reference age group, it is necessary to record data by member as well as by household. As demonstrated in Figure 4.6, the relationship between the household, mother, and child is based on the questionnaire number and the order number of the mother and child.

CHAPTER 4

Figure 4.6: Interlinked tables in MS Access



When creating the data entry application with sub-forms, keep in mind that it is easier for the data entry operator to enter the data if the data entry page contains only those questions that are on the page of the actual questionnaire. In other words, the data entry page should mimic the questionnaire page exactly.

2. Location Codes

The analyst may want to classify households based on their location, in which case a location code is essential. Although enumerators will record the cluster or village ID on the questionnaire, data entry operators must be allowed to enter only the valid location codes selected for sampling. This will minimize entry error. Annex 9 offers a detailed discussion on creating data entry masks to minimize entry errors. Moreover, as the names of the administrative levels of the cluster are already defined, it is not necessary for the data entry operators to record the type of location (e.g. region, province, district, or village), as this will only slow down entry speed.

3. Limiting Data Entry

A key role of a data entry tool is to minimize entry errors by controlling for impossible and unlikely responses. The data entry mask in Microsoft Access can control impossible values and outliers, flag possible data entry errors during the entry process, and block the entry of values if a filter question for those values has previously been asked. Depending on how significant the error is, the application can alert the entry operator, restrict the values that can be entered, or refuse the entry of a value. “Smart” programming is particularly important in entering a large amount of survey data. The five main types of controls that can be employed when developing a data entry mask are discussed in Annex 9. Although the examples are not exhaustive, the types of solutions they offer can be applied to other situations.

4. Handling Missing Data

For numeric data, missing values can mean that either the answer is “0” or “not applicable” or the data was not available at the time of the interview. For example, if a household cultivated maize in the past year but could not report the total expenses involved in maize cultivation, then the data is unavailable. If the household did not cultivate maize in the past year, then the question of “total expenses for the activity” does not apply to the household. In this case the response recorded on the questionnaire should be “not applicable.” If, on the other hand, the question is about how many shovels or vehicles the household owns, then a blank response may mean that the household does not own any of these items.

The person handling “missing data” should do so on a case-by-case basis, carefully reading through the completed questionnaire to understand why the data is missing and how to handle this. A well-designed questionnaire will typically assign codes for “not applicable” cases, and for household answers of “I do not know.” (Refer to the example for filters in Annex 9.)

4.3.4 Data entry

Once the data entry programme is written, it can be used to enter data. Data entry is a very important step in the data management process, and should be carefully monitored. As with data collection, usually many people are involved in data entry, and the quality of the final data set depends on the quality of their work.

4.3.4.1 Testing the data entry programme

Testing the data entry programme is an extremely important step, to be undertaken prior to actual data entry. This should be done with filled-in questionnaires. The entered data should then be cross-referenced with the questionnaires to make sure the data in the fields are appropriately ordered. Once the data entry is declared to be error-free, the data entry process can begin.

4.3.4.2 Managing data entry operators

Managing the data entry operators is a key step to ensuring an accurate database. Continuous monitoring of data entry work from the beginning to the end can significantly reduce entry errors. The time it takes to completely enter the data from a household survey depends on a number of factors, including the design of the data entry application, the number of variables in the questionnaire, design of the questionnaire, the key stroke speed of the entry operator, the number of operators involved in data entry, and the overall management of the entry process. Key steps in managing the process are.

a. Hire entry operators with prior experience of data entry. Familiarity with Microsoft Access is an added advantage but is not required. Key stroke speed, attention to detail, common sense, initiative for solving problems, and a willingness to seek guidance when necessary are all important qualities of a good data entry operator.

b. Provide training to data entry operators on the questionnaire, survey objectives, and data entry application (much like enumerator training). The training should allow the operators to enter data and identify mistakes. Keep in mind that data entry operators will be better able to identify and solve errors if, in addition to the data entry template and data masks, they understand the survey, its objectives, and the meaning of the questions. One day of training is usually sufficient, followed by a half-day to one full day of practice data entry, where the supervisor reviews the quality and accuracy of their work.

c. Remember that it is normal for data entry operators to work slowly at first. It is critical to focus on accuracy rather than speed. As the operators get to know the programme, the questionnaire, and the typical responses, their speed will increase. The daily target for amount of data entered should be set only when the supervisor is pleased with the quality of work after several days of data entry.

d. It is good practice to assign a particular computer to one (or two, if they are working in shifts) data entry operators. This way it is easier to monitor the quantity and quality of each operator's work.

e. To ensure the accuracy of entered data, the supervisor should be responsible for checking a certain number of questionnaires per day (usually about 5 to 10 percent of all questionnaires) against the entered data.

f. Keep a daily record of how many questionnaires per day each data entry operator has completed. It is advisable to pay the entry operator by day rather than by page or by questionnaire, as the latter discourages the operator from taking time to solve any problems he/she encounters, and could increase the number of errors in the final database. However, after the first or second day of data entry, it is a good idea to set daily quotas for questionnaires entered in the database by each data entry operator.

4.3.5 Paper questionnaire management during data entry

Tracking the filled-in questionnaires is extremely important during the entire survey process, ensuring that none is lost or entered more than once.

a. The filled-in questionnaires should be organized by cluster, in boxes or in envelopes. Clearly write the name of the cluster on the envelope or box. Identify an area where questionnaires will be stored prior to data entry, and a second area, clearly separate from the first, where questionnaires will be stored post–data entry. Identify a third area, near the data entry computers used, where the envelopes/boxes of questionnaires currently being entered can be stored. Finally, identify a fourth area where questionnaires with problems can be stored until their problems can be resolved.
b. Prepare a register in which all of the clusters and the numbers for the filled-in questionnaires in each of the clusters are recorded. This can be done by recording the name of the cluster and the questionnaire numbers contained in that cluster (e.g. 130052 through 130075). When the data entry operator selects a cluster for entry, she/he should sign and date the register to assume responsibility for that cluster's questionnaires. This will also help to track the questionnaires (which are also all clearly numbered).
c. Each data entry operator should work on one envelope of questionnaires at a time, and should be responsible for all of the questionnaires in that envelope. As data entry for each questionnaire is completed, the questionnaire should be returned to its envelope/box. This will prevent the loss or misfiling of questionnaires.
d. When the data entry operator starts to enter information from a questionnaire, he/she should record his/her name and/or code on the cover page of the paper questionnaire. If desired, this code can also be entered into the database during data entry, allowing easy identification of data entry operator during analysis.
e. When all information from a questionnaire is entered, a clear mark (e.g. a large checkmark or a slash) should be made across the entire front page using a highlighter. Once all questionnaires in an envelope/box are entered, the same mark should be made on the envelope/box.
f. When the data entry operator puts the envelope/box containing the already entered questionnaires in its designated area (see point a.), she/he should sign and record the date in the register.

4.4 HOUSEHOLD DATA ANALYSIS AND PROCESSING

4.4.1 Objective

This section is designed to help analysts (who should already have an in-depth knowledge of statistics) to analyse the household survey data generated by CFSVAs. These guidelines do not elaborate on common statistical or data management techniques, as this information is beyond the scope of these guidelines and must be acquired through academic course work and/or on-the-job training and supervised experience.

Why analyse primary data?
The CFSVA Food and Nutrition Security Conceptual Framework describes how various factors influence the food security situation and vulnerability of households. Using information obtained from various sources, the analyst describes and evaluates household food security status, the factors that influence household food security, the livelihood strategies employed, and the health and nutritional status and other livelihood outcomes at the household level.

Information generated by the CFSVA is used to explain how different households are exposed to risk and how they manage to cope. This information is combined with data obtained from secondary sources to describe the geographic, economic, and social context and explain the risk factors that influence the extent of vulnerability and the capacity to cope with shocks.

4.4.1.1 A note on statistical software
Because WFP uses SPSS for most of its quantitative data analysis, the guidance presented here focuses on that programme. However, experienced statisticians may choose to use other software packages.

For most cluster analysis, and often for principle component analysis, WFP-VAM uses ADDATI,[62] but SPSS can process this analysis, too.

For anthropometric z-score calculations of under-5s (stunting, wasting, underweight), WHO ANTHRO 2005[63] is used. Epi Info is essentially obsolete unless the ENA add-on for EPI Info is used.

4.4.2 Preparation for the analysis

4.4.2.1 Hierarchical data structure
CFSVAs consist primarily of household data. However, information gathered at the household level often includes data on individual household members, such as age, sex, children's education, nutritional status of mothers and children (under 5), women's childcare practices, and women's knowledge of HIV/AIDS. This may result in multiple "units," or cases, from each household. Additionally, data may be gathered at the village or community level that is pertinent to each household in the community (such as presence of schools and health clinics).

These data need to be organized into several data files, one for each unit of analysis, corresponding to the level at which the data were collected. For example, the member-level information (e.g. age, sex, children's education) should be saved into one file, while household-level information (e.g. assets, expenditure, current food consumption) should be saved into a separate file. There should be a separate file for anthropometric

62. This software can be downloaded for free at: http://cidoc.iuav.it/~silvio/addawin_en.html
63. This software can be downloaded for free at: http://www.who.int/childgrowth/software/en/

information for children (sex, age in months, height, and weight). Meanwhile, a different file should be used for child-care data. Similarly, if village-/cluster-level information is collected, a separate data file would be needed.
For CFSVAs, data can generally be organized in up to five data sets:
1. Village
2. Household
3. Individual
4. Mother
5. Child

It is essential to develop a data management plan before data entry.[64] Data entry application may automatically produce the five data sets (if designed to do so) or one large data set that needs to be reorganized into several data sets.

To obtain information about each of these levels, each of these data sets needs to be analysed. However, the analyst may desire to combine information from different data sets into one combined data set. Using SPSS, queries cannot be made between individual data sets; therefore, data sets must be merged in SPSS using a "many-to-one" relationship.

The merging of different data sets needs to be done so that member-level data sets can add to the information gathered from the household. Data analysis should in general be done only at the lowest level contained in that data set (i.e. member level), as described in Box 4.8.

Box 4.8: Relating child nutrition with other indicators

An analyst wants to look at child malnutrition as it relates to the household water source. There can be multiple children in each household, so the relationship between these two data sets (child and household) when merging is many children to one household. Using SPSS, this means merging the household data set into the child data set. This preserves all child information and keeps the number of children in the data set the same. However, in this merging process, some household information is lost (e.g., those with no children under 5) and some is duplicated (e.g. where there is more than one child in a household).

The resulting data set is used only for child-level queries (e.g. to answer the question "What percentage of wasted children live in households with unsafe drinking water supply?")

This merged data set cannot answer the question "What percentage of households have an unsafe water supply?" because some households were duplicated (i.e. those with more than one child) and others were deleted (i.e. those without any children). This question should be analysed within the household data set.

This merged data set cannot answer the question "What percentage of households have a wasted child as one of the members?" Analysis in this direction (from higher to lower aggregation level) is uncommon, and generally not recommended. To answer such questions, merging in an alternate direction would be required.

64. See Section 4.2.5

4.4.2.2 Organization of the database

Organizing the database is an important step to getting a clear idea of the variables the analyst is going to consider. It is also helpful to manage and analyse the data by different individuals. **It is a good practice to make a copy of the database and keep it in a separate folder.** The following are key aspects of database organization:

a. Verify that all **variable names** clearly identify the question in the questionnaire. This can be easily done by using the question's code. Do not change the variable names unless it is absolutely necessary. Changing a name may complicate the identification of the particular variable when comparing it to the original raw data (from MS Access or another data entry tool), especially for other analysts who might access the data later. Additionally, if additional cases are to be appended to the database, differing variable names will impede the process.
b. Often variable names are cryptic; therefore it is necessary to enter **variable labels** to clarify what each variable is. A well-designed data entry programme, properly exported to SPSS, will already have appropriate labels for all variables, but this should be carefully checked. If the labels are clearly written and correctly spelled, it will be easier and quicker to create tables for reporting.

Box 4.9: Example of variable names and labels

In the questionnaire:

HQ5.1b What is the main source of drinking water for your family? 1…, 2…, 3…

The variable name could be HQ5.1b.
The variable label could be "drinking water source."

c. The variable type should also be correctly identified (usually "string variable" for letter/word values, and "numeric variable" for numbers). For categorical variables, it is necessary to enter the value labels for each variable, following the coding from the questionnaire. This information is essential for analysing the data and also for cleaning the categorical variables in the data set (see section 4.4.2.3 on data cleaning).
d. Identifying the measure (scale, ordinal, or nominal) of a variable is a key part of database organization. This information enables the software to conduct appropriate analyses with specific variables.
e. During data cleaning it is also important to specify whether a variable has one or more missing values. It is not uncommon for an analyst to be the first person to discover that a variable is missing values. Coding missing data can be done in several ways. However, each variable, and possibly even different analyses of the same variable, will have their own specific needs, and so there is no cut-and-dried rule for dealing with missing data.
f. Another step involves data recoding. This process is particularly useful for categorical variables. For example, yes/no questions are best coded as 1/0 (not 1/2). The 1/0 option is preferred because a simple mean illustrates the frequency, and in case of regression analysis, the yes/no questions are already recoded as

binary variables, with the sign of the regression coefficient pointing in the intuitively correct direction. If the data entry programme is well designed, this should not be necessary. Boolean variables from MS Access translate automatically in SPSS into 1 (yes/present/true) and 0 (no/absent/false). A good and essential rule is not to lose data during recoding. For example, do not recode and simultaneously replace a continuous variable with a categorical variable. Instead, keep the original variable and create a new categorical variable.

g. A good practice is to keep only those variables meaningful to the analysis in the final dataset. If the analyst creates too many working variables before arriving at the final working data set, these variables will need to be deleted after computation, otherwise the size of the database will increase exponentially and become difficult to manage.

Figure 4.7 gives an example of all the fields that must be organized before data analysis.

Figure 4.7: Example of a database

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	hhid	Numeric	7	0	household ID	None	None	8	Right	Nominal
2	Date	Date	20	0	Date	None	None	17	Right	Scale
3	IntID	Numeric	4	0	IntID	None	None	7	Right	Nominal
4	IntName	String	200	0	IntName	None	None	17	Left	Nominal
5	SupID	Numeric	4	0	SupID	None	None	8	Right	Nominal
6	State	Numeric	2	0	State	None	None	6	Right	Nominal
7	STATEID	Numeric	8	0		None	None	8	Right	Scale
8	LEVELS	Numeric	8	0	levels of analysis	{1, rural population}...	None	8	Right	Nominal

4.4.2.3 Data cleaning

Data cleaning is an essential step in data analysis. Every dataset contains some errors, and a significant amount of time in data analysis is spent "cleaning" the data.
Data cleaning can commence once the data are organized into different files. Data cleaning does not mean simply confirming that the data recorded on the paper questionnaires is the same as that in the dataset. It also entails several iterative steps of checking the dataset(s) to ensure that the data are credible.

Usually the cleaning of CFSVA data sets is done in several stages. The initial part of the data cleaning can be done with the software used for data entry[65] (most often MS Access). Cleaning should not be done as an automatic process but, rather, as a critical, well-thought-out series of recorded decisions.

UNIQUE ID

The first step in the data cleaning process is to ensure that the total number of households in the data set equals the total number of filled-in paper questionnaires. It is important to carefully review the data sets to confirm that all questionnaires have been entered only once and that all unique IDs are truly unique. This step ensures that data sets can later be merged and other household specific variables added, if necessary.

65. See section 4.3.3.

In the more recent versions of SPSS, there is now an option, in the "Data" menu, to "Identify Duplicate Cases," which makes this process very simple. If this option is not available, running a "frequency of the household ID" can be useful for detecting the presence of duplicate households. Household IDs resulting in a frequency of 2 or more have duplicates.

Additionally, it is a good idea at this point to take a random selection of questionnaires and compare them to the database (this should also happen as part of the quality control step during data entry). This verifies that all questions are being entered correctly and that variables are not being mislabelled.

Check the variables

The next step in data cleaning become more subjective, thus it is important not to make any permanent changes to the data unless you are absolutely confident in the decision. Regularly save backups of the database (without replacing earlier backups) so that any changes made can be undone at any time. These steps include checking for outliers and checking for errors/inconsistencies.

Check for outliers

An outlier is an observation that is numerically distant from the rest of the data. Statistics derived from data sets that include outliers will often be misleading. In most samplings of data, some data points will be further away from their expected values than what is deemed reasonable. In the presence of outliers, any statistical test based on sample means and variances can be distorted. Estimated regression coefficients that minimize the sum of squares for error (SSE) are also sensitive to outliers. Outliers can be caused by data collection/data entry errors or by extreme observations that for some legitimate reason do not fit within the typical range of other data values (High 2000).

Check the distribution of data values by levels of a categorical variable, if available. This procedure should always be one of the first steps in data analysis, as it will quickly reveal the most obvious outliers. For continuous or interval data, visual aids such as a dot plot or scatter plot are good methods for examining the severity of any outlying observations. A box plot is another helpful tool, since it makes no distributional assumptions, nor does it require any prior estimate of a mean or standard deviation. Values that are extreme in relation to the rest of the data are easily identified.

Running a frequency table or simple descriptive statistics could also be useful for detecting outliers. Working with outliers in numerical data effectively can be a rather difficult experience. Neither ignoring nor deleting them is a good solution. If nothing is done with the outliers, the results will describe essentially none of the data – neither the bulk of the data nor the outliers. Even though the numbers may be perfectly legitimate, if they lie outside the range of most of the data, they can cause potential computational and inference problems (High 2000). Outliers are not "missing," just too high or low given our expectations; hence they should not be recoded as missing data. There are a couple of ways to deal with outliers.

- Since means are sensitive to extreme values, median values can be used instead of means.
- Maintain the "raw" version of the data, which retains the outliers, but create a "processed version" in which new variables are created that, for example, replace outliers with medians. Create a variable that denotes whether an outlier has been replaced by a median. That way, no data are discarded.

In both cases, it is crucial to report how outliers were managed.

Box 4.10: Possible effects of outliers

- Bias or distortion of estimates (especially of the arithmetic mean)
- Inflated sums of squares (which make it unlikely to partition sources of variation in the data into meaningful components)
- Distortion of p-values (statistical significance, or lack thereof, can be due to the presence of a few, or even one, unusual data value)
- Faulty conclusions (it is quite possible to draw false conclusions if irregularities in the data have not been investigated)

Check for errors/inconsistencies

Impossible values are often found in data sets, in spite of the filters used in the data entry programmes. Sometimes the values are absolutely impossible or contradictory to the information given in prior questions.

Box 4.11: Example of inconsistent values

- The number of the household members is 50.
- People with little or no land had a considerable harvest.

Once an inconsistent value has been identified, the data should be checked on the paper questionnaire to exclude the possibility of data entry error. If the data was entered wrongly in the paper questionnaire, the analyst should be able to decide whether to exclude the value, based on his/her experience and the type of variable. The rule is to change "absolutely impossible" values into "missing values," that is, if there is no way to determine if it is too time-consuming to investigate for the correct value. However, this is a subjective choice and should be approached with absolute caution.

Sometimes, even if the answer appears clean when you compare it with another variable, there is still an evident contradiction. In this case, the rules for an efficient data cleaning will be first to check the original questionnaire; if the answer is not there, look at other variables that can have a connection to those contradictory variables. If even this solution does not yield results, record the value of both the variables as missing.

Box 4.12: Example from Laos database

In the child database there is information about child demographics and household demographics.

Section 1, question 1.7: the enumerator should complete the information about the household demographics (number of people in the different sex and age groups).
Section 10, questions 10.3 and 10.4: the enumerator entered the information about the child's age and sex. In the database, the analyst found many inconsistencies that were difficult to solve, including households where the number of people in a specific age group does not match the number of children measured.
In this situation, the analyst, after cross-checking with the paper questionnaire, should try to find the truth in other variables (e.g. by looking at the variables related to education to see if the child was included in the wrong age group) or exclude the case from the analysis.
In other cases, the sex of children is different in the two sections. In this case, the information in Section 10 should be more accurate because the children were present during the measurements. So the analyst changed the variable 1.7 based on the information collected in 10.4.

Usually a questionnaire is developed with the flow of the questions kept in mind. In many cases there are skips in the questionnaire that allow the interviewer to bypass questions not applicable to a particular respondent. The data should be entered accordingly. For example, if a household did not cultivate any land, questions regarding harvest and crop types are not applicable to them. However, a well-designed data entry programme should automatically skip the fields that are not applicable to a household based on the information entered for the filter question.

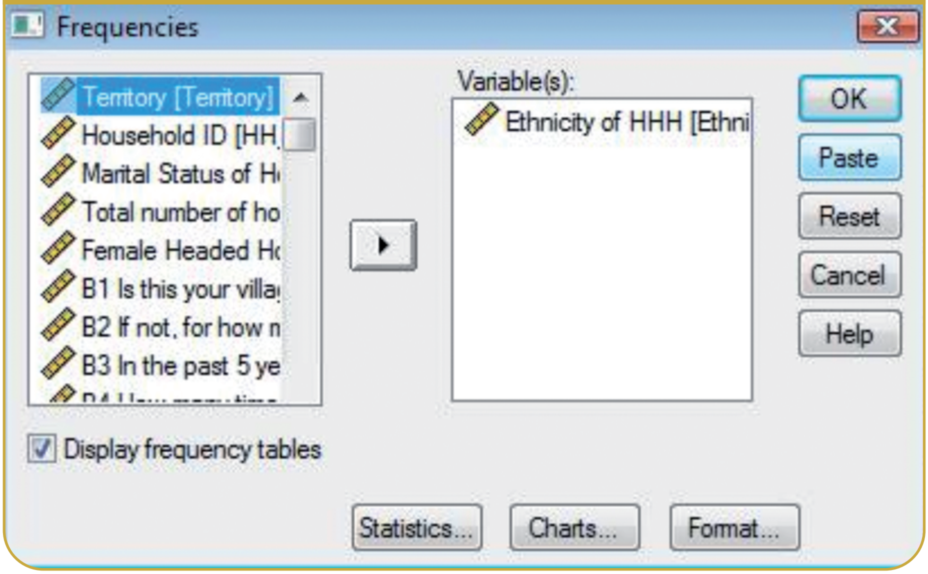
4.4.2.4 General rules for data cleaning

- Do not start guessing, predicting, or assigning values. Even if a value seems obvious, do not make a change unless it is supported by clear evidence and the change is recorded.
- Bear in mind that most often the database will be managed/analysed by different people after the cleaning. Prepare an easy-to-read, clean database in which all the variables have their basic information; this will reduce mistakes and minimize time spent trying to comprehend the data set.
- Save a copy of the unchanged database before making any changes to it.
- Recode the history of changes in a syntax file. A record of the changes will be invaluable and can help replicate the same cleaning in different backups.
- In case of contradiction/inconsistency/doubts in the database between variables:
  1. Check the original questionnaire.
  2. Check the validity of the data by comparing them to other variables in the database.
  3. Change the value only if you are 100 percent sure.
  4. Consider the case as missing data.
  5. Keep the syntax for recoding automatic changes (if any).

4.4.3 Data analysis

4.4.3.1 Standard practices

Figure 4.8: Pasting syntax



This section discusses several standard practices that are useful in CFSVA data analysis. By following these general guidelines, CFSVA analysts will have more compatible methods of organizing their analyses.

Use of syntax

In SPSS, the syntax is a log of all transformations and procedures used. Syntax can easily be generated in SPSS. Using syntax is a matter of personal experience and preference, ranging from only minimal use for data cleaning to use for conducting all analyses.

In most of the interactive menu for a transformation or procedure in SPSS, there is a "Paste" option. This will save the syntax of the command or transformation in the **most recently** opened syntax file (see Figure 4.8).

It is recommended to keep several syntax files: one for data cleaning, one for key complex transformations, and others for the main analytical steps. Keeping the syntax of transformations and key analysis is good practice. That way, if you are ever uncertain about how a variable was created or corrected, you can go back and check.

Syntax can be copied, pasted, and edited in the SPSS Syntax Editor. It is a good practice to write a brief description of the procedure before each syntax command. Separating different parts of the analysis is also considered good practice.

Advantages of using syntax

There are several reasons for using syntax. It can improve efficiency and transparency, and also save time.

A lot of transformations for different variables are similar. A clever use of the Copy/Paste and eventually the Find/Replace function in the Syntax Editor can avoid repetitive chores, save time, and even reduce the chance for error in transformations. Some analytical procedures might take more manipulation than others. This can easily be done by using the interactive menus in SPSS. However, if an analyst wants to change the procedure after the analysis, he/she will have to rebuild the entire procedure, reset all options, and reselect all the variables. Clearly it takes a lot of unnecessary extra time before the same procedure with the same results can be reproduced. If the syntax is saved, however, no time is wasted at all.

At a later date, the analyst or a colleague might want to review the transformations performed or the exact way a procedure was conducted. Use of syntax will enable a review of the draft analysis and make necessary corrections before finalizing the report. It is not uncommon to identify inconsistencies in the data or problems in previous procedures. Moreover, others should be able to see exactly how new variables were defined. Using syntax will give instant access to the formula used to create the new variables. It is recommended to keep a complete syntax of key transformations and procedures.

Data backup

Frequently save backups of the data set. An erred transformation of a variable or a manipulation of the data file can result in lost data. Each successive version of a database should be independently saved. This allows the retrieval of original data if such a mistake occurs.

Labelling variables and values

The cleaned database should have complete variables and values recorded. When creating new variables, be sure to enter suitable variable labels and values of categorical variables. This will enable another analyst to easily and quickly interpret the variables in the data set. As described in section 4.4.2.3, it is recommended that the variable names reflect the questionnaire number, and the variable label provide a more detailed and widely understood name. Calculated variables should have an appropriate name, and a label that clearly identifies the variable to future analysts.

4.4.3.2 Types of variables

The different types of data collected in a household survey, including age, sex, income, assets, and names of districts/provinces, can be categorized according to their measurement scale. Four measurement scales are generally used in statistics: nominal, ordinal, interval, and ratio. Nominal and ordinal variables are considered to be **categorical variables**, while interval data and ratio variables are considered **continuous variables.**

Categorical variables

A categorical variable is one for which each response can be put into a specific category. The categories are usually labelled and coded. These categories must be both mutually exclusive and exhaustive. Mutually exclusive means that each possible survey response should belong to only one category, whereas, exhaustive requires that the categories cover the entire set of possibilities. If the age categories are 0–6, 7–12, 13–18, they are mutually exclusive, as a person can never be in two of these categories at the same time. To be exhaustive, we have to add a category ("19 or more") so that all possible cases are covered. Categorical variables can be either nominal or ordinal.

Nominal: A nominal variable describes a name or category. Contrary to ordinal variables, there is no "natural ordering" of the set of possible names or categories. Sex, household status, and type of dwelling are examples of nominal variables. Another example is type of crop, which could be categorized as 1 = wheat, 2 = rice, 3 = maize, 4 = sorghum, 5 = millet, 6 = other. Nominal variables cannot be analysed using means; the mode can be used.

Ordinal: Ordinal variables order (or rank) data in terms of degree. Ordinal variables do not establish a numeric difference between data points. They indicate only that one data point is ranked higher or lower than another (Shawna, J. *et al.* 2005). They are not customarily analysed using means. The variable "food consumption group," for example, is ordinal because the category "acceptable food consumption" could be considered better than the category "poor food consumption." There is some natural ordering, but it is limited since we do not know by how much "acceptable food consumption" is better than "poor food consumption."

Example: Variable: Food consumption group, where:

1 = poor food consumption
2 = borderline food consumption
3 = acceptable food consumption

Binary: A special type of categorical variable is a dichotomous/binary variable, which is a nominal variable consisting of only two categories (or levels). Observations can be classed into two groups: male/female, group 1/group 2, true/false, yes/no. All cases having a certain characteristic. For example, "household has female head" could be coded with a value 1, for "yes", while all other cases without that characteristic could be coded with a value 0, for "no". Coding 1/0 allows calculations, which are normally not possible with a nominal variable. For example, the percentage of cases having the given characteristic (e.g. percentage of female household heads) corresponds with the average of the variable.

Continuous variables
A variable is said to be continuous if it can take an infinite number of real values. Continuous variables are interval or ratio variables.

Interval variables
The numbers assigned to objects have all the features of ordinal measurements, and in addition, equal differences between measurements represent equivalent intervals. That is, differences between arbitrary pairs of measurements can be meaningfully

compared. Operations such as addition and subtraction are therefore meaningful. The zero point on the scale is arbitrary; negative values can be used. However, ratios between numbers on the scale are not meaningful, so operations such as multiplication and division cannot be carried out directly. For instance, the phrase "today it is 1.2 times hotter in degrees Celsius than it was yesterday" is not very useful or meaningful; in degrees Fahrenheit it might be 1.4 times hotter. Stating that the birth year of person A is 5 percent higher than the birth year of person B is also not useful or meaningful.

The central tendency of a variable measured at the interval level can be represented by its mean, median, or mode, with the mean giving the most information. Variables measured at the interval level are called interval variables, or sometimes scaled variables, though the latter usage is not obvious and is not recommended. Examples of interval measures include temperature in Celsius scale or Fahrenheit scale.

Ratio variables

A ratio variable, has all the properties of an interval variable, but also a clear definition of 0.0. When the variable equals 0.0, there exists none of that variable. Variables such as height and weight are ratio variables. Operations such as multiplication and division are therefore meaningful. The zero value on a ratio scale is non-arbitrary. The central tendency of a variable measured at the ratio level can be represented by its mode, its median, its arithmetic mean, or its geometric mean, as with an interval scale.

Table 4.22: Summary of statistical measures by type of variable

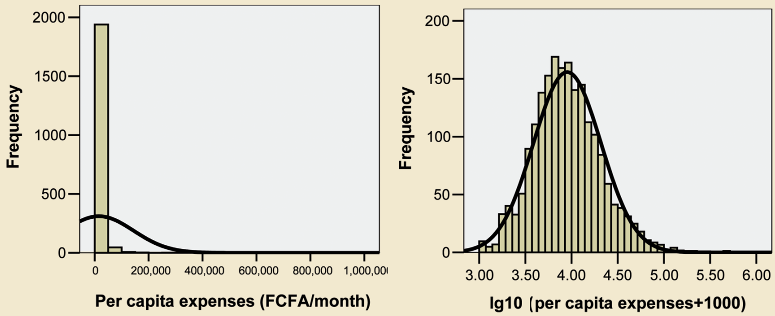
Okay to compute	Nominal	Ordinal	Interval	Ratio
Frequency distribution	Yes	Yes	Yes	Yes
Median and percentiles	No	Yes	Yes	Yes
Add or subtract	No	No	Yes	Yes
Mean, standard deviation, standard error of the mean	No	No	Yes	Yes
Ratio, or coefficient of the variation	No	No	No	Yes

A continuous variable can be categorized into a categorical variable to facilitate months: 24–35 months, 36–47 months, and 48–59 months. The actual age of the children collected in a survey can be categorized into these groups by giving a code such as 0–5 months = 1, 6–11 months = 2, and so on.

Distributions of continuous variables

The distribution of continuous variables should be considered during analysis. Most common procedures and statistics assume that variables are normally distributed. When the distribution of a continuous variable is highly non-normal, alternative data summaries should be used (e.g. median and not mean), and tests of significance that do not assume normality should be used (non-parametric tests). Sometimes, variables can also be transformed to have a more normal distribution. This transformed variable can then be used, for instance in a regression or a principal components analysis (PCA).

Figure 4.9: Distribution of per capita expenditure



Expenditures, revenue, and production typically have a skewed distribution, as in this example from the Cameroon CFSVA. The mean is 15,500 FCFA/month per capita, however, a few households have monthly per capita expenditures reaching 500,000 to 1,000,000 FCFA/month.

Using the per capita expenditures in a regression model or correlation, or running a t-test or ANOVA to compare means, would violate the assumption of normal distribution. Hence a logarithmic transformation was applied, and the resulting distribution resembles the normal. The transformed variable can now be used in correlations, regressions, PCA, etc.

Constructing indicators: Transformations

Simple or complex transformations are needed to create many of the key CFSVA indicators. Mathematical operations in the compute procedure, categorizing of values in the recode procedure, and other transformations possible in SPSS allow for the combining of different variables or the reconfiguration of variables into desired indicators.

Changing values of a variable

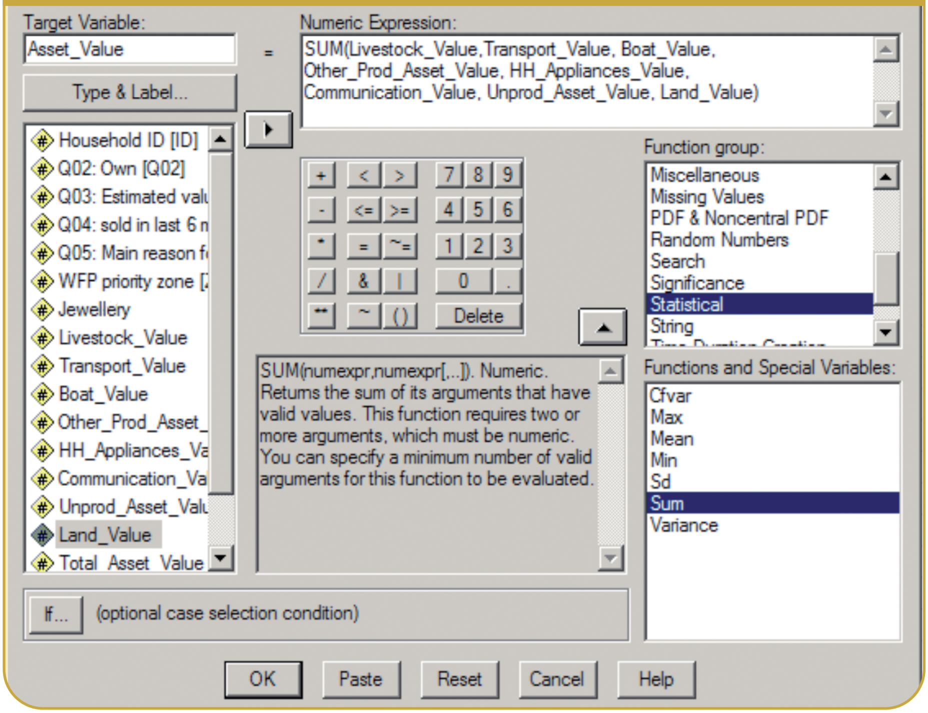
The “Recode” command in SPSS is generally used to change the values in a variable. Variables can be recoded into the same variables or into different variables. Generally, it is recommended that recoding be done into a different variable so that the original data is not lost, particularly in the case of an error in recoding. This command can be used to recode one categorical variable into a new categorical variable, or one continuous variable into a new categorical variable.

It is worth aggregating categories or transforming a continuous variable into a categorical when the list of possible answers is very long (e.g. the relationship of the HH members with the HH head), and/or when some answers have been chosen by few households.

Calculating new variables

New variables can be calculated using the “Compute” command in SPSS. To help compute new variables, SPSS has a number of mathematical operations, including addition, subtraction, multiplication, and division. However, the “Addition” command does not work if the variables added contain missing values. Using the “Sum” command (Figure 4.10) addresses the problem.

Figure 4.10: Computing a new variable using "sum" command



If the analyst adds variables that contain missing values, it is recommended that he/she not use the "Addition" command (+). If it is used, the sum will have a missing value every time there is a missing value in one of the added variables. The command "Sum" would treat all the missing values as if they were "zero," thus not increasing the number of missing values in the variable "Sum."

"Logical operators" can be used to set up conditions ("If" command in SPSS) to create a new variable. Box 4.13 includes a list of the logical operators available in SPSS.

With careful use of these operators, new variables can be constructed for CFSVA analysis. For example, CFSVAs often collect age of children, which is a continuous variable. However, analysts typically create age categories to generate cross tables with other variables like enrolment or drop out.

Box 4.13: Commonly used logical operators in SPSS

< less than
> greater than
<= less than or equal to
>= greater than or equal to
= equal to
~= not equal to
| or
and logical "and"

Box 4.14 demonstrates the use of logical operators to create age groups from a variable called "childage."

Box 4.14: Use of logical operators in SPSS

```
If (childage <= 7) agegroup = 1
If (childage >7 and childage <= 12) agegroup = 2
If (childage >12) agegroup = 3
```

However, be careful of missing value codes. For example, if "no answer" is coded as 99, this could be miscategorized as 3, meaning a child more than 12 years old.

The compute command can also use more advanced mathematical functions. For instance, the square root or the logarithmic transformation could be used to normalize a skewed distribution, and TRUNC or RND can be used to categorize continuous variables.

Box 4.15: Some useful mathematical functions in SPSS

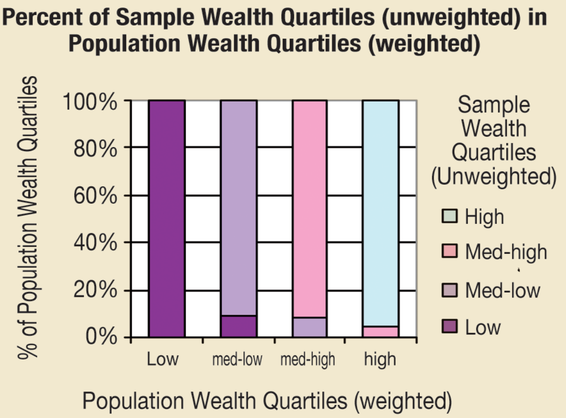
ABS(var)	the absolute value of a variable ABS 13.8=13.8; abs(-13.8)=13.8
RND(var)	the rounded value of a variable: RND(13.8)=14
TRUNC(var).	the truncated value of the variable: TRUNC(13.8)=13
SQRT(var)	the square root of the variable
lg10(var)	the base 10 logarithm of a variable
ln(var)	the natural logarithm of a variable
exp(var)	e raised to the power of the variable

Calculating n-tiles

N-tiles (usually quintiles) can be calculated automatically using SPSS. Under "Transform," use "Rank cases." Under "Rank types," select "n-tiles," and indicate the number of tiles desired, then continue. To deal with tied ranks, usually the mean is used (under "Ties," select "Mean"), although there may be circumstances where other methods should be used.

Most CFSVA data sets use probability weights, hence n-tiles should be calculated with weights on, so that (in the case of quintiles) 20 percent of the weighted sample lies in each quintile, allowing the quintiles to be applied to the population they are representing.

Figure 4.11: Example of incorrect classification of HHs into wealth quartiles



Based on various assets, a wealth index was calculated in the Cameroon CFSVA. When creating quartiles of this index, using the sampling weights, the results can be extrapolated to 25 percent of the population being classified in each quartile. If, however, the weights are turned off when creating the quartiles, 25 percent of the sample household are classified in each "sample wealth quartile," which does not coincide with population quartiles.

The example from the Cameroon CFSVA shows that if the weights are not turned on, 6 percent of rural households are classified in the wrong category.

COMPUTING RATIOS

At household level

Ratios are simply calculated using the "compute" function. Particular care must be taken with 0 values (division by 0 generates a missing value), and that missing values (such as 99,888) are coded and recorded as "missing" in the variable view.

At aggregate level

Although most ratios can be calculated at the household level, it makes more sense for some to be computed at the aggregate level. The enrolment rate at the household level, for example, is often 1 or 0 or missing (when there are no school-age children), whereas the enrolment rate of an entire subgroup is more meaningful. Remember, a statement for a ratio calculated at the household level is different from the same statement for a ratio calculated at the aggregate level, and should be reported as such.

Box 4.16: Example of aggregate ratio vs. household ratio

Calculating dependency ratio in both ways:
Household A has 4 children and 2 productive adults. The dependency ratio is 4 to 2, or 2.
Household B has 1 child, 1 elderly person, and 8 productive adults. The dependency ratio is 2 to 8, or 0.25.
The AVERAGE household dependency ratio = ((2 + 0.25)/2) = 1.125
The AGGREGATE dependency ratio = (sum of all children and elderly/sum of all adults) = (6/10) = 0.600

In order to calculate a ratio at the aggregate level, the values of the variables for the denominator and the variable for the numerator are first aggregated. For example, if calculating school attendance, the sum of all the children attending school in the sample needs to be calculated, and then the sum of all the school-age children needs to be calculated. Then the ratio for that level of aggregation is calculated, using the aggregates in the denominator and numerator. Alternatively, the average for the sample of the numerator and the denominator can be calculated independently, and then these averages are divided to achieve the aggregate ratio. The confidence interval of certain rates can be wide, especially for subgroups. It is therefore important to estimate the error correctly.

Calculation of anthropometric indicators (z-scores)

Epi Info, from the CDC, and Anthro, from WHO, are the two most commonly used applications for anthropometric data analysis. The current standard recommendation for CFSVAs is to calculate and report under-5 anthropometry (stunting, wasting, underweight) using both the NCHS and the new WHO references. Analysis of nutrition should then use the WHO reference data. Only the software WHO Anthro 2005 can currently calculate z-scores using both reference scores.

4.4.3.3 Descriptive statistics

Descriptive statistics are those used to describe characteristics of a sample or population, and they involve exploring the distribution of one variable (frequency) or the distributions between two or more variables (cross tabs). In SPSS, descriptive statistics can be most easily produced by the Frequencies command. Together with simple graphics, they form the basis for virtually every quantitative analysis of data.

Means

a) Simple mean of continuous variables

The (arithmetic) mean is the sum of all the values divided by the number of cases (considering only valid cases and excluding missing cases). It can be used for continuous data (that is, data measured on an interval or ratio scale). The mean is a measure of the variable's central tendency. Statistics such as mean (and standard deviation, defined further on) assume a normal distribution and are appropriate for quantitative variables with symmetric, or normal, distributions. Typical means calculated in CFSVAs include number of household members and monthly income.

b) Value codes of binary variables

While defining the coding of variables, it is recommended to use 1 for "yes" or "present," and 0 for "no" or "not present." In a household survey, the population mean of a variable coded in this way corresponds with the proportion of households that answered "yes" or where the reply was "present," and when multiplied by 100, gives the percentage prevalence of "yes."

Similarly, one could agree to specify 0 for "male" and 1 for "female." However, for the sex of children, to be used in anthropometric data transformations,[66] "male" should always be male = 1 and female = 2. In this case, means cannot be used easily to calculate proportions and percentages.

Medians

The median is the middle of a distribution when the values are ranked from highest to lowest, meaning half of the values are above the median and half are below the median, the fiftieth percentile, i.e. the middle value of a set of observations ranked in order.

66. This is the convention for Epi Info from the CDC and Anthro from WHO.

If there is an even number of cases, the median is the average of the two middle cases when they are sorted in ascending or descending order. The median is a measure of central tendency not sensitive to outlying values, unlike the mean, which can be affected by a few extremely high or low values.

The median is a robust statistic, appropriate for quantitative variables that may or may not meet the assumption of normality. It is preferable to use the mean when the data are not normally distributed. For example, some expenditure and income data have very skewed distributions, and the median may be a better summary than the mean. The median can be used with measurement scales that are at least ordinal (that is ordinal, interval, or ratio).

Modes

Modes are rarely used in describing CFSVA indicators. However, they may occasionally be of use. The mode is defined as the most frequent variable value. In the set of values 1, 2, 3, 3, 3, 4, 5, 5, 6, 6, 7, 8, 9, the mode is 3, as it appears more frequently than any other value. There can be more than one mode. Using the mode is more appropriate when there are only a few possible values of the variable (for instance, to describe the household size, the mode could be used). The same is true for categorical variables (for example, "farming" is the most common livelihood strategy of the households.

PERCENTAGES/PROPORTIONS

Frequencies

Frequencies is one of the more common descriptive functions used in the analysis of CFSVAs. They are most commonly used to produce global prevalence (prevalence for the whole data set). SPSS gives two prevalence results: percentage and valid percentage. Percentage includes the missing cases (system missing and those coded as missing) in the denominator, whereas valid percentage excludes the missing cases and includes only the cases with data in the denominator. The analyst needs to interpret which of these to report. If the missing values are assumed to be not different from the values with valid responses, the "valid percentage" statistic should be given; if the missing values are different (nonexistent options, not applicable), the analyst should consider reporting the "percentage" as compared to the total population. Valid percentage is the most common; however, in certain cases percentage may be more relevant.

Box 4.17: Example highlighting the difference in interpretation

In this example from Laos, households were first asked if they cultivated land in the last year. If they responded "no," then the enumerators skipped the rest of the agriculture section, leaving the questions blank. If the response was "yes," then the following question was asked: "What was your main crop cultivated in the past year?" In the data entry, this question on the main crop was left blank (system missing) if there was no response. A frequency of the main crop cultivated results in the following table:

(cont...)

(...cont)

Main crop cultivated

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Glutinous rice	2827	72.0	82.1	82.1
	White rice	365	9.3	10.6	92.7
	Maize	65	1.7	1.9	94.6
	Beans	4	0.1	0.1	94.7
	Cassava	3	0.1	0.1	94.8
	Vegetables	11	0.3	0.3	95.1
	Fruits	19	0.5	0.5	95.6
	Tobacco	1	0.0	0.0	95.7
	Groundnuts and other nuts/seeds	6	0.2	0.2	95.8
	Other	143	3.7	4.2	100.0
	Total	3444	87.7	100.0	
Missing	System	482	12.3		
Total		3926	100.0		

When looking at glutinous rice, two statements can be made. Using the percentage, it can be stated ”72 percent of all households cultivated glutinous rice as their main crop in the past year.” This statement is of ALL households. Using valid percentage, it can be stated “of the households practicing agriculture in the last year, 82 percent cultivated glutinous rice as their main crop.”

Cross tabulations

Cross tabulations are another way of exploring frequencies, and are one of the most common descriptive tools used in CFSVA analysis.

Unlike frequencies, cross tabs include only the valid cases. When calculating prevalence in SPSS, three options are commonly used: percentage rows, percentage columns, and percentage total. This will determine how SPSS calculates the prevalence in each cell. For each cell, the numerator remains constant: the number of valid cases belonging to the two groups. The denominator in percentage rows is the total number of cases in the row cells. The denominator in percentage columns is the total number of cases in the column cells. The denominator in total percentage is the total number of cases in ALL cells, which is equal to the valid number of cases in the data set.

Understanding this difference between percentage rows, columns, and choosing the correct one to report is critical. As a general rule, the “independent variable” should be put in the columns and the “dependent” in the rows; column percentages are more important than row percentages. For instance, we can look at the influence of wealth on food consumption. If we put wealth in the column, we should focus on the column percentages and compare them across the rows (food groups).

In Table 4.23, three food consumption groups were cross tabulated with the quintiles of wealth score (the first being the poorest, the fifth being the richest).

Table 4.23: Example of cross tabulation

			Quintiles of wealth score					
			1	2	3	4	5	Total
Food consumption groups	Poor Food Consumption	% within Food consumption groups	67.1%	13.4%	8.5%	3.7%	7.3%	100.0%
		% within Quintiles of wealth score	7.1%	1.4%	0.9%	0.4%	0.8%	2.1%
		% of Total	1.4%	0.3%	0.2%	0.1%	0.2%	2.1%
	Borderline Food Consumption	% within Food consumption groups	38.4%	26.8%	17.8%	11.9%	5.1%	100.0%
		% within Quintiles of wealth score	20.3%	14.2%	9.3%	6.3%	2.7%	10.6%
		% of Total	4.1%	2.8%	1.9%	1.3%	0.5%	10.6%
	Acceptable Food Consumption	% within Food consumption groups	16.6%	19.2%	20.6%	21.4%	22.2%	100.0%
		% within Quintiles of wealth score	72.6%	84.4%	89.8%	93.3%	96.5%	87.3%
		% of Total	14.5%	16.8%	18.0%	18.7%	19.4%	87.3%
Total		% within Food consumption groups	20.0%	19.9%	20.1%	20.0%	20.0%	100.0%
		% within Quintiles of wealth score	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	20.0%	19.9%	20.1%	20.0%	20.0%	100.0%

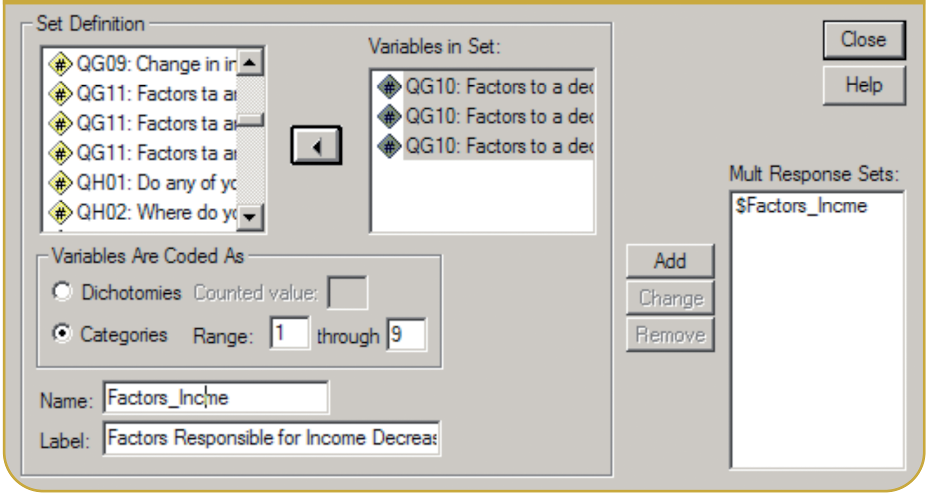
Table 4.23 shows the relationship between the food consumption of households and wealth status. It shows that food consumption increases stepwise by wealth quintile. More than 67 percent of households in the first wealth quintile reported poor food consumption, while 22.2 percent of households in the fifth wealth quintile and 21.4 percent of households in the fourth wealth quintile reported good food consumption.

4.4.3.4 Analysing multiple responses

A number of questions in CFSVAs call for multiple responses from the respondents. For example, a typical CFSVA questionnaire includes questions about major crops cultivated, main income sources, and food sources. All these questions generate multiple answers. A household might have obtained its food from different sources, including food bought from the market, produced, received as food for work, and borrowed from neighbours. It is a common practice to create separate variables for each of these answers, and these answers are mutually exclusive

The multiple response feature in SPSS allows one to analyse variables taking multiple responses into account. The first step is to define variable sets. All the variables containing (multiple) responses should be inserted into the "Variable in Set." The next step is to select the type of variables included in the variable set and their range of values. After giving the name of the "Multiple Response Set," add the variable set to the Multiple Response Set. Now the multi-response set is ready for analysis. To analyse, go to "Multiple Response," then "Analyse," and select the desired analysis.

Figure 4.12: Define multiple response sets



It is a good practice to paste the command to a syntax file for future reference. This will also help the analyst if she/he wants to regenerate the table.

Box 4.18 presents syntax to compute percentages. It essentially computes percentages of responses and percentages of cases.

Box 4.18: Syntax to compute percentages

```
mult response
groups=$Income_decline_factors ‘Factors responsible for income decrease’ (QG10_V1
QG10_V2 QG10_V3 (1,9))
/frequencies=$Income_decline_factors.
```

The output of the analysis is presented in Table 4.24. The column that presents percentage of responses calculates percentage of total responses. For example, in this case 623 households responded to the question “What are the factors that led to a decrease in your household income?” A household could answer 3 different responses from a list of 9. Sixty-five households identified 3 factors, 194 households identified 2 factors, and 364 households identified only 1 factor responsible for income decrease.

Altogether, 623 households responded to this question with 947 responses. The percentage of responses column calculates percentage of responses and the percentage of cases column calculates the percentage of households that responded to this question (623 in this example).

Table 4.24: Output of the analysis

		Responses		Total
		N	Percent	
Factors responsible for income decrease	Loss of employment	124	13.1%	19.9%
	Loss of crop/animal	65	6.9%	10.4%
	Prolonged illness of income earner	152	16.1%	24.4%
	Death of income earner	48	5.1%	7.7%
	Decrease in remittance income	8	0.8%	1.3%
	Loss of asset	147	15.5%	23.6%
	Exposure to natural disaster	124	13.1%	19.9%
	Market failure	157	16.6%	25.2%
	Other	122	12.9%	19.6%
Total		**947**	**100.0%**	**152.0%**

4.4.3.5 Measures of variation

Variance and standard deviation

Variance is a measure of dispersion around the mean of a continuous variable, equal to the sum of squared deviations from the mean divided by one less than the number of cases (degrees of freedom). The variance is, therefore, the average squared distance between the mean and the observations made (and so is a measure of how well the model fits the actual data). However, the variance is measured in units that are the square of those of the variable itself. The **standard deviation** is obtained by calculating the square root of the variance.

The **standard deviation** of a distribution, a measure of dispersion based on a deviation from the mean (which are squared, summed, and averaged and then the square root is taken), has the same unit as the original observations and can be used for data measured on an interval or ratio scale. A large standard deviation (relative to the mean; also called coefficient of variation) indicates that the data points are distant from the mean. In this case, the mean may not be an accurate representation of the data. A standard deviation of 0 would mean that all the scores were the same.

In a normal distribution, 68.27 percent of cases fall within one standard deviation on either side of the mean, 95.45 percent of cases fall within two standard deviations, and 99.73 percent fall within three standard deviations. For example, if the mean age is 45, with a standard deviation of 10, 95 percent of the cases would be between 25 and 65 in a normal distribution.

Confidence intervals

Confidence intervals enable analysts to make statements on the precision of their estimates. For example, if 30 percent of households were observed in the representative sample to be female headed, confidence intervals could be added to state that "the analysts are 95 percent sure that between 25 and 37 percent of households (in their sampling universe) are female headed." Confidence intervals should always be used in CFSVAs when reporting highly standardized indicators such as stunting, wasting, underweight, low BMI, and low MUAC. Confidence intervals can be used in CFSVAs when reporting key indicators such as percentage of food insecure, percentage of poor food consumption. Confidence intervals are not typically reported for descriptive indicators in the text of a CFSVA; however, one should strive to include them in annex tables.[67]

4.4.3.6 Tests of significance

Significance is a statistical term that indicates how sure the researcher is that a difference or relationship exists. Tests of significance help the researcher and the audience know if differences between groups are real or by chance. When a statistic is significant, it simply means that one can be very sure that it is reliable and can be referred to the entire population. It does not mean the finding is important or that it has any decision-making utility. This significance, produced by the statistical tests discussed here, is referred to as the p-value (probability value).

The p-value can be interpreted as the probability of a difference occurring by chance alone. If all other biases are eliminated or accounted for, then one can assume that when this p-value is small, the differences are due to a factor other than chance. The cut-off for significance most often used is 0.05. If a p-value is less than 0.05, then assume that the relationship observed is real not by chance. Usually p-values are reported by their actual value to three decimal places, or as >0.05, <0.05, <0.01, or <0.001. Significance levels, when appropriate, are usually reported in the body of the report and in the annex tables.

In this section, some of the more commonly used statistical tests are presented. However, there is a wealth of further statistical tests, many available in SPSS. For a more complete guide to tests of significance, see Discovering Statistics Using SPSS,[68] or any other statistical manual or textbook.

It is very important to note that CFSVAs employ cluster sampling methods, which require special analytical approaches to calculating significance levels and confidence intervals. Standard packages such as the basic SPSS package, do not compute accurate p-values for surveys that are sampled using a cluster design. The appropriate statistical analyses can be obtained using the SPSS Complex Samples module or other special software.

Table 4.25 provides some guidance on what tests of significance to use when comparing different types of data. Keep in mind that this is a generic list and should be used only as a guide. It is the analyst who should decide which test is appropriate for what analysis based on a number of factors.

67. Automatic applications exist for computing confidence intervals for percentages. It is good practice to report them, at least for key indicators.
68. A. Field, 2005, *Discovering Statistics Using SPSS*, 2nd ed., London: SAGE Publications Ltd.

Table 4.25: Example of test of significance for different types of variables

	Dependent variable	Independent variable	When to use	Example	Procedure
Independent T-test	Continuous	Categorical binomial	To compare differences in the means of two groups (identified by the categories of the binomial variable) To see if the difference is statistically significant ($p<0.05$)	Compare the mean z-scores of male and female children	Run the independent samples T-test; Report the two means; Check if the T value is statistically significant ($p<0.05$)
One-way ANOVA: Post-hoc Multiple Comparisons	Continuous	Categorical	To compare differences in the means of three or more groups (identified by the categories of the categorical variable)	Compare the mean z-score by residence status (IDP, refugee, or resident HHs)	Run the One-Way ANOVA post-hoc procedure Check if the categorical variable explains in a significant way some of the observed variation through the F-test. Check which differences are statistically significant ($p<0.05$) through the post-hoc tests (e.g., REGWQ, Tukey HSD, Games-Howell, etc.)
Chi-square	Categorical	Categorical	To detect whether there is a statistically significant association between two categorical variables	Explore the association between food consumption groups and ethnic groups	Compute the Chi-square and report the value Check if the value is statistically significant ($p<0.05$) (The Chi-square helps determine whether the association is statistically significant)
Bivariate Correlation	Continuous	Continuous	To assess the general association between two variables (i.e.,one variable increases/decreases when another increases/decreases)	Correlation between children's height and weight	Compute the Pearson Correlation Coefficient and report the value Check if the correlation is statistically significant (two tailed tests) ($p<0.05$)
Simple Linear Regression	Continuous	Continuous/ Categorical binomial (0/1 values)	To measure how the dependent variable changes with a one-unit increase in the independent variable	Regressing food consumption score by wealth index	Run Simple Linear Regression Model Report R^2 adjusted, B value Check and report if B is statistically significant ($p<0.05$)
Multiple Linear Regression	Continuous	Two or more continuous/ categorical binomial (0/1 values)	To measure how the dependent variable changes with a one-unit increase in the independent variable (controlling by the other variables in the model)	Regressing food consumption score by wealth index and gender of the HH head	Run Multiple Linear Regression Model Report R^2 adjusted, B values Check and report if B values are statistically significant ($p<0.05$)
Multivariate General Linear Model (GLM)	Continuous	2 or more continuous variable and/or 2 or more categorical variables	GLM combines ANOVA and Regression to analyse the effects of more than one independent variable on the dependent variable (and to see how these independent variables interact)	Analyse the effects of ethnic group, province, and wealth index on the food consumption score	Run a Multivariate GLM Interpret the output from main ANOVA table Report R^2 adjusted, B values
Logistic Regression	Categorical	Two or more continuous variables and/or two or more categorical variables	To predict the probability of an event occurring for a given household	Predict which households are more likely to be food insecure according to the province of residence and WI	Run a Logistic Regression Check the overall fit of the model Check which variables significantly predict the outcome (in SPSS, check table “Variables in the equation”)

CHAPTER 4

In a typical CFSVA, the most commonly used tests of significance include the Chi-square test, z-test, t-test, and the ANOVA. For further information on tests of significance, consult a statistics manual.

4.4.3.7 Multivariate analysis

Multivariate analysis in statistics describes a collection of procedures involving analysis of more than one statistical variable at a time. In design and analysis, these techniques are used to perform studies across multiple dimensions while taking into account the effects of all variables.

Regression

Regression analysis is a technique used for the modelling and analysis of numerical data consisting of values of a dependent variable (response variable) and of one or more independent variables (explanatory variables). The dependent variable in the regression equation is modelled as a function of the independent variables, corresponding parameters ("constants"), and an error term. The error term is treated as a random variable. It represents unexplained variation in the dependent variable. The parameters are estimated so as to give a "best fit" of the data. Most commonly the best fit is evaluated by using the "least squares method," but other criteria have also been used.

Data modelling can be used without there being any knowledge about the underlying processes that have generated the data; see http://en.wikipedia.org/wiki/Regression_analysis—cite_note-Berk-0#citenote-Berk-0; in this case the model is an empirical model. Moreover, in modelling, knowledge of the probability distribution of the errors is not required. Regression analysis requires assumptions to be made regarding probability distribution of the errors. Statistical tests are made on the basis of these assumptions. In regression analysis, the term model embraces both the function used to model the data and the assumptions concerning probability distributions.

Regression can be used for prediction (including forecasting of time-series data), inference, hypothesis testing, and modelling of causal relationships. These uses of regression rely heavily on the underlying assumptions being satisfied.

Regression analysis is complex, and therefore cannot be adequately covered in these guidelines. However, a few key concepts and guiding principles are presented here. CFSVAs often, but not always, use regression analysis. Nutritional analysis makes frequent use of regression techniques.

a) When is regression used for CFSVAs, and why?

In CFSVA, regression is primarily used to understand causal relationships between the variables that are important for decision-making purposes. For example, to explore the relationships between stunting (height-for-age, an indicator of chronic malnutrition) and dietary quality while controlling for sanitation facilities, access to potable water, mother education, and household income, the analyst may want to set up an OLS (Ordinary Least Square) model, where height-for-age is a dependent variable and all of the other variables mentioned could be explanatory variables.

However, before estimation of the model, the analyst has to test for:

- **multicollinearity:** one or a combination of explanatory variables is strongly correlated to another explanatory variable;
- **heteroskedasticity:** the variance of the error terms changes across observations;
- **specification error:** the model is wrong, by missing important explanatory variables or by having other incorrect assumptions; and
- **endogeneity:** when an explanatory variable is itself a function of the dependent variable.

Necessary correctional measures need to be taken if the tests identify multicollinear variables, heteroskedasticity or omitted variables. If any of the explanatory variables is found to be endogenous, the variable has to be replaced by suitable instrumental variables. It is important to understand that a simple regression involving only the dependant variable and one independent variable is similar to a Pearson correlation.

b) Why control for other factors – confounders?

Confounders refer to factors that relate both to the dependent (outcome) and independent variable of interest. For example, we can hypothesize that children under 5 tend to be more underweight (low weight for age z-score) in female-headed households. We can run a simple compare means, or a t-test, to explore the differences in mean z-score between male- and female-headed households. However, the critical question is whether the sex of the household is the only factor responsible for the nutritional status of the children in the household? In a regression analysis, one could enter both education level and sex of the head of the household. It may be found in this analysis that the head of household no longer has a significant effect on underweight z-score, but that education level does. This could likely arise because female heads of household, in this example, often have lower education levels than male heads of households. The regression analysis controls for the difference in sex when it estimates the effect of education, and vice versa. Hence we can say "controlling for education, sex of household head is not significantly related to underweight z-score."

Female-headed households is still important as an identifier of vulnerable households for targeting, since gender may be more easily identified than education level, even if we know that the real reason is that most female household heads have a low education level.

c) Why explore interactions?

Interactions illuminate how a cause of food insecurity might be modified by another variable. For example, if we look at sanitation, water source, and underweight status of children under 5, we might find that improving sanitation has no effect unless in the presence of a safe source of water. In the regression, the two variables (sanitation and water) are simply multiplied together to give the interaction term. Environmental variables, economic factors, education, and age are common effect modifiers (variables that result in statistical interactions).

d) Coefficient of determination

In linear regression models, R^2 is a statistic that gives some information about the merit or fit of a model. R^2 is the square of the correlation coefficient between the dependent variable and the estimate it produced by the independent variables, or equivalently defined as the ratio of regression variance to total variance. It is a measure of determination of how well the regression line approximates the real data points. An R-squared of 1.0 indicates that the regression line perfectly fits the data, and 0.0 indicates that one term does not help to know the other term at all. For regression applied to household surveys, it is normal to find an R^2 between 0.15 and 0.25, or (exceptionally) a little higher.

Principal component analysis and cluster analysis[69]

This section discusses the following two multivariate analysis techniques:
- Principal component analysis (PCA) – which belongs to the factor analysis family; and
- Cluster analysis – which belongs to the classification family.

Both techniques can typically be used to reduce the complexity of the data set for exploratory purposes: factor analysis, to reduce the selected variables into a lesser number of factors; and cluster analysis, to group all cases into a number of groups.
Detailed explanations of how PCA and cluster analysis techniques work are beyond the scope of these guidelines; interested readers should refer to specialized textbooks for more information. This section presents a simplified summary of the statistical ideas on which PCA and cluster analyses are currently used in CFSVA and FSMS, as well as of the terminology used throughout those analyses.

In addition to SPSS, several commercial statistical software packages perform both PCA and cluster analysis. With the support of WFP and FAO, VAM typically also uses a software developed explicitly for socio-economic and food security analysis (ADDATI, or the brand new Windows version ADDAWIN). This software was designed for the use of food security specialists. It includes preselected algorithms proven to be suitable when analysing socio-economic and nutrition data for food security and vulnerability analyses. It uses the output of the PCA as the input for the cluster analysis and facilitates the final interpretation of the outputs providing the cluster results in terms of the original input variables.

ADDATI/ADDAWIN cannot perform factor analysis with rotation. For this type of multivariate analysis, the software normally used in VAM is SPSS.

Principal component analysis

Factor analysis is used to study the patterns of relationships among many dependent variables, with the goal of discovering the underlying variations that affect them. The inferred underlying variables are called factors. Principal component analysis (PCA) uses a factor extraction method to form uncorrelated linear combinations of the observed variables. The first component explains maximum variance. Successive components explain progressively smaller portions of the variance and are all uncorrelated with each other.

PCA is one technique of multivariate analysis that applies to continuous variables. The objective of PCA is twofold:

69. This chapter strongly benefits from inputs and quotations from: WFP/VAM, *Household Food Security Profiles (Thematic Guidelines)*, April 2005; S. Griguolo, *ADDATI Users' Manual*, July 2003, IUAV; S. Landau and B.S. Everitt, *A Handbook of Statistical Analyses Using SPSS*, 2004; Chapman & Hall/CRC, Andy Field, Discovering Statistics using SPSS, 2005; SAGE ADDATI help; and SPSS help.

- to identify and describe the underlying relationships among the variables by creating new indicators (called “factors” or “principal components”) that capture the essence of the associations between variables; and
- to reduce the complexity of the data, saving a limited number of these new variables that is sufficient to keep the most relevant aspects of the description with a minimal loss of detail.

PCA yields as many principal components as there are initial variables. However, the contribution of each principal component to explaining the total variance found among all variables will progressively decrease from the first principal component to the last. As a result, a limited set of principal components explains the majority of the matrix variability, and principal components with little explanatory power can be removed from the analysis. The result is data reduction with relatively little loss of information.

It is recommended to use the rotated solution in most circumstances. Rotation of the result will give a new solution: the new factors explain the same variance and can be much better interpreted as the underlying dimensions. The analyst can understand the “meaning” of each dimension and then decide which of the uncovered dimensions he wants to use in subsequent analyses.

Uses of PCA

Two of the most common uses of PCA in CFSVA analysis are briefly described here:

Scoring on the first principal component (single factor solution)

PCA might be used to build a synthetic composite index moving from more than one variable. In this case, the first principal component is taken as the new variable on which statistical units might be measured and ranked. For example, variables from a household survey that are associated with wealth (quality of housing, assets owned by the household, etc.) can be used to perform a PCA. Since what those variables have in common is related to wealth, the first component can be interpreted as a **wealth index**.

While the single factor solution is very straightforward and easy to interpret, its use should be limited to specific cases. If the variables are uni-dimensional, then an exploratory analysis showing a single factor grouping, where all items “hang together,” is supportive. If the variables used in the analysis have multiple facets or dimensions, a single factor solution would not be able to maintain a minimum description for all of the analysed statistical units. In this case, trying to capture the underlying variations of all of the input variables through a unique index would undermine the instrument’s construct validity. That is because the use of a small selection of measuring items (i.e. the information maintained by one factor only) could lead to false and confusing results that would not reflect the complexity of the original data.

Using principal component(s) as input variable(s) for follow-up analysis (specifically cluster analysis)

PCA creates a set of new variables or components that, being perfectly uncorrelated, explain different portions of the original total variance.

As one of the main purposes of PCA is to reduce the dimensionality of the data set, components are ranked by their decreasing contribution to explaining the total variance. It is hence possible to remove components with little explanatory power. If the main purpose of PCA is to describe statistical units on the basis of the relationships among selected variables, data reduction is a secondary objective. Furthermore, if the final aim is to cluster units based on those relationships, it is recommended that analysts keep as many principal components needed to capture up to 80 percent of the total variance. Such a high level of consistency with the original complexity of the data would ensure a good reflection of the relationships among variables. It would also guarantee that particular combinations of variables' values were maintained and not smoothed too much through a high data reduction approach.

When data reduction is the primary objective, the analyst may want to remove more components. One rule of thumb is that the Eigen value of each extracted component should be higher than 1; an alternative is to keep as many factors (after rotation) that still have a readily understandable meaning. For a subsequent clustering, the analyst can even exclude factors irrelevant for the clustering activity to be undertaken.
For example, average decadal rainfall, Normalized Difference Vegetation Index (NDVI), and other climatic data were used to conduct a factor analysis in Sudan. Four underlying, meaningful factors were retained. The third factor, related to rainfall during the dry season, was not used for clustering, since rain during that season has little economic importance.

Cluster analysis

Clustering data is a common technique for statistical data analysis. Clustering is the classification of similar objects into groups or, more precisely, the partition of a data set into subsets (clusters) so that the data in each subset share some common features, often proximity according to some defined distance measure.

Note that cluster analysis is an exploratory data analysis tool that aims at sorting different objects into groups such that the degree of association between two objects is maximal if they belong to the same group, and minimal otherwise.

Clustering is one of the key parts of many large data set analyses. In fact, it is not possible to analyse and describe the situation of each statistical unit (be it a household, a village, a district, a region, etc.) separately, since there might be too many. There is a clear need to identify main patterns of similar characteristics. Clustering involves some kind of subjectivity based on analysts' choices of specific methods for clustering.

In conducting cluster analysis, analysts are often faced with two questions: How many clusters are there in the data set? and What is the compactness (inertia) of each cluster?

How many clusters in the data set?

A given data set does not contain a definitive number of clusters. First, because cluster analysis involves a series of iterations performed by statistical software, there will be some variance in the number of clusters and the assignment of particular households to clusters each time the analysis is performed, depending on the initial "cluster seeds." Second, several different methods and algorithms can be used to produce clusters, and the number of clusters produced will vary depending on the type of clustering method used. For very large data sets, the partition method, with a random selection of the initial centres, seems to be most appropriate. Specific algorithms to improve the quality of a partition are implemented, being different in different software packages.

What is the compactness of each cluster (inertia)?

The measurement of the dispersion or compactness of each cluster is called **inertia or internal variance of the cluster.** The degree of inertia within and among clusters provides a useful means of determining the final number of clusters that best fits the data.

There are no standard thresholds indicating what level of inertia is good, acceptable, or poor, and the final decision remains with the analyst. However, the ratio between the inertia of the overall cloud (the dispersion found among all units in the dataset) and the inertia within each cluster should be maximized. Doing so ensures that the similarity among units belonging to the same cluster (e.g. within clusters) is high, while the similarity between clusters is very low (e.g. maximizing intra-cluster homogeneity and inter-cluster heterogeneity).

One of the strengths of ADDATI (or ADDAWIN) is that it incorporates a specific formula (objective function) to calculate the intra- and inter-cluster inertia as a measurement of partition optimality with a given number of clusters. In other words, it measures how compact a set of clusters is.

In addition, the clustering in ADDATI displays a graph that plots how the value of the objective function decreases when the number of clusters retained is increased. By inspecting this graph, the user can focus on one or more promising partitions, with a number of clusters within the range he/she would like to obtain and a value of the objective function sufficiently high. This tool indeed helps the analyst find a number of final clusters, which is a fair **trade-off** between the level of synthesis achievable (few clusters are always more convenient) and a significant level of homogeneity of characteristics within the clusters (provided by the value of the objective function that represents the rate of information maintained).

One of the common uses of cluster analysis in CFSVA is to create groups of households with similar food consumption patterns, for profiling purposes or for further analysis. Typically the principal components are used to create the clusters.

Cluster analysis is also used to categorize households that share similar livelihood strategies into **livelihood groups**. The aim of CFSVA livelihood grouping is not to replace a comprehensive livelihood analysis but to utilize livelihood strategies as a basis for classifying populations.

4.4.4 Key references: Household data analysis

- Chapman and Hall/CRC Andy Field. 2005. *Discovery Statistics Using SPSS*. Sage ADDATI help; and SPSS help.
- Griguolo, S. 2003. ADDATI User's Manual. IUAV, July.
- High, R. 2000. *Dealing with Outliers: How to Maintain Your Data's Integrity.* Computing News. UO Computing Center. University of Oregon.
- Landau, S., and B. S. Everitt. 2004. *A Handbook of Statistical Analyses Using SPSS*, http://cc.uoregon.edu/cnews/spring2000/outliers.html.
- Shawna, J., K. Marcus, C. McDonald, T. Wehner, and M. Palmquist. 2005. *Introduction to Statistics.* Writing@CSU. Colorado State University Department of English. Retrieved 12/31/2007 from http://writing.colostate.edu/guides/research/stats/.
- United Nations, Department of Economic and Social Affairs, Statistics Division. 2005. *Household Sample Surveys in Developing and Transition Countries.* Studies in Methods, Series F No. 96.

CHAPTER 5

Qualitative and community-level data in CFSVAs

This chapter is not intended to provide a general background on qualitative data. Instead, it focuses only on how to better use qualitative approaches in the context of a CFSVA. (General guidance on qualitative methods can be found in the EFSA Guidance Sheets numbers 10 and 11, available from WFP OMXF). Neither does it aim to provide a comprehensive list of topics/issues that can be explored with communities. Rather, it concentrates mainly on those topics that:
- are particularly relevant for food security analysis;
- fit well in the household Food and Nutrition Security Conceptual Framework of the CFSVAs; and
- can be studied better at the community level rather than the household level.

Qualitative techniques are used in a number of ways during a CFSVA, including in developing the household survey, collecting some types of community-level data, interpreting quantitative findings, and evaluating intervention options.

5.1 TYPES OF COMMUNITY-LEVEL DATA IN CFSVAs

Three instruments are typically used in CFSVAs to collect community-level data: focus groups discussions (FGDs), community discussions, and key informant interviews. Other techniques, such as observations and transect walks, are often used in EFSAs, but these are less structured and require a different kind of training.
This section provides a broad description of focus groups, community discussions, and key informant interviews. The specific role of these tools in the context of CFSVAs is also explained. Since focus groups and community discussions are qualitative tools, this section includes a general introduction to qualitative methods.

Box 5.1: Qualitative methods

Qualitative methods are typically used to collect in-depth information on the perceptions, judgments, and opinions of individuals. These methods undergo a continuous process of revision as an understanding of local perceptions develops.

The fundamental difference between qualitative and quantitative techniques is in the type of information they provide. While quantitative information is uniform and comparable, providing statistical breadth, it does not provide the level of contextual detail that enables an understanding of meanings, processes, and reasons. Qualitative information should therefore not be considered simply as an isolated supplement to be inserted into an assessment after quantitative data has been collected and analysed, but rather as an essential complement within the analysis, providing the interpretive data necessary to make sense of the descriptive data produced in the quantitative analysis. The use of this kind of mixed-methods approach allows for triangulation of qualitative and quantitative information regarding food security and vulnerability, and the associated programming constraints and opportunities (TANGO 2002).
When applying qualitative methods, the researcher becomes the instrument of data collection. Thus, results may vary greatly depending upon the researcher involved. Hypotheses and additional follow-up questions are generated during data collection and analysis, and measurement tends to be subjective. Therefore, by their very nature, the methods are often not objectively verifiable. Because qualitative methods involve direct communication with target populations, they typically employ a participatory approach, and rely on open-ended techniques. Interviewers and facilitators must be well trained to collect data with flexibility and respect.

- ***Focus group discussions***

Focus group discussions are a form of group interview. They differ from other group interviews in that both the selection of participants and the discussion are determined by a particular topic. FGDs are typically guided by a specific outline that is structured to be flexible enough to allow themes relevant to the community to emerge. Participatory techniques (e.g. mapping, listing, ranking) are frequently used during focus groups to ensure that all participants are involved in the conversation.

In the context of CFSVAs, FGDs are used to elicit information on themes related to food security. They involve a small number of community members (i.e., six to eight participants) selected according to criteria consistent with the CFSVA's theme.

For CFSVAs, focus groups are preferable to observation and open-ended interviews. FGDs are not as informal or as unstructured as simple observation (although for the latter, simple guides can be prepared that indicate what to look at and how to record it in a systematic way). Instead, they allow the team to record a large amount of relevant information within a limited period and are conducted according to a prearranged topical outline.

- ***Key informant interviews***

Key informant interviews are typically conducted using a prepared list of questions designed to elicit information from one or more individuals chosen for their in-depth knowledge of a specific issue concerning the community. Key informants often include traditional community leaders (chiefs, religious leaders, etc.), health workers, teachers, business people (in the case of market analysis), and local government representatives. Most of the key informant questionnaires used in CFSVAs use open-ended questions (i.e. questions to which informants cannot respond with a simple yes or no) and are used to gather data relevant to food security on which key informants will have particular knowledge (infrastructure, institutional relationships, agricultural trends, etc.).

- ***Community discussions***

A community group discussion involves a mixed group that includes men, women, and young people from all subgroups within the community: village, camp, urban neighbourhood (EFSA Handbook 2005). It is essential that a diverse array of individuals be represented in the discussion. Community discussions differ from focus groups in that the goal of a community discussion is to get answers on a subject related to the whole community and for which the CFSVA team would like an opinion that reflects that of the community. Such meetings are therefore open to all members of the community.

- ***Moderating team***

Focus groups and community discussions require the presence of one facilitator (or moderator) and one assistant, who together constitute the moderating team; it is essential to have both for a successful interview. The facilitator is in charge of conducting the discussion with the group/community. Ideally, she/he should also be involved in the data analysis. The assistant should be responsible for taking the notes during the meeting.

5.2 ROLE OF COMMUNITY-LEVEL DATA IN CFSVAs: AN OVERVIEW

The Food and Nutrition Security Conceptual Framework adopted by CFSVAs considers food availability, food access, and food utilization as core determinants of food security and links them to households' asset endowments; livelihood strategies; and political, institutional, and economic environment.

The themes to be investigated in community-level data collection should be selected based on the need to delve into an issue of relevance that is insufficiently covered by the household-level data. Community-level data collection is also used to gather data that would be repetitive among interviewed households belonging to the same community/village. The theme should also fit in the conceptual framework used by CFSVAs as described earlier in these guidelines. Table 5.1 gives an overview of the role of key informant interviews, community discussions, and focus group discussions in the CFSVAs.

Table 5.1: The role of community-level data in a CFSVA

Key informant interview	Focus group discussion	Community discussion
• Infrastructure (roads, health centres, schools, etc.) • Market information, especially through a trader's checklist on supply, demand, reach, prices, etc. • Migration; other coping strategies of households in the community • Agriculture and livestock • Community priorities	• Risk analysis (e.g. identification of shocks and responses, role of external factors, risk management strategies and weight of coping strategies for the CSI) • Livelihood analysis (e.g. identification of inputs and constraints for the main livelihood activities, identification of assets relevant to the wealth index) • Any other theme of interest for the CFSVA	• Coping strategies and capacity of the community and households to face a shock (social network, main external support received, ongoing projects, use of remittances, etc.) • Seasonality • Can be applied to all underlying causes of food insecurity to explain which households and communities are affected

If sufficient time and human resources are available, focus groups and community discussions can be used both before and after the assessment (see Box 5.2).

Box 5.2: Focus groups and community discussions: pre- and post-survey functions

Pre-survey: Before the CFSVA household questionnaire is designed, focus groups and community discussions can be used to get an overview of the situation. This will help determine what topics to include in the survey, how questions can be asked to elicit relevant information, and which alternatives to include in the pre-coded boxes. The qualitative interviews can lead to the formulation of hypotheses to be tested in the quantitative interviews (e.g. "after a shock that has brought about a food shortage, villagers are more likely to send young female family members to the city to earn money than they are under normal circumstances"; the quantitative analysis can help determine to what extent and under which conditions this statement is true).

Post-survey: Sometimes the findings of a quantitative survey raise new questions that cannot be answered by quantitative data alone, or there may not be resources or time to run a new quantitative data collection to explore these issues. In this case, qualitative data collection, for example, can be undertaken to explore the reasons behind the surprising results (e.g. increased school attendance after school feeding programmes are stopped), to get more information on a group of positive or negative deviants, or to get more detailed knowledge on what lies behind some of the pre-categorized answers.

5.3 MAIN CHALLENGES

This section draws attention to challenges that frequently arise while collecting qualitative data through focus group discussions and group interviews at the community level.

• ***Lack of expertise***

Qualitative data collection requires interviewers with excellent skills and a very good understanding of the ultimate objectives of the research (see section 5.4.1, "Selection of the facilitators"). In the context of CFSVAs, it may be difficult to procure the services of people with the necessary expertise.

• ***Subjective relationship***

One of the main risks of collecting data at the community level concerns the subjective relationship between the moderating team and the participants. Each brings a set of assumptions about the other. Accordingly, care must be taken to recognize (and minimize) the degree to which a facilitator's belief system or biases could influence the conversation. Conversely, team members should be sensitive to the possible preconceptions held by participants, and acknowledge those preconceptions as part of their data collection. Participants' preconceptions/behaviours may have serious effects on the collecting of community-level data:

- Facilitators may be asked to provide money (or food) in exchange for information, or for the right to talk to some members of the community (e.g. women, elderly); and
- Input from the community may be biased by their expectation of receiving food or other assistance in return for their participation.

• ***Time***

In the context of CFSVAs, lack of time can be the greatest constraint to the collection of community-level data, especially for FGDs and community discussions. Such data is difficult to collect and analyse quickly, and CFSVAs seldom allow researchers the optimum timeframe for either.

• ***Flexibility***

Data collection seldom occurs exactly according to plan. Potential participants may be absent; and participants may exhibit difficult and unexpected behaviour during the discussion (e.g. total misunderstanding of the topic, disagreements during the conversation). Facilitators should be able to recognize when the data collection is completely biased by these elements and decide when and how to adapt their approach to such elements.

5.4 SELECTING AND TRAINING THE MODERATING TEAMS

This section reviews several practical issues related to the collection of community-level data. In particular, it provides guidance for the selection and training of facilitators for FGDs and community discussions. It also draws attention to challenges related to data collection.

5.4.1 Selection of the facilitators

- ***Who?***

Despite the effort to be impartial, the facilitator cannot be entirely neutral. She/he is a member of an ethnic group, a gender, and an age category. Any of these elements can inhibit openness within the group or influence responses and results.

For large-scale analyses such as CFSVAs, identifying skilled facilitators and assistants for focus groups and community interviews can be extremely difficult. As a general rule, facilitators should have previous experience with qualitative data collection, strong interview skills, and an understanding of group dynamics. In the context of CFSVAs, they should also be familiar with key concepts related to food security (e.g. wealth, food consumption, health, hygiene, shocks, and coping strategies) and have a good knowledge of the Food and Nutrition Security Conceptual Framework adopted by the CFSVAs.

To a certain extent, skills in qualitative data collection can be enhanced through training. However, training will yield good results only if the facilitators have some experience with qualitative data collection methods. If skilled enumerators cannot be found, it may be preferable not to include FGDs and community discussions in the CFSVAs. Box 5.3 provides a detailed list of the characteristics of good facilitators and assistants.

Box 5.3: Key characteristics of good facilitators and assistants

A good facilitator should have:
- Previous experience in qualitative data collection;
- Previous experience in collecting data related to food security and vulnerability;
- Strong interview skills;
- An understanding of group dynamics;
- Interest in the subject and participants' views (curiosity);
- The ability to understand and elaborate on participants' opinions;
- Familiarity with key concepts related to food security;
- Familiarity with the household Food and Nutrition Security Conceptual Framework of the CFSVA; and
- Knowledge of the local language and an ability to accurately translate into the language of the study (if this is not possible, the utilization of focus groups and community discussions becomes tedious and questionable).

A good assistant should have:
- The ability to understand and elaborate on participants' opinions;
- Familiarity with key concepts related to food security;
- Familiarity with the Food and Nutrition Security Conceptual Framework of the CFSVA;
- Knowledge of the local language and an ability to accurately translate into the language of the study; and
- Sufficient self-discipline to refrain from participating in the discussion and yet remain objective and open-minded; the assistant is expected to talk only if invited by the moderator.

• ***Working with partners***

Partners may have more robust experience in collecting data at the community level. It is advisable to contact them and to be aware of their capabilities. If they can bring an added value to the collection of community-level data, it is advisable to involve them in data collection.

• ***Gender implications***

Team composition should also be appropriate for the cultural values and norms of the population. This often requires that teams be composed equally of men and women. It may also require attention to age, ethnicity, or other social divisions. Multidisciplinary and gender-balanced teams not only strengthen the diagnostic process and encourage cross-fertilization of ideas, but they also facilitate the inter-organizational sharing and learning that enhance problem analysis and future programming (TANGO 2002). Capacity in gender analysis skills is an additional important consideration. Where team members cannot be found who satisfy the optimal composition appropriate to a population, attention should be paid to the potential effects of this on the data.

• ***Number of teams***

Given the degree of subjectivity and variability in qualitative data collection, it is advisable to use no more than four teams (four facilitators and four assistants) for each CFSVA. Limiting the number of qualitative research teams increases the efficiency of the training and helps maintain consistency in data collection and participant selection. In some cases, limiting qualitative data collection to two teams of four people can lead to improved results.

5.4.2 Training facilitators

The process of collecting community-level data differs from that of collecting household-level data. Facilitators rely not on a set of predetermined questions to be asked in sequence, but rather on a checklist of topics to guide the interview and a set of techniques (i.e. participatory tools) to delve deeper into chosen topics, pursuing relevant material as it emerges during the process.

Facilitators must be trained to solicit topical data while remaining open to new and unforeseen information that may be relevant to the subject in question. The most effective training in qualitative data collection is that conducted concurrently with quantitative training. Ideally, the core parts of the qualitative training should be undertaken directly after the quantitative data collection training. This will allow the facilitators of the FGDs, key informant interviews, and community discussions to be fully aware of the scope and objectives of the CFSVA, including the specific role of community-level data. Box 5.4 provides a brief list of topics that should be included in training for community-level data collection as part of a CFSVA.

Box 5.4: Training topics for qualitative data collection in CFSVAs

1. Introduction to CFSVAs

Participants should receive a general introduction to key concepts related to food security and the Food and Nutrition Security Conceptual Framework adopted by the CFSVAs. Familiarity with such concepts is crucial for collecting community-level data (especially through focus groups and community discussions), as the facilitators cannot use a precise list of close-ended questions.

2. Community-level data in the CFSVAs: Overview and role

Participants should receive a general overview of the tools that will be used to collect community-level data and the role of facilitators. They should also be made aware of the main objective of this information in the context of the CFSVA, and of which kind of information will be extracted using these tools.

Team members should understand that community-level data has an important role in the CFSVAs and that collecting this type of data is not an optional activity, to be conducted only if time permits. If the team does not perceive the real value of such a difficult and time-consuming component of the research, the quality and usefulness of the data collected will inevitably suffer.

3. Tools: Participants, questions, and participatory techniques

At this stage of the training, facilitators and assistants should:

- Know who is supposed to participate in the discussions (e.g. focus groups, community discussions, or key-informant interviews);
- Be familiar with the outline of questions and understand the purpose of each question;
- Know how to use participatory techniques (if they are to be used); and
- Know how to facilitate appropriate responses from the participants (e.g. answers that are pertinent to the topic).

4. Reporting and analysing

Facilitators and their assistants should know how to summarize the information during and after the discussion. In particular they should be trained in identifying the central issues emerging from the discussion and report them in a systematic and standard way. If standard blank templates are to be used in community data collection, facilitators and assistants should be trained in how to complete them.

5. Organization of the discussions/focus groups

Both focus groups and community discussions require some preliminary activities (described in section 9.5.5). Facilitators and assistants should be trained in such steps. It is advisable to train facilitators in organizational issues. Some organizational issues are related strictly to the type of participants involved in the discussions.

6. Field test

A sufficient amount of time should be dedicated to practicing the use of the tools, either among team members or with a randomly selected audience. Field tests are essential for identifying the most skilled facilitators, revising the outline of questions, amending the participatory techniques, and sharing good practices.

After each test, facilitators and assistants should provide feedback to all participants. In this way, trainees improve their skills and continue to develop them once in the field. To enable consistent improvement, feedback should be provided not only during training, but also during actual facilitation in the field.

5.5 FOCUS GROUP DISCUSSIONS DURING A CFSVA

This section describes the necessary steps in properly organizing and facilitating focus groups. It includes the following topics:
- Role of FGDs/identification of themes
- Development of questions/techniques
- Field test
- Selection of communities
- Selection of participants
- Taking notes and making summaries
- Analysing
- Reporting

5.5.1 Role of focus group discussion/identification of themes

The process of conducting a FGD starts by establishing the themes and specific issues of interest. Within the context of food security and vulnerability analyses, qualitative data methods are typically used to capture the variations in livelihood strategies that determine exposure to risk.

- ***Livelihood analysis***

Focus group discussions can be used to understand the link between assets contributing to livelihood activities and major sources of household income, as well as to identify constraints to the existing income-generating activities.

If focus groups are used as a pre-survey tool (see Box 5.2), they can be useful in identifying the relevant assets and housing characteristics for the computation of the wealth index. (An example of wealth ranking is included in Annex 3: Examples of Participatory Tools for Risk Analysis.)

- ***Risk analysis (shocks, responses, external factors)***

Focus groups can be used to identify the major shocks faced by households in the community (or by specific groups), the effects of the shocks on household livelihood strategies, and household responses and external actors assisting the communities. This type of discussion offers valuable input for programme recommendations. In particular, it can be used to identify a community's needs, and gaps in the community's formal and informal safety nets.

Focus groups can be used to compare the current situation with the past; determine if the community is experiencing a period of stress/emergency; and, if so, characterize the magnitude and nature of that stress.
As recommended in section 4.2.4.2, Coping Strategies Index Module, focus groups can be used to contextualize the coping strategies included in the Coping Strategy Index and modify/update the weights.

It is also possible to focus the discussion on other themes, such as specific shocks and food utilization. However, the discussion should always consider the consequences on food security. Annex 1 ("Example of Topical Outline for Use in Qualitative Assessments") gives more precise guidance on the potential themes to study.

5.5.2 Development of questions/tools

The purpose of qualitative tools in a CFSVA is to explore and understand the context specificities related to food security. Once the main themes have been identified, specific questions and techniques should be developed. These should take into account the themes under discussion as well as the audience, time constraints, and cultural context. Box 5.5 includes some general recommendations on how to formulate and ask effective questions. In general, participatory techniques are tools of communication that encourage the expression of community perspectives. During a CFSVA, participatory tools can be used to bring together the knowledge of the vulnerable groups of the community in a more effective way.

Several participatory tools have been tested for their usefulness in assessments, food assistance management and distribution. WFP's Guide on Participatory Techniques and Tools, and OMXF guidance materials offer a comprehensive overview of such tools.[70] The identification of the most suitable tools depends upon the issues under study. Table 5.2 gives an example of the tools used for FGDs on livelihood, risks, and shocks. The suggested examples are found in Annexes 2 and 3. They should not be seen as prescriptive, given that every CFSVA should assess the tool that is most appropriate to its particular communities and objectives.

Box 5.5: General recommendations on questions

- Undertake a comprehensive situation analysis, a review of secondary data, and dialogue with the country office and partners while preparing the questions.
- Questions should be clear; if they do not work during the training or the pilot testing, they should be revised or eliminated.
- Questions should be asked in a conversational manner. The wording should be direct, straightforward, comfortable, and simple.
- The wording should be neutral and the questions should be translated into the local language. Translation should be done by several bilingual members of the team or outsiders to ensure it is faithful to the original meaning.
- Focus group questions should be developed to cover a period of between 60 and 120 minutes.

For risk analysis, the table suggests that the identification of the major shocks is the first issue to address. Yet some countries may want to focus on a specific (well-known) shock and explore its consequences and responses. Table 5.3 offers ideas for issues that can be explored for FGDs related to HIV/AIDS.[71] It highlights the variety of sub-topics that emerge and the importance of identifying the most appropriate participatory techniques and participants for each.

70. Technical Guidance Sheets Nos. 10 and 11 are available at www.wfp.org/food-security. Participatory qualitative techniques selected for EFSAs, for example, include transect walk, seasonal calendar, time line, proportional piling, pair-wise ranking, Venn diagrams, community maps, wealth (or well-being) ranking, ranking and scoring, and a rapid visual assessment of crop/livestock conditions. See *Emergency Food Security Assessment Handbook* (chap. 11). These participatory techniques may also be incorporated into a CFSVA.
71. See also the VAM thematic guidelines, *HIV/AIDS Analysis: Integrating HIV and AIDS into Food Security and Vulnerability Analysis,* and *Food Assistance Programming in the Context of HIV*.

Table 5.2: Examples of themes and participatory tools for a CFSVA

Themes and issues under study	Participatory tools	Examples
Livelihood analysis		
• Identify major sources of income (livelihood activities)	Boxes for identification of income source Livelihood ranking	• Annex 2, step 1 • Annex 2, step 2
• Identify the assets contributing to livelihoods, and the constraints to the existing livelihood activities	Boxes for livelihood inputs and constraints	• Annex 2, step 3
Risk analysis		
• Identify the major shocks faced by the households	Boxes for identification of shocks Shocks ranking	• Annex 3, step 1 • Annex 3, step 2
• Identify the effects of the shocks on the household livelihood strategies	Risk Matrix for the identification of affected groups	• Annex 3, step 3
• Identify the responses to different shocks and assess their efficiency in reducing, mitigating, or coping with the shocks	Risk Matrix for the identification of community response	• Annex 3, step 4
• Identify external actors helping vulnerable households avoid being adversely affected by shocks	Institutional analysis	• Annex 3, step 5
• Identify elements of community notions of wealth to enable ranking of individual households	Wealth ranking	• Annex 3, step 6

Table 5.3: Topics/issues related to HIV/AIDS addressed during a focus group discussion

TOPIC	Objective(s)	Suggested technique(s)*	Recommended participants (where appropriate)
Impact and response	Identify HIV/AIDS impact on households and response (e.g. impact on livelihood assets and strategies, food security, roles).	Problem tree Problem and solution technique Story with a gap	People living with at least one chronically ill adult member
Support	Identify institutions, formal and informal groups, and places that play a role in HIV/AIDS prevention and care. Understand their importance and interrelation.	Checklist	Key informants (e.g. health personnel)
	Measure community's awareness of the HIV/AIDS-related services in the community.	Problem and solution technique	Perceptions held by all the community's members are important. In this case, it is not appropriate to focus only on the affected households or key informants.
	Identify factors that prevent people affected by chronic illness from seeking care (i.e. using existing services).	Force field analysis	People living with at least one chronically ill adult member

*A detailed description of these techniques can be found in UNAIDS, 2004, *Techniques and Practices for Local Responses to HIV/AIDS*, Amsterdam: KIT Publishers.

Questions and participatory tools are most effective when they elicit the information the researcher wants to collect, while remaining simple enough to be understood by the participants. In order to ensure this is the case, considerable attention should be paid to developing effective questions and tools. It is also important that the facilitators are aware of good practices and follow some basic rules. Box 5.6 provides a brief description of principles that should be followed not only while conducting FGDs, but throughout the implementation of CFSVAs.

Box 5.6: General principles for conducting focus group discussions

- **Location:** Interviews should be conducted in a relatively private place, where participants feel comfortable speaking.
- **Introduction:** Facilitators should open the session with a traditional greeting and other local meeting conventions (e.g. a prayer), explain who they are and whom they work for, the purpose of the data collection, what will be done with the information, and who will have access to it. Facilitators should then explain that they do not make decisions about assistance or interventions (if that is the case).
- **Flexibility:** Facilitators should begin the discussion in general and concrete terms. They should be flexible, while making sure that all topics get covered (following an outline). Granting participants space and time while guiding the conversation in the right direction is a difficult skill. Although facilitators should not appear rushed, clear time limits should be set for each session; one hour is recommended for CFSVAs.
- **Participation:** Facilitators should be constantly aware of their own biases and the ways in which they may be perceived by participating communities. The tone of interviews should be informal, not interrogatory. Although care should be taken to ensure that all participants in a group context contribute to the discussion, the interviewer should respect sensitive issues and the right of participants not to respond. Interviewers should avoid implicit and explicit value judgments while conducting interviews.
- **Timing:** Fieldwork should be sensitive to time issues: the schedule of the study (e.g. the need to have data before a certain date); the time of the day the team arrives in the community (e.g. key informants may be busy, women or men may be working in the field); and the season of the year (e.g. unavailability of a certain group of the population due to seasonal migration or work in the field). It is important that each team knows how to handle various problems that may be experienced during the fieldwork. For example, the team will have to be aware of security concerns, as discussions could happen after dark.
- **Reviews:** During data collection it is advisable to conduct periodic random reviews of data collection forms.

5.5.3 Field testing

Field testing is used mainly to assess:

- The appropriateness of the questions to the local context, and the wording of key concepts in the local language;
- The effectiveness of the participatory techniques in stimulating the conversation and coming up with rankings;
- The ability of the teams to moderate the focus groups, take notes, and summarize the answers into a standard template; and
- The time needed to conduct interviews and the amount of information that can be realistically covered in a given period.

After the field test, facilitators' feedback should be used to modify questions, participatory techniques, and standard templates. The experience should also be used to improve facilitators' ability to conduct discussions.

5.5.4 Selection of communities

Since focus groups require time to be organized and conducted, it is not possible (nor is it necessary) to conduct them in every community visited during a CFSVA. Focus groups collect data from a group of individuals within a subset of communities. This section discusses the methodology for selecting communities and participants for the focus groups.

The selection of the communities entails the following steps:
1) Establish the number of communities where focus groups will be conducted.
2) Select from the list of sampled communities those where focus groups will be conducted.
3) Determine practical rules for replacing communities, if necessary.

• ***Establish the number of communities where focus groups will be conducted***
The number of communities where focus groups are to be conducted must be established before the data collection starts. Usually only one focus group is conducted in each village/community, although two are necessary when participants are divided by gender. In general, the total number of focus groups depends upon:
• the number of areas of estimation (AoEs);
• the number of skilled facilitators (human resources); and
• the time allocated for data collection (i.e. the time the teams can spend in the village).

There is no strict rule regarding the total number of FGDs to conduct. Ideally FGDs should be stopped when the facilitators/researchers believe that any "new" information garnered from additional focus groups would not be enough to justify the organizational effort additional FGDs would require. However, this strategy is applicable only if data collection and analysis (at least preliminary analysis) are conducted concurrently.

As a general rule, CFSVA teams should undertake a **minimum of three FGDs in each AoE** and to increase the number according to the size and heterogeneity of the AoE. **It is not advisable to conduct more than five focus groups** in an AoE (even in the larger areas). Indeed, if the AoE being surveyed shares the same socio-economic and/or livelihood patterns (which is the purpose of stratification of the sampling frame), increasing the number of FGDs will probably not bring additional information.

• ***Select from the list of sampled communities those where focus groups will be conducted***
Within each AoE, communities for the focus groups can be sampled randomly (e.g. simple random sampling or systematic sampling) or through purposive sampling.

Random sampling avoids introducing a systematic bias in the data collection. It ensures that within each AoE all the villages/communities have the same probability of being selected.

Purposive sampling can be done by asking the key informants to identify villages that are typical of the livelihood zone. Villages most affected by recent/current problems should be avoided, unless they reflect the typical condition of the livelihood zone or present great interest for the theme(s) under study.
Ideally, focus groups should be undertaken in all the areas within an AoE. However, if the purpose is to explore a theme or issue relevant only in some areas of the country, it is better to concentrate the focus groups on those areas.

- ***Determine practical rules for replacing communities***

If, for any reason, it is not possible to organize FGD in a selected village, the team is expected to replace it with another village/community. A simple replacing strategy is to pick up a village from the same region (or AoE) that the missed village belongs to. By doing so, the geographic distribution of the FGD across the country will not be affected, nor will the total number of FGDs be different.

5.5.5 Selection of participants

As a general rule, the number of participants in an FGD should be between 6 and 10. The higher the number of participants, the more difficult it is to facilitate the discussion. Defining and contacting participants for the FGDs requires the following actions:
1) Decide whether the theme under study concerns all the people living in the village or only a specific group.
2) Decide whether the theme or the context requires separate focus groups for men and women.
3) Establish a strategy for contacting the participants and forming the groups.

- ***Does the theme under study regard all the people in the community?***

If there is a good reason to think that the theme under study regards the whole community, there is no need to select a specific group/category. However, some shocks (e.g. chronic illness, orphanhood, drought) may concern a specific group within the community (e.g. households living with chronic illness, orphaned children, farmers). In such a case, the profile of the most appropriate participants for the theme under study should be clearly defined during the preliminary steps of the survey design.

- ***Does the theme or the context require separate focus groups for men and women?***

Undertaking separate focus groups for men and women duplicates the effort. Gender-separated groups are particularly appropriate if there is a good reason to believe that perception of the problem is different for men and women or where communication between the two groups on certain issues may be difficult or inappropriate. In many situations, regardless of the subject under study, men tend to dominate the discussion. It is therefore essential to consider the added value of separating men and women. If the facilitator finds him/herself in this situation, it is still possible to separate men and women even after an FGD has begun.

- ***How should the participants be contacted and the groups formed?***

Care should be taken to ensure that interviews are conducted with the full permission and understanding of the participants. For community discussions, participants can be identified, selected, and gathered in two ways:

1. Through the leader of the community, especially if she/he is informed in advance of the upcoming data collection: Community leaders should be informed in advance of the visit and should know that the survey includes a discussion with a small group of people living in the community. If possible, before the enumerators arrive at a village they should select a group of people and identify a location (and appropriate time of day) for the interviews. Community leaders can be extremely helpful in setting up the discussion. However, they can also introduce a bias while selecting participants. This "interference" can lead to the dominance of a certain group/category in the discussion.

2. By visiting and approaching people at the place where they usually meet: This strategy can be very effective for some groups, such as households living with chronic illness and orphans and people working at the market. If the focus groups require the participation of a specific group, and community leaders are expected to help in the selection process, it is important to explain clearly to them which characteristics the participants must have.

Box 5.7 provides some examples for contacting households living with HIV/AIDS.

Box 5.7: Contacting households living with HIV/AIDS for focus groups

Facilitators can contact households living with HIV/AIDS through:

Health centres: Participants can be found either at a generic health centre providing services related to HIV/AIDS (e.g. centre providing tuberculosis treatment, antiretroviral treatment, a prevention of mother-to-child transmission programme, a milk bank, assistance to people living with AIDS, family planning and/or HIV/AIDS prevention programme, home-based care service for chronically ill people). Health personnel can help identify participants.

Community leaders or key informants: Participants can be identified through a community leader or during the key informant interviews (especially if health personnel are involved in the interview).

5.5.6 Taking notes and making summaries

Taking notes

Participatory tools generate a great deal of information. Listening and remembering a conversation is crucial. There are several ways to remember the content of a conversation:

- **Take notes during the discussion:** This should be done by the assistant moderator or process documenter. The facilitator can also take notes on what he/she considers the "key message" or items that require follow-up or more in-depth questioning or that are not clear.
- **Use recording devices:** This technique is widely used in the private sector. Recording has the advantage of making the transcription of results more precise. However, sometimes recorders should not be used in the field because of noise or cultural sensitivity to them.
- **Use flip charts or other tools:** Flip charts and other paper supports are often used to summarize ideas, enhance conversation, and write key messages to which the participants can refer. During the CFSVAs, in certain contexts, this technique may be limited due to the level of education/literacy of the audience.

Making summaries

Summarizing answers is an important preliminary step to do before doing the "actual" analysis. Once the FGD is over, facilitator and moderator should write down all important messages as soon as possible (while they are still "fresh"). Annexes 2 and 3 propose using a standard table to summarize the ideas discussed during the conversation. The more similar the structure of the summaries, the easier it will be to consolidate them into a report.

5.5.7 Analysis: Consolidation of summaries and linking with the quantitative data

At this stage, the analyst is expected to pull together the summaries from the focus groups and produce a global picture. In CFSVAs, the analyst (or the person responsible for writing the report) may be a different person from the focus group facilitator. She/he may not be familiar with the language of the discussion. It is therefore crucial to have the facilitators and the analyst sit down together for a proper translation of the summaries.
If the note taker and facilitator took notes and summarized the results following a standard template, the analysis can proceed with the following steps:

1. Consolidate focus group responses: Pull into one table the summaries of the focus groups. Read all the tables and recode the answers: give the same label to similar answers. For instance, if one focus group reported "food vendor" and another reported "vegetable vendor," recode; if judged similar in the local context, these answers can use the same label (e.g. "food vendor" for both).

2. Arrange the rows for each table separately so that all the responses of one focus group are in the same row: Do this for all the tables (focus groups). At the end, all the tables should be transformed into rows and the final qualitative data table should have as many rows as the number of focus groups.

a) **Pull together all the lines** (see Table 5.4)

b) **Examine the results:** The approach is similar to one commonly used in the analysis of data from a household questionnaire. It is crucial to report:
  - The sources of livelihood in order of importance (use "Source" and "Contribution to total income" columns);
  - For the most important sources, report the more recurrent inputs and constraints (use "Inputs" and "Constraints" columns); and
  - Identify differences across the country (use "Village and "Area" columns).

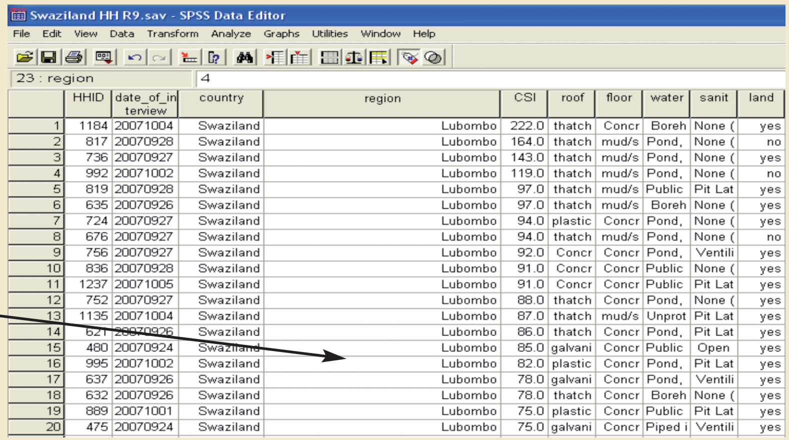
c) **Link with the quantitative information:** Most of the time, results from the FGDs are reported at the community/village level. This is evident in Table 5.4, where each line reports the results of one village. The community-level information can be easily aggregated at a higher level (e.g. region, livelihood zone) as required. In the household data set, each household is linked to a region or livelihood zone of residence. It is therefore easy to link the two data sets. Community-level information can be used to contextualize and enrich the household-level data. Figure 5.1 visualizes the action of linking the two sets of information.

Table 5.4: Example of summary matrix of qualitative data from multiple focus group discussions

Village	Area	Source	Rank	Average contrib. to total income (%)	Inputs	Constraints	Source	Rank	Average contrib. to total income (%)	Inputs	Constraints	Source	Rank	Average contrib. to total income (%)	Inputs	Constraints
Machi	Caprivi	Domestic help	1st	38	–	–	Taxi driver	2st	19	Saving for car rental	Expensive official license	Food vendor	3st	18	–	–
Chipa	Lubombo															
Kiwi	Lubombo															
Ghyia	Caprivi															

Figure 5.1 Linking community-level and household-level data

Area	Source	Rank	Average contrib. to total income (%)	Inputs	Constraints
Caprivi	Domestic help	1st	38	–	–
Lubombo	Taxi Driver	1st	50	–	–


Swaziland HH R9.sav - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Window Help
23 : region 4

	HHID	date_of_interview	country	region	CSI	roof	floor	water	sanit	land
1	1184	20071004	Swaziland	Lubombo	222.0	thatch	Concr	Boreh	None (	yes
2	817	20070928	Swaziland	Lubombo	164.0	thatch	mud/s	Pond,	None (	no
3	736	20070927	Swaziland	Lubombo	143.0	thatch	mud/s	Pond,	None (	yes
4	992	20071002	Swaziland	Lubombo	119.0	thatch	mud/s	Pond,	None (	no
5	819	20070928	Swaziland	Lubombo	97.0	thatch	mud/s	Public	Pit Lat	yes
6	635	20070926	Swaziland	Lubombo	97.0	thatch	mud/s	Boreh	None (	yes
7	724	20070927	Swaziland	Lubombo	94.0	plastic	Concr	Pond,	None (	yes
8	676	20070927	Swaziland	Lubombo	94.0	thatch	mud/s	Pond,	None (	no
9	756	20070927	Swaziland	Lubombo	92.0	Concr	Concr	Pond,	Ventili	yes
10	836	20070928	Swaziland	Lubombo	91.0	Concr	Concr	Public	None (	yes
11	1237	20071005	Swaziland	Lubombo	91.0	Concr	Concr	Public	Pit Lat	yes
12	752	20070927	Swaziland	Lubombo	88.0	thatch	Concr	Pond,	None (	yes
13	1135	20071004	Swaziland	Lubombo	87.0	thatch	mud/s	Unprot	Pit Lat	yes
14	621	20070926	Swaziland	Lubombo	86.0	thatch	Concr	Pond,	Pit Lat	yes
15	480	20070924	Swaziland	Lubombo	85.0	galvani	Concr	Public	Open	yes
16	995	20071002	Swaziland	Lubombo	82.0	plastic	Concr	Pond,	Pit Lat	yes
17	637	20070926	Swaziland	Lubombo	78.0	galvani	Concr	Pond,	Ventili	yes
18	632	20070926	Swaziland	Lubombo	78.0	thatch	Concr	Boreh	None (	yes
19	889	20071001	Swaziland	Lubombo	75.0	plastic	Concr	Public	Pit Lat	yes
20	475	20070924	Swaziland	Lubombo	75.0	galvani	Concr	Piped i	Ventili	yes

5.5.8 Reporting

Findings from the focus groups on the theme under study (e.g. livelihoods or shocks) are particularly valuable if they are disaggregated by specific characteristics of the villages (e.g. geographic location) or participants (e.g. livelihood strategies, gender). Ideally, results from the FGDs should be incorporated in the main CFSVA reports and used to triangulate findings from the quantitative data and/or provide a better insight into the problem under study. The report should clearly state when results are from the focus groups, to ensure transparency with results.

It is difficult to suggest a procedure for incorporating these findings. As a general rule, the person in charge of the CFSVA report should be aware of the issues addressed during the FGDs and should use the results from the FGDs to enrich the analysis from the household survey.

The report "Risk, Vulnerability, and Livelihoods in Monrovia: An Analysis of Qualitative Data from 10 Urban Settlements" provides a good example of how to summarize findings from FGDs. It not only outlines key risk factors, outcomes, and responses, but also offers insights into local contexts and in-depth information on how livelihoods contribute to reducing vulnerability and food insecurity. Box 5.8 summarizes the methodology used, clearly stating that the answers from the focus groups are stratified according to the main livelihood activity of the village/community where the discussion took place.

Box 5.8: Linking livelihoods and risk management: Approach adopted in Monrovia, 2003

"The primary livelihood strategy – as reported in men's and women's focus groups — then served as the basis for re-clustering the sample for analysis purposes. [..] The following typologies were generated:

• Petty trade-base: Findings from women's focus groups in this typology indicate that their petty trade activities require few specialized skills and a small amount of financial investment... An important component of women's petty trade activities is informal credit. Men's focus groups also identified unskilled, labour-based activities as forming the basis of their petty trade.

• Fishing-based livelihoods: Both men and women rely on fishing-based activities as the primary source of livelihood. Violent storms and heavy rains, therefore, place seasonal constraints on this type of livelihood activity for communities. Gender roles are defined rather traditionally, with men doing the actual fishing and women handling the responsibility of cleaning, drying and selling the catch... Petty trade supplements fishing in terms of income sources and also serves as the primary source of income for those who cannot engage in fishing activities.

• Urban agriculture and gardening-based livelihoods: men's and women's focus groups identified economic and natural risks as the main problems affecting their food and livelihood security. Natural risks are related to flooding, pest infestations and soil being waterlogged as a result of heavy rainfall... Vegetables and greens cultivated through gardening are for household consumption as opposed to sales. Both men and women engage in gardening, whereas men are in the majority in respect to urban agriculture."

(*Source: WFP VAM, 2003*, "Risk, Vulnerability and Livelihoods in Monrovia: An Analysis of Qualitative Data from 10 Urban Settlements.")

5.6 COLLECTING DATA THROUGH KEY INFORMANTS

Key informant interviews and community discussions are sometimes confused in CFSVAs. Key informant discussions are usually used to gather, from "experts," more in-depth information about a certain topic. The level of conversation and the information gathered are therefore different from that obtained in an FGD or community discussion. Examples of key informant interviews, the limitations of the approach, and the information key informants can provide are all detailed in this section.

5.6.1 Sampling

For CFSVAs, key informant interviews will typically be conducted in each randomly selected village for the household interview. Key informants are interviewed in each village where quantitative data is gathered. The sampling of communities follows the standard procedures used in CFSVA (see Chapter 4 on sampling). Therefore, no additional sampling procedures are needed.

Each village should be geo-referenced using GPS. Each team visiting the village/community/camp should be equipped with a GPS unit and should know how to use it. The same applies for household data collection.

5.6.2 Design of questionnaire and selection of participants for key informant interviews

Key informant questionnaires for CFSVAs can be short and focused primarily on previously identified priority information. Their length is determined by the information already available (secondary data), the objective of the survey, and the time and resources available for each team collecting the data. A sample topic outline is provided in Annex 1.

5.6.2.1 Main sections

The questionnaires touch upon topics related to the community and/or to the households in the community. Wherever possible, questions should be open-ended. The inclusion of topics depends on information needs identified according to the Food and Nutrition Security Conceptual Framework and available existing information. Examples of themes commonly used in community questionnaires include:

Community information. Most of the key informant questionnaires used in CFSVAs start with collecting information on the current population. For example, in the Ghana CFSVA, information was gathered on ethnicity and religion in the community and on the number of female-headed households in each village. This information can be used to triangulate information on female-headed households and other demographic information collected at the household and focus group level.

Temporary and seasonal migration patterns. Community questionnaires could be used to gather information on seasonal migration and on movements within the community. This information could be used to compare migration patterns at the household level and could also give an indication of the community labour market (which can be useful in an analysis of livelihoods at both household and focus groups levels). In the Rwanda and Nepal CFSVAs, key informants were asked about the age of the migrants and the type of work the migrants sought.

Community infrastructure. In this section, questions on health and education infrastructure may be grouped, but the topics can also be treated as stand-alone sections. The objective of this section is to collect data on community access to education, health, transport, and market infrastructure. Such information is useful when compared to results from the household survey (especially data on education and health), and risk analysis (data on institutions and the services they provide in the event of a shock).

Education and health. The community questionnaires often include questions for school and health officials on infrastructure, access to health and education, school attendance, main diseases, and costs of services. The data can be combined in the same manner as described in the previous paragraph. The collected information is useful for school feeding and HIV/AIDS programmes, and for partners involved in CFSVAs, such as UNICEF and WHO.

Agriculture and livestock. Depending on the context and the partners involved (especially FAO and the Ministry of Agriculture), this section can be split into two parts. The household survey usually collects data on crops and harvest and on ownership of livestock. The questions asked depend on the available information and the data gathered at the household level, but predominantly involve community assets. This data can facilitate the creation of a seasonal calendar, as was done in the Sudan CFSVA. In that case, key informants (agriculture extension officers) were asked to draw a calendar, by month, for a typical year and for the year preceding the survey.

Market information. In addition to traders, interviews that give information on prices, availability of products, and market reviews are important. The community questionnaire is a good tool for looking at market access and the selling and buying habits of each community. In Mauritania, the community questionnaire pointed out the selling and credit practices in markets. In Uganda, the community questionnaire pointed out the available products and the frequency with which they are on the market. It also highlighted the main problems in the marketing of certain products.

Community priorities and food assistance. Interviewing key informants and community members allows for a broader picture of the challenges and constraints affecting the members of a village/community. This section of the questionnaire should focus on assistance received by the community (typically over the past year or last six months) and what the community sees as their priority. This information can help the authors of the CFSVA put together relevant recommendations. Open-ended questions are used to determine community priorities, which are often divided into two groups: immediate and long term. One has to be careful with this section, as the answers are often biased and should ideally be triangulated with other information, such as observation or focus groups on risk analysis.

5.6.2.2 The interview and the choice of key informant

The interview

During a CFSVA the villages visited for the household survey are all targeted for key informant interviews. Ideally, the arrival of survey teams in the village is advertised and should accommodate the schedule of key informants (e.g. midday for teachers might be the best time, and high season for malaria might not be an easy time for health officials).

The inclusion of many open-ended questions makes the interview longer. The selected questions should focus on obtaining priority information. The closed-ended questions serve as introduction and transition questions, providing the open-ended questions with a brief background, so they do not appear to come out of nowhere.

Key informants may very well be important members of the community, and one might have to respect certain rules before being able to get information from them. They may also portray their community in a particularly biased way. The use of multiple key informants is therefore strongly suggested, to obtain different opinions on the same subject.

The choice of key informant

One might feel that the questions asked, while within the informant's field of knowledge, are too "difficult" for him/her to answer. It is important to remember that ideas that might be alien to the facilitator are often the very basis of how rural households survive. When talking to key informants, particularly at the village level, one is actually talking to experts. Key informants should be chosen on the basis of their knowledge about an issue or set of issues relevant to the CFSVA. It is self-defeating to ask key informants about issues they do not have first-hand information on. If key informants are uncomfortable answering a question, or do not know an answer, take them at their word.

5.6.3 Analysis and analysis tools

The data collected through key informants at the community level is of two types: quantitative and qualitative. Methods for analysing both quantitative and qualitative data are discussed here.

Quantitative data

The community data as currently gathered in CFSVAs is mostly quantitative. As a result, the analysis follows the same principles as that of household survey data. Data entry is done mostly using MS Access and SPSS. Frequencies of responses are then run, and the data can be considered representative for a certain part of the territory (villages have usually been selected through a stratified random sample, and the number of observations is much lower than for households; however, there is no design effect; see Chapter 4 on sampling). This also allows linkage between the village- and household-level data sets. It is interesting to see the degree to which data collected at the community level reinforces, completes, and contextualizes data collected at the household level.

Qualitative data

The data gathered in response to the open-ended questions and the discussions generated should be carefully recorded following a particular format; at the questionnaire's design stage, sufficient space should be allowed for the recording of such data. Depending on how a question is phrased and on the information obtained, the interviewer should be aware of the basic qualitative data collection; a special session on this topic should be included in CFSVA training. The discussions and summaries could follow the guidance presented in section 5.5.7, on focus groups.

5.7 DESIGN OF QUESTIONNAIRE AND SELECTION OF PARTICIPANTS FOR COMMUNITY DISCUSSIONS

Although called a discussion, a community group interview is often more of an information-sharing event designed to build a baseline understanding of a particular village in a particular region. These discussions ideally involve around 10 participants and are held as an entry point into a community. They can be used to identify potential participants for focus groups.

One has to remember that community discussions regroup all members of the community, not only members considered key informants. Key informants, however, can be part of the community discussions.

5.7.1 Questionnaires

As most of the questions for community discussions are open-ended, there is no need for a formal questionnaire (as is the case for key informant interviews). For focus groups, the facilitating team for the discussion will need a checklist of the themes and questions they want to ask. These could be divided into different sections, with some close-ended question starting the discussion. For example, one section could be as in Box 5.9.

Box 5.9: Example of health section of a focus group discussion questionnaire

HEALTH, EDUCATION, AND SCHOOL FEEDING
What are the main diseases in this village?
Specify, for example:

- Fever
- Diarrhoea
- Respiratory infections
- Pneumonia
- Diabetes

Have there been changes during the past six months compared to other years (at this period of the year)? Yes/No
If yes, what is different, and why?
Are there more children not attending school this year compared to last year? Yes/No
If more are not attending, why?
Do children in this village receive food at school or to take home? Yes/No
If yes, does it make a difference in the decision of families to keep their children at school? Yes/No
If no, or partially, what else would be needed to keep children at school?

Summary Table – Health, Education, and School Feeding

Main diseases	Actions when sick	Most affected by disease	Reasons for not attending school	If more not attending, why?

Additional information:

As is the case for focus groups, a summary table can be used to report the main information to be used in the analysis. Nonetheless, it also requires detailed notes by the note taker. The checklist should therefore include enough space for the note taker and the facilitator to include all comments they need to add.

All the topics described in section 5.6 for the key informant questionnaire can be explored during community discussions – once again, depending on needs, available information, and the data to be gathered through the household, key informant, and other qualitative tools used in the CFSVA. The topics included must follow the analysis plan identified for the CFSVA and the Food and Nutrition Security Conceptual Framework described in Figure 1.1, section 1.3

Seasonality

It is important to study the seasonality of the underlying causes of food security. Box 5.10 gives an example of a seasonal calendar that can be used for CFSVA. Community discussions are often considered the best tool for developing a reliable seasonal calendar.

Box 5.10: Example of a seasonal calendar

Using a large sheet of paper, prepare an empty table using the model here. Write down the months, starting with January or with the month of the main crop harvest (whichever is easier for people to refer to as the starting point for yearly events).

Fill in the rows as follows:
- months of highest/lowest food prices in shop or market
- months of highest/lowest family income
- months of highest/lowest family food stocks (wheat, potatoes)
- months of sale of own food production (wheat, potatoes, etc.)
- months of sale of cash crop production (cotton, horticulture, etc.)
- months of highest/lowest sale of animals or animal products
- months of highest/lowest possibilities to find wage labour in agriculture
- months of highest/lowest possibilities to find wage labour outside agriculture

Summary Table: Seasonal Calendar

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Food prices (high/low)												
Family income (high/low)												
Family wheat stocks (high/low)												
Family potato stocks (high/low)												
Sale own food production (high/low)												
Sale cash crops (high/low)												
Sale animals/products (high/low)												
Wage labour in agriculture (high/low)												
Wage labour in other sectors (high/low)												
Migration out of the village (high/low)												

5.7.2 The discussion

The community discussion differs from the key informant interviews in that it tries to bring together members of the community, including key informants, to discuss relevant issues. The main differences are the terms *interview* and *discussion* and the choice and number of participants.

For a community group discussion, the facilitating team will try to bring together a group that represents the community, by gathering members from all age and social groups. Ideally, the discussion group should not exceed 10 to 12 participants, but is often larger. The rules and challenges of facilitating such discussions are the same as for focus groups; keep in mind that the bigger the group, the harder it will be to get everyone to participate.

In order to collect as much relevant information as possible, the team will have to consider facilitation techniques, such as a rough community mapping that identifies basic landmarks and socio-economic indications. Each of these participatory approaches may lead to interesting discussions regarding the vulnerability of individual households and market access.

Other techniques may include a timeline of shocks over the past five years and a seasonal calendar showing when people tend to migrate, when the harvest takes place, and when prices are higher (Box 5.10). This information is then combined with the rest of the collected qualitative data (focus groups, key informants, household data, etc.) in preparation for analysis.

5.7.3 Analysis of community discussion data

As discussed in section 5.5 (on FGDs), the information gathered through the community discussion is qualitative. Therefore, the techniques that apply to focus groups apply also to analysing community discussions. The main difference lies in the data gathered: while the community discussions provide information on topics of general interest to the community, the focus group provides information on a particular theme.

As a consequence, the community information can be linked to the information gathered through key informants, either to confirm data or to obtain greater detail on themes such as times of migration, which migrations are considered normal, and which are considered coping strategies. The information collected can also be cross-tabulated with much of the information from the household questionnaires. Examples include common livelihood strategies, strategies that have proven most effective, social networks in place at difficult times, and the reasons why children are not sent to school.

5.8 CROSSCUTTING THEMES

5.8.1 Gender and community-level data

Key informants

Because key informants offer inherently subjective information, it is important to interview a diverse mix of informants and compare the findings among them. From a gender perspective, ensure that both female and male key informants are interviewed. In addition to the information gained from these interviews, separately discussing key issues with community leaders reduces their opportunity to dominate discussions with community members.

Focus group discussions

Although subgroups may be defined using a number of criteria (wealth groups, livelihoods, age groups, etc.), holding separate discussions with men and women within each group defined by one or more criteria provides an opportunity for different voices and perspectives to be heard.

Questions that can be incorporated into FGDs or interviews with key informants to ensure that a gender perspective is incorporated should address the following areas and how they pertain to both men and women:

- **Assets and land ownership:** access and control of men and women, inheritance practices, etc.;
- **Access to markets, credit, services:** who buys, sells, borrows; level of interest rates, extension access, etc.;
- **Activity profiles/division of labour:** productive and reproductive activities, and related time allocation;
- **Livelihoods/employment:** major economic activities, wage labour opportunities, earnings, reproductive work activities, security constraints, and access to water;
- **Risk and coping strategies:** different types of security problems and coping mechanisms;
- **Intra-household decision-making processes and resource allocations:** extent of political and social participation in the communities;
- **Education:** various reasons for not attending;
- **Food consumption, nutrition, and health:** cultural dietary restrictions, food utilization/intra-household allocation of food; and
- **Migration/displacement:** reasons for migration, impact of refugees/immigrants on host community, consequences of migration on women's position within the community (i.e. what happens when the men leave).

5.8.2 HIV/AIDS community-level data

In countries with high prevalence of HIV/AIDS, it can be useful to collect community-level data related to the epidemic. Within the context of a CFSVA, interviews with key informants can be used to identify issues related to access and (formal/informal) support for people living with HIV/AIDS; focus groups can be used to enhance the understanding of household responses to HIV/AIDS.

Table 5.5 outlines HIV/AIDS-related issues relevant to food security and vulnerability analysis that can be studied through a qualitative approach. For each specific objective, the table identifies an appropriate technique and a proper group of respondents.[72] It is not possible to prescribe a universal rule regarding who should participate in a focus group discussion on HIV/AIDS. Much depends on the context and the specific purpose of the exercise. However, some issues can be better explored by involving specific groups of people, such as health personnel or people living with a chronically ill adult member. Table 5.5 links some specific objectives to a recommended group of participants, considered to be the best informants.

Table 5.5: HIV/AIDS topics/objectives that can be addressed with communities

TOPIC	Objective(s)	Suggested technique(s)	Recommended participants (where appropriate)
Impact and response	Identify the impact of HIV/AIDS on the households as well as household responses (e.g. consequence on livelihood assets and strategies, impact on food security, and changes in roles within the household).	Problem tree Problem and solution technique Story with a gap	People living with at least one chronically ill adult member
Support	Identify institutions, formal and informal groups, and places that play a role in HIV/AIDS mitigation, prevention, and care.	Checklist	Key informants (e.g. health personnel)
	Measure community awareness of the HIV/AIDS-related services present in the community. Determine their importance and interrelation.	Pie chart	Community members
	Identify factors that prevent people affected by chronic illness from seeking care.	Force field analysis	People living with at least one chronically ill adult member

Source: VAM/WFP, 2008, *"Integrating HIV/AIDS in Food Security and Vulnerability Analysis."*

If data collection requires the participation of people living with a chronically ill household member, it is crucial to adopt a quick and easy strategy for identifying and contacting these people.

- Participants can be more easily found at a generic health centre that provides services related to HIV/AIDS (e.g. TB treatment, anti-retroviral (ARV) treatment, a prevention of mother-to-child transmission programme, a milk bank, assistance to people living with AIDS, family planning, an HIV/AIDS prevention programme, and home-based care service for chronically ill people). Health personnel can help identify participants.
- Alternatively, participants can be identified through the community leader or during the key informant interviews, especially if doctors or health personnel are involved in the interview.

72. The suggested techniques are selected from the toolkit "*Techniques and Practices for Local Responses to HIV/AIDS*," developed by UNAIDS and the Royal Dutch Tropical Institute, which presents practices distilled from local responses worldwide (http://www.kit.nl).

5.9 KEY REFERENCES: USING QUALITATIVE DATA TECHNIQUES IN A CFSVA

- Action Aid International. 2003. *Participatory Vulnerability Analysis: A Step-by-Step Guide for Field Staff.* London.
- Carcalho. S., and H. White. 1997. *Combining the Quantitative and Qualitative Approaches to Poverty Measurements and Analysis: The Practice and the Potential.* Washington, World Bank.
- Center for Refugee and Disaster Studies. 2000. *Training in Qualitative Research Methods for PVOs & NGOs (and Counterparts)*. The Johns Hopkins University School of Public Health.
- Krueger. R. A. 1998. *Developing Questions for Focus Groups: Focus Group Kit 3.* London: Sage Publications, Inc.
- ibid. 1998. *Moderating Focus Groups: Focus Group Kit 4.* London: Sage Publications, Inc.
- Pretty, N., et al. 1995. *Participatory Learning and Action: A Trainer's Guide*. London: IIED Participatory Methodologies Series.
- TANGO International. 2002. *Household Livelihood Security Assessments: A Toolkit for Practitioners.* CARE USA, PLHS Unit.
- University of Illinois Extension Service, Office of Programme Planning and Assessment. 1992. *Key Informant Interviews, Needs Assessment Techniques.*
- WFP. 2009. *Emergency Food Security Assessment Handbook*, second edition. Rome.
- WFP and Save the Children (UK). 2000. *An Introduction to the Food Economy Research in Southern Sudan,* 1994-2000.

CHAPTER 6

Food security analysis in a CFSVA

6.1 ANALYTICAL STEPS

6.1.1 Introduction

This chapter presents a chronological flow of analytical steps to be followed in the CFSVA. Adhering to this order is not compulsory; some steps naturally precede others; some can be implemented simultaneously.

The results of the analysis should be integrated in the report in such a way that the report reflects the various dimensions of the Food and Nutrition Security Conceptual Framework. Accordingly, the report will reflect not the chronology of the analysis but rather the logic of the framework. Important links between the various components in the framework should be highlighted in the report.

Secondary data analysis, literature review, qualitative information (especially from focus group discussions), and primary quantitative data are all used in a CFSVA to:

- Describe all the components of the Food and Nutrition Security Conceptual Framework, highlighting the particular characteristics of the country studied and the important differences among regions and population groups;
- Explore how local food security outcomes are influenced by various underlying and basic causes;
- Combine information to identify regions and population groups most at risk of food insecurity;
- Create a benchmark for food security assessments in the event of a crisis or shock;
- Design and implement food security monitoring systems that track key trends and regions within a country;
- Create a "global food security baseline" against which the impact of specific WFP projects/programmes can be compared; and
- Derive key findings, conclusions, and recommendations for improving food security in the short and long term (WFP 2007).

The analysis of food security in a CFSVA is grounded in the Food and Nutrition Security Conceptual Framework. By looking at livelihood assets and strategies and linking them to food access and utilization, the Household Food Security Conceptual Framework combines a livelihood approach with the UNICEF Nutrition Framework.[73]
The framework describes the basic underlying causes of food and nutrition security, by exploring how households use natural, physical, human, financial, and social assets to develop livelihood strategies within the context of the socio-economic and political environment.
These analytical steps are outlined here:

Pre-survey steps: Desk study

1. Hazard analysis
2. Food security zoning
3. Description of the food security context of the country

73. See Chapter 1, Introduction and Concepts.

4. Analysis of trends and seasonality
5. Market analysis

Post-survey steps

1. Analysis of focus group discussions
2. Computation of a wealth index and wealth n-tiles
3. Creation of livelihood groups
4. Analysis of household food consumption
   a. Computation of Food Consumption Score (FCS) and food consumption groups
   b. Analysis of food consumption patterns
5. Utilization of the FCS as a proxy for food for food security
6. Contextualization: incorporation of seasonality and trends
7. Nutrition analysis
8. Market analysis (primary data)
9. Food security profiling and determinants of food insecurity
   a. Determinants of household food security
   b. Profiling of household food security
10. Analysis of household vulnerability and risk, and scenario building
11. Conclusions and recommendations

Results of the desk study and of FGDs must be integrated throughout the quantitative household data analysis, and not in separate chapters of the CFSVA report. Brief reports (of a few pages) should be prepared on each of the subtopics, synthesizing the results of the literature review, before proceeding to the other analytical steps. These will be integrated in the final report, triangulated with other data.

6.1.2 Pre-survey steps

These steps should generally take place before primary data collection. They will both inform and complement primary data collected as part of the CFSVA. These steps rely on the collection and analysis of secondary data. See Chapter 3 for detailed information.

Although data quality and usefulness vary from source to source, CFSVA studies should attempt to locate and utilize existing secondary sources, when they:

- are up to date;
- contain applicable coverage and/or aggregations of data; and
- adhere to accepted data collection methods and techniques.

Any meaningful information that can be related to the Food and Nutrition Security Conceptual Framework should be reviewed and synthesized, and the relevant parts reported. Literature review is often the main source of information on the political and socio-economic environment.[74]

74. See Chapter 3 for more information on secondary data analysis.

Some of the most common sources of secondary data include:
- Government documents
- Municipal development plans
- Official statistics
- Technical reports
- Project reports
- Baseline studies
- Project evaluations
- Professional and academic journals
- Reference books
- Research organizations
- Public and private universities
- Public and private libraries
- Computerized databases
- Internet websites

6.1.2.1 Hazard analysis

Information regarding the type of hazards that threaten food security in the country, their level of intensity, and the temporal and spatial dimensions of the hazards should be collected.

6.1.2.2 Food security zoning

Food security zoning is the construction of geographical areas based on the physical distribution of basic causes of food insecurity (e.g. ecological, agricultural, socio-economic factors). Food security zones can be determined from analysis of secondary data. Through multivariate techniques (PCA and clustering), areas with similar ecological, agricultural, and socio-economic characteristics are grouped into a zone. Geographical areas with similar underlying causes show similar food security characteristics. Within zones, the resulting food security situation is expected to be rather homogenous, though there can be large differences between zones.[75] The resulting zoning is used for sampling and analysis.

Food security zoning may or may not be undertaken. Typically, the country is already stratified into geographic areas that differentiate unique ecologies associated with food security. These may or may not coincide with administrative boundaries, but more frequently they do coincide with agro-ecological zones declared by ministries of agriculture. Hence, such agro-ecological zones or, alternatively, pre-existing food security zones such as FEWS livelihood zones, can be used.

All strata, whether food security zones created for the survey, administrative boundaries, or existing zones, should be reviewed and cleared by all stakeholders.

6.1.2.3 Analysis of trends and seasonality (secondary data)

Seasonal factors influence household food consumption patterns, particularly for the food insecure. Many basic and underlying causes of food insecurity such as climate,

75. A factorial model (GLM), with the zoning system as an explanatory variable and a food security indicator as the dependent variable, systematically proves this.

availability of food and food prices, prices of agricultural produce, household purchasing power, household stocks, as well as the household food security status itself, health and hygiene conditions, and individual nutritional status, have a cyclical nature and long-term changes. Accordingly, it is essential to include trends and seasonality analysis in a CFSVA in order to contextualize the results.

Information on both long- and short-term cyclical trends, such as rainfall data, agricultural production, trade volumes, and fluctuating food prices can be obtained from secondary sources. The seasonal calendar, Normalized Difference Vegetation Index (NDVI), and rainfall estimate (RFE) patterns can be used to determine how the food security situation at the time of data collection may improve or deteriorate during various times of the year (WFP 2007).

Determining when the CFSVA data collection took place in relation to the seasonal calendar is the first step toward placing findings in a larger context.

6.1.2.4 Market analysis (secondary data)

"The main purpose of the market analysis in the CFSVA is to provide baseline information, analyse market effects on food security and vulnerability, and analyse possible market shocks" (WFP 2007). The market plays a major role in guaranteeing household access to food. Volumes, seasonality, real prices of staples, prices of cash crops and livestock (as appropriate), determinants of demand and supply, and trade patterns are best analysed through the use of secondary data.

6.1.3 Analytical steps (post-survey, primary, and secondary data)

A major component in a CFSVA consists of collecting data at the household and community level. After primary data have been verified and cleaned, the steps for analysis outlined in section 6.1.1 should be followed. Some analysts find it useful to produce output tables at an early stage in the analysis. This step will depend on complex indicators and strata described in section 4.1.2. The mean values of all continuous indicators are displayed for various administrative units, food security zones, livelihood groups, food consumption groups, and gender. For categorical variables, the frequency of main categories can be tabulated for each of these units.

6.1.3.1 Analysis of focus group discussions

If FGDs were included as part of the primary data collected, they should be analysed early on to help direct subsequent steps. FGDs may provide insight into the root causes of food insecurity and vulnerability and the reasons for resilience, exposure, and vulnerability.

This information must be combined with other data in relevant parts of the report.

6.1.3.2 Computation of a wealth index and wealth n-tiles[76]

Household wealth is an important contributor to food security, resilience, and nutritional outcomes. Additionally, it can be a long-term indicator of successful

76. See section 8.2.2.6. for detailed information

livelihood strategies. An asset wealth index can be computed with the use of principal components analysis (PCA).

6.1.3.3 Creation of livelihood groups[77]

Households that share similar livelihood strategies tend to be affected by the same shocks or risk factors. The livelihoods group to which a household belongs is therefore an important determinant of food security outcomes. Livelihood grouping is achieved through "a cluster analysis preceded by principal components analysis (PCA)." Primary livelihood strategies can be recognized through the PCA. Livelihood groups are then composed based on their reliance on various strategies (WFP 2007).

6.1.3.4 Household food consumption analysis

Food consumption analysis is the cornerstone of the CFSVA. The analysis of diet diversity and frequency provides key information on food consumption at the household level. It is the basis for classifying households according to food security status. See Section 6.2.1 for detailed information.

Computation of food consumption score and food consumption groups

The construction of a food consumption score (FCS) allows for a standardized interpretation of households' diet diversity and frequency. Standard thresholds help categorize households into different food consumption groups (poor, borderline, adequate).

Analysis of food consumption patterns

PCA and cluster analysis are exploratory techniques and are used only to describe typical dietary patterns in the population, providing valuable information on the quality of food consumption.

6.1.3.5 From food consumption to food security

The FCS has been a reliable indicator of food insecurity in all CFSVAs and EFSAs where it has been applied. The FCS is one of the only food security indicators that can be equally collected in many different regions, countries, and settings, though other indicators may be more context-specific. It is important to check during a household survey that all food security indicators point in the same logical direction. Therefore, the FCS, or the FCGs, are compared with the food consumption patterns and other typical food security indicators (wealth index, household expenditures, household food production, coping strategies index, etc.). If the FCS shows a contradiction with other indicators, further investigation may reveal the reasons,[78] and the results should be presented taking these reasons into account. This will help to properly interpret and contextualize the FCS. If all indicators confirm the FCS (that is, the FCS reflects the current household food security situation), we can confidently state that the FCGs are equivalent to "current food security groups" (poor, borderline, adequate). The more universal nature of the FCS and its cut-offs allows us to come up with a more standardized food security classification.

77. See section 1.2 detailed information on livelihoods.
78. For instance, some households received food aid; some regions have different seasonality.

6.1.3.6 Contextualization: Incorporation of seasonality and trends

The timing of the household survey influences the values of some key indicators. If data are collected immediately after the harvest, results will be very different from those collected during the lean season. Equally, results will differ if the agricultural season after which the survey is conducted is an exceptionally good one, or an exceptionally difficult one.

Therefore, the result indicator levels have to be put in context by:

- A qualitative description of the current situation as it relates to seasonal calendars; and
- The incorporation of historical data on production, prices, or other indicators, in order to make statements about the current year as compared with past years.

6.1.3.7 Nutrition analysis[79]

Food insecurity is one of the immediate causes of malnutrition, and undernourished populations are more vulnerable to food access and utilization problems. The prevalence of chronic and acute malnutrition in various regions and subpopulations (such as FCGs) is very relevant information. Be sure that sampling methods and confidence intervals are considered and included when reporting anthropometric variables.

6.1.3.8 Market analysis (primary data)

Aside from secondary data and data from the traders and market surveys, we can collect data at the community and household level that provides information on how various types of households are involved in markets, how vulnerable households are to market price changes, and which market-related factors are linked with food security. Market analysis may also be part of seasonality/trend analysis for contextualization, as well as part of risk analysis.

6.1.3.9 Food security profiling and determinants of food insecurity

The next two steps are complementary. Each feeds into the other, and so their order is irrelevant. Exploratory analysis done for household food security profiling may highlight key possible causal factors of food insecurity, and the causal analysis will provide more details for food security profiling.

Household food security profiling

Through simple cross-tabulations and group means, we can describe many indicators (demography, human assets, livelihoods, nutrition, access indicators, etc.) for each of the food security groups.

This step addresses the questions:

- Who are the food insecure (livelihoods, nutrition, wealth groups, access, etc.)?
- Where are the food insecure (geospatial analysis)?
- How many food insecure are there currently?

This step also begins to address the question:

- Why are they food insecure?

79. See section 6.2 for more information on the importance of nutrition in analysis.

We can find the absolute numbers in food-insecure and borderline households for various population groups, administrative zones, and agro-ecological areas by multiplying population numbers by prevalence found in the various strata.

Determinants of household food security
To study causality, one should set up controlled experiments to see the effect of one factor – an impossible task in the context of the CFSVA. Although causality "strictu sensu" cannot be observed in a household survey, we can measure the "net association" of many immediate and underlying causes of food security or nutrition using a general linear model. This, together with information from qualitative sources, helps explain observed differences in food security, highlighting important determinants and excluding some suspected causes of food insecurity.

This step explores with depth the question:
- Why are the food insecure food insecure?

6.1.3.10 Household vulnerability and risk analysis, and scenario building[80]
The current "static" situation of relative food security or food insecurity can change based on households' sensitivity to shocks to the livelihoods on which they depend and on their resilience to the effects of those shocks. Vulnerability and risk analysis bring a more dynamic element to the CFSVA.

The vulnerability analysis identifies population groups and regions most vulnerable to various shocks. A vulnerability map is very useful at the time a shock occurs: vulnerability information combined with information on the extent and intensity of the shock provides an initial calculation of the expected impact on food security.

If households vulnerable to a particular hazard live in areas where the chance for such a hazard is high, they are at risk of becoming food insecure.

Risk analysis will allow agencies involved in food security to prioritize areas where they should concentrate, and will allow WFP, in particular, to focus on these high-risk areas and be prepared for the emergency operations.

In particular, this step attempts to address the questions:
- What are households vulnerable to?
- What is the probability of key hazards occurring?
- What are the effects of these hazards on household food security for different population groups?

6.1.4 Conclusions and recommendations

From all preceding analytical work, significant conclusions about food insecurity can be made. As important immediate, underlying, and basic causes of situations are identified, recommendations can be made that tackle, in the short and long term, various sources of food insecurity.

80. See sections 3.3.5 and 6.2.4 for detailed information on risk analysis.

6.1.5 Key references: Analytical steps

- CARE. 2002. *Household Livelihood Security Assessments: A Toolkit for Practitioners*. Atlanta: CARE USA.
- WFP. 2007. *CFSVA Methodology Workshop Report*. 17–20 April. Rome.

6.2 KEY FOOD SECURITY ANALYSIS IN A CFSVA

6.2.1 Food security classification

6.2.1.1 Food consumption groups

Food consumption indicators are designed to reflect the quantity and/or quality of people's diets. In CFSVA, the most commonly used food consumption indicator is the food consumption score. This proxy indicator represents the dietary diversity and energy and macro and micro (content) value of the food people eat. It is based on dietary diversity (the number of food groups consumed by a household over a reference period), food frequency (the number of times, usually in days, a particular food group is consumed), and the relative nutritional importance of different food groups. The FCS is calculated from the types of foods and the frequency with which they are consumed over a seven-day period.

Although it provides essential information on people's current diet, the FCS is of limited value for in-depth analysis of food consumption patterns, for the following reasons:

- It is based on a seven-day recall period only. This is insufficient for a full analysis of food consumption over longer periods, which is likely to vary according to season, for example.
- It provides no indication of the quantity of each foodstuff consumed.
- It does not give information on intra-household food consumption, such as who eats first and last.
- It does not show how food consumption has changed as a result of the crisis, unless previous food consumption scores for the same types of households are available.

More information is needed if food consumption practices and trends are to be fully understood. For example, questions regarding customary food consumption should be asked to complement the seven-day household FCS.

The calculation of the FCS is explained in Box 6.1[81] and Table 6.1.

81. For further information on the application of the FCS, see *Food Consumption Analysis: Calculation and Use of the Food Consumption Score in Food Consumption and Food Security Analysis*, Draft, WFP Vulnerability Analysis and Mapping Branch, August 2007.

Box 6.1. Calculation of the food consumption score

In the household questionnaire

Households are asked to recall the **foods they consumed in the previous seven days** (see the list of items in Table 6.1). Each food item is given a score of 0 to 7, depending on the number of days it was consumed. For example:

- If potatoes were eaten on three of the last seven days, they are given a frequency score of 3.
- If potatoes were eaten on three of the last seven days, even if they were eaten twice on each of those days, at two meals, they are still given a frequency score of 3.

In the analysis

Food items are grouped according to **food groups** (see Table 6.1) and the frequencies of all the food items surveyed in each food group are summed. Any summed food group frequency value over 7 is recoded as 7.

Each food group is assigned a **weight** (see Table 6.1), reflecting its **nutrient density**. For example:

- Beans, peas, groundnuts, and cashew nuts are given a weight of 3, reflecting the high protein content of beans and peas and the high fat content of nuts.
- Sugar is given a weight of 0.5, reflecting its absence of micronutrients and the fact that it is usually eaten in relatively small quantities.

For each household, the **household food consumption score** is calculated by multiplying each food group frequency by each food group weight, and then summing these scores into one composite score.

The household score can have a maximum value of 112, which implies that each of the food groups was consumed every day for the last seven days.

The household score is compared with pre-established **thresholds** that indicate the status of the household's food consumption. WFP finds the following thresholds to be applicable in a wide range of situations:

- Poor food consumption: 0 to 21
- Borderline food consumption: 21.5 to 35
- Acceptable food consumption: > 35

These thresholds can be adjusted if there is clear justification for doing so, for example, in populations where the consumption of sugar and/or oil is frequent among nearly all households surveyed, even when the consumption of other food groups is rare and the food score is otherwise low. In these cases if this base diet of oil and sugar is combined only with frequent consumption (7 days' worth) of a starch base, the score already arrives at 21. However, this clearly cannot be classified as even a borderline diet. For this reason the thresholds can be raised from 21 and 35 to 28 and 42 (by adding 7 to each threshold, this accounts for the daily consumption of oil and sugar, which gives 7 points to the FCS).

If the consumption of oil and sugar for the overall population is higher, the FCS thresholds should be changed as follows:

- Poor food consumption: 0 to 28
- Borderline food consumption: 28.5 to 42
- Acceptable food consumption: > 42

Table 6.1: A completed food consumption score template

Food item	Food group	Weight (A)	Days eaten in past 7 days (B)	Score A x B
Maize, rice, sorghum, millet, bread and other cereals Cassava, potatoes, and sweet potatoes	Cereals tubers, and root crops	2	7	14
Beans, peas, groundnuts, and cashew nuts	Pulses	3	1	3
Vegetables, relish, and leaves	Vegetables	1	2	2
Fruits	Fruit	1	0	0
Beef, goat, poultry, pork, eggs, and fish	Meat and fish	4	0	0
Milk, yoghurt, and other dairy	Milk	4	1	4
Sugar and sugar products	Sugar	0.5	4	2
Oils, fats, and butter	Oil	0.5	2	1
Composite score				26

Since the FCS is a continuous variable, standard statistics such as the mean and variance can be calculated, and trends of means over time and across categories can be determined. For FCGs, frequencies and cross-tabulations can be determined.

Box 6.2: Validation of the food consumption score

Recent research conducted by the International Food Policy Research Institute (IFPRI) has attempted to validate the use of the FCS for the classification of household food security status with the help of survey data from three countries – Burundi, Haiti, and Sri Lanka. The study found the usefulness of the dietary diversity and food frequency indicators encouraging. There are positive and statistically significant associations with calorie consumption per capita, particularly when small quantities are excluded from food frequencies. However, the cut-off points currently used by WFP to define poor, borderline, and adequate FCGs correspond with energy intake, which is considerably below the usual average 2,100 kcal per capita per day benchmark, which is often used to define undernourishment. Hence, the poor FCG corresponds with extreme undernourishment, and even some households belonging to the "acceptable food consumption group" have consumption below 2,100 kcal per capita per day.

Table 6.2: Food consumption groups with corresponding food consumption score thresholds and corresponding energy intake levels

Food consumption groups	Burundi		Haiti	
	Food consumption score	Corresponding energy consumption in kcal/capita/day	Food consumption score	Corresponding energy consumption in kcal/capita/day
Poor	≤23	≤1550	≤28	≤1600
Borderline	>23 and ≤37	>1550 and ≤1800	>28 and ≤42	>1600 and ≤1900
Acceptable	>37	>1800	>42	>1900

These data reinforce the notion of context specificity in formulating FCS.
In CFSVAs it is recommended that the current cut-offs illustrated in Box 6.1 be used.

CHAPTER 6

6.2.1.2 Description of the current food security situation

A key part of describing household food security in an EFSA or CSFVA is conducting a short-term household food security classification based on the "current" food consumption in the household as a proxy for "current" food security. This gives a snapshot depiction of the situation of the household at the time of the data collection.

This is an essential step in both the EFSA and CFSVA processes and is the starting point for situation analysis and scenario-building exercises (see section 6.1.3.10).
In short, households are classified according to the FCS (poor, borderline, or acceptable). For some households, the FCS may not reveal their current food security situation. Information on how these households access food and their sources of food is crucial and will allow the households to be reclassified.

The basis of the current household food security description is hence the food consumption score (FCS) and its thresholds, as described in Table 6.2. This usually[82] means that households with an FCS of 21 or below have poor food security; households with an FCS between 21.5 and 35 have borderline food security, and households with an FCS higher than 35 have acceptable food security.

The household food consumption classification is meant to serve as a standardized, objective, and replicable tool for describing **short-term food security.** The standardization of this classification is made possible by using the household FCS as a basis for comparison. Although differences in context must be considered while interpreting the FCS, this method of standardization is acceptable given that the FCS is well-defined[83] and objectively measurable. Similarly, if FCS thresholds are appropriately defined, resulting food consumption groups relate to a certain degree with corresponding levels of food intake.[84]

If the FCS does not properly reflect the food security situation of the moment, **this short-term household food security description needs to be adjusted**. This is the case for households whose sources of food are not sustainable or whose strategies for accessing food are uncertain or damaging their future livelihoods, or whose strategies are so severe they are endangering household members' health. A typical example are food aid recipients, who may benefit from acceptable food consumption at the time of the assessment, but who, without food aid, most probably will have "poor" food security. The FCS-based classification as a description for the current food security situation should be triangulated with other food security indicators, such as the Coping Strategy Index (CSI) and income and production indicators.

Finally, a CFSVA and an EFSA should always go beyond the description of current food security as described in this section. For a complete situation analysis, the analyst will make statements about the evolution of the overall food security context and about critical livelihood factors related to resilience, coping mechanisms, and how income generation and food production will define future access to food.

82. Thresholds 21–35 can be adjusted up to 28–42, based on the consumption of oil and sugar; see Table 6.2.
83. It will often be necessary to establish country-specific questionnaires to ensure that evaluation of household food consumption using the standard FCS is appropriate to the local context.
84. See Box 6.2. – IFPRI validation study.

The description of the current household food security situation, therefore, serves only as a starting and reference point. The sections in the EFSA manual on “Conducting a situation analysis” and “Conducting a forecast analysis” go beyond this “snapshot analysis” and include projections for the future.

The situation analysis will start from the current household food security description, which is based on the FCS and has to be complemented by a livelihoods analysis to become a true, forward-looking food security classification. Adjustments are based on the judgement of the analyst, who, based on other food security indicators and a livelihoods analysis, will concentrate on how households will be able to access food in the immediate or long-term future. He will ideally draw on a combination of quantitative indicators and qualitative information. Context-specific indicators used for household food security classification are:

- Income sources, non-sustainable or harmful coping strategies, debt, distress indicators;
- Production, stocks, reserves;
- Food sources, including aggregate food supply; and
- Asset ownership and access to natural resources.

Therefore, analysis based on household food consumption alone should not serve as a simplistic approach to targeting food assistance during programme implementation. In fact, the description of “current” household food insecurity does not automatically equate to a food assistance requirement: not everyone with poor food consumption at the time of data collection needs assistance, and some households with currently good consumption may need assistance later. It is therefore indispensable to define assistance, and to have a good understanding of how households access food, their livelihoods and the effects of shocks thereon, and the macro trends for the future.

6.2.2 Nutrition analysis

In the Food and Nutrition Security Conceptual Framework (based on the UNICEF framework), presented in section 1.3, the nutrition status is the final outcome, at the individual level, of all factors mentioned in the framework. The two immediate causes of malnutrition are the heath status and individual food intake, which are intimately linked to food security. Hence, assessing and understanding the reasons behind the nutrition status are important parts of food security analysis. Unfortunately, many food security analyses focus solely on food **availability** (production, imports) and food **access** (market access, purchasing power), and largely ignore food **utilization** (the degree to which individuals can convert food accessed by the household, into energy for maintaining nutritional status and growth).

In this section, nutrition is discussed primarily in terms of anthropometric and micronutrient malnutrition outcomes. As described by the Food and Nutrition Security

WFP defines malnutrition as “a state in which the physical function of an individual is impaired to the point where he or she can no longer maintain adequate bodily performance processes such as growth, pregnancy, lactation, physical work and resisting and recovering from disease.”

Conceptual Framework, which is based on the UNICEF Conceptual Framework (the standard for nutrition), nutritional status is an outcome of both health status and dietary intake. Dietary intake is itself an outcome of food security status and care practices. Therefore, **nutrition status is an outcome of food security status.**

While the "utilization" aspect of food security is sometimes addressed using anthropometric outcomes, CFSVAs often do not describe these outcomes in terms of "food security." This is due to the fact that in some cases it may be difficult to determine whether poor nutritional status is caused by food insecurity, inadequate care and feeding practices, disease, or any combination of these factors. It is therefore critical for CFSVAs to analyse the possible underlying causes of poor nutrition in order to identify the interventions that have the greatest potential for improving nutritional status among WFP beneficiaries. The conceptual framework presents a basic outline of the many related factors that often contribute to malnutrition at the individual level.

6.2.2.1 Underlying causes of malnutrition

Looking at the Food and Nutrition Security Conceptual Framework, Figure 1.1, we see that the most immediate determinants of malnutrition are insufficient dietary intake and disease. However, continuing down the causal chain, we find that these causes are themselves the result of multiple underlying causes (WFP 2005).

Household food insecurity

Consumption-related malnutrition is typically the result of an inability to access adequate food to meet dietary needs through either household production or purchases. In nutrition assessments, **wasting** (low weight for height) is often interpreted as a sign of inadequate food consumption over the short term or of acute disease, such as diarrhoea (both of which can be due to inadequate care practices). Longer periods of inadequate food consumption or repeated disease, heavy parasite load, micronutrient deficiencies, and so on may result in chronic malnutrition as measured by **stunting** (low height for age) (CDC/WFP, 2005). An understanding of both the quantity and the quality of diet in terms of macro- and micronutrients is important for analysing the relationship between food security, as measured by consumption and malnutrition. In emergency contexts, global acute malnutrition also must be assessed because of the importance of oedema in these settings. Global acute malnutrition combines the prevalence of wasting with the presence of oedema. In emergencies, the prevalence of oedema can be quite high, so it is imperative that Global Malnutrition Rate (GMR) be presented, together with the presence of wasting and oedema.

Nutritionists have also long recognized dietary diversity as a key element of high-quality diets. Dietary diversity is the sum of the number of different foods or food groups consumed by an individual or household over a specific time period. Increasing the variety of foods is thought to ensure adequate intake of essential nutrients and thus promote good health. Lack of dietary diversity is a serious problem among poor populations in the developing world. Because their diets are based predominantly on starchy staples and often include little or no animal products and few fresh fruits or vegetables, nutrient inadequacy is a common outcome. Measures of dietary diversity are relatively simple to collect, and research shows that high measures of dietary diversity are associated with both nutrient adequacy[85] and adequate nutritional status.[86] In addition, dietary diversity at the household level tends to increase with income and wealth (WFP 2005; CDC 2005; Hoddinott and Yohannes 2002; Ruel 2002; Swindale *et al*. 1999).

Dietary diversity (number of different foods or food groups eaten by an individual or household over a specific time period) is recognized as being one determinant of diet quality. A varied diet is more likely to ensure adequate micronutrient intake. Poor dietary diversity is often a problem among poor households in the developing world. Measures of dietary diversity include the FCS and other standard indicators.

Inadequate health services and unhealthy environment

The health environment of a household or individual has a direct impact on their ability to prevent and control disease. Necessary conditions for adequate nutrition include access to and utilization of:

Good quality health services
Health and nutrition are linked in a vicious circle in which the presence of disease contributes to malnutrition, and malnutrition makes individuals more susceptible to disease. Common afflictions that contribute to malnutrition include diarrhoea, acute respiratory infections, measles, and malaria. Access to health clinics, pre- and post-natal care, vaccinations, and vitamin supplements are central elements of well-functioning health services. Vitamin A supplementation and measles immunizations often serve as reasonable proxy indicators of access to health services.

Safe water supplies
Clean and safe water is essential for maintaining health and achieving adequate nutritional status. Diseases such as diarrhoea from unsafe water sources and vector-borne diseases are common problems in developing countries strongly associated with negative nutritional outcomes.

Adequate sanitation
Availability of and access to adequate sanitation services (latrines, washing facilities, etc.) are essential for maintaining a healthy environment. Personal hygiene practices and the correct disposal of faeces both have a direct influence on nutritional outcomes.

Adequate shelter
Adequate shelter has always been recognized as a basic human need. In the absence of adequate housing, multiple aspects of personal well-being are negatively affected, including nutritional status.

Inadequate maternal and child care

Malnutrition can occur even when access to food and health care is sufficient and the environment is reasonably healthy. The social context and the care environment within

85. IFPRI/WFP, 2008, Validation of Food Frequency and Dietary Diversity as Proxy Indicators of Household Food Security, http://documents.wfp.org/stellent/groups/public/documents/ena/wfp186895.pdf.
86. Several CFSVAs show significant correlations between low FCS and wasting and underweight.

the household and the community also directly influence nutrition. Factors influencing nutrition status include:

- Breastfeeding practices;
- Weaning practices;
- Maternal hygiene behaviour;
- Relationships between morbidity and water and sanitation;
- Pregnancies and antenatal care;
- Pre-natal, peri-natal, ante-natal health of the mother; and
- HIV/AIDS.

6.2.2.2 Integrating health and nutrition information into a CFSVA

To enable achievement of WFP goals related to the prevention of severe malnutrition and the provision of nutritional support to moderately and severely malnourished individuals, it is essential that CFSVA provide reliable information on the scale, scope, and distribution of malnutrition (CDC/WFP 2005). Where possible, CFSVA should provide information on two distinct types of malnutrition:

- **Macronutrient malnutrition:** acute or chronic deficiencies in protein and energy. The groups most vulnerable to macronutrient malnutrition are pregnant and lactating women, children under 5 years of age, and women of reproductive age.
- **Micronutrient malnutrition:** deficiencies in key vitamins and minerals. Micronutrient deficiencies are widespread in developing countries and affect all age groups. However, given the technical requirements of accurately measuring micronutrient malnutrition, most CFSVAs will be limited to assessing the prevalence of stunting, wasting, and underweight, which are as much outcomes of macro- as of micronutrient deficiencies.

Previously, nutrition surveys have often attempted to include a food security component based on recommendations provided from WFP and NGO partners, including Action Against Hunger, CARE, and World Vision. However, given their focus on obtaining a precise estimate of malnutrition within a well-defined population, the findings of these surveys are often of limited use in explaining the wider issues related to household food security, since the scope of nutritional surveys often leaves less time to spend with a household to gather an equal breadth of food security–related information.

The preferred method of integrating analysis of nutritional status and household food security is to include health and nutrition modules in a household food security and vulnerability survey questionnaire. The majority of questions in these modules are adapted from Demographic and Health Surveys (DHSs) or from Multiple-Indicator Cluster Surveys (MICSs). The benefit of using these pre-formulated questions is that they help to standardize the information collected to enable comparison across diverse groups and regions, and they ensure that analysis of nutrition- and health-related indicators is consistent.

When collecting information on nutrition- and health-related indicators, it is important to keep two key objectives in mind, both of which have significant implications for the design of the CFSVA:

Primary objective: To link food security with nutritional outcomes, controlling for other influences (health/hygiene, caring practices).

Secondary objective: To provide indicative levels of key nutritional outcomes by zone, administrative boundary, or other grouping, as appropriate.

If nutrition information is collected to provide data on prevalence of malnutrition by certain strata, then strict sampling protocols will need to be observed. If nutrition information is collected only to link with measures of food security, sampling restrictions will likely be less significant.

Specific objectives for the collection and analysis of nutrition data as part of CFSVA will depend on the programming context, the availability of reliable secondary data, and a variety of other factors. Given the implications for sampling and other technical considerations, it is critical that the objectives of nutrition data collection be clearly stated, agreed upon by WFP and partner organization staff, and technically supported prior to the initiation of field activities.

6.2.2.3 Anthropometric indicators of malnutrition

A number of well-established anthropometric indicators exist for measuring physical status, growth, etc., whether from macro- or micro-related malnutrition.[87] These measures typically require collecting data on weight, height, or length; gender; and age of a subset selected from the population of interest.

When accurately taken, these measurements serve as reasonably accurate indicators of chronic malnutrition reflecting past growth failure, and indicators of acute malnutrition reflecting current macronutrient nutritional well-being (WFP 2005). For an in-depth review of these and other key nutrition and health indicators and definitions, please refer to A Manual: Measuring and Interpreting Malnutrition and Mortality.[88]

Indicators for measuring malnutrition among children[89]

For children, the following four main indicators of malnutrition are frequently used in CFSVAs (WFP 2005). They are all measures of growth (i.e. stunting measures linear growth; wasting and MUAC measure thinness due to lack of growth or actual tissue loss, and underweight is a composite of both).

- **Stunting** (low height for age, "shortness for their age") is a measurement of chronic malnutrition characterized by a slowing in the growth of a foetus or child and resulting in a failure to achieve expected length in comparison to a healthy, well-nourished child of the same age. Stunting is an indicator of past growth failure, and is associated with a number of long-term factors, including chronically inadequate levels of protein and energy intake, micronutrient deficiencies, frequent infection, inappropriate feeding practices over a sustained period, and household poverty. It is not, however, an accurate measurement of short-term changes in nutritional status.

87. *Anthropometry* is the term used to describe measurement of the human body.
88. Published by CDC and WFP, July 2005, available at http://docustore.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp097839.pdf
89. Children 0 to 59 months of age.

- **Wasting** (low weight for height, "thinness") is a measurement of acute malnutrition characterized by considerable weight loss or failure to gain weight, which results in a child's weight being substantially below what would be expected of a healthy child of the same length or height. Wasting is an indicator of current malnutrition and is associated with inadequate food intake, incorrect feeding practices, disease, infection, or a combination of these factors. Wasting in individuals and population groups can change quickly, showing marked seasonal patterns associated with changes in food availability and access, as well as disease prevalence. Because we do not need the age of the child to calculate wasting, the difficulties encountered in contexts where exact age is difficult to determine do not affect the accuracy of wasting as a proxy indicator of acute malnutrition.

- **Underweight** (low weight for age) is a composite measurement of stunting and wasting as it is influenced by both height and weight. Although underweight is a good indicator for assessing changes in malnutrition over time, care must be taken in interpreting this indicator, as it reflects both chronic and acute malnutrition, which is both its strength and weakness as an indicator. Low weight for age is a composite measure of stunting and wasting, as it could be influenced by both height and weight (an underweight child is too thin and/or too short for his age, thus does not weigh what a child his age should). Hence, underweight does not provide the information needed to distinguish between the two.

- **Mid-upper arm circumference (MUAC)** is a measurement of acute malnutrition. Although wasting is the preferred measurement of acute malnutrition, MUAC offers a quick and easy predictor of immediate risk of death due to macronutrient malnutrition. As such, it provides a useful tool for screening cases of acute malnutrition among children 12 to 59 months, particular in emergencies. In some cases, MUAC for age offers a more refined version of this indicator of acute malnutrition. Recent thinking indicates that MUAC may be less of a blunt tool than previously thought, and could be used as a basic anthropometric measurement.

Age (all **children 0 to 59 months**) rather than the more common age range of 6 to 59 months should be included in the survey target population if there is reason to believe that children from 0 to 5 months are unusually vulnerable to malnutrition.[90]

Indicators for measuring malnutrition among adults

For adults, there are two main indicators of (macronutrient) malnutrition:

1. **Chronic energy deficiency (CED)** is a measurement of underweight for non-pregnant adults determined by using the **body mass index (BMI)**.[91] An adult (18 years of age or older) with a BMI below 18.5 kg/m^2 is considered malnourished. CED can be used as a measurement of well-being and is a good proxy indicator for overall adult health. Low BMI is correlated with a large number of health-related outcomes, including early onset of chronic conditions and an increased risk of premature mortality.

90. Taken from: *A Manual: Measuring and Interpreting Malnutrition and Mortality*, CDC, WFP 2005.
91. Calculated by dividing weight in kilograms by the square of height in metres.

Pregnant women and those with oedema are excluded from surveys assessing BMI because of the bias introduced by weight gain not related to nutritional status.[92]

2. **Mid-upper arm circumference (MUAC)** can also be used for assessing acute adult malnutrition.

While the decision of which anthropometric measurement to use in a particular CFSVA will depend on a range of factors, some general guidance should be followed:

- ***Simple anthropometric measurements***

Where resources are limited, but nutritional indicators are considered critical to an assessment of food security, indicators such as MUAC of children under 5 (and possibly women of reproductive age) can be collected. These require less material cost, less training time for enumerators, and less time in the field than full measurements of weight and height.

- ***More complete anthropometric measurements***

Where time, resources, need, and expertise allow, it is recommended to collect:

- Age (children 0 to 59 months)
- Weight (children 0 to 59 months)
- Height/length[93] (children 0 to 59 months)
- Weight (women of reproductive age)
- Height (women of reproductive age)

MUAC of children and women of reproductive age may also be gathered, if this data is deemed useful by the country office and implementing partners.

Table 6.3a: Classification of severity of malnutrition in a community for children under 5 years of age[94]

Severity of malnutrition	Acute malnutrition (%) (weight for height) < -2 z-scores	Chronic malnutrition (%) (height-for-age) < -2 z-scores	Underweight (%) (weight for age) < -2 z-scores
Acceptable	<5	<20	<10
Poor	5–9	20–29	10–19
Serious	10–14	30–39	20–29
Critical	⩾15	⩾40	⩾30

Table 6.3b: Classification of malnutrition for individual children for mid-upper arm circumference (MUAC)[95]

Well-nourished	>13.5 cm
Mild malnutrition	12.5–13.5 cm
Moderate malnutrition	11.5<12.5 cm
Severe malnutrition	<11.5 cm

92. Taken from: *A Manual: Measuring and Interpreting Malnutrition and Mortality*, CDC, WFP 2005.
93. Children 2 years and older are measured standing, while children under 2 years are measured recumbent.
94. This table is taken from *A Manual: Measuring and Interpreting Malnutrition and Mortality*, CDC, WFP, 2005. The information in the table is itself taken from WHO publication, *The management of nutrition in major emergencies*, Geneva, 2000.
95. Taken from *The revised selective feeding guidelines for the management of malnutrition in emergencies*, IASC, November 2008 (draft).

Table 6.3c: Classification of adult malnutrition using body mass index (weight in kg/height in m^2)[96]

Mild thinness	17–<18.5
Moderate thinness	16–<17
Severe thinness	<16

- ***Other anthropometric indicators***

Although not typically included in CFSVAs, other anthropometric indicators such as BMI of men or other growth indicators of children above 5 years of age could be included. This is generally not recommended or necessary, however, and should only be done in partnership with other organizations specializing in nutritional surveys.

- ***No anthropometric measurements***

In some contexts, it may not be feasible to collect any anthropometric indicators, due to the existence of recent nutrition surveys that WFP does not want to duplicate (such as a DHS or MICS), the non-interest of partners, or restrictions in budgets. In these cases, secondary data will be used in the analysis and presented in the report. However, using only secondary data will limit the ability to make subsequent linkages between household food security and nutritional outcomes.

6.2.2.4 Indicators of micronutrient malnutrition

Micronutrient malnutrition indicators are more difficult to measure and diagnose than anthropometric measures of growth, and rely primarily on the presence of clinical symptoms. However, biological measures are more accurate indicators of malnutrition, and advances in technology now allow for field-based diagnostics to take place. It is especially important to monitor changes in micronutrient utilization among populations heavily dependent on food rations or in situations where individuals may be particularly sensitive to changes in ration composition (e.g. high HIV-prevalence areas). The three most common and preventable micronutrient deficiencies are those of iodine, iron and vitamin A. Table 6.4 describes each of these micronutrient deficiencies, its symptoms, and the common proxy indicators used to determine the micronutrient status of populations.

Table 6.4: Common indicators of micronutrient malnutrition

Micronutrient	Deficiency disease	Symptoms	Food source	Proxy indicators
Vitamin A	Xerophthalmia	Night blindness, Bitot's spots	Red, orange, and yellow fruits and vegetables, pulses, green leafy plants, liver, fish oil	Vitamin A supplementation for children and lactating women, consumption of vitamin A–rich foods
Iron	Anaemia	Clinical signs of pallor, tiredness, headache, and/or haemoglobin test	Meat, liver, green leafy vegetables	Iron supplementation for pregnant women
Iodine	Goitre, cretinism	Swelling of thyroid gland in the neck	Iodized salt, seafood	Iodized salt testing
Vitamin C	Scurvy	Painful joints, swollen and bleeding gums, slow healing of wounds	Citrus fruits, green leafy vegetables	Prolonged absence of fresh foods

96. Taken from the *Food and Nutrition Handbook*, WFP (no publication date).

6.2.2.5 Indicators for assessing the health environment, care, and feeding practices

As mentioned earlier, the health environment and access to health services are closely linked with nutritional outcomes (see Figure 1.1, Food and Nutrition Security Conceptual Framework). Determining the prevalence of the primary public health illnesses and the available options for treatment are important for assessing vulnerability, and for determining the underlying causes of malnutrition. Feeding practices and food preparation are also key determinants of infant and child health, growth, and susceptibility to illness. Depending on the context, each of the following indicators may be appropriate for inclusion in a nutrition component of a CFSVA (Ruel 2002).

- ***Disease prevalence***

Often there is a correlation between illness and malnutrition, especially among children. CFSVAs should use a standard recall period of two weeks for determining the prevalence of diarrhoea, fever, and acute respiratory infection (coughing with faster than normal breathing).

- ***Hygiene practices***

Research indicates that domestic hygiene practices such as disposal of faeces and hand washing at critical times (e.g. after defecation, before meal preparation, before eating) are important for child health and nutrition.

- ***Treatment of illness***

CFSVAs can incorporate questions regarding the use of professional medical treatment to treat child illnesses. It is also important to inquire about other health-seeking behaviours and home treatment of illnesses, such as increasing breast milk feeding or the use of oral rehydration solutions (ORS) when a young child has diarrhoea.

- ***Immunization***

Immunization is one of the most important and cost-effective interventions that health systems can provide. CFSVAs can use household surveys and, in some cases, hospital immunization records to determine access to and utilization of child immunization services. CFSVAs could seek information on immunization coverage for the six major vaccine-preventable diseases: diphtheria, tetanus, pertussis (DPT), childhood tuberculosis (BCG), polio (OPV), and measles). When there is no health card readily available, this information takes a considerable time and effort to collect, so often the CFSVA gathers only measles as a proxy for immunization in general.

- ***Feeding/breastfeeding practices***

Information gained from household surveys on the types of food consumed by children and the frequency of feeding should be used to analyse causal links with nutritional outcomes. CFSVA questionnaires may also include questions related to the timing of initial breastfeeding, the frequency and exclusivity of breastfeeding, the use of bottles, and the duration of breastfeeding and weaning practices.

- ***Food preparation and storage***

Where possible, CFSVAs may obtain valuable information on food preparation and storage. This information can provide insight into whether specific social or cultural issues affect food use. Information could come from secondary data or, when required, primary data collection (focus groups, household surveys, etc.).

6.2.2.6 Issues to consider when conducting health and nutrition surveys as part of a CFSVA

There are many valid reasons for incorporating health and nutrition components into a CFSVA. Most important, collecting information on health and nutrition allows a more holistic analysis of food security by investigating food utilization in addition to food availability and access. In addition to the considerations for primary data collection mentioned in sections 4.1 and 4.2, several specific factors need to be considered when conducting an analysis of nutrition and health at the household level.

Skills

An experienced nutritionist or other qualified person in the team must assist with the study design and the training and supervision of the enumerators.

Purpose of nutritional survey

When designing and implementing a nutritional survey as part of CFSVA, it is important to remember its value as a programme design tool. The survey can also serve as a baseline for nutrition surveillance or as a monitoring round during the lifespan of an intervention. Consistent monitoring and reporting of standard indicators of nutritional status not only improves the management of food assistance programmes, but also uses anthropometric measurements to provide valuable insight into programme impact. Still, attribution will remain difficult (i.e. are the changes observed due to the project or to other factors?).

Sampling

Detailed information on sampling is provided in section 4.1 of these guidelines. There are, however, additional considerations when conducting nutritional surveys as part of a CFSVA.

- ***Avoid sampling bias***
  In conducting sampling for nutritional surveys, particularly those involving collection of anthropometric data, it is critical to take precautionary steps to avoid sampling bias. In CFSVAs, using the standard sampling approach, households are selected whether or not they have children under 5. In households that have children under 5, information is recorded for all children. Selecting only one child (the oldest, at random, etc.) in a household with multiple children under 5 will bias the sample.

- ***Determine sample size***
  For surveys designed to measure changes in specific indicators over time or differences in indicators between distinct groups, the required sample size will be calculated based on the sample size formula in section 4.1.3.1 and Figure 4.2.
  For prevalence, 30-by-30 design is often used: 30 children in 30 clusters per reporting domain. These measurements of 900 children would, at a malnutrition rate of 50 percent and with a design effect of 2, allowing for some non-respondents, give a precision of +/- 5 percent.

References and standards

Nutritional surveys need to follow standardized methods in order to provide accurate data that can be compared across groups and over time. These methods must be assessed to ensure that the methods followed are in compliance with accepted standards. Evaluating each of the factors in Table 6.5 will allow for accurate interpretation and use of the nutrition and health surveys.

Table 6.5: Reliability checks for nutritional surveys

Factor	What to check
Sampling procedure and size, and sampling bias	Stunting, wasting, underweight (children) or MUAC or BMI (adults): • Ensure that a representative sample is taken (see Box 4.5). • Check for sampling biases, such as not all children in a household being selected, high non-response. • Check that the sample size and sampling methodology result in an acceptable 95 percent confidence interval of the reported prevalences. • Check that the confidence intervals are reported and the design effect is accounted for and that sampling conditions will not largely inflate the design effect. Standard procedures strictly applied everywhere: • All selected households visited, low non-response • All subject members of each selected household measured/interviewed; low non-response due to temporary absence from the home or other reasons
Measuring error	No systematic errors due to faulty measuring equipment, techniques, or recording: • Accurate scales read to 0.1 kg • Well-made height/length boards read to 0.1 cm • Accurate MUAC tape reads to 0.1 cm • Age data is not systematically rounded to the nearest year or 6 months. Reliable age estimation techniques used where birthdates are not reliably known/recorded. Ensure that missing and flagged cases are reported, and that the % of missing/flagged is relatively low (MICS cites less than 10% missing as a general rule)[97].
Measure to reduce bias and ensure accuracy and precision	All survey personnel trained following standard procedures and good practice guidelines: • All survey personnel already competent, or received adequate supervised practical field training • Competent and experienced trainers used • Supervisors verify the standard measuring and recording

Working with partners

In addition to building the capacity of WFP staff, collecting nutritional information for CFSVA should be seen as an excellent opportunity to improve collaboration with implementing partners and specialized technical agencies (international or national NGOs, UNICEF, WHO, MoH, etc.). This is especially true given that other organizations will frequently be responsible for collecting nutrition information or will be well positioned to support WFP efforts related to nutrition. The process of designing nutrition surveillance instruments provides an opportunity for WFP staff to engage in dialogue with partners on data collection issues before surveys are undertaken. Similarly, the analysis of findings will provide information on the causes of malnutrition that can be used to persuade partners to focus their activities in ways supportive of or

97. UNICEF, 2006, *Multiple-Indicator Cluster Survey Manual 2005: Monitoring the situation of children and Women*. Division of Policy and Planning, New York.

complementary to WFP food security activities. Finally, involving partners in nutritional assessments enables WFP to ensure that information collected by partners conforms to WFP reporting needs, and is systematically collected using statistically appropriate methods (CDC/WFP 2005).

Training of enumerators

Providing adequate training for enumerators in conducting household interviews and collecting anthropometric data is critical for ensuring accurate analysis of the links between household food security and nutrition. As with other aspects of a CFSVA, it is essential that all enumerators involved in nutritional data collection receive the same training, to ensure the quality and consistency of the data. At a minimum, training should include:

- Presentation of the objectives of the nutritional survey and explanation of the sampling method;
- Description of information requirements:
  - Overview of anthropometric measurements;
  - Achievement of a common understanding of all health and nutrition questions; and
  - Distribution of a written survey protocol guide to all enumerators.
- Demonstration of measuring techniques;
- Observation of repeated practice (using dolls, other enumerators, and children, when appropriate) of measuring techniques;
- Standardization exercise:
  - All enumerators should collect anthropometric measurements from the same set of children (usually around five). Measurements are then collated and compared to determine outliers and reasons for inaccuracies. This exercise should be repeated until an acceptable level of variation is achieved between enumerators.
- Field-testing:
  - It is important that pre-survey field testing is conducted (in areas not covered by the CFSVA) in order to practice household selection, anthropometric measurement, and completion of questionnaires; and
- Review:
  - Based on the results from the field test, survey supervisors should work with enumerators to review individual performance, relevance of the questionnaire, survey methodology, and logistics.

Equipment

The validity of anthropometric measurements ultimately depends on the accuracy of the data collected. It is therefore essential that all enumerators have both the skills and equipment needed to maintain consistent measuring techniques as well as to accurately read and record anthropometric data. For instance, it is critical that the age of children is accurately determined and recorded in order to determine whether the child should be measured standing (24 to 59 months) or recumbent (0 to 23 months) for height/length and for converting height and weight into the standard indices.

Measuring boards for length/height can typically be obtained from NGOs or UNICEF, or they can be constructed locally. They should measure up to 120 cm for children (0 to 59 months) and be accurate to within 0.1 cm. Ideally, they should also be lightweight, durable, and have few moving parts. Each team of enumerators should have their own board for measuring the length/height of children. Measuring tapes for collecting data on adult height can typically be obtained locally.

Hanging scales are often the most practical method of weighing children. However, where possible, WFP advocates the use of electronic scales. UNICEF and others have also found electronic scales to be durable and easy to use, especially given the option of weighing both mother and child. In addition to making the weighing process less stressful for infants and young children, electronic scales have the advantage of obtaining the weight of the mother for body mass index (BMI) calculations. Each field team should have their own scale, whether hanging or electronic. When UNICEF is present in the assessment area, there is often an agreement in place to allow WFP to borrow scales for use in anthropometric data collection.

Learning from experience:
- It can take months to purchase scales.
- Hanging scales need to be carefully zeroed/calibrated throughout the data collection.
- Bathroom scales are NOT acceptable for children.

6.2.2.7 Analysing anthropometric data

Once anthropometric and other nutritional data have been collected, they must be carefully entered and appropriately analysed. Each of these tasks requires specific skills, which means it is preferable that they are conducted by individuals with previous experience. Those responsible for entering data must have the skills necessary to accurately enter health information collected at the field level into computer files for analysis. As such, they should have sufficient experience to ensure that health- and nutrition-related information is entered accurately. Section 4.3 of these guidelines provides more detailed information on data entry (CRS 2005).

It is critical that personnel responsible for data analysis have previous experience with the statistical software used to analyse health data and produce tables and graphs depicting averages and proportions. It is particularly useful if they have previous experience with at least one of the other software packages used for analysis of nutritional data described in Box 6.3. Data analysts will be expected to calculate confidence intervals and z-scores using analysis software, and must have the ability to account for the sampling methodology (design effect) when conducting the analysis. These complex sampling and analysis issues are covered in Chapter 4 - Household-Level Data in CFSVA.

All CFSVAs should use the new WHO reference standards. From the WHO website, download the software needed to calculate z-scores using a variety of statistical packages, including STATA and SPSS.

Box 6.3: Nutritional analysis software for calculating z-scores

WHO Anthro

This software is recommended for calculating both the NCHS and the new WHO reference curve z-scores. The WHO Anthro 2005 software has been updated and released as WHO Anthro, with updated manuals for PC and MD. The software consists of the following modules:

- Anthropometric calculator
- Individual assessment
- Nutritional survey

The software also flags outliers (usually the result of incorrect measurements, coding errors, or incorrect ages). The software produces complete tables of anthropometric outcomes, including confidence intervals of prevalences, which can take into account the sampling methodology. WHO Anthro can be downloaded free from:

http://www.who.int/childgrowth/software/en/

Epi Info

The Centers for Disease Control and Prevention (CDC) and WHO have developed a free software package called Epi Info that can be used for both the calculation of anthropometric indices and some of the necessary analysis.

The software transforms the age, gender, weight, and height data of children into indices and z-scores (NCHS reference only), and flags outliers (similar to WHO Anthro). Once the anthropometric indices have been calculated, they can be easily presented in simple tables using the specified cut-offs and age categories consistent with normative standards. The Epi Info software and manual can be downloaded free from www.cdc.gov/epiinfo.

Nutri-survey

This Windows-based software is available from www.nutrisurvey.de. The programme was designed specifically for nutrition surveys by the Work Group on International Nutrition of the University of Hohenheim/Stuttgart, in cooperation with the German Agency for Technical Cooperation (GTZ). The purpose of the programme is to integrate all steps of a nutrition baseline survey into a single programme. The programme contains a standard Nutrition Baseline questionnaire that can be easily customized for the specific site, a function for printing out the questionnaire, a data entry unit that controls the data being entered, a specially adapted plausibility check, a report function, and a graphics section. For further statistical evaluation, the data can be exported to SPSS or other statistical programmes. However SPSS, SAS, Stata, and other packages are not used for calculating z-scores unless using the specific scripts from WHO.

A final note on collection of nutritional information for CFSVAs

While the nutritional indicators described in this chapter are powerful, useful, and well recognized measures, they are also difficult, time-consuming, and costly to collect. When included in CFSVAs (or any WFP data collection exercise), shortcuts should not be taken that will compromise data collection or analysis. Such compromises can result in data that is unreliable or, worse, that misdirects interventions.

Analysts of nutritional data will be expected to calculate average z-scores for height for age (HAZ), weight for age (WAZ), and weight for height (WHZ). They must also be able to calculate the proportion of children whose z-scores are -2 and -3. For each of these values, they will also need to calculate 95 percent confidence intervals. There are currently two different references commonly used to calculate z-scores for WAZ, HAZ, and WHZ. The first and most common are those established by the U.S. National Center for Health Statistics (NCHS). Alternatively, analysts may use the more recently established references developed by the WHO. Current recommended practice in

CFSVAs is to report basic results from both references and run analyses based on the new WHO reference. While Epi Info can only calculate comparisons to the NCHS reference, Anthro is capable of calculating comparisons to both the NCHS and the new WHO reference.

Again, the steps involved in analysis of nutritional data are relatively complex and involve rigorous statistical methods to ensure accuracy. This requires the direct involvement of individuals with significant experience in the management of nutrition and health information. Box 6.4 identifies several reference documents that may prove useful in incorporating assessments of nutrition into CFSVAs.

Box 6.4: Additional sources of information on analysis of nutritional data

World Food Programme (WFP)
Thematic Guidelines: Nutrition and Health, 2005
http://www.wfp.org/food-security

Centers for Disease Control and Prevention (CDC)/World Food Programme (WFP)
A Manual: Measuring and Interpreting Malnutrition and Mortality, 2005
http://www.unhcr.org/publ/PUBL/45f6abc92.pdf

Food and Nutrition Technical Assistance (FANTA)
Anthropometric Indicators Measurement Guide, revised edition, 2003
http://www.fantaproject.org/downloads/pdfs/anthro_2003.pdf

Key references: Food security analysis

- CDC/WFP. 2005. *A Manual: Measuring and Interpreting Malnutrition and Mortality*.
- Cogill, Bruce. 2003. *Anthropometric Indicators Measurement Guide*. Revised edition. Food and Nutrition Technical Assistance (FANTA).

6.2.3 Market analysis

This section provides suggestions on the relevance of market information to the CFSVA process and what type of market information needs to be collected and analysed for a CFSVA. The goal is that such information will lead to a better understanding of the role of markets in food security in a specific country and will serve as a baseline for other outputs, such as emergency food security assessments (EFSA), food security monitoring systems (FSMSs), and local procurement assessments, and will contribute to better identifying recommendations for response options during crises.

As an integral part of CFSVA, market information and analysis is more likely to be scattered under different headings in the standard outline of a CFSVA, though a core part will remain in a market chapter.

6.2.3.1 Importance of market information in food security and vulnerability analysis

Very few, if any, households produce all their food needs. Most households in urban areas and many (often vulnerable households) in rural areas, acquire much of their food and other necessities through market purchases or barter exchange,[98] depending on one market or another to gain income. In selected countries, WFP household surveys suggest that the majority of households consider markets as a main source of food, especially during the lean season (Table 6.6). Households with borderline food consumption tend to devote a larger proportion of their expenditures to food, compared to other food consumption groups. Vulnerable households are thus likely to be affected by changes in market prices, terms of trade, or wage rates, and they face difficulties regarding access to food, with the risk of falling into poor food consumption groups.

Table 6.6: Household dependence on markets for food, in selected countries

Country	WFP household (HH) surveys			USDA estimates
	Food expenditure/total expenditure (%)	Market as a major source of food (% of HH)	Source and survey season	Food expenditure/total expenditure (%)
Mali	National average: 52 Borderline HH: 55	70	WFP/CFSVA (2006) HH survey in post-harvest (2,074 HH)	53
Nepal	National average: 50 Borderline HH: n/a	n/a	WFP/CFSVA (2006) HH survey in lean season (1,676 HH)	58
Niger	National average: 63 Borderline HH: 72	> 70 (excluding milk)	WFP/CFSVA (2005) HH survey in lean season (1,800 HH)	n/a
Lao, PDR	National average: 65 Borderline HH: 68	< 40 (cereals and pulses)	WFP/CFSVA (2007) HH survey in harvest season (3,926 HH)	n/a
Liberia	National average: 66 Borderline HH: 72	> 80 (cereals)	WFP/CFSVA (2006) HH survey in post-harvest season (5,409 HH)	n/a
Rwanda	National average: 55 Borderline HH: 75	65	WFP/CFSVA (2006) HH survey in lean season (2,786)	n/a
Tanzania	National average: 63 Borderline HH: 64	66	WFP/CFSVA (2006) HH survey in lean season (2,772 HH)	73
Timor-Leste	National average: 55 Borderline HH: n/a	59	WFP/CFSVA (2006) HH survey in lean season (1,700 HH)	n/a

Baseline market information is considered very useful for providing answers to the main questions generally addressed by a CFSVA about vulnerable people: the "who, how many, where, why, and what" questions. Market analysis within a CFSVA has the following main objectives:

98. In exceptional cases, such as refugees living in closed camps, food aid may account for the total supply of food, but even in such cases, internal markets quickly emerge within the camps.

- Provides baseline information prior to emergencies against which the impact of various shocks can be compared.
- Analyses the effects of markets on food security and vulnerability.
- Predicts possible responses of markets to potential shocks in an emergency.
- Indicates the potential role of markets for food security interventions.

Market prices and the level of effective demand (purchasing power) can provide short-term incentives for commercial trade flows (imports) and, in the medium term, production. Merchants' stocks often serve as one form of local reserve (e.g. grain stores). Ideally, trade flows should enhance local market supplies and help stabilize prices when needed. Barriers to market functioning, however, can limit the effectiveness of markets, even in the presence of unmet demand, and markets can be manipulated for private gain (e.g. by hoarding). In addition, assistance interventions almost inevitably have an impact on markets, except when they are effectively timed and targeted to households that have no effective demand or when the aid volumes are very small compared to the regular trade flows.

Based on the analysis of market information, the findings can contribute to decision-making by:
- Better targeting needy households by distinguishing areas of good availability and functioning markets from other areas;
- Using the results of shock-response simulations to identify key market variables requiring monitoring in view of emergencies; and
- Improving situation analysis and appropriate decision-making during emergency needs assessments.

6.2.3.2 Key market issues to be addressed by a CFSVA

Baseline information on markets is required to:

1. Analyse:
   a. **market environment:** external factors that affect market conditions;
   b. **market structure:** characteristics of the markets that significantly affect the behaviour and interaction of buyers and sellers; and
   c. **market conduct and performance:** the dynamic of market actors' behaviour and particularly price dynamics.

2. Identify the relationships between aggregate food availability, different categories of markets, and vulnerable people's food access (i.e. the physical access of households to food they purchase on the market and their economic access, which is determined by their purchasing power); and

3. Determine the potential threats/risks to markets' capacity to supply the necessary food and other basic goods at prices affordable by households.

As an integral part of a CFSVA, market baseline information will target the key aspects of these issues that may be relevant to food security and decision-making. In this context, key questions and their interpretation can be summarized in Table 6.7.

Table 6.7: Key market issues to analyse

Key issue	Interpretation
Market structure, conduct, and performance How do markets function, and what are the main characteristics of market actors?	• Establishes marketing channels through which food flows. • Evaluates whether markets are integrated. • Identifies market power and interests of actors.
Food availability Are markets capable of supplying food commodities in sufficient quantity all the time?	• Evaluates various sources of food supply and other essential goods (production, stocks, imports, food aid, etc.), the regularity of supplies, and impediments to food availability on markets. • Identifies potential threats that would affect food availability on markets.
Food access How do households depend physically and economically on markets?	• Evaluates whether households can afford food commodity prices on the markets, using their income sources and how price changes affect their effective demand (purchasing power) throughout the year. • Anticipates potential threats that would affect households' access to food sourced from markets.
Shock response How is the market likely to respond to a shock, and what are the likely consequences for different types of household vulnerability?	• Simulates the impact of shocks on household food security and vulnerability. • Identifies significant market variables that may be disrupted by an emergency and therefore may affect household food security and vulnerability.

6.2.3.3 Market data collection and analysis

Market data collection and sampling issues

From a food security analysis perspective, market data collection is most useful in areas where a significant proportion of the community is partly or largely dependent on market transactions, as is now the case in most areas.

Market data collection usually takes place through various sources:

- At a macro-level, **secondary data** are useful for time series (trend and seasonality) analysis of production, prices, and the macro-economic background (trade policies, regulations, institutions).

- At a meso-level, trader questionnaires help capture specific issues related to market structure, conduct, and performance as defined in Table 6.7. Trader interviews can help give an idea about whether and why markets are not working. In surplus areas, market data can provide useful information on whether there is local purchase capacity. In deficit areas, it provides indications on whether non-food options such as cash and vouchers can be considered, provided markets are functioning (i.e. food flows in to meet demands, and prices are stable) notwithstanding operational constraints. "Open" key informant interviews — such as with large importers or cereal traders — are useful for triangulating information on market structure and performance.

- At a micro-level, **community interviews** are key for explaining other factors affecting market environment (infrastructure, transport, communication, physical access, seasonality), demand, and supply issues faced by communities.

- **Household surveys** are a part of the micro-level analysis of markets that helps capture households' interactions with markets (actual and potential buying and selling behaviour, marketable surpluses, and incomes). Knowledge on households' market participation can help in an analysis of their market dependence with regard to their food security status. Such knowledge can also inform local purchase decision-making, especially when purchases are targeted to smallholders. In order to meet such objectives, household survey questionnaires are expected to include questions on production quantities for the main food commodities (cash or staples) and livestock, share of production stored for either consumption or seeds, share of production used to reimburse debts, share of production for sale, quantity of food commodity purchased, and periods of sales and purchases. Household surveys usually include questions on income and food sources (including from markets).

It is desirable to code the location of households, communities, traders, and markets so that the results can be linked with each other.

While sampling methods are clearly established for community and household questionnaires through representative random sampling, it is rather difficult to fully randomize trader samples because there is hardly any record of the number and type of private traders operating in each market. Expert judgement and key informants are therefore crucial to building trader samples.

Purposive sampling is generally used for trader surveys to ensure that predefined categories are included (e.g. small traders/retailers versus wholesalers; grain traders versus fruit/vegetable/livestock traders) according to survey objectives. Market types are also generally targeted accordingly. For instance, rural or village markets are likely to coincide with small traders and assemblers who interact mostly with farmers and consumers, whereas urban markets (at district and regional levels) are likely to incorporate more stakeholders starting from farmers, assemblers/collectors, wholesalers, and retailers.

In rural settings, trader surveys take place preferably on the main market day of the week. In urban settings, trader surveys take place on almost any day of the week. It is therefore useful to have some knowledge of where markets are physically located so that several markets and different market stakeholders (i.e. retailers, wholesalers, transformers, assemblers, transporters or collectors) can be visited. This will ensure that the diversity of the region surveyed is represented.

The sample scheme of the trader survey is generally stratified *ex ante* by type of trader, but the sample size is large enough to offer sufficient representativeness within the limited financial and logistic means of a study. There is no predefined sample size. Once a sample size is chosen, including market and trader types, enumerators are asked to interview traders randomly, following the instructions given by the team leader. Box 6.5 provides an example of sample size constructed in Timor-Leste.

Box 6.5: Trader survey in Dili, EFSA 2007: Sampling locations and sample sizes

A purposive sampling method was used in the market survey following the geographic distribution of dominant traders (i.e. wholesaler dry food, wholesaler fresh food, retailer, and petty trader). There are five major market locations/clusters in Dili (Halilaran, Audian, Comoro, Mercado Lama, and Taibesi). The sample sizes and target groups in each cluster were drawn purposively, taking into account the limited number/availability of the target groups. The total number of respondents in the survey was 117 traders. Had there been a great number of wholesalers, their total number would have been increased, as the ultimate goal was to interview a significant number of traders to be able to make more sense out of their perceptions.

Type of trader	Number of market locations					Total
	Halilaran (wholesale market)	Audian (wholesale market)	Comoro (retail market)	Taibesi (retail market)	Mercado Lama (retail market)	
Wholesaler dry food (cereals and pulses)	8	4	0	0	0	**12**
Wholesaler fresh food (fruits and vegetables)	15	0	0	0	0	**15**
Retailer (small shops)	14	0	6	6	12	**38**
Petty trader (vegetables, fruit, fish, meat, chicken)	14	0	15	14	9	**52**
Total	51	4	21	20	21	**117**

A structured and pre-tested questionnaire was used in the survey. Three teams of three people each (a total of nine people) from the Department of Statistics collected the market data over two days. A two-day training session was given to enumerators prior to data collection. Subjects were randomly selected and interviewed in the sampling locations.

Market data analysis in a CFSVA

When analysing market data within the Food and Nutrition Security Conceptual Framework, it is important to make a clear distinction between cause/effect relationships. Empirical evidence indicates that households' capacity to buy food depends on variables at the household and macro level (national, regional, and sub-regional).[99] The macro dimension is mainly about food availability and supplies. At the microeconomic level, households and individuals have at their disposal assets such as labour, human capital, physical capital, social capital, common and public goods. These assets are used to generate income in various forms, including through production, earnings, and returns to assets, sale of assets (including livestock), and remittances. Households have access to food through markets, such as sale of agricultural produce.

99. See literature review in *World Hunger Series: Hunger and Markets*, WFP (due for publication, 2009).

Annex 11 provides an indicative outline, detailing the issues mentioned in 6.2.3.3, with some suggestions on the analytical contents, indicators, tools, and source of secondary and primary data. In addition, it proposes the location of such information in the general outline of a CFSVA. The most common topics suggested for inclusion in the analysis concern:

- The **economic background and market environment**, including trade policies, regulations, and institutions;

- **Availability** and diversity of major food crops (especially cereals) and livestock that are normally traded nationwide and at the sub-national level;

- **Market structure**, including marketing chains of main commodities, trade flows (including cross-border trade), and traders' selling and buying behaviour (including seasonal patterns, i.e. during harvest, post-harvest and lean seasons);

- **Market conduct and performance**, including producer and retail prices of major commodities (e.g. staple food, cash crops, animals and livestock products) and price trends analysis (e.g. import parity prices and seasonality); market integration analysis (e.g. price differentials and transaction/transport costs between regions). Annex 12 provides further details, especially on market integration and price analyses;

- **Market participation of households** (including the seasonality thereof), including household selling and buying behaviours and household purchasing power (e.g. skilled and unskilled wage rates, agricultural incomes, transfers and remittances, terms of exchange of cash or livestock into staple food); and

- **Market-related risks analysis** (income and price shocks).

Macro environment

Information on macro-economic performance and policies affect food security at the national as well as the household level through markets. A brief overview of such information is important for helping us understand, first, the nature of and changes in policies affecting the incentives of the actors (traders and consumers) and their ability to respond to those incentives (e.g. key factors such as trade policy measures can undermine food availability to households due to high transaction costs), and, second, the impact of business environment changes on domestic commodity markets and the food security situation of vulnerable population groups. Notwithstanding emergency situations (natural or man-made disasters), market failures generally occur when the market environment is distorted. Food may therefore be available, but not necessarily where households need it the most, even if they can afford the price.[100]

100. Bonnard P., 2001, *Improving the Nutrition Impacts of Agricultural Interventions: Strategy and Policy Brief*, Washington, D.C.: Food and Nutrition Technical Assistance (FANTA) Project, Academy for Educational Development.

Policy measures that affect markets and food security are multi-dimensional:

- **Sector-specific policies**: labour market (wage rate and training policies), capital market (credit and interest rate policies), land market (property rights and rental markets), social sectors (government social protection and safety net policies), and government public investment in roads and transport.

- **Macro-economic policies**: exchange rate policies (e.g. effects of exchange rate misalignment), trade and border policies (tariffs, non-tariff and subsidy measures), domestic market interventions (price interventions on domestic markets, e.g. taxes, subsidies, government food purchases and price controls), and general price level and price stability (measures affecting inflation).

- **Regulatory environment**: food quality (rules and regulations, implementation of both product processing and conservation standards and enforcement), food health and safety (sanitary standards, foodborne diseases).

Aggregate food availability

The analysis of food availability, especially at the sub-national level, captures surplus/deficit areas. Whether food is available to households depends not only on production and stocks, but also on whether markets make food available in a particular region through flows from other regions (domestic markets) or through imports, including aid (international or regional markets).

It is therefore desirable to analyse production patterns in combination with stocks, national reserves, imports, internal trade flows, and food aid. Adequate food availability at the aggregate level is considered a necessary condition for household food security, but insufficient for achieving adequate food access at the household level.

Market structure, conduct, and performance

The ways in which markets work to smooth out supply/demand pressures and the impacts of the shocks or seasonal patterns can be understood through the market conduct and performance analysis. The functioning of markets is generally based on information on market chains, competition, and price patterns (price differentials between regions, seasonality, and volatility margins). This analysis can help us understand issues related to the movement of food commodities. For example, baseline information on market integration can help identify the main functions of markets that can be affected by a crisis.

In a crisis situation, comparison with the baseline can help actors decide whether the scope of the impacts distorts markets to the point that interventions are required. The analysis of price patterns (differentials, seasonality, cycles, and volatility) can provide an understanding of the potential impacts of price shocks on household food security (see 6.2.4).

Market participation of households

Household market participation can be analysed through income sources, purchases from markets, reported food sources, sales on markets, and periods of sales/purchases. Households' market participation is determined both by marketable surplus/deficit status (net buyer/seller) and their ability to physically and economically access markets. Physical access indicators such as distance, travel time, means of transport, and transport costs to markets are factors that can undermine household income opportunities and food access through market purchases. Households' economic access to markets depends mainly on their purchasing power, which is influenced by prices and incomes (generated from production, assets, labour, transfers, and remittances) and the availability of own production and stocks. All those factors, and thus economic access, show seasonal patterns. Prices are determinants of a household's real income and directly influence the level of food purchases.

To the extent possible, participation in the (casual and permanent) labour market can also be covered because of its role as an income source as well as a coping mechanism. An example of household casual labour market participation is given in Box 6.6.

Box 6.6: Labour market participation of rural households in Lao, PDR

The Lao PDR CFSVA finds that casual labour is an important part of income sources of villages, especially during the lean season and for non-farm activities. The existing demand is for unskilled labour, mostly agricultural work (bush cleaning, paddy land maintaining, or land fencing) for neighbours. The practice of agricultural casual labour or labour exchange is common in the wet season and mainly during the lean season to meet household food needs. In general, payments are made either in kind, with some 5 kg of paddy rice per day, or in cash. According to the community survey, the daily wage rate of unskilled labour is kept at between 10,000 and 20,000 Kip, with large variations among villages in each province. The highest rates are paid in Vientiane and Bolikhamxay provinces, while the lowest are paid in Pongsaly and Luang Prabang provinces.

Non-farm activities offer limited employment opportunities and cash income but are less preferred than permanent agricultural activities because they are considered by villagers as temporary and risky, hence low paying relative to the effort expended. In general, the lack of skills put villagers at a disadvantage in both finding and keeping employment opportunities in non-farm activities such as mining, logging, and road and house construction.

Source: WFP, 2007, CFSVA Lao, PDR.

Markets and household vulnerability and risk

Household vulnerability refers to how a household's livelihood system would be exposed if a market shock occurred and to the capacity of that household to cope with or withstand the effects of the shock. Market analysis is expected to highlight risk factors (possible shocks and hazards related to markets), provide some indicators of household capacity to cope, and the implications for their food security situation. The analysis of supply shocks (e.g. drought, floods, and pests), demand shocks (e.g. price and income seasonality and volatility), market-related coping mechanisms (e.g. sales of assets, daily labour, purchases using credit, debts), and their actual impact on household food security can inform decision-making. If a household is vulnerable and, at the same time, likely subject to a market shock, it is considered "at risk for food security."

On the demand side, price changes have a direct impact on households' real incomes, all things being equal. Notwithstanding data limits, the analysis of the impact on household vulnerability is crucial. Box 6.7 provides arguments on the importance of price analysis as part of household vulnerability analysis.

Box 6.7: Potential impacts of price increases on household vulnerability

Price shocks have direct implications on food-insecure populations. Prices are determinants of a household's real income and directly influence the amount of food it can purchase. Food price volatility is cause for concern because of its direct impact on household food security. Low food commodity prices are generally seen as a stress on incomes of farmers, who are net sellers. Price increases are expected to propagate both the incidence and depth of poverty and food insecurity as many households are net buyers of food, including small farmers. Usually both urban and rural households depend on markets, though to a greater degree for urban households. Rural households depend on income-generating activities, such as subsistence agriculture, seasonal agricultural wage labour, and various forms of off-farm self-employment. The resources generated through these activities are often very low, thus contributing to food insecurity. When under stress, various coping mechanisms are employed to reduce the severity of deprivation and avoid more irreversible processes of destitution.[101] Low incomes, combined with the volatility of agricultural incomes and production that depends on weather conditions, prompt rural households to resort to markets to meet their food needs. They can therefore face deteriorating conditions due to price increases.

In the poorest households, there is a direct and immediate negative impact on their food security. The degree of being affected by food price fluctuations is to a large extent dependent on livelihood, gender, the number of dependants to be supported by economically active household members, the degree of dependence on markets, the degree of integration into social networks, and income levels. In general, the poorest households consume predominantly cereals and only very small quantities of other foods. Higher prices can therefore result in a poorer diet and nutritional status, as families allocate more of their incomes to purchases of energy-dense cereals and less expenditures on nutritious food and health care.

	Poor food basket (%)	Borderline food basket (%)	Fairly good food basket (%)	Good food basket (%)	Total per region (%)
Assaba	0.0	5.8	0.5	0.1	**6.4**
Adrar	0.0	0.3	0.0	0.2	**0.5**
Brakna	0.0	0.2	2.1	7.2	**9.5**
Gorgol	0.7	3.7	0.9	0.7	**6.0**
Guidimakha	0.2	9.3	1.2	3.7	**14.4**
Hodh El Charghi	0.0	0.0	0.3	0.1	**0.4**
Hodh El Ghardi	0.0	0.0	0.1	0.0	**0.1**
Inchiri	0.0	0.0	0.0	0.0	**0.0**
Tagant	1.1	2.7	0.1	0.1	**4.1**
Trarza	0.0	1.0	2.0	11.4	**14.4**
Nouakchott	NA	NA	NA	NA	**NA**
Total population (%)	**0.2**	**2.5**	**0.9**	**3.2**	**6.8**

101. Davis B. et al., 2007, *Rural Income Generating Activities: A Cross-Country Comparison*, EFSA Working Paper 07-16, FAO, Rome.

For example, using the 2005 CFSVA data for Mauritania, a study by WFP concluded that about 7 percent of the total population is at risk, as they would not be able to afford the same food basket as in 2005 with December 2007 food commodity prices (see table above). The most affected households are primarily food-deficit farmers who are vulnerable to erratic weather conditions (droughts and floods) and with low incomes from livestock. This estimate is based on the cost-of-food-basket approach, which estimates the proportion of vulnerable households (HH), i.e. that cannot afford the cost of a baseline food basket anymore as a result of higher food prices. The CFSVA usually classifies households into food consumption groups (poor, borderline, acceptable, and good). This grouping is based on diet diversity and a frequency score (food consumption score). The underlying rationale of the cost-of-food-basket approach is that households might not be able to afford the cost of their previous food basket anymore and are at risk of dropping from one food consumption group to a lower one. This would happen if their current real food expenditure was above the baseline figures as a result of higher food prices.

6.2.4 Risk analysis

Individuals, households, communities, and even nations face multiple hazards from different sources. Risks are the combination of the probability or frequency of occurrence of a defined hazard and the magnitude of the consequences (Benson 2005). Hazards often cannot be prevented, and if they materialize, they can generate a shock that affects individuals, households, and communities in both predictable and unpredictable ways (TANGO, 2004).

In the last two decades, more than 1.5 million people have been killed by natural disasters. Worldwide, for every person killed, about 3,000 more people are exposed to hazards. Natural disasters and other hazards have a significantly greater impact on lesser-developed countries than on more developed countries.[102]

The objective of food security risk analysis is to identify populations and regions likely to experience serious declines in their future food security status (i.e. food consumption, livelihoods disruption) due to effects of a particular hazard. Within the context of CFSVA, risk analysis combines hazard analysis with vulnerability analysis. There are both advantages to and limitations in performing risk analysis as part of a CFSVA.

Risk analysis and scenarios can:
- Identify geographic areas and populations at risk, enabling decision-makers to define proper strategies and interventions;
- Highlight the key factors contributing to increased vulnerability among households; and
- Evaluate the potential effects of these factors on households.

Risk analysis and scenarios cannot:
- Accurately predict or quantify the magnitude or extent of a shock;
- Accurately predict or quantify the effect on households and their capacity to cope; or
- Accurately predict or quantify the resulting impact on food security.

102. UNDP Bureau for Crisis Prevention and Recovery, 2004, *A Global Report: Reducing Disaster Risk: A Challenge for Development,* New York.

Food security risk analysis consists of the following steps:

- Identify the main risks a given community faces;
- Assess the probability and intensity of these hazards occurring;
- Assess the vulnerability of households and communities becoming food insecure because of these hazards; and
- Assess multi-hazard risks and hazard hotspots at the sub-national level.

Combining information on multiple hazards and vulnerability gives an indication of the likelihood that each of these hazards may result in food insecurity.

6.2.4.1 Definitions relevant for WFP risk-reduction activities

The degree of vulnerability depends on the nature of the risk, how it would impact community and household resilience, or the "ability to bounce back or recover after adversity or hard times, to be capable of building positively on these adversities" (Mission 2005).

A household's resilience often is related to:

- The magnitude of the shock a household or community can absorb and still remain viable;
- How well a household or community can self-organize after the exposure to the hazard to maintain an acceptable level of functioning and structure;
- How well a household or community can learn from these difficult circumstances; and
- The household's characteristics, notably its assets and livelihood strategies.

When exposed to a shock, households often see a decline in welfare levels, which can include overall reductions in food availability, loss of or diminished access to food, fluctuation in income streams, and poor nutrition. The greater the share of household resources devoted to food and health service acquisition, the higher the vulnerability of the household to food and nutritional insecurity.

Households who are able to minimize the loss of welfare resulting from shock exposure by employing livelihood and risk management strategies are likely to be **less vulnerable** than those who cannot cope with, or mitigate, the effects of risks. Therefore, livelihoods are secure when households have secure ownership of, or access to, resources (both tangible and intangible) and to income-earning activities that can off-set risks, ease shocks, and meet contingencies. When households are able to acquire, protect, develop, utilize, exchange, and benefit from assets and resources, livelihoods are secure (CARE 2002). To some degree, certain shocks – such as annual flooding and drought – can be anticipated by households. In these situations, households engage in risk-reduction strategies such as rehabilitation of flood banks, planting drought-resistant crops, or seasonal migration.

Other types of risks preclude preventive planning. Policy failures or economic downturns, civil conflict, and epidemics are generally unpredictable. In such circumstances, households typically engage in **risk-mitigation** and **risk-coping strategies**. Examples of the former include greater deployment of adult household labour, use of informal credit systems, and reduction of non-essential expenditures. Risk-coping measures are used in extreme situations, when strategies are not feasible or possible. Common examples include drastic reductions in food intake, selling/mortgaging of productive assets (livestock in particular), borrowing from neighbours, and deploying child labour.

The following definitions may prove useful in incorporating risk analysis into CFSVAs and have been adapted from the United Nations International Strategy for Disaster Reduction (UN-ISDR) terminology.

Food security risk analysis: A methodology to determine the nature and extent of risk to food security by analysing potential hazards and evaluating existing conditions of vulnerability that could pose a threat to household food security.

Hazard: The potential to cause harm; also, the probability of a potentially damaging phenomenon occurring within a given time period and area. Hazard can be expressed mathematically as the probability of occurrence of an event of certain intensity, in a specific site, and during a determined period of exposure.[103]

Hazard analysis: Identification, studies, and monitoring of any hazard to determine its potential, origin, characteristics, and behaviour.

Vulnerability to food insecurity: The conditions that increase the susceptibility of a household to the effect of hazards on its food security. Vulnerability is a function of a household's exposure to a specific hazard (e.g. flood, drought) and its coping capacity (or the direct impact of the hazard on the household, mitigated by its coping capacity).

Coping capacity: The means by which households use available resources, and their ability to face adverse consequences that could lead to a disaster.

Risk to food insecurity: The probability of food insecurity resulting from interactions between a natural or human-induced hazard and vulnerable conditions. The probability of a loss of food security depends on the hazard, vulnerability of households, and the number of households in the affected area. This relationship can be represented by the following equation:[104]

$R = H \cdot Pop \cdot Vul$

Where

R is the **risk** for food insecurity (in number of HH/year in a sub-area)

H is the **hazard**, which depends on the probability of a given hazard of a certain intensity in that area (in percentage)

Pop is the **population** living in a the area (in number of HH)

Vul is the **vulnerability** of the population (a function of the exposure of lives and livelihoods and household resilience with regard to the effects on their food security) (as a percentage of all HHs in the sub-area)

The risk to food insecurity from multiple concurrent hazards is obtained by aggregating the risks from each individual hazard.

103. United Nations University, Institute for Environment and Human Security (UNU-EHS).

104. Adapted from: UNDP, 2004, "*Reducing Disaster Risk: A Challenge for Development*."

6.2.4.2 Hazard analysis

Individuals, households, communities, and even nations face multiple hazards from various sources. Hazards often cannot be prevented, and if they materialize, they can generate a shock that hurts individuals, households, and communities in both predictable and unpredictable ways (TANGO 2004). Hazard analysis is an assessment of the level and type of hazard exposure experienced by a population. These hazards can come from exogenous climatic or environmental sources, institutional constraints, economic changes, political edicts, and conflict or social change. Valid hazard analysis is central to conducting an accurate and informative vulnerability analysis. Key to this type of analysis is consideration of:

- Policy-relevant categorization of hazards;
- The history of hazards, which includes the frequency of occurrence and severity of impact;
- The type of exposure to such hazards; and
  - Covariate (collective): these include many natural hazards (droughts, floods), but also price changes, where most households or communities are exposed; covariate hazards affect groups of households, communities, regions, or nations;
  - Idiosyncratic: these hazards have a more random and individual distribution, such as accidents, injury, theft, or human lifecycle events such as (non-communicable) disease and death; they may occur for individual households but not whole communities.

Whether the hazard is collective or individualistic will determine the type of risk-management strategies required to manage the risk.

Hazard/Risk Inventory

A hazard inventory identifies the types of shocks that have a possibility of occurring in a given region. For each shock, the frequency, magnitude, duration, timing, speed of onset, correlation, and geographic location are specified. Different hazards may result in similar shocks, which means that the particular challenges faced by populations may come from a variety of unrelated sources. (Table 6.8, on page 249, presents four categories: environmental, social, economic, and conflict hazards/risks).The list of hazards/risks should be as comprehensive as possible and should reflect the hazards/risks most likely to produce the shocks that incur the most severe impacts.

The **frequency** of a particular hazard, its **spatial dimensions**, and its **timing** need to be assessed. Some hazards, such as earthquakes, may strike infrequently, while a hazard of low rainfall may be experienced almost every year. Timing refers not only to periodicity but also to how the hazard relates to impacts during the course of a year. For example, floods/droughts will have differential impacts on an agricultural population depending on when they occur in relation to the agricultural calendar. In addition, there may be periods of the year when exposure to several shocks increases, simultaneously increasing the likelihood that a household will experience difficulties.

Example of Drought Hazard Analysis: Water Balance Model

An appropriate example of drought hazard analysis is provided by the FAO Water Balance Model, which is particularly applicable in countries relying on rain-fed agriculture where insufficient availability of water presents a major constraint to food security.[105] The water balance model is typically used to estimate yield reduction due to water stress.

Water balance follows the principle of supply and demand: water is supplied to a crop in the form of precipitation or irrigation soil water retention; demand is created by requirements for water in the crop, soil, and atmosphere. When demand exceeds supply, the crop does not have enough water for its optimal development and is stressed. In order to accurately gauge crop water requirements, FAO developed the water balance model.[106]

The most important inputs to the model are precipitation and potential evapo-transpiration (PET). FEWS NET at the United States Geologic Survey (USGS) calculates daily PET values for Africa at 1.0 degree. Rainfall estimate (RFE) images for the African continent are obtained from NOAA at 0.1-degree (~10 km) spatial resolution. In addition, the Water Requirement Satisfaction Index (WRSI) model uses relevant soil information from the FAO (1988) digital soils map and topographical parameters from Digital Elevation Model (DEM)–derived data. Key for drought risk analysis is that a reduction in WRSI leads to reduced productivity of a crop.[107]

Box 6.8: EM-DAT and natural hazard statistics

Since 1988 the WHO Collaborating Centre for Research on the Epidemiology of Disasters (CRED) has been maintaining an Emergency Events Database, EM-DAT (http://www.emdat.be/). EM-DAT was created with the initial support of WHO and the Belgian Government.

The main objective of the database is to serve the purposes of humanitarian action at national and international levels. Three groups of disasters are distinguished in EM-DAT: natural disasters, technological disasters, and complex emergencies. For a disaster to be entered into the database, at least one of the following criteria must have been met:

- 10 or more people were reported killed;
- 100 people were reported affected;
- a state of emergency was declared; and
- there was call for international assistance.

Risks of mortality and economic losses are calculated as a function of the expected hazard frequency and expected losses per hazard event. The database obtains global hazard data on cyclones, drought, earthquakes, floods, landslides, and volcanoes from a variety of sources.

Drought Hazard Identification

Based on the definition of hazard given in section 6.2.4.1, drought hazard could be defined as the probability of occurrence of a drought of a certain intensity, in a specific site during the growing season. The approach followed is to look at the historical WRSI for a number of points in the country, obtain a frequency distribution, and compute the probability that a certain threshold has not been reached.

The WRSI can be computed for points on the map for every year for which sufficient input data are available. For example, in Sudan, it is calculated for the period 1996–2006, for points 0.5 degrees apart. A higher-density grid is recommended if computing power allows. If we defined a drought as a season when the WRSI was less than 80, we could obtain for every location the frequency that such a drought season occurred over the number of years studied. We could then assume that this frequency equalled the probability that drought would occur in the future in this geographical location. Values for areas in between these points are obtained by interpolation (kriging), and a

105. *Crop Forecasting Philosophy of FAO: An Overview,* http://80.69.76.153/wiki/index.php?title=Chapter2&redirect=no.
106. See Frere and Popov, *"Agrometeorological Crop Monitoring and Forecasting."*
107. Based on FEWS NET.

map can be constructed giving the probability that the WRSI will remain below 80. The probability ranges from 0 percent (zero of the years observed were wet, with a WRSI of "always >80") to 100 percent (every single year observed was too dry, with a WRSI "always <80"). Using the same process, a probability map can be constructed for "severe drought" (defined as a season when the WRSI is less than 50).

Box 6.9: Calculation of Water Requirement Satisfaction Index (WRSI) (FEWS NET)

WRSI calculation requires a start-of-season (SOS) and end-of-season (EOS) time for each modelling grid cell. The onset of rains is determined using a threshold amount and distribution of rainfall received in three consecutive decades. SOS is established when there is at least 25 mm of rainfall in one decade followed by a total of at least 20 mm of rainfall in the next two consecutive decades. The length of growing period (LGP) for each pixel is determined by the persistence, on average, above a threshold value of a climatological ratio between rainfall and potential evapo-transpiration. Thus, EOS was obtained by adding LGP to the SOS decade for each grid cell. The WRSI model is capable of simulating different crop types whose seasonal water use pattern has been published in the form of a crop coefficient. Such crops include maize (corn), sorghum, millet, and wheat.

At the end of the crop growth cycle, or up to a certain decade in the cycle, the sum of total actual evapo-transpiration (AET) and total water requirement (WR) is used to calculate WRSI in a Geographic Information System (GIS) environment at 0.1 degree (about 10 km) spatial resolution. A case of "no deficit" will result in a WRSI value of 100, which corresponds to the absence of yield reduction related to water deficit. A seasonal WRSI value less than 50 is regarded as a crop failure condition (Smith 1992).
http://earlywarning.usgs.gov/CentralAmerica/dailywrsirm.php

Box 6.10: Useful Internet resources for drought hazard analysis

http://www.hoefsloot.com/agrometshell.htm
http://earlywarning.usgs.gov/Global/index.php
http://www.em-dat.net/
http://www.fao.org/ag/agl/agll/gaez/index.htm
http://edc.usgs.gov/products/elevation/gtopo30/hydro/index.html
http://igskmncnwb015.cr.usgs.gov/adds/
http://www.preventionconsortium.org/?pageid=17

6.2.4.3 Vulnerability analysis

Vulnerability to food insecurity is defined as increased susceptibility of a household to the effect of hazards on its food security. Vulnerability is a function of a **household's exposure** to a hazard and its resilience or **coping capacity**. Through vulnerability analysis we want to arrive at a reasonably accurate estimate of how many households in a certain area would become food insecure if a specific hazard event occurred in that area.

The first task in conducting a vulnerability analysis is to assess the types of shocks communities and households are typically exposed to and how effective their strategies for managing risk are. Much of this information can be obtained through qualitative interviews with community leaders and households. Table 6.8 offers a partial list of the impacts of various shocks on the range of assets that typically support household livelihood strategies.

Table 6.8: Examples of the association between hazards, shocks, and various forms of livelihood capital

Hazards/risk:	Drought, storms, flooding, land degradation, pests, animal disease, earthquakes, volcanoes	Disease epidemics (malaria, cholera, dysentery), AIDS, injuries	Policy changes, discrimination, unequal access to resources	Macro-economic policies, market and trade, recession	War, violence, discrimination
Sector:	**Environmental**	**Health**	**Social**	**Economic**	**Conflict**
Livelihood capital			**Shocks**		
Physical capital	Climatic activity destroys physical infrastructure	Asset divestiture, loss of capacity to provide public services		Asset divestiture	Conflict leads to destruction of physical infrastructure, assets stolen or destroyed
Natural capital	Destruction of land and resources		Appropriation and loss of common property resources, increased theft	Price shocks, rapid inflation, food shortages	Conflict leads to loss of land, assets, and theft
Social capital	Recurring environmental shocks break down ability to reciprocate	Morbidity and mortality affect networks	Breakdown of labour reciprocity, breakdown of sharing mechanisms, lack of social cohesion, reduction in safety net support	Shift to institutional forms of trust, stricter loan collateral requirements, migration for employment	Communities displaced by war, theft leads to breakdown in trust
Economic capital	Seasonal climatic fluctuations reduce employment; morbidity and mortality of income earner; loss of crops	Employment policies, declining subsidies or inputs, poor investment in infrastructure, taxes		Unemployment, falling real wages, price shocks	Marketing channels disrupted by war
Human capital		Declining public health expenditures, user charges, declining education expenditures	Breakdown in community support of social services, unequal access to services	Privatization of social services, reduction in labour opportunities	Conflict destroys social infrastructure, mobility restrictions

Once the inventory and categories of risk are determined, the level of risk exposure must be assessed. Risks are not felt evenly among all households, and there are several ways the population may be disaggregated to look at specific exposure levels. Exposure to various types of hazards can often be identified according to livelihood group, but also by socio-economic group, ethnic group, or gender, among others. Risk exposure can be explored in questionnaires at the household level, especially through the livelihoods section or with qualitative instruments such as key informant interviews, focus group discussions, or structured community discussions.

6.2.4.4 Risk management analysis

Risk management strategies can be thought of as falling into one of three basic categories. Activities undertaken by households and communities to reduce the likelihood of exposure to a shock are defined as **risk-reduction** strategies. Those taken to minimize the impacts of a shock before it occurs are defined as **risk-mitigation** activities. **Risk-coping** responses are the strategies initiated post-shock. The post-shock responses initiated by households are usually referred to as coping strategies. Post-shock responses initiated by communities are informal safety nets, while those carried out by governments and NGOs are formal safety nets. A thorough understanding of risk-management responses identifies entry points for appropriate programming that supports and enhances the positive activities already taking place and fills in gaps where additional support is needed. This approach re-emphasizes the key role of risk reduction and mitigation due to its functional relationship with vulnerability.

It is important to collect information on risk management at all three levels (risk reduction, risk mitigation and risk coping). Information on risk management at the household level is gathered through household surveys focusing on particular activities and behaviours. Risk-reduction and -mitigation activities are best understood as they apply to particular livelihood strategies. Household coping strategies generally begin with the short-term strategies and transition to longer-term strategies as the impact of the shock continues.

The latter type are often referred to as distress strategies; they are more detrimental to household livelihoods and will slow the process of recovery. Households using a larger array of coping strategies and/or strategies considered more severe in their impacts are usually more vulnerable.

Box 6.11: Household coping strategies

Short-term coping strategies
- Migration of household members to look for work
- Searching for wild foods
- Selling non-productive assets
- Reducing number and size of meals
- Changes in diet to less preferred or less nutritious foods

Unviable (distress) coping strategies
- Selling productive assets
- Household dissolution
- Theft
- Prostitution
- Mass migration
- Begging

The degree of vulnerability depends on the nature of the risk and a household's resilience, or "ability to bounce back or recover after adversity or hard times, to be capable of building positively on these adversities" (Mission 2005). A household's resilience for a given magnitude of shock often is related to:

- How well a household or community can self-organize after the exposure to the hazard to maintain an acceptable level of functioning and structure;
- How well a household or community can learn from these difficult circumstances and adapt; and
- The household's characteristics, notably its assets and livelihood strategies.

The risk of livelihood failure determines a household's level of vulnerability. When exposed to shock, households often see a decline in welfare levels, which can include overall reductions in food availability, loss of or diminished access to food, fluctuation in income streams, and poor nutrition. The greater the share of resources devoted to food and health service acquisition, the higher the vulnerability of the household to food and nutritional insecurity. Households able to minimize loss of welfare resulting from shock exposure by employing livelihood and risk management strategies are likely to be less vulnerable than those who cannot properly cope with or mitigate the effects of the risk. Therefore, livelihoods are secure when households have secure ownership of, or access to, resources (both tangible and intangible) and income-earning activities that can offset risks, ease shocks, and meet contingencies. When households are able to acquire, protect, develop, utilize, exchange, and benefit from assets and resources, livelihoods are secure (CARE 2002).

Information on risk-coping strategies at the household level can be obtained through surveys by including questions that enable calculation of a **coping strategies** index (see section 4.2, "Household Questionnaire"). Risk-coping responses at the community level (informal safety nets) can also be gathered during village surveys, using focus group discussions. Information on formal safety nets can be gathered through secondary data and from government and NGO key informants.

In general, community informal safety nets deal better with idiosyncratic shocks. Covariate or collective shocks tend to overwhelm community-level resources and often require an external response in the form of a formal safety net from the government or NGOs. Since households often turn to neighbours and communities for help in times of need, the diminished capabilities of the community may leave households more vulnerable as they struggle to cope with the shock on their own.

A number of community-level institutions operating in any given locale can play a role in helping manage risk. These institutions can be either formal or informal groups or organizations and include religious groups, social clubs, savings and credit groups, funeral societies, and service delivery institutions focusing on health or education. The focus of the CFSVA is not simply to document the existence of these groups but to determine to what extent they contribute to the ability of households to manage risk.

External or formal means of risk management can address both idiosyncratic and covariate forms of risk. More frequently the formal management activities deal with covariate shocks. The types of activities may range from formal safety nets, to deal with shocks such as crop loss, to market subsidies of staple crops, to address soaring prices. At the same time, policy programmes that encourage ready credit opportunities help individuals respond to idiosyncratic shocks such as illness or death.

Determining Trends to Understand Vulnerability

Vulnerability to food insecurity is a forward-looking concept related to people's proneness to future acute loss in their capability to acquire food. For this reason, it is important to identify trends in livelihood strategies and changes occurring in internal household dynamics related to risk management. For example, are households pursuing certain livelihood strategies more likely to sell off assets in the event of a shock? In addition, it is important to determine the role of social networks, institutions, and inter/intra-community dynamics and the ways in which each of these influence the ability of individual households to cope with a particular shock. For example, HIV has had a significant negative impact on community informal safety nets in southern Africa. The breakdown of these informal safety nets will make HIV-affected households more vulnerable. In many countries, repeated surveys permit assessment of change over time. Several rounds of DHS and/or poverty surveys may have been conducted. Some data from routine information systems permit time series analysis; this includes price series, climate data, and sometimes anthropometric data from clinics/health facilities.

Opportunity Analysis

In addition to analysing risks and vulnerabilities, it is also important to take into account the opportunities available to communities, households, and individuals within the programme setting that can contribute to risk management. For example, many households devise positive responses to constraints that could form the basis for interventions aimed at increasing food security among the target population. These households are often more resilient. Such households can be identified through the use of a positive deviance approach, which seeks to promote adoption of the successful practices of resilient households among the wider food-insecure population.

Opportunities for risk management may also be derived from efforts being promoted by community-based organizations and local NGOs. Such groups may be operating effective programmes that future projects can build upon. Enabling conditions for supporting such opportunities may exist at the policy level through changes promoted by the Government. Finally, a coalition of organizations can collaborate in a complementary way to scale up risk-management approaches. Positive deviant households are typically identified through a combination of qualitative approaches, including focus group discussions, key informant interviews, transect walks, and community ranking exercises.

Quantifying household vulnerability

Incorporating quantitative methods into an analysis of vulnerability allows for more precise description and appropriate targeting of populations vulnerable to food security. In order to objectively quantify vulnerability, information on proxy indicators related to risk exposure and risk management must be collected. By combining this information with expected future food security trends, CFSVA teams can arrive at reasonable predictions regarding which households are more likely to be affected by a specific shock. Box 6.12 provides an example of how quantitative analysis was used to predict household food insecurity in Burundi.

Box 6.12: Food Security Monitoring System (FSMS) data from Burundi

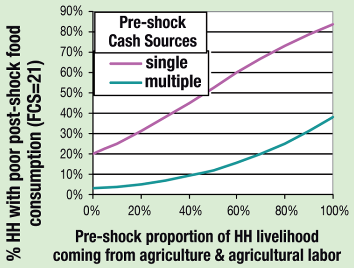
In 2005–2006, the Plateau Nord and Sud Ouest areas of Burundi were affected by severe drought. In the immediate wake of the drought, data from 135 agricultural households was collected and analysed in order to gauge the effect of the drought on household food security. Pre-shock data (2004–2005) was used to estimate household wealth and sources of livelihood. The FSMS monitored the food consumption of the same households before and during the crisis (April 2006).

Through logistic regression (R^2 = 0.42), the research team was able to determine the elements that contributed to increased levels of household food insecurity resulting from the drought.

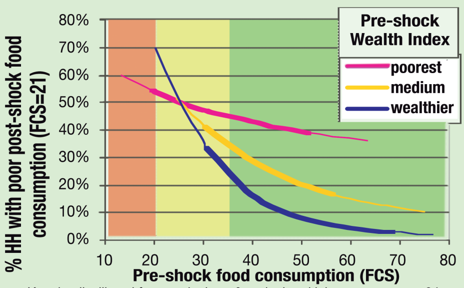
The team found that the more households' livelihoods depend on agricultural production and agricultural labour (pre-shock), the higher the probability they will suffer poor food consumption during the post-shock crisis. This reflects the dimension "exposure." The "risk management aspect" is shown by the fact that (all other factors being equal) households with only a single source of cash income rarely manage to keep good food consumption after the shock, whereas households with several cash sources fare much better.

	B	S.E.	Wald	df	Sig.
FCS_Oct05	-0.06	0.02	10.64	1	0,001
Wealth index	0.70	0.69	1.03	1	0,310
(FCS_Oct05*wealth index	-0.03	0.02	4.00	1	0,045
Part from agri/agri labour	2.89	1.23	5.54	1	0,019
More than 1 source of cash	-2.04	0.52	15.71	1	0,000
Constant	1.22	1.04	1.37	1	0,241

Households with acceptable pre-shock food consumption *and* higher pre-shock wealth were less likely to fall into poor post-shock consumption (all other factors being equal). These households clearly have a better capacity to mitigate and cope with the effects of a shock. Relative wealth seemed to be especially important for households with adequate pre-shock food consumption, to avoid poor consumption after the shock.



Keeping the coping capacity constant: medium wealth and FCS



Keeping livelihood from agriculture & agricultural labour constant at 64

Exposure of household livelihoods

Based on the qualitative analysis, one can foresee which livelihood activities are typically affected by a certain hazard. For instance, high price inflation would have a significant effect on households that purchased their food and less on those producing their own food; drought during the growing season affects farmers, but livestock may be affected only in the case of extreme drought. These "qualitative" findings will be used to estimate, using the livelihoods section of the HH questionnaire, how much the primary effect of a shock would be.

For instance, we assume that a drought would reduce agricultural production by 50 percent (based on the Water Balance Model, a WRSI of about 80 percent leads to a harvest of only 50 percent). A household depending 70 percent on agriculture would hence, as a primary effect, lose 35 percent of its livelihood because of the drought. Using SPSS, this can easily be estimated for all households.

Risk management

The capacity of households to manage (prevent, mitigate, and cope with) a shock depends on the use of more "resistant" strategies, alternative sources of income, reserves, savings, social networks, etc. Some households will cope with a certain loss better than others. Risk management can be very complex, and the analyst will use proxy indicators to estimate risk management in a simplified way, assuming that he/she can capture the big picture. The Burundi example shows that with only a few determinants, already much of the food consumption outcomes can be explained.

In the vulnerability analysis, we are considering several intertwined elements. First, it is less likely that households with adequate food consumption (as measured by the **food consumption score**) will slide into poor food consumption due to a shock than those households with borderline food consumption. Second, some households have better risk-prevention, -mitigation, and -coping strategies to help them avoid reducing food consumption than others. Although this is hard to measure directly, the **asset wealth index** turns out to be a good proxy of this.

Based on the relationship between asset wealth and coping capacity, the wealth indices appear to be strong determinants of coping capacity. Tercile groupings have been commonly used; however, the specific categories should be determined based upon exploratory analysis. As the Burundi example illustrates, other indicators, such as multiple sources of cash income, help households better cope with the effects of drought than others.

The cut-offs could be set as in Box 6.13.

Box 6.13: Effect of shocks on household access to food and coping strategies

The fact that a certain household reports a covariate shock indicates how poorly it can deal with the same situation compared with a neighbouring household, with the same livelihood strategies, who did not report the situation as affecting its customary ability to consume food. In every region in Sudan, and for every livelihood group, poor households (as measured through asset wealth) are much more likely to report shocks than rich ones. From this we conclude that asset wealth is a good proxy for the risk management capacity of households, since rich households manage risk so well that they do not even consider the situation as a "shock," in comparison to what their neighbours report. This relation still holds when controlled for other factors, such as the region and the livelihood group the households belong to.

Factor analysis reveals three different categories of coping strategies:

1. The first category is an inferior category and could be called "suffering the consequences of the shock, because no positive action can be taken, or deploying desperate ways to obtain additional food." These are compensation strategies employed by the households, such as: fasting one day, eating fewer meals, eating less and lower quality of food, taking no action at all, scavenging for food, or working for food only.
2. The second category of coping mechanisms could be labelled as "generating extra income, including credit, or cutting expenditures." This category includes migration of household members, working for money, purchasing food on credit, borrowing food, and cutting health or education expenditures.
3. We call the third category of coping "reliance on own household reserves." This typically involves using the household savings, selling or slaughtering livestock, and selling "other" assets.

With increasing asset wealth, the coping strategy of choice would typically include less severe measures. Fasting an entire day or scavenging for food is almost exclusively done by households belonging to the asset poor wealth deciles. It is also the poorer deciles who often report eating less. Selling assets and spending savings, on the other hand, is a strategy chosen by households with higher levels of assets as measured by the asset index. This again highlights the fact that asset wealth is a good proxy indicator of coping capacity. The poorer a household is, the more desperate and ineffective its measures for compensating for a shock are.

Conclusion: wealthy households do not only utilize effective coping strategies more often, but they also only rarely feel that a certain situation affects their customary ability to procure or consume food. In short, they are more resilient.

Estimation of vulnerability

Vulnerability of a household to a certain hazard is defined by how much the household's livelihood would be affected (first proxy indicator: part of total revenue lost) by the hazard taking place and how well the household could cope with the effects of the shock (second proxy indicator: the food consumption score; and third: the wealth index).

Table 6.9 shows how this information is typically combined to define a household as vulnerable.

Table 6.9: Percentage lost revenue (in cash or in kind) that defines various wealth groups as vulnerable to a shock

	Current food consumption		
	Poor FCS<=21* (%)	**Borderline 21<FCS<=35 (%)**	**Adequate FCS>35 (%)**
Poor coping strategies† (poorest wealth group)	10	20	25
Mediocre coping strategies (middle wealth group)	10	20	35
Good coping strategies (highest wealth group)	10	20	55

* Refer to Box 6.1 and Table 6.1.

† In the Sudan example – poor coping strategies: households from wealth deciles 1–4; mediocre coping strategies: households from wealth deciles 5–7; good coping strategies: households from wealth deciles 8–10; additional indicators, such as the diversity of income sources, could be used.

The thresholds set in the table signify that households with good coping strategies and adequate current food consumption are considered vulnerable to a hazard only if 55 percent of their usual revenue is lost because of that hazard. Households with poor coping capacities and borderline food consumption are already considered vulnerable if 20 percent of their revenue is lost. The table is based on observations in Burundi; however, these relationships could well be different in other contexts. Households whose current food consumption is already "poor" would most likely still have poor food consumption after a shock. We can consider them chronically food insecure (if data were collected during the "good season") or at least cyclically food insecure (if data were collected during the lean season).

One could debate the levels of these thresholds. Setting them higher or lower would result in fewer or more households identified as vulnerable to a certain shock. Hence, the absolute number of households found vulnerable through the analysis is only an indicative one. On the other hand, since these thresholds are systematically applied to all households in all regions, this approach to vulnerability can be used very well to observe regional differences and differences between livelihood groups vulnerable to a certain shock.

6.2.4.5 Bringing it all together: risk to food security

Risk to food security is defined as "The probability of food insecurity resulting from interactions between natural or human-induced hazards and vulnerable conditions." The questions to be answered are: "Is it likely that vulnerable households will be confronted with the hazard they are vulnerable to?" And "How many such households are there (proportionately or in absolute numbers) in a particular region or sub-population?"

As described in the previous sections, vulnerability analysis can indicate if a household is vulnerable to a specific hazard; hazard analysis can indicate the likelihood of a hazard occurring in a certain location. By merging both data sets, one can identify households that are both vulnerable to a specific hazard and living in a hazard-prone region and thus at risk of becoming food insecure.

The ideal way would be to multiply, for each household, the probability that the household will be vulnerable to a shock (see Burundi example: this can be calculated if adequate panel data is available) by the probability of a specific hazard's occurring during a certain time period: Risk = VulnerableProb * HazardProb. Practically speaking, a less subtle approach is used to classify a household as either "vulnerable" or "not vulnerable." Using SPSS, an expression as below would then be used to identify a household as being at risk for the hazard. A region where the specific hazard strikes at least once every ten years is, in this case, considered hazard prone:

If (Vulnerable = 1 and HazardProb> = 0.1) At_Risk = 1.

The expression selects all households that will become food insecure at least once every ten years because of a specific hazard. Prevalence of households at risk can be studied by region or by group, such as livelihood or ethnic background. A similar result can be obtained visually, by overlaying a vulnerability map with a hazard map. Areas in the overlaid map that are hazard prone and where many vulnerable households live (proportionately or in absolute numbers) are considered at risk: food insecurity will likely occur for these households in the not so distant future. For more on GIS and remote sensing data used for risk analysis, refer to Annex 10.

6.2.4.6 Key references: Risk analysis

- FAO. 1996. Agro-Ecological Zoning Guidelines. FAO Soils Bulletin No. 73. Available online at: http://www.fao.org/docrep/W2962E/w2962e00.htm#P-2.
- ibid. 1992. Summary of Proceedings and Decisions, CGIAR Mid-term Meeting, Istanbul, Turkey, 19–22 May 1992. http://www.fao.org/Wairdocs/TAC/X5756E/x5756e00.htm#Contents
- Frere, M., and G. F. Popov. 1979. *Agrometeorological Crop Monitoring and Forecasting*. FAO Plant Production and Protection Paper No. 17. Rome: FAO.
- Gommes, R. and P. Hoefsloot. 2006. *Crop Forecasting Philosophy of FAO: An Overview.* September. Obtained online at: http://80.69.76.153/wiki/index.php?title=Chapter2&redirect=no.
- International Strategy for Disaster Reduction: *Terminology: Basic Terms of Disaster Risk Reduction*. Obtained online at: http://www.unisdr.org/eng/library/lib-terminology-eng%20home.htm
- Mukhala, E., and P. Hoefsloot. 2004. *AgroMetShell Manual.* Harare, Zimbabwe: SADC Food Security Programme, in collaboration with FAO.
- Senay, G. 2004. FEWS NET: *Water Requirement Satisfaction Index (WRSI).* October. Obtained online at: http://earlywarning.usgs.gov/CentralAmerica/dailywrsirm.php
- UNDP Bureau for Crisis Prevention and Recovery. 2004. *A Global Report: Reducing Disaster Risk: A Challenge for Development.* New York.

6.2.5 Analysis of gender-disaggregated information

6.2.5.1 Quantitative data analyses

Food security and vulnerability analyses rarely use gender as the sole criteria for defining groups for analysis (although such estimates are likely to be generated for reporting purposes), because the distinction between male/female or male-headed household/female-headed household rarely defines homogenous groups. Rather, gender should be viewed as a cross-cutting theme that can be combined with other criteria for comparative purposes. However, analysis of anthropometric data should be disaggregated by sex to understand gender dimensions in food consumption, access to health services, and so on. To understand the determinants of malnutrition, a regression model could be applied using sex as an exogenous variable.

6.2.5.2 Qualitative data analysis

While quantitative data may be used to test hypotheses concerning the causes of negative outcomes and the extent to which gender appears to be a causal factor (e.g. incidence of malnutrition is higher among girls than boys), qualitative data is critical for generating hypotheses about the underlying causes of food security and vulnerability and the relationship of gender to these outcomes. Continuing with the malnutrition example in the previous paragraph, qualitative data may provide some indication of whether and how boys and girls from the same household adopt different coping (food adjustment) strategies in the event of household-level food shortages.

6.2.6 Including HIV/AIDS in WFP food security and vulnerability analysis

In countries with high prevalence of HIV/AIDS, geographic targeting should combine data on food insecurity with data on HIV prevalence to identify food-insecure areas where HIV prevalence is particularly high. However, including HIV/AIDS in the food security and vulnerability analysis goes far beyond simply overlapping these data.
When HIV/AIDS indicators are integrated into a food security and vulnerability analysis, findings can strengthen the country-specific empirical evidence on the interaction between HIV/AIDS and household food security and enhance the understanding of the role of food assistance in mitigating the impact of HIV/AIDS. In particular, data can be used to:

- explore the relationship between HIV/AIDS, food consumption, and nutritional outcomes;
- compare the livelihood assets and strategies both of households living with HIV/AIDS and of non-affected households; and
- highlight the coping strategies frequently adopted by affected households.

When information is collected on children orphaned or made vulnerable by HIV/AIDS, data can be used to:

- identify OVC shortcomings in education and nutrition and determine whether they are more likely to be involved in working activities (child labour); and
- compare the food security status, livelihood assets, strategies and coping mechanisms of households with orphans and households without orphans.

Not only can such findings increase the country-specific evidence on the link between HIV/AIDS, livelihood, and food security, but they can be used to guide other steps of targeting (e.g. development of beneficiaries' selection tools).

6.2.6.1 General challenges in analysing the impact of HIV/AIDS

While WFP CFSVA has proven to be a useful tool for identifying traditional causes of food insecurity, a number of critical challenges arise in assessing the bi-directional relationship between vulnerability to food insecurity and HIV/AIDS. These challenges include:

Lack of longitudinal data

A deep understanding of the inter-relationships among HIV/AIDS, livelihoods, and food security can be reached only by analysing longitudinal data. CFSVAs are not meant to provide panel data. Even in countries where food security assessments have been conducted several times, survey instruments are not necessarily developed to provide repeated cross-sectional data. As a consequence, CFSVAs are more suitable for capturing key disparities in livelihood assets, strategies, and outcomes at a certain point in time rather than exploring dynamics.

Lack of country-wide results

CFSVAs are usually conducted in rural areas. The exclusion of urban areas represents a serious challenge to HIV/AIDS analysis because HIV prevalence is often higher in urban settings, and services for people living with HIV/AIDS are more frequently provided in such areas.

Inference to the population

As a general rule, food security and vulnerability analyses sample villages through a probability-proportional-to-size approach and then select households randomly. The percentage of affected households that can be expected in a CFSVA depends on the HIV prevalence in the country or region surveyed. In countries with low or medium prevalence, the number of households in the sample affected by HIV/AIDS can be very low. This increases the chance of findings that are not statistically significant and limits the possibility of conducting multivariate analysis.

6.2.6.2 Chronic illness a proxy indicator for HIV/AIDS

Most of the empirical studies on HIV/AIDS consider a household as affected by HIV/AIDS if:

- one or more adult members are infected by HIV/AIDS; and/or
- one or more adult members died recently (over the past 12 months) from AIDS.

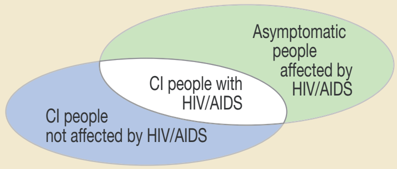
There are practical constraints to identifying households affected by HIV/AIDS. In particular, (1) it is difficult to test for HIV antibodies; and (2) it is challenging to ask an individual directly if she/he is sero-positive. (People either do not know or prefer not to disclose their status.) Because of these problems, proxy indicators are necessary for identifying households living with infected individuals.

In high-prevalence countries, chronic illness (CI) is typically used to identify HIV/AIDS-affected households. In particular, a household is considered affected by HIV/AIDS if:

- there is at least one chronically ill adult (e.g. aged 18 to 59; age range can be context-specific) in the household; and/or
- the household experienced recently (over the past 12 months) the death of one or more adult members (e.g. aged 18 to 59; age range can be context-specific) from chronic illness.

The term chronic illness easily can be misunderstood; it is therefore crucial to agree on a definition. An individual is usually considered “chronically ill” if she/he lives in a condition or with a disease that has prevented her/him from being fully functional for at least 3 out of the past 12 months. Using chronic illness as a proxy indicator for AIDS leads to inclusion and exclusion errors (see Figure 6.1).

Figure 6.1: Chronic illness as proxy for HIV/AIDS



Alternative ways of identifying people living with AIDS continue to emerge, but the results are not appropriate for a CFSVA.[108] In the absence of clear guidance from such alternative methods, chronic illness remains the most common proxy indicator for HIV/AIDS.

As a general rule, chronic illness is appropriate if:

- HIV/AIDS prevalence in the country is high (e.g. official thresholds do not exist, but in most of the southern African countries, HIV prevalance justifies using chronic illness); and
- there are no other factors or illnesses leading to high incidence of chronic illness.

If these circumstances are not met, chronic illness should be interpreted as an indicator of a health-related shock within the household, rather than as a proxy indicator for HIV/AIDS. From this perspective it is still useful to understand the impact of prolonged illness on food security (see Box 6.14).

Box 6.14: Assessing the validity of chronic illness

In some countries, Demographic and Health Surveys (DHSs) include blood tests for HIV antibodies and questions about chronic illness. This makes it possible to assess the validity of chronic illness as a proxy indicator for HIV/AIDS.

During the Cameroon CFSVA 2007, a secondary data analysis on DHS 2004 data showed that only 12 percent of the chronically ill adults (aged 15 to 49) were HIV-positive and 87 percent of the HIV-positive adults (15 to 49) were not chronically ill. As a consequence, chronic illness (and death from chronic illness) was analysed to explore the impact of prolonged illness on household food security.

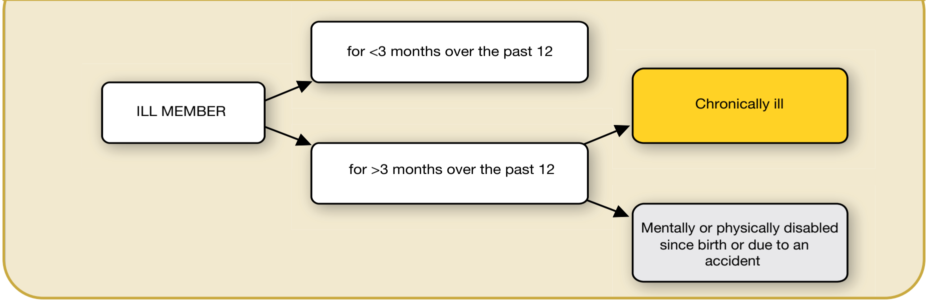
Inclusion and exclusion errors cannot be entirely eliminated, but they can be reduced by improving data collection. With this regard, it is important to:

- Be cautious in using methods based on a verbal autopsy or list of symptoms until solid evidence emerges that such methods are effective. Such methods deserve attention and need more investigation, but they require careful training of enumerators.
- Make sure the questionnaire distinguishes between individuals who are mentally or physically disabled (due to accident or since birth) and those who are chronically ill.
- Provide the enumerators with a clear definition of chronic illness. They should understand that different stages of chronic illness exist and that a CI individual can be bedridden; not bedridden, although completely unable to work; or not bedridden and able to work for a few hours or days. Also, months of inactivity can be non-consecutive.

108. Some instruments, such as the checklist-type tool developed by the Catholic Relief Services (CRS), are not suitable for a population-based survey. Other tools (e.g. the verbal autopsy method developed by Ainsworth and Semali, 1998) do not require medically skilled enumerators, but they do require careful phrasing of the questions and direct access to people who live or have lived in close contact with the chronically ill.

- Urge the enumerators to rephrase the term chronic illness during data collection, adopting terms more suitable to the local context. For instance, enumerators can substitute "to be fully functional" with other terms, such as "to engage in daily activities as usual" or "to be operational." The important thing is to convey the meaning of the term "fully functional," not to use the same terminology everywhere.

Figure 6.2: Avoiding misunderstanding of the term "chronic illness"



As mentioned earlier, it is important to determine if other factors exist that may cause a high incidence of chronic illness. If they do, they may compromise the effectiveness of the proxy indicators.

A more detailed list of illnesses has been adopted in some surveys.[109] Along with the distinction "ill for less than three months/chronically ill/disabled," some surveys have included specific illnesses in their list of possible answers, such as malaria, asthma, TB, cancer, and AIDS. The purpose is to improve the ability to distinguish between people suffering from HIV/AIDS and people who have been sick for more than three months due to other problems. The main challenge of this approach is that interviewees usually have no competence in medicine and are rarely aware of which illness they experience. If this approach is used, it is crucial to train the enumerators to describe the main symptoms of all the listed illnesses. Also, it is advisable in the analysis to consider individuals affected by TB as persons living with HIV/AIDS, since TB is frequently associated with AIDS.

The way we capture the presence and key attributes of chronically ill or deceased household members depends on how demographic data are collected during the household survey. Examples of data collection modules are provided on pages 30–36 of the HIV/AIDS technical guidelines summarized in section 4.2 of the present guidelines ("Household Data Collection").

6.2.6.3 Analysing food consumption in the context of HIV/AIDS

Proxy indicators for HIV/AIDS are typically reported at the household level (e.g. percentage of households with at least one CI adult member or percentage of households with one orphan). Food security and vulnerability analyses collect data at the household level on food consumption, food sources and stocks, livelihood assets, and strategies. As a consequence, **analysis of the interaction between HIV/AIDS, food security, and livelihoods has to be conducted at the household level**.

To understand the relationship between HIV/AIDS and food security, analyses should focus on the key issues usually considered in food security analysis: current food consumption, dietary diversity, sources of food, food reserves, etc. The purpose is to identify differences in food security between affected and non-affected households.

6.2.6.4 Analysing nutritional outcomes

The unit of analysis for nutritional outcomes (e.g. women's and children's nutritional status) is the individual. **The interaction between HIV/AIDS and nutritional outcomes can therefore be analysed at the individual level**. In general, household-level factors have an impact on individual status. Thus, the analysis can control for the presence of HIV/AIDS proxies to assess the impact of HIV/AIDS on women's and children's nutritional status.

6.2.6.5. Consideration of livelihood assets and strategies

In order to have a complete understanding of the impact of HIV/AIDS on household food security, it is crucial to consider livelihoods in the analysis. HIV/AIDS literature highlights the main effects of HIV/AIDS on livelihoods. Some indicators typically collected during CFSVAs can be used to explore the interaction between HIV/AIDS and livelihoods and analyse data through an HIV lens. These are listed in Table 6.10. The list can be useful also during questionnaire design, to help ensure that the HIV/AIDS component is included in the assessment. This list is the result of a comprehensive review of household questionnaires used in CFSVAs over the past two years, and community and household surveillance (CHS) questionnaires.

Table 6.10: Analysing HIV/AIDS impact: indicators available from CFSVAs

Human capital	• Household size • Household composition (gender) • Household headship (gender and age) • Orphans (presence of, number of) • Migration in (and reason) • Migration out (and reason) • Prevalence (%) of effective dependents • Decrease in amount of land cultivated • Reasons for not cultivating land (or for cultivating less land) • Children's enrolment in school • Children dropping out from school (and reason) • Children's absenteeism (and reason)
Financial capital	• Expenditure on health, medicines, transport, and funerals (e.g. per capita and household expenditures as percentage of total expenditure) • Expenditure on seeds, fertilizers, etc. • Income-generating activities and their contribution to total income • Access to credit • Borrowing money (primary reason and source of credit)
Social capital	• Material support received from relatives or friends (e.g. money, food, clothing, and agricultural inputs)
Natural capital	• Kind of irrigation system, fertilizers, and pesticides used • Quantity and kind of crops cultivated • Access to land • Proportion (%) of land cultivated (out of total land owned) • Proportion (%) of land rented or lent out (out of total land owned) • Selling land (and reason)
Physical capital	• Assets and livestock owned • Selling or bartering livestock (by kind of livestock and reason for selling or bartering) • Selling assets (and reason for)

109. See 2006 Mozambique Baseline Report (December 2006).

It is extremely important to note that:

- WFP food security and vulnerability assessments help identify differences between affected and non-affected households. Since CFSVAs are not meant to be longitudinal studies, they face challenges in understanding the vicious circle between HIV/AIDS and livelihoods. CFSVA household data alone cannot answer questions regarding processes.
- Some effects mentioned in the HIV/AIDS literature do not have a related indicator in the food security and vulnerability studies.[110]
- Some of the indicators reported in the table help identify negative consequences (e.g. disparities in labour force), while others highlight coping strategies (e.g. migration of household members). Still, coping mechanisms can serve as indicators for HIV/AIDS impact because as the impact becomes more severe, household strategies tend to become more desperate.

6.2.6.6 Capturing the process

As mentioned earlier, CFSVA data can be used only for a cross-sectional analysis. To obtain insight into the process it is necessary to include explicit questions about consequences and coping mechanisms. There are three options for how to do this:

1. Ask respondents to identify and rank strategies adopted by the household to minimize the effects of chronic illness and death.[111]
2. Include questions on specific coping strategies in the appropriate sections of the questionnaire and use those questions to compare actions taken by the affected and non-affected households.
3. Look at the Coping Strategy Index (see example in Box 6.15, page 264).

6.2.6.7 Shortcomings of orphans and other vulnerable children (OVC)

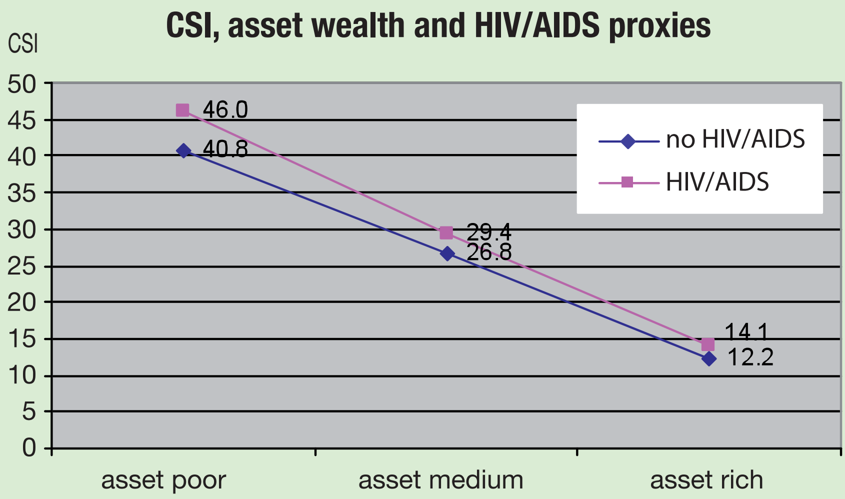
Within the context of a CFSVA, OVC are operationally defined as children under 18 years old who are orphaned, or who live in a HH with a chronically ill adult, or who live in a HH where an adult member has recently died from chronic illness. The presence of one of these attributes is sufficient to classify a child as an OVC. The analysis can focus only on orphans or be extended to vulnerable children. Vulnerable children can be included in the analysis only if proxy indicators for HIV/AIDS have been collected through the household questionnaire.

110. Indicators of human capital that are usually *not* available in the CFSVA are: “time available for domestic labour and child-care,” “loss of agricultural knowledge, practices and skills.” Indicators of financial and social capital usually *not* available are: “liquidation of savings accounts,” “pledging of future crops,” “increased reliance on community willingness to support educational/nutritional needs of orphans,” “decrease in time to participate in social/cultural activities.” Impacts on natural and physical capital usually *not* considered include “reduction in soil fertility,” “fallow land returning to bush,” and “appropriation of land by relatives.”

111. This kind of approach has been adopted by the food security group working at Michigan State University in several survey instruments (http://www.aec.msu.edu). Details and examples are reported in the technical guidelines *Integrating HIV/AIDS in WFP Food Security and Vulnerability Analyses.*

Box 6.15: Coping Strategy Index (CSI)

Households living with chronic illness or with a recent death from chronic illness are more likely to adopt strategies to cope with a decline in food consumption. This is true for all the asset categories and becomes more evident among asset-poor households.



Source: Regional analysis undertaken on the Community and Household Surveillance (CHS) survey data, Round 8, Lesotho, Malawi, Mozambique, Swaziland, Zambia, and Zimbabwe, 2007.

Most of the studies suggest analysing the impact of HIV/AIDS on OVC by considering education, labour, nutritional, and health status. Food security and vulnerability analyses also allow consideration of food consumption, livelihood assets, and strategies. Box 6.16 includes an example of findings related to orphans.

Data analysis can be conducted both at the individual and the household level. It is not advisable to impose a universal rule on what should be analysed at which level, as the choice depends upon the kind of indicators and the objective of the research.

As a general rule, individual-level analysis is more appropriate for exploring individual attributes (e.g. shortcomings in education and nutritional status and exposure to labour). Household-level attributes (e.g. food consumption, livelihoods, and coping mechanisms) can be analysed both at the individual and the household level.

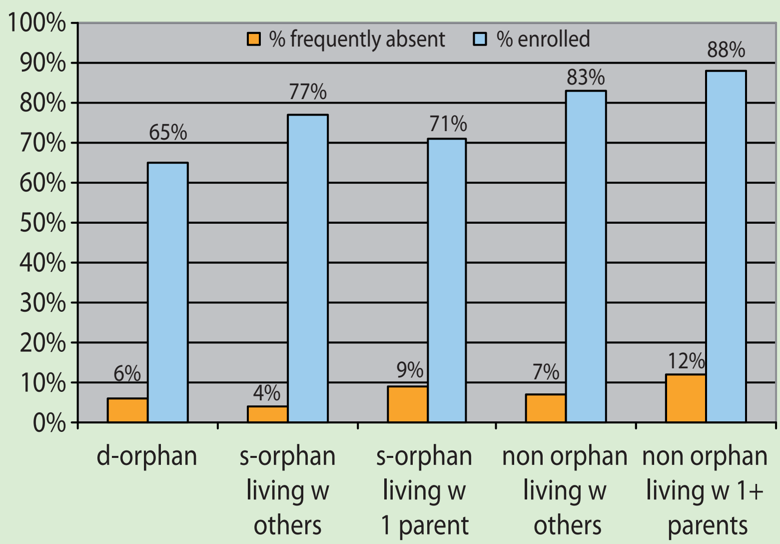
Box 6.16: Greater Monrovia Comprehensive Food Security and Nutrition Survey (CFSNS) - Education, by parental status

The CFSNS conducted in Greater Monrovia in 2006 found that 10 percent of children (ages 0–17) are orphans; 18 percent have both parents alive but do not reside with the biological parents, and 72 percent have both parents alive and reside with at least one of them. The high prevalence of non-orphans living with relatives or acquaintances is a consequence of the social disruptions Liberia faced during the civil war.

Overall, 85 percent of children (ages 6–17) are in school. The percentage decreases among orphans, and reaches 65 percent among double orphans. Single orphans living with others are more likely to be enrolled than single orphans living with the surviving parent.

The chance of being enrolled in school increases among non-orphans, and is as high as 88 percent among non-orphans living with at least one biological parent.

Even when taking into account their more frequent absences, non-orphans are more likely to attend classes than orphans.



6.3 COMMON KEY INDICATORS IN FOOD SECURITY ANALYSIS

1. Size of household

The number of people living in a household, with household defined according to country-specific standards.

Calculation: n/a

Key concerns: Outliers can include large extended families; definition of HH used during data collection must be considered.

Reporting: Mean number of household members (usually to the tenth); median number of household members. Also possible by percentage of HHs above a certain cut-off, particularly if there are many outliers. Rarely graphed.

Descriptive procedure: Descriptive statistics compare means.
Interpretation: Mean and median; care should be taken to evaluate the two (looking at S.D., frequencies, etc.) to arrive at the best descriptive statistic.
Type of variable: Continuous (can be transformed to categorical ordinal).
Importance to food security analysis: Larger households may be more vulnerable to food insecurity, although this is not a rule. The size of the household is sometimes used in calculating food rations. For a more vulnerability-related indicator, see item 2, "Age dependency ratio."

2. Age dependency ratio

Calculation: The age dependency ratio is equal to the number of individuals aged below 15 or above 64 divided by the number of individuals aged 15 to 64, expressed as a percentage.

A measure of the portion of a population composed of dependents (people too young or too old to work).

Dependents: no. of people < 15 and >64
Independents: no. of people ≥15 and ≤64
The formula is:

$$= \frac{\text{Number of dependents}}{\text{Number of independents}} \text{ X } 100$$

(Age 15-64)

The young dependency ratio includes children ages below 15 years, and the elderly dependency ratio focuses on those ages over 64 years. For example, if in a population of 1,000 there are 250 people under the age of 15 and 500 people between the ages of 15-64. The youth dependency ratio would be 50% (250/500).
Key concerns:
This indicator gives insight into the number of people of non-working age compared to the number of those of working age. A high ratio means that those of working age – and the overall economy – face a greater burden in supporting the young and the ageing population.

This indicator is useful at the aggregate level (country, region, state, etc.) but not at household level, because if a household has no independent people (denominator equal to zero), then the dependency ratio will be an invalid number. For this reason the percentage of dependent people (see following indicator) is used.
Reporting: Ratio and comparison between different levels (when possible).
Descriptive procedure: Compare means.
Interpretation: The higher the ratio, the higher the dependent population. If the ratio is 100 percent, then there is the same number of dependent and independent people. The interpretation of the different ratio should be contextualized with the analysis.
Type of variable: Continuous
Importance to food security analysis: Population groups or regions with a high age dependency rate are generally less food secure. A rising age dependency ratio is a concern in many countries facing an ageing population, since it becomes

difficult for pension and social security systems to provide for a significantly older, non-working population. In many countries where WFP is active, a high dependency ratio is due to a high number of children, including orphans, and to the mortality of adults due to violence or diseases such as AIDS.

3. Percentage of dependents[112]

Number of dependent people in a household based on household size.

Formula: (Number of dependents/household size) x 100
Key concerns: This indicator should be used at the household level.
Reporting: Compare means of the various food security or livelihood groups in order to describe the demographic profile.
Descriptive procedure: Mean; compare means.
Interpretation: A total of 0 percent means that no member in the household is of a dependent age, while a number equal to 100 percent means that all members of the household belong to the dependent age category. Table 6.11 provides an example of dependency ratio and percent of dependents.

Table 6.11: Example of proportion of dependents and dependency ratio calculation

No. of dependents	No. of independ-ents	Household size	Percentage of dependents	Dependency ratio (%)
1	2	3	33	50
2	1	3	67	200
4	2	6	67	200
5	6	11	45	83
6	5	11	55	120
2	2	4	50	100
13	11	24	54	118
4	2	6	67	200
5	1	6	83	500
8	4	11	73	200

Type of variable: Continuous, ratio.
Importance to food security analysis: This information can be considered a vulnerability indicator. If the percentage of dependents in the household is very high (>70 percent), then there is a serious vulnerability problem in that household.

112. In countries with high prevalence of HIV/AIDS, it might be interesting to look at the percentage of effective dependents instead of limiting the analysis to the percentage of dependents. The formula is: (no. of effective dependents/household size) x 100. The effective dependents include the dependents and the persons who are of working age but are either chronically ill or disabled.

4. Crowding Index

> The average number of people per room in a dwelling

Calculation: For each household, the crowding index is calculated as the number of household members per number of rooms in the dwelling.
Key concerns: Multiple-dwelling households (as with extended families, polygamy, etc.), nomadic populations living in tents, large huts without internal separations, and other atypical housing situations, may give results that are difficult to interpret.
Reporting: Mean occupation ratio; rarely graphed.
Descriptive procedure: Compare means.
Interpretation: Larger ratio means a larger number of people per room.
Type of variable: Continuous; may or may not be normal.
Importance to food security analysis: Can be an element in the construction of a wealth index.

5. Child-headed households or elderly-headed households

> Households with a household head aged less than 18 years or aged 60 years and older.

Calculation: Age of household head must be converted to a categorical (bivariate) variable.
Key concerns: Be sure the HH head age data is clean and that missing values (e.g. 99, 999, etc.) are not recoded into age categories. In some cultures the eldest (male) in the household is almost automatically considered the household head, making the indicator less meaningful.
Reporting: Reported as a percentage of households.
Descriptive procedure: Frequencies, cross-tabs.
Interpretation: As a simple percentage.
Type of variable: Categorical, bivariate.
Importance to food security analysis:
Child-headed or elderly-headed households are usually more food insecure.

6. Marital status of household head

> A person's marital status describes his/her relationship with a significant other. Some common statuses are: married, single, separated, divorced, widowed, cohabiting.

Calculation: n/a
Key concerns: Make clear distinction among marital status categories in the questionnaire according to the local context (e.g. include "cohabiting" if this is a relatively common practice; reporting cohabiting people as married or single would be misleading).
Reporting: Reported as a percentage of households.
Descriptive procedure: Frequencies, cross-tabs.
Interpretation: As a simple percentage.
Type of variable: Categorical, nominal.
Importance to food security analysis: Food insecurity is often found to be related to certain marital status categories. For example, in many contexts widows/widowers or single mothers are expected to be more prone to food insecurity. Marital status can be

used as a criterion for targeting programme options if CFSVAs give evidence that it affects household food security status.

7. Literacy rate

Being able to read and write a simple sentence.

Calculation: Simple calculation of percentage.
Key concerns: Be sure the proper population is selected. The question is typically applicable to all members aged 15 years and older in the household roster.
Reporting: Reported as a percentage.
Descriptive procedure: Frequencies, cross-tabs.
Interpretation: As a simple percentage.
Type of variable: Categorical (bivariate); nominal.
Importance to food security analysis: Literacy is often positively correlated with improved food security.

8. Physically challenged members of a household

With respect to an individual, a physical or mental impairment that substantially limits one or more of the major life activities of that individual.

Calculation: Simple calculation of percentage.
Key concerns: Physical challenge should be clearly defined to avoid different interpretations among enumerators and interviewees. Especially in countries with high prevalence of HIV/AIDS, it is crucial not to confuse chronic illness and mental/physical disability:

- Chronic illness is a prolonged illness that does not allow individuals to engage in normal activities for at least 3 months (over the past 12).
- Physical challenge limits the capabilities of individuals, but is the effect of an accident, illness, etc. An individual can be disabled from an illness even when the illness is over.

Reporting: Reported as a percentage (of people and of households with disabled heads/adult members).
Descriptive procedure: Frequencies, cross-tabs.
Interpretation: As a simple percentage.
Type of variable: Categorical (bivariate); nominal.
Importance for food security analysis: Disabilities, limiting an individual's capacity to perform normal activities, augment people's exposure to food insecurity. A high incidence of physically challenged individuals could indicate a deterioration in the surveyed population's capacity to participate in certain types of humanitarian assistance programmes such as labour-intensive food-for-work.

9. Sex of household head

The sex of the head of the household.

Calculation: Simple calculation of percentage.
Key concerns: There should be very little, if any, missing data. Miscoded cases should be set to "missing," when unable to correct. When calculating from the household roster, households reporting more than one head of HH should be set to "missing" when heads are of the opposite gender.

Reporting: Percentage female-headed households; percentage male-headed households.
Descriptive procedure: Frequencies, cross-tabs (may require data transformation from individual data to HH data).
Interpretation: Interpreted as a simple percentage.
Type of variable: Categorical, bivariate.
Importance for food security analysis: Sex of household head is often correlated with food security. In many cultures, female-headed households are found to be (but are not always) more food insecure, because these households lack an income earner. High levels of female-headed households may indicate some underlying problem (conflict, migration, etc.).

10. Age of household head

> The age of the household head (regardless of sex) should be in completed years.

Calculation: n/a
Key concerns: Outliers, missing data; definition of HH head.
Reporting: Mean age of household head, percentage of household heads under the age of x years (cut-off defined on a country-by-country basis, but is usually less than 18 and over 60).
Descriptive procedure: Compare means; descriptive statistics.
Interpretation: Mean (watch for outliers), possibly median.
Type of variable: Continuous (can be transformed to categorical, hierarchical)
Importance for food security analysis: Sometimes older or very young household heads are more vulnerable to food insecurity.

11. Years of schooling

> Total years of education, by individual member in the household roster.
> Often surveys collect this information only for household heads and their spouses.

Calculation: n/a if the years of schooling of the HH members are kept separated; however, sometimes calculated as the sum of years of education of the HH head and his/her spouse if summarizing the information into one indicator.
Key concerns: If the HH member repeated grades, the years of education increase, but not his/her level of education.
Reporting: *For the HH head:* (1) mean number of years of education; (2) median number of years of education; (3) percentage of HHs where the years of education of the HH head is above a certain cut-off (e.g. number of years required to complete the primary cycle).
For the spouse of the household head: Report mean and median number of years of schooling. If the years of education of the HH head and spouse are combined into one indicator, report: (1) the mean of the combined indicator; (2) the median of the combined indicator; (3) the percentage of HHs where the years of education of the HH head and her/his spouse is above a certain cut-off.
Descriptive procedure: Descriptive statistics (mean, SD, shape of the distribution, frequency distribution, etc.); compare means.
Interpretation: Carefully interpret the mean, the median, the shape of the distribution. Take care in identifying outliers and extreme cases.

Type of variable: Continuous (can be transformed to hierarchical).
Importance for food security analysis: HHs where the level of education of the HH head (and his/her spouse) is poor can be more vulnerable to food insecurity and are more likely to have children out of school or not completing the primary cycle (thus reinforcing the cycle of poverty and vulnerability).

12. Level of schooling

Level of schooling attained by the HH member. Usually this indicator is collected for the HH head and his/her spouse. Typically the following answers are provided: no education, did not complete primary cycle, completed primary cycle, did not complete secondary cycle, completed secondary cycle, had vocational training, did not complete university, completed university.

Calculation: n/a (if levels of schooling of the HH head and his/her spouse are kept separated); level of schooling of the HH head and his/her spouse can be combined into one indicator in order to have one proxy of the educational status of the HH.
Key concerns: Poor sensitivity of the indicator (e.g. behind a "did not complete primary cycle" we may find different levels: people who almost completed the cycle and people who attended for only one year). How to treat Koranic schools, non-formal education, and vocational education.
Reporting: (1) percentage of HHs by the level of schooling of the HH head, (2) percentage of HHs where the level of schooling of the HH head is above a cut-off (e.g. primary cycle completed), (3) percentage of HHs by the level of schooling of the spouse of the HH head, (4) percentage of HHs where the level of schooling of the spouse of the HH head is above a cut-off, (5) percentage of HHs by the combined indicator.
Descriptive procedure: Descriptive statistics (frequency distribution, mode, etc.), cross-tabulations.
Interpretation: Caution in interpreting the results of the cross-tabulations and Chi-square, especially with a low number of cases.
Type of variable: Categorical ordinal.
Importance for food security analysis: HHs where the level of education of the HH head (and his/her spouse) is poor may be more vulnerable to food insecurity and are more likely to have children out of school or not completing the primary cycle (thus reinforcing the cycle of poverty and vulnerability).

13. Net enrolment

For primary school: The number of primary school–age children enrolled in primary school out of the primary school–age children living in the area. For secondary school – if information is available: the number of secondary school–age children enrolled in secondary school out of the secondary school–age children living in the area.
Net enrolment range: From 0 (none of primary school–age children living in the area attend school) and 1 (all the primary school–age children living in the area attend school). NER can be expressed in decimals or with percentages.

Calculation:
Option A for each HH: The number of **primary school–age children** living in the HH who are enrolled in primary school divided by the number of primary school–age children living in the HH. This calculation gives the NER of each HH.
Option B at the aggregate level: The total **number of primary**

school–age children included in the sample who are enrolled in primary school divided by the number of primary school–age children of the area. This calculation gives the overall NER of the children.

Key concerns: The term enrolment is different from the term attendance. A child is enrolled if his/her name is on the school register. A child is attending if he/she is actually going to school. Considering that the main purpose of WFP School Feeding activities is to increase access to education, enrolment should be a priority instead of attendance. However, it is suggested not to use the term enrolment in the questionnaire because people may not understand its meaning. It is suggested to simply ask, “Is [...] going to school?” It is better to calculate NER at the aggregate level.

Reporting: The average NER of the HHs, the median NER of the HHs, the overall NER of the children.

Descriptive procedure: Descriptive statistics, compare means.

Interpretation: Take care in interpreting the mean, the median, and assessing the shape of the distribution. Distinguish clearly between the average (or mean) NER of the HHs and the overall NER of the children (different level of aggregation; both statistics are valid and they complement each other).

Type of variable: Continuous.

Importance for food security analysis: Poor educational status increases vulnerability to food insecurity.

14. Gross enrolment rate

Calculation: The total **number of children (disregarding their age)** living in the sample HH who are enrolled in primary school divided by the number of the population of the age group living in the area that officially corresponds to the primary education level.

> The number of children enrolled in primary school (disregarding their age) out of the population of the age group that officially corresponds to the level of education shown. Gross enrolment rate: GER is 0 if none of the children living in the area is going to school; GER could be higher than 1 if there are pupils still enrolled in primary school even if they are already of an age to be in secondary, or if there are those too young to officially enter primary education. GER can be expressed in decimals or with percentages.

Key concerns: The term enrolment is different from the term attendance. A child is enrolled if his/her name is on the school register. A child is attending if he/she is actually going to school. Considering that the main purpose of WFP School Feeding activities is to increase access to education, enrolment should be a priority instead of attendance. However, it is suggested not to use the term enrolment in the questionnaire because people may not understand its meaning. It is suggested to simply ask, “Is [...] going to school?” It is better to calculate NER at the aggregate level.

Reporting: The average GER of the HHs, the median GER of the HHs, the overall GER of the children.

Descriptive procedure: Descriptive statistics, compare means.

Interpretation: Interpreting gross enrolment is not as easy as interpreting net enrolment. Both high and low gross enrolment rates are undesirable. High gross enrolment levels (over 100 percent) indicate large numbers of overage children in primary school, indicating poor academic progress and a high level of repetition in the school system. Low gross enrolment rates reflect low net enrolment rates from lack of

school attendance because children either have poor access to schools or are kept away by their parents.
Type of variable: Continuous.
Importance for food security analysis: Poor educational status increases vulnerability to food insecurity.

15. Housing construction materials

Major material of house roof, wall, and flooring.

Calculation: n/a
Key concerns: Appropriate material definitions. The list of possible options must be tailored to the local context; try to identify materials that can distinguish households on the basis of their wealth.
Reporting: Reported as percentage of households.
Descriptive procedure: Frequencies, cross-tabs.
Interpretation: As a simple percentage.
Type of variable: Categorical, nominal.
Importance for food security analysis: Aside from a simple description of the household status from single house features, determined material categories could be used in the construction of the wealth index, an indicator that aims to synthesize different components of the household-level wealth (see Wealth Index in Annex 3).

16. House ownership

Legal title or other right to use the house. Three broad categories are usually present as minimal differentiation: ownership, rental payment, or staying for free without owning the house.

Calculation: n/a
Key concerns: The concept of ownership differs from context to context. Even if house ownership is more of a concern than speaking about land ownership (with diverse land tenure legislation in different countries and in different society organizations) house ownership, based on local knowledge, has to be clearly defined and commonly understood by survey designers, enumerators, and interviewees.
Reporting: As percentage of households.
Descriptive procedure: Frequencies, cross-tabs.
Interpretation: As simple percentage.
Type of variable: Categorical, nominal
Importance for food security analysis: In certain contexts, especially urban and peri-urban, where housing is commonly accepted to be a problem, house ownership provides a common measure of wealth and stability. House ownership can also be used in creating a wealth index.

From the food security perspective, house ownership should be analysed by taking gender into consideration. In some contexts, when the man dies, the wife loses both the land and house. Such inheritance practices may start a destitution process in which households headed by widows face strong challenges to maintaining their livelihood assets.

17. Toilet/sanitation

Improved sanitation facilities	Unimproved sanitation facilities	No facilities
Connection to a public sewer Connection to a septic system Pour-flush latrine Simple pit latrine Ventilated improved pit latrine	Public or shared latrine Open pit latrine Bucket latrine	

Access to sanitary means of excreta disposal is estimated by the percentage of the population using improved sanitation facilities. Improved sanitation facilities are those more likely to ensure privacy and hygienic use. WFP adopts the definition from UNICEF and WHO, which have teamed up to track progress on global water and sanitation goals through the Joint Monitoring Programme (JMP) for Water Supply and Sanitation.

Calculation: Usually collected using multiple options; categories are recoded into three main groups, improved versus unimproved sanitation facilities or no facilities.
Key concerns: Measuring "basic sanitation" is complicated. Ideally, the definition of this term would encompass critical components of what sanitation services should aim for: privacy, dignity, cleanliness, and a healthy environment. From a monitoring point of view, however, such characteristics are difficult to measure. To resolve these issues, the UNICEF/WHO JMP classifies sanitation facilities and water supply sources as either "improved" or "unimproved." In doing so, it makes the assumption that those classified as "improved" are likely to be more sanitary than "unimproved" ones.

Not all people with access to improved facilities or sources actually use them. Consequently, the JMP has adopted "use" as the primary indicator for monitoring progress in both water and sanitation. (http://www.unicef.org/wes/mdgreport/monitoring1.php).
Reporting: Coverage estimates are expressed as percentage of households using improved sanitation facilities.
Descriptive procedure: Frequencies, cross-tabs.
Interpretation: As simple percentage.
Type of variable: Categorical, nominal.
Importance for food security analysis: "Use of improved sanitation" is a commonly used indicator to assess hygiene at the household level. Access to a sanitary latrine is a proxy indicator to measure vulnerability to diarrhoea and other diseases; hence this indicator is a proxy measure of food utilization. It is crucial to collect this indicator whenever anthropometry is included as part of the survey, as the relationship between level of malnutrition and poor hygiene can be explored. The use of selected improved sanitation facilities can also be used in the construction of the wealth index.

18. Water source (drinking, other)

Improved drinking water sources	Unimproved drinking water sources
Household connection Public standpipe Borehole Protected dug well Protected spring Rainwater collection Tubewell/hand pump	Unprotected well Unprotected spring Rivers or ponds Vendor-provided water Bottled water* Tanker truck water

*Bottled water is not considered improved due to limitations in its potential quantity, not quality.

Access to safe drinking water is estimated by the percentage of the population using safe drinking water sources (definition from UNICEF/WHO Joint Monitoring Programme for Water Supply and Sanitation).

Calculation: Usually collected using multiple options; categories are recoded into two main groups, improved versus unimproved drinking water sources.

Key concerns: Existing surveys do not provide information on the quality of water, either at the source or in households. Improved water sources may still contain harmful substances, and water can be contaminated during transport and storage. Although "safe drinking water sources" provides a good indicator for progress, it is not a direct measure of it. Dangerous levels of chemicals, such as the arsenic and fluoride, increasingly found in groundwater in South and South East Asia, are of growing concern, along with infectious or other toxic substances. The proportion of the population using safe drinking water is therefore likely to be lower than that using improved drinking water sources.
(http://www.unicef.org/wes/mdgreport/monitoring1.php).
See number 17, "Toilet/sanitation," for discussion on improved/unimproved water sources and concept of use versus that of access.

Reporting: Coverage estimates are expressed as percentage of households using improved drinking water sources.

Descriptive procedure: Frequencies, cross-tabs.

Interpretation: As simple percentage.

Type of variable: Categorical, nominal.

Importance for food security analysis: The use of unsafe drinking water is directly related to water-related diseases such as diarrhoea, cholera, and typhoid. These types of diseases are often found to be a cause of malnutrition. The use of selected improved water sources can also be used in the construction of the wealth index.

19. Distance from water source

Time or distance from the house to the main water point.

Calculation: If time or distance is collected as a continuous variable, values can be recoded into categories. Cut-offs should be decided inspecting the shape of the distribution in order to avoid having disproportionate categories.

Key concerns: Time or distance can be reported as "one way" or "round trip." If the aim is to estimate the overall time spent to access water (exclusively dedicated to fetch water), "round trip" seems to be more appropriate. If the aim is to carry out a spatial analysis between water point distance and other food security indicators, the indicator should be collected as "one way."

Reporting: Mean and median of the distance/time. Pay attention when calculating mean of time, if values are recorded in hours. The values should be recorded in the same way consistently by all enumerators (e.g. 1h30, 1h15, etc. or 1.5 h, 1.25 h, but not a combination of the two). The data entry programme can, if necessary, convert to the decimal notation. If the indicator is collected using categories, it is reported as percentage of households.

Descriptive procedure: Descriptive statistics, compare means; frequencies and cross-tabs, if categories are used/constructed.

Interpretation: Take care to carefully interpret the mean, the median, the shape of the distribution. Take care also in the identification of outliers and extreme cases. Take care

in the appropriate selection of cases, when values are entered as missing (e.g. 999) or non-applicable (e.g. water on premises 888, which would mean 0 minutes as timing or 0 metres as distance). Appropriate selection must also be performed according to the aim of the analysis: if the aim is to estimate the overall average time in a certain location (village, district, province, country), households that have water on premises (time to go to fetch water = 0) should be included as part of the population of that area. If the aim is to estimate the time needed to fetch water for households that have to walk/move to get water, households with water on the premises should be considered as non-applicable and hence excluded from the calculation.
Type of variable: Continuous or categorical (ordinal).
Importance for food security analysis: Access to water is a key element for food security and human life in general. The time household members have to invest to obtain water cannot be allocated for other activities such as household care or income generation.

20. Source of light, cooking fuel

Main source of light and main cooking fuel used.

Calculation: n/a
Key concerns: These indicators are fairly straightforward, with limited possibility of misunderstanding. Attention must be paid during the construction of the option lists: options must be limited but also vary enough to differentiate among households.
Reporting: Reported as percentage of households.
Descriptive procedure: Frequencies, cross-tabs.
Interpretation: As simple percentage.
Type of variable: Categorical, nominal.
Importance for food security analysis: The source of light could be an indicator of social status. However, use of certain facilities depends not only on household well-being, but also on those facilities' general availability. For example, if electricity is not present in a rural area, using it as a category to identify better-off households would be futile.

The source of cooking fuel could be a wealth indicator at the household level, but is probably more useful at an aggregated level, when the extensive use of particular items, such as firewood, could damage the natural environment. Pay particular attention to the use, availability, and deterioration of stock of natural cooking fuel when population density is high or suddenly increases, as in areas with IDPs or refugees, both in and outside camps, or urban and peri-urban areas.

Selected categories from both source of light and cooking fuel could be used in the construction of the wealth index.

21. Assets (household, physical)

Number of wealth and production assets.

Calculation: Ownership is collected in numbers. The asset index can be calculated in different ways. However, the most common method is either to collect the value of each asset and then monetize the assets, or to assign a weight to each of the assets based on its relative value. The final stage is to sum them.

Key concerns: Listed assets must be typical for the context but allow for inter-household differences to be captured. Knowing the condition of the asset can significantly help to put a correct monetary value on it. Sometimes, what is generally considered an asset to the household could, instead, be a liability. For example, a recent study in the West Bank reported that having cattle is considered a liability for Bedouin households, as cattle feed is more expensive than the price realized when cattle are sold. In this type of situation, cattle should not be included in the asset or wealth index.

Reporting: Mean and median value of asset or mean/median index value of asset, and percentages of households owning important productive assets (e.g. land, plough, tools, net, boat).

Descriptive procedure: Important productive assets – frequency and cross-tabs; asset or wealth index – compare means.

Interpretation: Key productive asset ownership interpreted as simple percentage. The asset index can be constructed in different ways. Generally, the higher the score, the wealthier the household.

Type of variable: Categorical (bivariate: yes/no) or continuous when number of each single asset is recorded.

Importance for food security analysis: Ownership of selected assets may be a sign of wealth or social status. Also, certain items, such as jewellery, household assets, land, and livestock, may be sold in case of difficulties. Consistent depletion may cause the household to slide into deep poverty, preventing them from coping with future emergencies.

22. Household shocks

Shocks are defined as events that have negative consequences for individuals, households, or communities. They can be of a natural, economic, political, or social nature. Usually analysis differentiates between covariate and idiosyncratic shocks. Covariate refers to shocks that affect a number of households, whole communities, or geographically defined areas, such as natural disasters, pandemics, or civil insecurity. Idiosyncratic shocks affect individual households and their members, such as loss of employment or the illness/death of a household member.

Calculation: n/a

Key concerns: Shock depends on household perception, and intensity is not measured.

Reporting: Percentage of households reporting specific shock(s)

Descriptive procedure: Frequency, cross-tabs.

Interpretation: As simple percentage.

Type of variable: Categorical (bivariate: yes/no).

Importance for food security analysis: Shocks, by definition, imply a negative consequence on individual, household or community welfare status. Shocks can lead to increased vulnerability and a decrease in food security at the individual and household levels.

23. Coping strategies

> The ways a community, household, or individual adjusts their livelihood strategies in response to a shock or risk.

Calculation: n/a
Key concerns: The concept of coping strategies is sometimes not well distinguished from activities undertaken in normal day-to-day life. Coping strategies should be reported when something extraordinary (a shock) occurs that worsens the ability of households to maintain their usual standard of life. An effort should be made to distinguish livelihood activities from coping strategies in both questionnaire design and enumerator training. For example, seasonal migration in some communities is a part of a livelihood strategy; hence it should not be considered a coping strategy. In other communities, seasonal migration is not a common livelihood strategy and should be considered a coping strategy.
Reporting: Percentage of households reporting specific coping strategies.
Descriptive procedure: Frequency, cross-tabs.
Interpretation: As simple percentage.
Type of variable: Categorical based on frequency.
Importance for food security analysis: Coping strategies may involve short-term changes in behaviour, such as switching diets, consuming less expensive foods, or borrowing money. When normal coping and response strategies are exhausted, households will use negative crisis strategies, such as selling productive assets (e.g. female livestock). Repeated shocks and the use of crisis strategies to manage their effects can lead to increased vulnerability and a decrease in food security at the individual and household levels.

24. Coping Strategies Index (CSI)

> CSI scores are often used as a proxy variable for food insecurity. The CSI is based on the frequency and severity of coping strategies for households reporting food consumption problems. Higher CSI scores indicate a more serious food security situation, and lower scores, a better one. Typical coping strategies include "changing the diet to less preferred food types," reducing portions, and reducing the number of meals.

Calculation: Refer to the manual, *The Coping Strategy Index: Field Methods Manual.* (CARE/WFP, 2008).
Key concerns: Proper weighting of coping strategies by severity is required.
Reporting: Mean score; presented as a table, or bar graph.
Descriptive procedure: Compare means.
Interpretation: A higher score indicates more frequent and/or use of more severe coping strategies.
Type of variable: Continuous; normality of the distribution has to be checked before performing statistical analysis.
Importance for food security analysis: The CSI is a particularly powerful tool for monitoring the same households or populations over time.

25. Access to agricultural land

Legal title or other right to use agricultural land. Usual response options are: ownership, renting land, tenant farmer (who gives a share of the crop to the landlord), borrowing land, or use of land for free.

Calculation: n/a
Key concerns: Rules and regulations to access land depend on the local context. Uniformity across a country has to be confirmed when a country-wide survey is planned. Possible options must be clearly defined and commonly understood by survey designers, enumerators, and interviewees, on the basis of local knowledge.
Reporting: (1) percentage of households with access to land, and (2) percentage of households accessing land through each specific option.
Descriptive procedure: Frequency, cross-tabs.
Interpretation: As simple percentage.
Type of variable: Categorical, nominal.
Importance for food security analysis: The way people access land may influence their use of it. For example, in many contexts, people are more likely to invest in their land (irrigation, terracing, land maintenance, etc.) if they have ownership or secure long-term use of it. In many countries, access to land is one of the key determinants of food security for rural communities.

26. Amount of agricultural land/farm size

Area of land cultivated regardless of tenure status.

Calculation: Area of land information is collected by asking about land cultivated under each crop. In order to arrive at the total cultivated land size, a simple sum is required.
Key concerns: Farm size could differ by agricultural season, which should be taken into account.
Reporting: Mean score; presented as a table, or bar graph.
Descriptive procedure: Compare means.
Interpretation: Generally speaking, in rural areas, the more land available, the better the household wealth status. However, this parameter must be carefully interpreted, together with other information on livelihood activities. Average land cultivated in a defined geographical area is also used to estimate the population pressure on available land.
Type of variable: Continuous.
Importance for food security analysis: Land availability, under the definition of "owned" and "cultivated," is a key factor for many people in rural areas in terms of food availability.

27. Main crops cultivated

Main crops cultivated as declared by the households.

Calculation: n/a
Key concerns: There are two possible ways to collect this information. One, list all possible crops and have the households rank, in descending order of importance, those they cultivate. According to experience, few households report more than three or four crops, and the additional information is usually of little value because of the small prevalence of extra crops. However, if the aim of the data collection is to carry out an agriculture survey, this approach may be worth following.

The second option: limit the number of main crops reported to three or four.

Reporting: (1) percentage of households that cultivate crop "xx" as the main crop, second crop, third crop, etc.; or (2), using a multiple response analysis, to report (a) percentage of households that cultivate a specific crop regardless of its ranking (percentage of cases), and (b) which crops are most often reported as cultivated (percentage of responses) regardless of their importance.

Descriptive procedure: Frequency, cross-tabs, multiple response.

Interpretation: As simple percentage.

Type of variable: Categorical, nominal.

Importance for food security analysis: Information on the main cultivated crops is important and must be linked with other data on livelihood activities. Growing food crops versus cash crops can indicate different income strategies/opportunities. Cash crops and food crops are generally specific to each country. Examples of commonly grown cash crops include jute, cotton, sugar cane, chili, onion, coffee, cocoa, banana, and oranges. Examples of commonly grown food crops include paddy, wheat, maize, cassava, sorghum, millet, and barley.

28. Agricultural production

Produced quantity of each cultivated crop.

Calculation: n/a

Key concerns: Quantities are self-reported and may not be accurately measured. Underestimation/overestimation might be an issue. Also, estimating quantities of crops that are not harvested in bulk at a specific time of the year, such as cassava, which is dug out from the ground all year round when needed, is particularly difficult, and reported quantities are therefore questionable. Outliers have to be carefully checked against other variables, such as cultivated land size.

Reporting: Total household crop production, per capita crop production.

Descriptive procedure: Mean.

Interpretation: Average produced quantities are usually compared among geographical areas, livelihood groups, male/female-headed households per each reported crop. Attention has to be paid to the selection of appropriate households (all households versus households that cultivate that particular crop only), according to the CFSVA's needs. For example, if the aim is to compare the average production of maize by region, households that do not produce maize have to be included in the "compare mean" computation, because the regional figure requires the inclusion of all sampled households selected in each region. If the aim is to compare the level of productivity of one crop among two groups of households that grow that crop, the number of valid cases (households) in each of the two groups has to be reported.

Type of variable: Continuous.

Importance for food security analysis: Agricultural production is one of the pillars of food availability, both at the geographical and household levels, even if it is not the only availability factor. Crop production has to be analysed together with price trend and other market information.

29. Food stocks

> Usually recorded as number of months a particular crop harvest is supposed to last for household consumption.

Calculation: n/a

Key concerns: Duration of stock is self-reported. Under- and overestimation might be an issue. Rarely, duration of stock is calculated comparing an average needed per capita per day quantity and figures from crop production. The key concerns reported for indicator 28, "Agricultural production," apply here. Another problem arises when duration of stock is enquired about for more than one crop. Global stock duration at the household level is rarely the simple sum of the duration in months of the two (or more) crops.

Reporting: Number of months crop stock is expected to last, as self-reported by households.

Descriptive procedure: Mean.

Interpretation: The longer the period of duration of the food stock, the better for a household's food security. However, this interpretation could be considered fairly rigid, as it does not take into account (a) trading part of the harvest (e.g. selling it right after the harvest, selling it when prices are higher), or (b) the possibility that households rely on different kind of livelihoods (e.g. commerce, salary).

Type of variable: Continuous (or categorical ordinal).

Importance for food security analysis: Food stocks may represent a safety measure for household food security in terms both of self-consumed food and of valuable items that can be sold in a time of crisis.

30. Ranked main and secondary livelihood activities

> Primary, secondary, tertiary, and main, second, quaternary income activity as declared by the households.

Calculation: n/a

Key concerns: When reporting, it is imperative that interpretation take missing values for secondary and tertiary activities into account. The accompanying table here illustrates that 86 percent of households were engaged in a secondary activity. Because of the missing values, there is a great difference between reporting the percentage and the valid percentage. The percentage values refer to the percentage of households engaged in a particular secondary activity of all sampled households. The valid percentage refers to the percentage of households engaged in a particular secondary activity of only the households engaged in secondary activities. This is explained more in depth in "Reporting."

Reporting: Percentage of households reporting a certain activity as main, secondary, tertiary, etc.

Example from the accompanying table: Agriculture was the main livelihood activity for 57 percent of the households. Twenty-nine percent of households reported wage labour as their main livelihood activity. Eighty-six percent of households reported undertaking a second activity. There are two possible ways of reporting the figures: (1) "50 percent of HH (of the ones engaging in a second activity) are engaged in wage labour, 33 percent in agriculture," and so on; or (2) "while 14 percent of HHs are not engaged in a second livelihood activity, 29 percent of all HHs engage in wage labour, 29 percent in agriculture," and so on.

Livelihood activities	Main activity		Second activity	
	%	Valid %	%	Valid %
Agriculture	57	57	29	33
Wage labour	29	29	43	50
Fishing	14	14	14	17
Subtotal	**100**	**100**	**86**	**100**
Missing	0		14	
GRAND TOTAL (including missing)	**100**		**100**	

Descriptive procedure: Descriptive statistics – frequencies.
Interpretation: Many different activities could be listed as a potential option. However, only the ones that record reasonable significant percentages are worth being mentioned.
Type of variable: Categorical, usually many categories.
Importance for food security analysis: Understand which activities are more important in a certain population and natural and social context. Specifically, these activities are ranked.

31. Contribution from the different livelihood activities (average percentage, average absolute value)

Contribution from the different livelihood activities.

Calculation: Average percentages or average absolute income values per each activity.
Key concerns: Outliers, missing data, activities repeated in more than one activity variable, percentages in the activity variables that do not add up to 100 percent.
Reporting: Average percentages or average absolute income values per activity.
Descriptive procedure: If the information is collected using four variables (first, second, third, and fourth activity), the data must be restructured creating one variable per each livelihood activity. In other words, a new variable needs to be created that is the sum of the percentage contribution or amount of income earned from each livelihood activity. Activity values should be percentage contribution or absolute value, in the case of income. The new variable computation would be a sequence of logical operations. The following example is an extract from a longer SPSS syntax:

```
compute act01 = 0 .
IF (Activity1 =1) act01 = act01+Activity1_Value .
IF (Activity2 =1) act01 = act01+Activity2_Value .
IF (Activity3 =1) act01 = act01+Activity3_Value .
IF (Activity4 =1) act01 = act01+Activity4_Value .

compute act02 = 0 .
IF (Activity1 =2) act02 = act02+Activity1_Value .
IF (Activity2 =2) act02 = act02+Activity2_Value .
IF (Activity3 =2) act02 = act02+Activity3_Value .
IF (Activity4 =2) act02 = act02+Activity4_Value .
```

And so on . . .
The syntax will create as many variables as listed livelihood activities.
As is easily seen, if an activity is listed in more than one activity variable, its values would be summed up, and not lost as if overwritten.

Interpretation: Interpret as a simple percentage or average. Information may be presented in a table or a graph (pie or bar chart for percentages, bar chart for absolute values).
Type of variable: Continuous.
Importance for food security analysis: Understand main livelihood patterns as combinations of activities undertaken. The values could be presented for the entire data set or for stratification, which would be a means comparison among categories of a certain categorical variable (e.g. provinces, districts, HH head's sex). This type of analysis is done to detect differences among categories in terms of livelihood activities' importance.

32. Livelihood groups

Groups of households characterized by different livelihood activities or combination of activities.

Calculation: Cluster analysis (could be preceded by principal component analysis) performed on a set of different livelihood activities structured as different variables. Each activity is a separate variable; it can be binomial (done/not done) or continuous (if the relative importance of the activity to the total income is collected).
Key concerns: Outliers, missing data, activities repeated in more than one activity variable, percentages in the activity variables that do not add up to 100 percent; manageable number of groups (not too high) versus internal homogeneity of the single groups (more groups would reflect a better homogeneity of characteristics). The number of final groups is not fixed but should represent an intrinsic property of the set to be clustered. The decision on the number of groups to ask should be guided by statistical similarity. Some groups might be merged at a later stage (e.g. "food crop farmers" and "cash crop farmers" into a general "farmers" category). The decision of regrouping is subjective, because, statistically, software packages cannot recognize that two different variables might be interpreted as providing similar (hence merge-able) information. This can be solved by merging the similar livelihood activities variables before clustering.
Reporting: Percentage of households in each livelihood group; cross-tab with other key variables.
Descriptive procedure: If the information is collected using four variables (first, second, third, and fourth activity), data must be restructured creating one variable per each livelihood activity, reporting percentage contributions or absolute income values. A cluster analysis is performed. Frequency on the categorical variable is obtained as a result of the cluster analysis.
Interpretation: Interpret as a simple percentage. Information may be presented in a pie chart to show the relative size (importance) of groups. A bar chart could be drawn, with absolute values showing the different income availability of different livelihood groups.
Type of variable: Input variables are continuous. Output variable is categorical.
Importance for food security analysis: Understand the main livelihood patterns as combinations of activities undertaken. Households with different livelihood strategies have different food security status and are also vulnerable to different types of shocks.

Household expenditures and access to credit

The next group of indicators refer to household access to credit and household expenditure patterns.

33. Household expenditure indicators

- **Total cash expenditures on food items:** the total amount of cash allocated to specific food items on a monthly basis
- **Total cash expenditures on non-food items:** the total amount of cash allocated to specific non-food items on a monthly basis
- **Total monthly household cash expenditures:** the total amount of cash used to purchase both food and non-food items on a monthly basis
- **Total credit expenditures on food items:** the total amount of credit borrowed to purchase food items on a monthly basis
- **Total credit expenditures on non-food items:** the total amount of credit borrowed to purchase non-food items on a monthly basis
- **Total monthly household credit expenditures:** the total amount of credit used to purchase both food and non-food items on a monthly basis
- **Per capita total expenditure (always cash):** the average amount of cash a household can allocate/make available for each of its members
- **Per capita expenditure quintiles:** the distribution of expenditures across a sample population based on five groupings – with each group representing 20 percent of the population (weighted sample). The first quintile is considered to be the poorest 20 percent – in terms of the value of per capita expenditures – whereas the fifth quintile can be considered the richest 20 percent of the population.
- **Expenditure shares:** the proportion (or percentage) of expenditures devoted to (a) food; and (b) non-food items. Shares of expenditures have two dimensions. The first is the overall shares of household expenditures going towards food versus non-food; when both are combined, it equals 100 percent. The second dimension is the proportion of expenditures on specific food and non-food items (e.g. cereals, meat, education).

Calculation:

1. **The absolute values of all *expenditures collected on a semi-annual basis*** (both in cash and credit) should be divided by six. This step transforms the six-month cash (or credit) outlay on those non-food items into one-month values – allowing for comparison between food and non-food expenditures.
2. **Total *cash* expenditures on food items:** The sum of all the expenditure values for specific food items (e.g. wheat + rice + n = total food expenditures in cash).
3. **Total *cash* expenditures on non-food items:** The sum of all the expenditure values for specific non-food items (e.g. kerosene + education + health + n = total non-food expenditures in cash).
4. **Total monthly household cash expenditures**: Total cash expenditures on food items + total cash expenditures on non-food items.

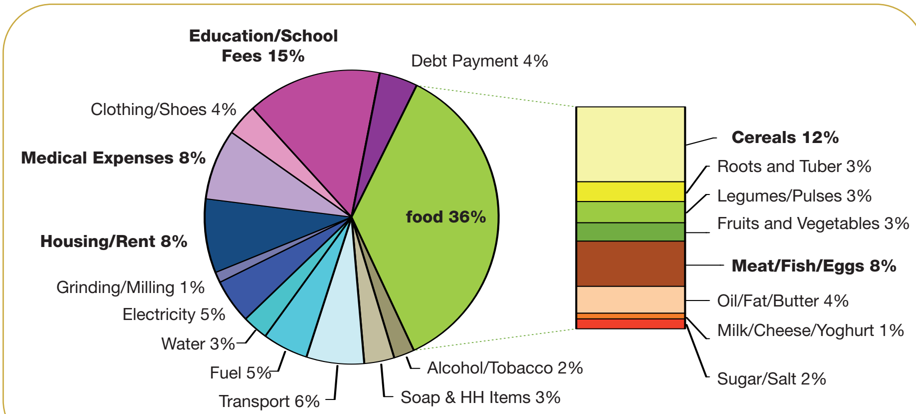
The same formulae apply to the calculation of credit-based expenditures:

- **Per capita total expenditure:** Total monthly household cash expenditures (step 4) divided by total number of household members.[113]
- ***Per capita expenditure quintiles:*** Transform the variable “per capita total expenditures (cash)” using “rank” and “Ntiles” commands in order to obtain the quintile (or class/category) – from 1 to 5 – to which the household belongs based on their per capita expenditure values.

113. An alternative approach is to divide the expenditure by the number of adult equivalents in order to take into account economies of scale for large households and the lower needs (for food, housing space, etc.) of children.

- ***Overall share of expenditures on food:*** Total expenditures on food items (cash or credit) divided by total monthly household expenditures (cash or credit).
- ***Overall share of expenditures on non-food:*** Total expenditures on non-food items (cash or credit) divided by total monthly household expenditures (cash or credit).
- ***Share of expenditures on specific food items:*** The expenditure value of a specific food item (cash or credit) divided by the total expenditure on all food items (cash or credit).
- ***Share of expenditures on specific non-food items:*** The expenditure value of a specific non-food item (cash or credit) divided by the total expenditure on all non-food items (cash or credit).

Key concerns: The quality, reliability, and accuracy of expenditure data have often been questioned. As is the case with income data, respondents may "under-report" their real expenditures.
Expenditures on food – be they with cash or on credit – vary seasonally (e.g. bulk purchases of staple foods). Given that the current module does not capture these seasonal dimensions, interpretation of data needs to be conducted carefully – spelling out constraints.
Cleaning/checking expenditure data is very important – especially when it comes to identifying outliers.
Reporting: Mean (or median, if the distribution is not normal) for all the created expenditure variables that are continuous, i.e. total and per capita expenditure (food, non-food, total; cash, credit, total) and share of expenditure (food versus non-food; and specific food items).
When reporting expenditure quintiles, specify range values associated with each quintile (minimum and maximum).
Figures can be reported in a table or graph. Pie charts are recommended when reporting shares (or percentages) of particular expenditures because (a) one unique graph displays expenditure distribution by food and non-food and by specific item (see accompanying example); and (b) it reminds one that the total is 100 percent.



Note that data on alcohol and tobacco – while collected as food expenditures – are normally reported as non-food outlays.

Descriptive procedure: Means comparison for continuous expenditure indicators; frequency and cross-tabs for quintiles.

Interpretation: Generally speaking, the higher shares of total expenditures going towards food, the greater the likelihood a household has poor food access. This is especially true if the household depends mainly on purchases for its food. Thus, for households with low levels of income who cannot produce enough food for themselves, buying food becomes, *de facto*, the main priority. As such, household resources will go towards ensuring that a minimum level of food is acquired to meet household needs. This, when compared to outlays on non-food priorities, will naturally result in a higher proportion of resources allocated to meet these food needs.

Care must be taken when interpreting outputs from food expenditure analyses as some households have a lower share of food expenditure because they rely on their own production. Similarly, better-off households may spend a high proportion on costly food items (e.g. meat), which might increase the percentage they spend on food. Therefore, it is important not only to understand the generalities of expenditure patterns, but also to investigate the types of items being prioritized.

Type of variable(s): Expenditure indicators can be continuous (total and per capita expenditures, shares of expenditure) and categorical (per capita expenditure quintiles).

Importance for food security analysis: Expenditures are useful as a proxy for wider purchasing power, which is an important component of food access. Moreover, understanding expenditures on specific items also allows the analyst to determine how households allocate scarce resources and give priority to competing needs.

34. Access to credit, sources of credit

The types of credit to which households have access; whether households use credit in order to purchase food. For CFSVAs, the following questions are used:

- Do you have access to credit or a source to borrow money from? (Circle all that apply):
  1. Yes – Relatives/Friends
  2. Yes – Charities/NGOs
  3. Yes – Local moneylender
  4. Yes – Bank
  5. Yes – Cooperative
  6. No access to credit
- Do you often purchase food on credit or borrow money to purchase food?
  0 = No
  1 = Yes
- If yes, in the last three months, how often did you borrow money or use credit to purchase food?
  1 = On one occasion
  2 = On two occasions
  3 = On three occasions
  4 = On more than three occasions

Calculation: Given that households can have more than one source of credit, the calculation is a simple count of how many different types of credit are accessible to a household.

Key concerns: There are two main concerns: (1) while it might be easy to gather data on the sources of credit, there is no information collected with regard to the "interest rates" charged to borrowers; and (2) there is usually no estimation of the amount borrowed.

Note also that in the first question, while options 1 to 5 are not mutually exclusive, option 6 indicates that the household does not have access to any credit source. Therefore it cannot be circled together with other options.

Reporting: there are four main ways of reporting outputs:

1. Percentage of households with access to specific types of credit.
2. Percentage of households with no access to credit.
3. Percentage of households that purchase food on credit.
4. Percentage of households that have purchased food once, twice, thrice, or more than three times in the last three months.

Descriptive procedure: Simple frequencies are the most common descriptive method. However, these frequencies can also be created for specific groups of households – e.g. those sharing the same livelihoods or food consumption patterns.

Interpretation: There are two main ways of interpreting credit data. The first is to look at whether the source of credit is formal or informal. Formal sources are usually banks and/or cooperatives. Informal sources are local moneylenders and/or friends and family. Poorer and food-insecure households tend to rely on informal sources of credit, as they do not always have the necessary collateral to borrow from banks. The second aspect of interpretation is to see how many times a household borrows on credit in a three-month period. If the frequency is high, then it is likely that this household will have problems both with food access and food consumption.

Type of variable: Categorical, bivariate.

Importance for food security analysis: Access to credit is important insofar as it can help households make productive investments and/or allow for asset accumulation. Households that do have access to credit are usually expected to be able to pay back the debt. At the same time, however, taking on credit – especially for food – can also translate into medium- and long-term debt if households are not able to generate sufficient income from their investments.

Nutritional status of children, women, and other nutrition-related indicators[114]

These indicators usually come from the section on maternal and child health, section 4.2.4.3, and are individual-level indicators. These indicators are discussed further in the CFSVA guidelines on nutrition, section 6.2.2.

35. Stunting

Low height for age for children over 2 years; low length for age for children under 2 years; a measure of chronic malnutrition characterized by a slowing in the growth of a foetus or child and resulting in a failure to achieve expected length in comparison to a healthy, well-nourished child of the same age.

Calculation: Measure a child's height and compare to a reference population; the most commonly used cut-off point with z-scores is -2 standard deviations (SD). For example, a child with a z-score for height for age of -2.56 is considered stunted, while a child with a z-score of -1.78 is not. If the z-score is less than -3 standard deviations, the child is considered severely malnourished.

114. Information in this section was obtained from Cogill, B., 2003, *Anthropometric Indicators Measurement Guide,* Washington, D.C.: FANTA Project.

Key concerns: Stunting is a measure of chronic malnutrition and therefore is not an accurate measurement of short-term changes in nutritional status.
Reporting: Results can be presented in table or bar graph format, disaggregated by age group and including frequency and percentage of children below the cut-off point of -2 SD.
Descriptive procedure: When using an anthropometric software package, there are three procedures to perform. First the raw data should be entered. Next the researcher(s) should combine the raw data on the variables (age, sex, length) to compute height for age. Third, the software should transform this data into z-scores so that the prevalence of stunting can be determined.
Interpretation: The percentage of children with z-scores of -2 SD or less represents the prevalence of moderate/severe chronic malnutrition in the sample population.
Type of variable: Continuous.
Importance for food security analysis: Stunting is an indicator of past growth failure and is associated with a number of long-term factors, including chronically inadequate levels of protein and energy intake, micronutrient deficiencies, frequent infection, inappropriate feeding practices over a sustained period, and household poverty.

36. Wasting

Low weight for height; a measure of acute malnutrition characterized by considerable weight loss or failure to gain weight, which results in a child's weight being substantially below that expected in a healthy child of the same length or height.

Calculation: Measure a child's weight and height and compare to a reference population; the most commonly used cut-off point with z-scores is -2 standard deviations (SD). For example, a child with a z-score for height for age of -2.56 is considered malnourished, while a child with a z-score of -1.78 is not. If the z-score is lower than -3 SD, the child is considered severely malnourished.
Key concerns: Because wasting is a response to short-term influences (see Importance for food security analysis below), it is highly sensitive to seasonality, and consequently should not be used for evaluation purposes.
Reporting: Results can be presented in table or bar graph format, disaggregated by age group and including frequency and percentage of children below the cut-off point of -2 SD.
Descriptive procedure: When using an anthropometric software package, there are three procedures to perform. First the raw data should be entered. Next the researcher(s) should combine the raw data on the variables (age, sex, length, weight) to compute weight for height. Third, the software should transform this data into z-scores so that the prevalence of wasting can be determined.
Interpretation: The percentage of children with z-scores of -2 SD or lower represents the prevalence of moderate/severe acute malnutrition in the sample population.
Type of variable: Continuous.
Importance for food security analysis: Wasting is an indicator of current malnutrition and is associated with inadequate food intake, incorrect feeding practices, disease, infection or a combination of these factors. Wasting in individuals and population groups can change quickly, showing marked seasonal patterns associated with changes in food availability and access, as well as disease prevalence. Because

wasting does not incorporate child age, the difficulties encountered in contexts where exact age is difficult to determine do not affect the accuracy of wasting as a proxy indicator of acute malnutrition.

37. Underweight

Low weight-for-age is a composite measure of stunting and wasting as it is influenced by both height and weight.

Calculation: Measure a child's weight and compare it to a reference population; the most commonly used cut-off point with z-scores is -2 standard deviations (SD). For example, a child with a z-score for height for age of -2.56 is considered malnourished, while a child with a z-score of -1.78 is not. If the z-score is lower than -3 SD the child is considered severely malnourished.

Key concerns: Although underweight is a good indicator for assessing changes in malnutrition over time, care must taken in interpreting this indicator as it reflects both chronic and acute malnutrition and does not provide the information needed to distinguish between the two.

Reporting: Results can be presented in table or bar graph format, disaggregated by age group and including frequency and percentage of children below the cut-off point of -2 SD.

Descriptive procedure: When using an anthropometric software package, there are three procedures to perform: First the raw data should be entered. Next the researcher(s) should combine the raw data on the variables (age, sex, length, weight) to compute weight for age. Third, the software should transform this data into z-scores so that the prevalence of underweight children can be determined.

Interpretation: The percentage of children with z-scores of -2 SD or less represents the prevalence of moderate/severe acute malnutrition in the sample population. It is important to remember that one cannot draw any conclusions with regard to whether or not the malnutrition is acute or chronic based on this indicator.

Type of variable: Continuous.

Importance for food security analysis: Prevalence of underweight children is commonly used as the indicator to assess changes in the magnitude of malnutrition over time.

38. MUAC

The circumference of the middle of the upper arm, measured in centimetres.

Calculation: Determine the middle of the upper arm by dividing the measured length of the upper arm from the tip of the shoulder to the tip of the elbow. Measure the MUAC at this point.

Key concerns: Inaccurate measurements.

Reporting: Findings can be reported in a bar graph or table.

Descriptive procedure: In general, for a child aged 6 to 59 months, if the MUAC is less than 11.0 cm, the child is severely malnourished. If the MUAC is between 11.0 cm and 12.5 cm, the child is moderately malnourished, and anything greater than 12.5 cm is considered good health. Generally, MUAC is recoded according to these categories and assigned a colour: red for severe malnourishment, orange for moderate malnourishment, and green for good health. Analysing the frequencies determines the prevalence of malnourishment in the sample.

Interpretation: The percentage of children with z-scores of -2 SD or less represents the prevalence of moderate/severe acute malnutrition in the sample population. It is important to remember that one cannot draw any conclusions with regard to whether or not the malnutrition is acute or chronic based on this indicator.
Type of variable: Continuous initially, recoded as categorical.
Importance for food security analysis: MUAC is a measure of acute malnutrition. Although wasting is the preferred measure of acute malnutrition, MUAC offers a quick and easy predictor of immediate risk of death due to macronutrient malnutrition. As such, it provides a useful tool for screening cases of acute malnutrition among children 12 to 59 months, particular in emergencies. In some cases, MUAC-for-age offers a more refined version of this indicator of acute malnutrition.

39. BMI of reproductive-age women

Index used to determine chronic energy deficiency (CED), a measure of underweight for non-pregnant adults[115]

Calculation: Weight in kilograms divided by height in metres squared. An adult (18 years of age or older) with a BMI below 18.5 kg/m^2 is considered to be malnourished.
Key concerns: Accuracy of measurement, difficulties in measuring the sick or disabled
Reporting: Table or bar graph reporting the percentage of women with a BMI below 18.5 kg/m^2
Descriptive procedure: Calculate weight divided by the height of the woman squared.
Interpretation: An adult (18 years of age or older) with a BMI below 18.5 kg/m^2 is considered to be malnourished. The percentage of women with a BMI above this cut-off point represents the proportion of women with good overall nutrition.
Type of variable: Continuous.
Importance for food security analysis: CED can be used as a measure of well-being and is a good proxy indicator for overall adult health. Research has also found associations between adult BMI and agricultural productivity. Low BMI is correlated with a large number of health-related outcomes, including early onset of chronic conditions and an increased risk of premature mortality.

40. Mortality rate (global, under-5, etc.)

The frequency of deaths in proportion to a population in a given period of time.

Calculation: For global mortality rates (GMR), the calculation is the number of deaths in the sample in a certain time period divided by the total sample and multiplied by 1,000. For children under 5, the calculation is the number of deaths in a certain time period divided by the mid-period population of children and the number of time units and multiplied by 1,000.
Key concerns: Comparisons between countries need to take into account the age structure of the population. Countries with higher percentages of the population aged 65 years and over could possibly have a higher mortality rate than a developing country with a relatively younger population (i.e. fewer people older than 65 years).
Reporting: Mortality rates are usually reported as simple fractions. Most often the denominator is 1,000 people and the timeframe is one year. Be sure to report these figures in order to avoid any misinterpretation.

115. Calculated by dividing weight in kilograms by the square of height in meters.

Descriptive procedure: See "Calculation."
Interpretation: High mortality rates can be indicative of crisis situations, whether they are disasters, conflict, or drought. As noted above, age structure also influences mortality rates.
Type of variable: Continuous.
Importance for food security analysis: Mortality indicators are used to assess the severity of a crisis, identify needs, and prioritize resources. They are also used to monitor and evaluate the extent to which the relief system is meeting the needs of affected populations and to gauge the overall impact and performance of humanitarian assistance in a given situation.

41. Disease prevalence

Often there is a correlation between illness and malnutrition, especially among children.

Calculation: CFSVA should use a standard recall period of two weeks for determining the prevalence of diarrhoea, fever, and acute respiratory infections (coughing with faster-than-normal breathing). This should be disaggregated by men and women over the age of 18 years, boys and girls under the age of 18 years but older than 5 years, and children under 5 years.
Key concerns: Inaccurate reporting from sampled households.
Reporting: Means or frequencies can be reported in a table format or cross-tabulated with sex or age group. Bar graphs are also useful for indicating significant health issues within the community.
Descriptive procedure: There are two methods of analysing this data: The first entails recoding the data into binomial 0/1 variables. Health issues within the community are then illustrated by analysing the frequency of households that have experienced a particular illness or health issue. The second method of analysis entails analysing mean number of episodes (e.g. diarrhoea) by sex and age group. This will illustrate any gender or age discrepancies in disease prevalence.
Interpretation: Greater mean or higher frequency indicate a significant health risk, depending on the disease or health issue; this often will relate to or impact nutrition issues.
Type of variable: Continuous initially; can be changed to categorical.
Importance for food security analysis: Health environment and access to health services are closely linked with nutritional outcomes. Understanding the prevalence of primary public health helps in determining the underlying causes of malnutrition.

Food Consumption Analysis

The next group of indicators covers food consumption analysis. A more detailed description of the food consumption score and its relation to food consumption clusters is found in http://vam.wfp.org/MATERIAL/FCS_Guidance.

42. Food Consumption Score[116]

A composite score based on dietary diversity, food frequency, and relative nutritional importance (see section 6.2.1.1) of different food groups.

Calculation:
Regroup all the food items into specific food groups. Sum all the food items of the same group and recode the value above 7 as 7.

- Multiply the value thus obtained for each food group by its weight,[117] creating new weighted food groups.
- Create the food consumption score by adding the weighted food groups.

Once created, the thresholds of the food consumption score should be determined based on the frequency of the scores and the knowledge of the consumption behaviour in that country/region.

The standard thresholds are:

FCS	Profiles
0–21	Poor
21.5–35	Borderline
>35	Acceptable

Using those thresholds, recode the variable food consumption score, from continuous to categorical.

Key concerns: There is no global agreement on thresholds or division into food groups and weights. This methodology should be validated and possibly adjusted in each country/situation based on the specific food consumption characteristics of each country or region, and culture.

The questionnaire should properly account for food items consumed in very small quantities. For instance, if a spoon of fish powder is added to the pot, this should be treated as a condiment rather than a day's consumption of fish. The same applies to a teaspoon of milk in the tea.

Reporting: Frequency of the profiles and diet composition of each profile (mean number of days); cross-tab with other key variables.

Descriptive procedure: Considering that we can have both a categorical and a continuous variable, there are many different possible analyses. The most common are: cross-tab, compare means, descriptive statistics.

Interpretation: Diet diversity is positively linked with adequacy of food intake. Hence, the smaller the value of the food consumption score, the worse the household food consumption.

Type of variable: Continuous, if considering only the food consumption score; categorical, if creating the food consumption profiles.

Importance for food security analysis: This can be considered the base of the food security analysis. The creation of the food consumption profiles is the first step in creating the food security profiles. It is a key variable because it helps us understand the food consumption gaps of specific groups.

116. Section 6.1.3 offers a detailed discussion on food consumption score.

117. See section 6.2.1.1, Box 6.1, and Table 6.1: Cereals, tubers and root crops = 2; Pulses = 3; Vegetables = 1; Fruit = 1; Animal protein = 4; Milk = 4; Sugar = 0.5; Oil, fats = 0.5.

43. Dietary diversity

The number of different foods or food groups eaten over a reference period (typically 24 hours), and without regard to frequency of consumption. The number of different food groups consumed over a given reference period is an attractive proxy indicator for the following reasons:

- A more diversified diet is an important outcome in and of itself.
- A more diversified diet is associated with a number of improved outcomes in areas such as birth weight, child's anthropometric status, and improved haemoglobin concentrations.
- A more diversified diet is highly correlated with such factors as caloric and protein adequacy, percentage of protein from animal sources (high-quality protein), and household income. Even in very poor households, increased food expenditure resulting from additional income is associated with increased quantity and quality of the diet.
- Questions on dietary diversity can be asked at the household or individual level, making it possible to examine food security at the household and intra-household levels.
- Obtaining these data is relatively straightforward. Field experience indicates that training field staff to obtain information on dietary diversity is not complicated, and that respondents find such questions relatively straightforward to answer and not especially intrusive or burdensome. Asking these questions typically takes less than 10 minutes per respondent.

For a detailed discussion on the dietary diversity indicator, visit the following website: http://www.fantaproject.org/downloads/pdfs/HDDS_v2_Sep06.pdf.

Calculation:
Regroup all the food items into specific food groups.

Food Groups

- Cereals
- Roots and tubers
- Vegetables
- Fruits
- Meat, poultry, offal
- Eggs
- Fish and seafood
- Pulses/legumes/nuts
- Milk and milk products
- Oils/fats
- Sugar/honey
- Miscellaneous

Key concerns: Dietary diversity score does not take into account the nutrient value of food items eaten.
The questionnaire should properly account for food items consumed in very small quantities. For instance, if a spoon of fish powder is added to the pot, this should be treated as a condiment rather than a day's consumption of fish. The same is true for a teaspoon of milk in tea.
Reporting: Mean dietary diversity score; compare mean between different groups.
Descriptive procedure: compare means; descriptive statistics.
Interpretation: Dietary diversity is positively linked with adequacy of food intake. Hence, a smaller value indicates poor quality of diet.

Type of variable: Continuous.

Importance for food security analysis: Quality of diet is equally as important as quantity of diet. It is one of the key variables because it helps us understand the food consumption gaps of specific groups.

44. Food consumption clusters

Clusters or classes based on the similarity of their food consumption patterns.

Calculation: You can base food consumption clusters on single food items or on the aggregate food groups.

Run **a principal component analysis** (PCA) to reduce the complexity of the data set, to uncover similar characteristics among households, and to reduce the importance of outliers without removing them from the analysis. The main purpose of a PCA in this analysis is to describe households on the basis of their consumption patterns of various food items.

To run a cluster analysis on the factors will create groups of households that share common underlying characteristics in their food consumption pattern. The number of clusters is not predefined and the analyst should determine it by:

- Maintaining a reasonable level of similarity within each cluster as measured by low cluster inertia; and
- Ensuring significant differences between clusters.

Once the number of the clusters is decided, the analyst can report on them separately, or categorize them into three or four food consumption groups.

Key concerns: PCA and clustering are exploratory techniques that help the analyst understand the data. The decision to categorize food consumption clusters into specific groups is still a subjective one, and the analyst will be able to do it if he/she has knowledge of local dietary habits. Specifically, there are no set guidelines for deciding which clusters are considered poor or good.

Table 6.12, opposite, from the Sudan CFSVA, 2007, looks at the average food consumption score for each cluster.

Reporting: Frequency among households and mean diet composition of each group; cross-tab with other key variables.

Descriptive procedure: Cross-tab, compare means, descriptive statistics.

Interpretation: The worse the category, the worse the food consumption.

Type of variable: Categorical variable.

Importance for food security analysis: It is a key variable because it helps us understand the food consumption patterns of specific clusters of households. It also indicates the specific food strategies of these households and allows for the building of scenarios.

Table 6.12: Food consumption profiles for Sudan

No.	%	Cereals and tubers	Pulses	Veg. and fruit	Meat	Milk	Oil	Sugar	Dates	Overali dietary profiles
1.144	5	2	1	1	1	1	0	1	0	Poor consumption
997	4.4	7	1	1	1	1	0	1	0	
1.357	5.9	7	1	1	2	7	5	7	1	Low protein
1.204	5.3	7	1	1	2	0	5	7	1	
1.010	4.4	6	4	6	2	1	1	1	0	Pulses, no oil
1.047	4.6	7	5	2	2	1	1	2	0	
599	2.6	2	3	2	2	5	6	7	1	Low cereals, but diversified diet
1.296	5.7	3	4	4	4	1	1	1	0	
1.038	4.5	7	1	1	6	3	1	2	0	No oil and sugar, but good consumption
1.527	6.7	6	6	5	6	1	2	2	0	
890	3.9	7	2	6	3	1	5	7	1	All good consumption
905	4	7	6	2	4	1	6	7	1	
1.420	6.2	7	2	2	7	6	6	7	2	
888	3.9	7	6	6	7	6	3	4	1	
1.052	4.6	7	6	6	5	1	5	6	1	
1.186	5.2	2	4	5	6	7	6	7	2	
974	4.3	7	4	6	2	7	6	7	2	
1.236	5.4	7	6	2	5	7	7	7	2	
1.123	4.9	7	2	6	7	7	6	7	2	
1.923	8.4	7	7	6	7	7	7	7	3	
22.816	100	6	4	4	4	4	4	5	1	

45. Sources of food

Source of each food consumed in the past seven days.

Calculation: Each source should be weighted on the frequency of consumption of that specific food item.

Each household is asked to report the main and secondary source of each consumed food item. All food items are coded as either "purchased," "produced," "gift," "borrowed," "food aid," and so on. The households can give two possible sources of food for the same food item: the main and secondary source of food.

- If only one source is mentioned for a food item, 100 percent of that item is attributed to that source.
- If two sources are mentioned, 60 percent of that food item is assumed to come from the main food source and 40 percent from the secondary source.

Usually, then, the number of responses for each source is weighted by the frequency of consumption of the foods accessed through a particular source. Subsequently the proportion of each source is calculated.

For example, over a week, if a household consumed:

7 times bread, from purchase and borrowing
2 times potatoes, from own production
6 times vegetable, from own production and purchase

7 times oil, from purchase
1 time meat from purchase

Purchase = bread (7*0.6) + vegetables (6 *0.4) + oil (7*1) + meat (1*1) = **14.6**
Borrowing = bread (7*0.4) = **2.8**
Own production = potatoes (2*1) + vegetables (6*0.6) = **5.6**

Total estimated consumption = **23**

Purchase = (14.6/23)*100 = **64%**
Borrowing = (2.8/23)*100 = **12%**
Own production = (5.6/23)*100 = **24%**

This computation will help to define the contribution of the various sources for the food consumption of each household. In some cases, when one or more sources are predominant, it could be useful to recode the source variable into a new categorical variable defining meaningful thresholds. For example, in the Darfur analysis, food aid as source of food was very remarkable. The analysis was focused on food aid. The percentage contribution from food aid was recoded into categories: no food aid, less than 50 percent food aid, and more than 50 percent food aid.
Key concerns: It is important to clean the data, especially when it is impossible to know the source of a specific food (e.g. meat from food aid).
The contribution of each food source is based on the frequencies of the different items. This is not necessarily the same as quantities or values.
Reporting: Tables, charts, cross-tabulation with other key variables such as the food consumption profiles or the livelihood groups

Table 6.13: Example - Mozambique Community and Household Surveillance, Round 6

	% purchase	% food aid	% own production	% gift	% casual labour	% barter	% borrow	% other
1	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
2	27	22	21	10	14	1	1	4
4	34	33	19	6	6	0	0	1
5	30	33	26	6	3	0	1	1

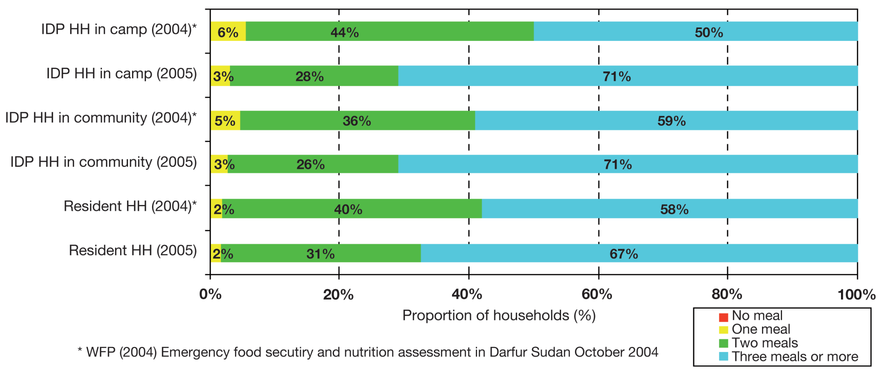
Descriptive procedure: Mean.
Interpretation: Sometimes some sources, such as food aid, borrowing, and gifts, are considered less reliable than others. There is no standard interpretation of these sources, but the analyst can use this information in his/her scenarios (e.g. "What will happen if the own production fails, or if purchases are impossible?"). It is important to check the sources with other key indicators, such as expenditure, agricultural production, or if the households actually received food aid.
Type of variable: Continuous or categorical.
Importance for food security analysis: This information is essential from the food security point of view because it gives information about food access.

46. Number of meals per day (adults and children)

> The number of meals consumed by the members of the household the day before the survey.

Calculation: This variable does not need a calculation; usually the data is collected separately for adults and children.
Key concerns: It is important to clean the data (e.g. if there are no children in a household, the value should be "not applicable," not "0"). Also of concern is how to define a meal (as compared to a snack). The number of meals can be different because of cultural reasons across countries, regions, and livelihood groups. For instance, some herdsmen might, even in good times, eat only one real meal a day.
Reporting: Tables (only if there are significant differences between the groups).

Figure 6.3: Meal frequencies among children 0-15 years of age



* WFP (2004) Emergency food secutiry and nutrition assessment in Darfur Sudan October 2004

Descriptive procedure: Compare means, frequencies.
Interpretation: It is often interesting to analyse this indicator by the food consumption groups/profiles.
Type of variable: Continuous.
Importance for food security analysis: This variable could help with recognizing households with extremely poor food consumption. However, do not use this variable alone to state the food security/insecurity level. This indicator could monitor changes over time.

47. Wealth index

> Wealth is the value of all natural, physical, and financial assets owned by a household, reduced by its liabilities. Although measuring wealth is possible, it requires making assumptions about the value of assets.
> The wealth index is a composite index that combines the ownership of key assets; it is used as a proxy indicator of household-level wealth.

Calculation: The calculation requires the various steps listed here. It is based on a principal component analysis.

1.Select the variables. Different kinds of variables can be included in the wealth index: productive assets, non-productive assets, household utilities, and others. The inclusion of productive assets/livestock and land ownership should be reconsidered if a high percentage of households do not practice agriculture or live in urban areas.[118]

Box 6.17 offers a list of the key variables usually collected and used in the wealth index.

Box 6.17: Typical variables used in a wealth index

Hand mill	Radio	Types of:	Persons per sleeping room
Sickle	Refrigerator		Land ownership
Axe	TV	- Water supply	Livestock
Livestock	Bicycle	- Toilet	
Hoe	Motorbike	- Flooring	
Tractor	Phone/cell phone	- Walls/house	
Plough	Chair	- Roof	
Etc.	Table	- Light source	
	Bed		
	Etc.		

2. **Explore the variables.** Running a frequency of each variable is useful to select the variables that should be included in the creation of the wealth index. The rule of thumb is that if a variable/asset is owned by more than 95 percent or less than 5 percent of the sample, it is excluded from the analysis. This is due to the fact that the wealth index is used to capture households with different wealth. For example, knowing that 98 percent of the sample owns a particular asset will not help the analyst characterize household asset ownership.
   Categorical variables should be recoded into one binary variable. For instance, if the light source has many options, such as "none," "wood fire," "oil lamp," "petrol light," "electricity," it might be appropriate in one country to recode this into "no/primitive," which would group the first two categories, and "with purchased energy source," which would regroup the remaining categories. In another country, the same options might be recoded differently: "no electric" versus "electric." The choice is based on the prevalence of both categories: if the prevalence is between 30 and 70 percent, then the indicator will probably help categorize more households than if the prevalence is only 5 percent. If necessary, two similar assets may be combined into one binary variable.

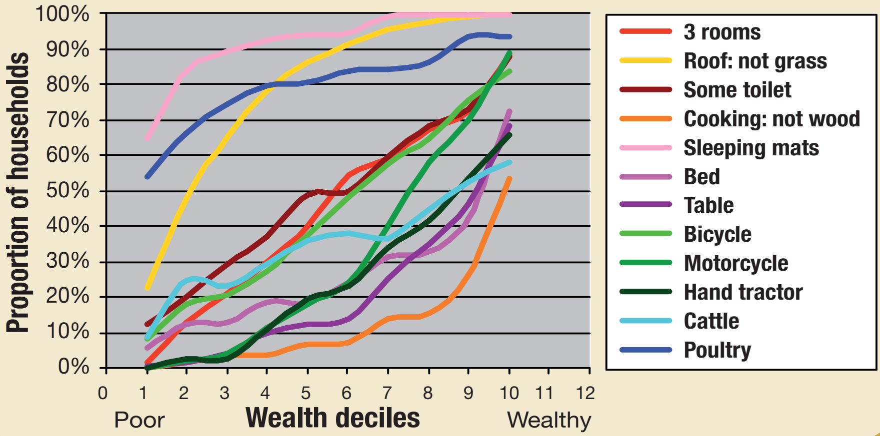
3. **PCA.** Once the variables have been selected, a principal component analysis must be run. For constructing the wealth index, the first principal component is taken to represent the household's wealth. This is an iterative process. Some variables that do not correlate well can be removed; if others do not behave as expected, one has to investigate this and maybe remediate the problem. For example, the possession of a radio turned out to be lower among the wealthiest in a country A, because these households often have radio-cassette players instead. The solution was to create a new variable: "radio or radio-cassette player."

118. The analyst should also bear in mind that if values are missing from any of the variables included in the PCA, the wealth index will not be calculated for that household.

4. **Recode the wealth index.** This variable (first principal component) can be used as a continuous variable — the higher the value, the higher the household's wealth index. It can also be recoded into categorical variable, usually terciles, quartiles, or quintiles ("ntiles").

5. **Graphing.** Graphing the "ntiles" by the variables included in the PCA will help the analyst understand if those variables are appropriate for the construction of the wealth index or if it is better to exclude/include other variables. The graph used for this is often called the "spaghetti graph," shown here:

Figure 6.4: Example of asset ownership in wealth deciles*



*Example from Laos CFSVA (2006). Percentage of households in each wealth decile owning various assets, including productive assets.

6. **Index.** The construction of the index requires several **iterations** before final results are obtained.

Key concerns: The analysis varies from country to country and the choice of the variables should be done after a careful exploratory analysis.
Reporting: Graph the "ntiles" with the variables included in the PCA; graph with other key variables, such as FCS, FCSg, coping strategies, and geographical strata.
Descriptive procedure: Correlate; compare means, cross-tab with other key variables
Interpretation: It is very difficult to interpret this variable alone. There is no standard interpretation because the factor coefficients of each original variable change with each PCA conducted. On the other hand, the indicator describes relative wealth comparisons within one country very well. Generally, the lower the value of the index, the less wealthy the household.[119]
For one country, we can keep the formula (the factor coefficients of each variable) as a baseline against which later measures can be compared.
Type of variable: Continuous or ordinal categorical (after construction of "ntiles").
Importance for food security analysis: This variable can provide an idea of the relative wealth situation of a household. Often, the wealth index can be used as proxy for vulnerability/resilience.

119. Assuming the variables are coded as suggested.

CHAPTER 7

Preparing conclusions and recommendations

This chapter outlines various options for food interventions based on the findings of the CFSVA. The response options are highly context-specific and should be tailored to address issues of access, availability, and utilization over the short, medium, and long term. When possible, interventions should be designed specifically for district or regional levels. Community input and priorities should be taken into consideration when designing an intervention and selecting programme activities.

Depending on the circumstances and practical possibilities, responses may be applied uniformly throughout the region or targeted to selected geographic areas. Geospatial analysis (discussed in section 3.5) will inform geographic targeting. Some may be targeted to specific types of households, while others, by their very nature, are targeted to specific types of individuals (while necessarily also benefiting the households of which those individuals are members). Assistance may be provided directly to households or individuals, or indirectly via a market intervention. Complementary action may be needed in some cases to address the food needs of individuals who may not be covered by measures directed to support households (WFP 2004).

7.1 RESPONSE OPTIONS

There are multiple ways in which a food security or malnutrition problem might be addressed. Some factors that affect the most appropriate response include (WFP 2004):

- The habits, priorities and culture of the affected population;
- The level of access to the particular area;
- The quality of infrastructure;
- Food availability and market conditions;
- The resources available (financial, human, logistical, etc.);
- The type of partnerships that are feasible (e.g. between government, United Nations, and NGOs);
- The political and economic environment; and
- The security situation.

Each response must be planned according to the particular circumstances and must be explicitly linked to the needs and gaps identified in the analysis. Table 7.1 outlines some steps specific to availability, access, and utilization that should be undertaken and reported:

Table 7.1: Problem and response analysis

	Problem Analysis	Response Analysis
Availability	• Seasonality • Distribution • Infrastructure	*Increase production* • Inputs/technology *Markets* • Increase imports • Trade flows & cross-border issues
Access	• Where are there access issues? Why? • Who has access issues? Why? • Seasonal?	• Cash vs. food interventions • Duration of intervention • Government policy, e.g. safety-net programmes • Role of other NGOs, United Nations agencies, etc. • Exit strategies/criteria • Income transfer
Utilization	• Prevalence of malnutrition • Causes of malnutrition	Response is contingent upon access/availability issues • Water & sanitation • Hygiene practices

The sectors and broad types of interventions that address food availability, access, and utilization factors include:

- **Food availability:** interventions to support agricultural production (crops, livestock), the movement of food between deficit and surplus areas, food distributions.
- **Food access:** interventions to support income generation (e.g. public works, food and/or cash for work), income transfers (cash/voucher distributions), food transfers (food distributions, school feeding), market intervention to support or reduce food prices, school feeding.
- **Food utilization:** interventions to improve health care, water, sanitation, shelter, nutritional knowledge (care practices), child care services.
- **Malnutrition:** interventions to improve food consumption (therapeutic and supplementary feeding programmes, school feeding, food distribution).

While short-term gains derived through relief aid, consumption-smoothing, or training are important for each potential intervention, it is important that they are seen in the context of longer-term household resiliency to shocks, productive capacity, and human capability. Cash-for-work and food-for-training activities are examples of medium-term activities, while other activities supporting livelihoods through provision of productive inputs can be seen as longer-term interventions (generally outside of WFP's mandate).

Various response options are outlined in Table 7.2. For more detailed descriptions, see Annex 15.

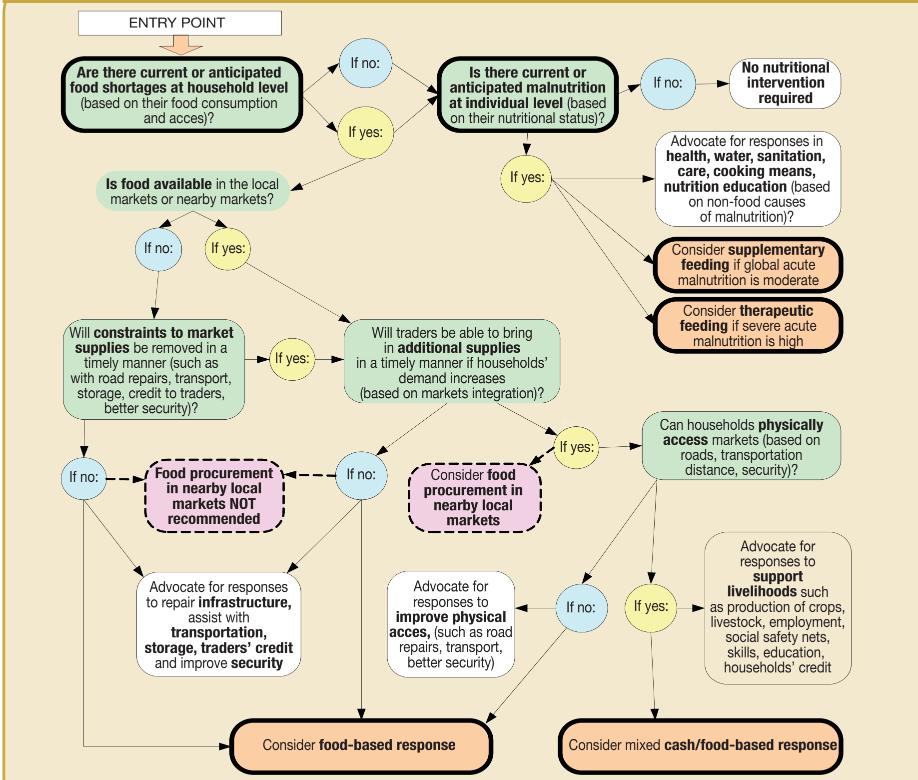
Table 7.2: Summary of response options

1. RESPONSES TO PROBLEMS OF FOOD **AVAILABILITY** AND/OR **ACCESS**	
a) **Food transfers** providing assistance to **households**	
Free food distribution	A distribution of free rations for households in need. Rations are designed to make up for household food access shortfalls. They may be general, that is, provided to all households in a particular area or population group, or targeted, to households in specific groups.
Food for work (FFW)	A food ration in payment for work, e.g. to rehabilitate or create the infrastructure necessary for specific livelihood activities (e.g. irrigation channels, fish ponds, rural roads, riverside jetties) or community services (e.g. health facilities). This may include incentives for work that requires little technical supervision in the aftermath of a disaster (e.g. general clean-up activities after a flood or cyclone).
Food for training **Exchange against produce**	Food provided as an incentive to enable (and encourage) individuals from food-insecure households to undertake skills training to increase their livelihood assets and their food production or options for earning income. Food given to affected rural households in exchange for their own produce (including livestock) for which there is temporarily no market locally.
b) **Food transfers** providing assistance to **individuals** (and also benefiting their households)	
Neighbourhood and home-based care programmes	Food provided to orphans and other vulnerable children (OVC) in the context of high prevalence of HIV/AIDS.
School feeding	A nutritionally balanced meal, or snack, for children/youths at school.
Food to other social service institutions	Food provided to orphanages; centres for unaccompanied children; homes for the elderly or handicapped people; hospitals and health centres providing in-patient care.
c) **Cash and other non-food transfers** providing assistance to **households**	
Cash transfer programmes	Cash distributed to target beneficiaries.
Cash for work (CFW)	A cash payment for work (similar to FFW).
Food vouchers	Beneficiaries receive vouchers that they can exchange for food in designated shops.
Non-food transfers	Non-food items or services (e.g. water, schooling, health care) provided free or at subsidized prices or through vouchers, thus sparing cash that could be spent on food.
Non-food support to livelihood activities	Productive inputs and/or services (or vouchers to obtain such inputs or services) to maintain, rebuild, or restore capital assets for food-insecure but economically active individuals and households. Productive inputs may include seeds, tools, fertilizer, irrigation, fodder or other livestock inputs, and tools and materials for artisans. Services may include veterinary care, extension services, improved access to pasture, and financial services such as emergency loans for productive activities.
d) **Market Interventions** to enhance availability and facilitate access for **households**	
Market assistance programme	Select (normally "second-choice") food commodities made available to traders and retailers to sell at controlled prices.
Market support	Reduction of logistic bottlenecks (e.g. repair of bridges or roads) or credit made available to traders.
2. RESPONSES TO PROBLEMS OF FOOD **UTILIZATION**	
Food preparation materials	Items required for preparing food, such as cooking sets, cooking fuel, and water. Such interventions are common for displaced and refugee populations.
Nutrition education, health, water and sanitation interventions	Interventions designed to improve feeding and care practices, prevent nutrient loss during food preparation, and prevent and treat diarrhoea or other diseases that affect nutrient absorption and utilization within the body.

3. RESPONSES TO **MALNUTRITION**	
a) Correcting high levels of **global acute malnutrition**	
Therapeutic feeding	Medical and nutritional treatment to save the lives of severely malnourished individuals. Treatment may be provided on site (in health centres or specially established therapeutic feeding centres [TFCs]) or — where cases are geographically dispersed — through a take-home ration with community-level follow-up by trained health workers (community-based therapeutic care [CTC]).
Supplementary feeding	The distribution of food to supplement the energy and other nutrients available in the basic diets of individuals with special nutritional requirements or who are malnourished. This may be either a take-home ration or a ready-to-eat food or porridge eaten on the spot. The food is in addition to the individual's share of the general ration, if any.
Public health measures	Measures to improve sanitation, water supplies, health care services and their use, measles vaccination, deworming, etc.
b) Correcting or preventing **micronutrient deficiencies**	
Food fortification	Foods fortified with specific nutrients (particularly vitamins and minerals), provided where the general diet is grossly deficient in these
Nutrient supplementation	Regular distribution of specific nutrient supplements (e.g. vitamin A capsules), when the general diet is grossly deficient in these

The decision tree in Figure 7.1 can be used to determine the type of response options that are most appropriate for a given context. Note that response options should take gender into account, although this is not illustrated in the chart.

Figure 7.1: Decision tree for response options



Source: WFP EFSA Handbook

Choosing No Response

In many complex situations characterized by widespread vulnerability, proposed food and non-food responses may actually jeopardize the well-being of the target population. In such situations, it is important to undertake a "Do No Harm"[120] analysis to determine the appropriateness of a food aid response. Aid can have important impacts on inter-group dynamics and conflict. A "Do No Harm" analysis helps to explain the complexity of conflict environments by:

- Mapping the interactions of aid and conflict;
- Analysing dividers and sources of tension;
- Analysing local capacities for peace/connectors; and
- Analysing potential intervention impacts on sources of tension and connectors.

7.1.1 SWOT analysis

After identifying potentially appropriate response options, analysis of the strengths, weaknesses, opportunities, and threats (SWOT) of each option can be conducted to determine the feasibility and appropriateness of proposed activities. Box 7.1, adopted from the WFP EFSA Guidelines, outlines SWOT analysis.

Box 7.1: SWOT analysis

SWOT Analysis is a way of systematically appraising different options. It is undertaken for each suggested response option.

Strengths and weaknesses – These reflect the **appropriateness** and **feasibility** of the response option.
The following **criteria** should be taken into account when assessing the **appropriateness** of a response option.
The response should:
- Address the factors that have been identified as contributing to risk;
- Reflect the needs and priorities of the affected population (disaggregated according to sex, age, etc.);
- Be compatible with local society and customs; and
- Be compatible with the interventions of the government or other agencies.

The response should not:
- Lead to dependency upon aid among any sector of the population;
- Have a negative impact on the local social, environmental, or economic situation (e.g. a large food distribution might discourage agricultural production);
- Divert people from other important tasks (e.g. productive activities, caring, collection of water and fuel);
- Expose the population or agency staff to security risks; or
- Stigmatize people (e.g. by explicitly targeting people with HIV/AIDS or from certain ethnic groups).

120. For more information on "Do No Harm" analysis, see Anderson, 1999, "Do No Harm: How aid can support peace or war," and Anderson, 2000, *"Options for Aid in Conflict."*

The following **criteria** should be taken into account when assessing the **feasibility** of a response option:

- Targeting criteria should be realistic, given social and cultural factors and the time available.
- It should be possible to undertake the response with the resources available. Consider financial, material, and human resources (including expertise).
- It should be possible to implement the response in a timely manner, given the urgency of the situation.

Opportunities and threats – These reflect the **external factors** that may affect the response. They are context specific. Some examples are given here:

Opportunities:
The introduction of new government policy that facilitates market functioning;
The end of the wet season and the improvement of transportation;
The signing of peace agreements; and
The harvest.

Threats:
Government policies that limit the scope of trade or aid programmes;
Reduction of donor interest in the country;
Deterioration of security; and
Lack of availability of key programme resources (e.g. fuel).

The **strengths, weaknesses, opportunities, and threats** are combined in a matrix. This helps when comparing response options, as each is described according to the same format. However, **judgement** still needs to be applied to decide upon the relative merits of different options.

In some cases, only one response option may be proposed. It is still advisable to undertake a SWOT analysis, in order to check the appropriateness and feasibility of the response.

7.2 TARGETING

Why Target?
In general, there are four main reasons for targeting assistance:

1. **Humanitarian.** Aid programmes should give assistance to those with the greatest need.
2. **Effectiveness.** The greatest impact can be achieved by focusing resources.
3. **Efficiency.** Efficient targeting improves the ratio of costs to benefits.
4. **"Do No Harm."** Untargeted aid – particularly food distributions – may damage the local economy or create dependencies.

Like most aid organizations, WFP must determine how best to use limited food and non-food resources. Options include narrowing the targeting criteria (e.g. excluding certain geographic areas or livelihood groups) and decreasing allocations. No option is ideal, or easy. When designing interventions, it is important to distinguish between targeting strategies based on assessment findings and those adopted to deal with limited resources.

Interventions should be targeted to the most vulnerable communities, and the most vulnerable households within those communities. In order to do so, it is essential to identify targeting criteria that are consistent with assessment findings regarding risk and vulnerability. These criteria should take into account the target population's vulnerability to food insecurity with respect to the three basic food security outcomes of availability, access, and utilization. Ultimately, groups targeted for assistance should include those at risk of food insecurity because of their physiological status, socio-economic status, or physical security, as well as people whose ability to cope has been temporarily overcome by a shock (TANGO 2004).

The process of targeting households will be undertaken by implementing agencies. The CFSVA will serve to guide the targeting process using the process outlined in section 7.2.1. Food aid programming requires targeting approaches that are flexible, both initially and throughout the course of interventions. In general, targeting involves defining the target group and identifying a target mechanism. Identifying the target group entails defining the criteria on which specific regions, populations, or livelihood groups, and households or individuals (i.e. levels) are given priority for the receipt of food or non-food assistance. Developing targeting tools or mechanisms ensures that only those who meet the criteria actually receive the benefits.

Targeting is inextricably linked to the assessment process. Successfully carrying out the targeting steps outlined in this section depends on accurate information. Targeting based on assumptions of need and levels of risk among particular groups may result in artificial targeting criteria and an inappropriate response.

7.2.1 Determining targeting level and defining target group

Depending on the context and intervention objectives, targeting occurs at multiple levels, including:

- ***Geographic targeting***
  Select areas or locations, ranging from countries or regions to villages and neighbourhoods, informed by geospatial analysis (e.g. agro-ecology, food economy zone, disaster areas, administrative zones).
- ***Group targeting***
  Population groups may be defined by livelihoods (e.g. smallholder farmers in food-insecure area), or groups that have lost normal access to food or livelihoods (e.g. IDPs or refugees).
- ***Household targeting***
  Within geographic areas and population groups, some families will be targeted, excluding others. Examples of target groups include the poorest or most food-insecure households.
- ***Intra-household or individual targeting***
  Selection of individual beneficiaries within households. Examples include children, women, livelihood activity, or those with the willingness to participate.

Geographic targeting should be combined with other levels of targeting to enhance the impact and cost-effectiveness of the intervention. In many contexts, targeting based on geographic level alone misses a large percentage of vulnerable people. Multi-stage targeting approaches combine geographic targeting with other targeting levels (e.g. group or livelihood system) and vulnerability criteria to capture the degree of variation in income and assets across households within regions.

7.2.2 Identifying targeting mechanisms

There are four basic types of targeting mechanisms:

- ***Market-based***
  Market interventions influence supply or demand of food or commodities sold by the target group in exchange for food (e.g. releasing strategic grain stocks to increase supply and lower or stabilize prices).
- ***Recipient self-targeting***
  Beneficiaries choose whether or not to participate, depending on costs and benefits.
- ***Administrative targeting***
  Beneficiaries are selected by outsiders using predefined criteria or indicators that are objective, measurable, and standardized (applied at any targeting level).
- ***Community-based targeting***
  Beneficiaries are selected by insiders – by community members themselves. Relies on participatory methods such as wealth-ranking, with variable input from government or project staff.

Each mechanism has pros and cons, which are summarized in Table 7.3. In practice, these four targeting mechanisms are often combined.

Table 7.3: Choosing a targeting mechanism

	Advantage/benefit	Interpretation
MARKET-BASED	• "Impersonal" selection-no corruption or bias • No administrative costs of direct selection (data and management) • Can recoup part of programme costs • Can avoid dependency and disruption to local economy (if well managed)	• Often benefits the better-off more than the poor (high inclusion error) • Can exclude the poor and vulnerable, who lack exchange entitlements (high exclusion error) • Needs good information and analysis of the market and the economic position of the target group • Can displace private-sector traders and/or discourage production (if poorly managed)
SELF-SELECTION	• No corruption or bias in selection - beneficiaries themselves decide whether to participate • No administrative costs of direct selection	• Needs good information and analysis to determine on the basis of which benefits and costs the intended people only will decide to select themselves • Programme must be able to accept everyone who self-selects, otherwise competition for resources will exclude the powerless and lead to possible bribery or pressure • Can result in high exclusion and inclusion errors • Not very effective in relief/crisis situations • Not effective with free distributions
ADMINISTRATIVE	• Objective, standardized, and verifiable (when successful) • Can be very accurate, therefore effective at minimizing errors (especially inclusion errors)	• High costs in management time, data collection and analysis, screening procedures (especially at household/individual level), and monitoring • Risk of bias, corruption, intimidation, theft, or error — requires monitoring/auditing • May not be feasible when administrative and information-gathering capacity is low • Identifying sensitive and specific indicators, and combining them to measure vulnerability, poverty, etc., can be technically very difficult
COMMUNITY-BASED	• Community members already know each others' situation (assets, income sources, household size, etc.), so no need for costly and difficult data collection/analysis • Community understands complex local interacting causes of vulnerability better than outsiders • Promotes participation in and ownership of the programme	• Risk of bias, corruption, intimidation, domination by powerful groups, etc. — requires monitoring/auditing • Community concepts of equity and vulnerability may not match donors' targeting priorities • Community may disagree with principle of targeting, and prefer to share aid among everyone • Developing and supporting community institutions needs significant staff time, skills, and resources • Difficult to standardize or compare targeting between different communities • "Community" may not include the most vulnerable groups, or in some situations (e.g. refugees) community may not actually exist • Costs to community decision-makers, in time and trouble, can be high

7.3 GENDER ANALYSIS AND PROGRAMME/INTERVENTION DESIGN

Integrating gender into qualitative and quantitative data analysis not only provides the information needed to develop comprehensive household food security profiles, but it also provides insight into how to design and implement gender-sensitive programming. For example:

- Food distribution – During general food distributions, entitlements are often issued in women's names. However, women do not always have control over the food once they have left the distribution site. Information on intra-household decision-making and control over resources will improve a general understanding of what happens once food is distributed, to what extent women are able to control it, and how it can potentially benefit the entire household.
- School feeding – In areas of wider gender gap in terms of school enrolment, understanding the reasons why one group is not sent to school is fundamental for assessing the appropriateness of take-home rations as an incentive for school attendance for pupils of a particular gender.

7.4. RECOMMENDATIONS RELATED TO HIV/AIDS

This section summarizes the content of the technical guidelines, HIV/AIDS Analysis: Integrating HIV/AIDS in Food Security Analysis, available online at the Food Security Analysis Service website (http://www.wfp.org/food-security).

7.4.1 WFP response to HIV/AIDS

In the early stages of its evolution, the HIV/AIDS pandemic was perceived primarily as a public health crisis. Since that time, the disease has been increasingly acknowledged as having contributed to the deterioration of human, financial, social, natural, and physical assets at the household, community, and national levels.

In response, humanitarian and development agencies have placed a growing emphasis on developing multi-sectoral programmes aimed at addressing both the short- and long-term impacts of HIV/AIDS. As part of this effort, WFP has adapted its activities to account for the complex relationship between food security, nutrition, and HIV/AIDS. In 2007, WFP established five key principles for programming in the context of HIV/AIDS. They clearly state that while WFP's primary focus remains vulnerability to food insecurity, it will promote and support the global fight against HIV/AIDS:[121]

1. The entry point for WFP involvement will always be situated in nutrition and food security. WFP's interventions will target beneficiaries based on their food security status, not just their HIV status.

2. When and where appropriate, WFP will take HIV and AIDS into account in all of its programming categories and in all assessments of needs.

121. FANTA and WFP, 2007, "*Food Assistance Programming in the Context of HIV.*"

3. WFP's HIV/AIDS response in specific countries will depend on the national strategy and will always fit within the government's framework for action.

4. In order to minimize the debilitating stigma and discrimination often associated with HIV and AIDS, WFP will support local non-governmental organizations and community-based organizations, including associations of people living with HIV/AIDS. WFP will use food aid to complement and scale-up existing government, United Nations and NGO partner activities in prevention, mitigation, and care for HIV-affected individuals and families.

5. WFP food assistance will place special emphasis on women and vulnerable children, in particular orphans, and will support the broader national and international response to HIV/AIDS to ensure that food aid is part of a larger package provided to HIV-affected households and communities.

Main WFP-supported programming options for people living with HIV/AIDS are shown in Box 7.2.

Box 7.2: Main programming options for people living with HIV/AIDS[122]

WFP-supported programmes provide food assistance to:

- Facilitate OVC access to education and support care (e.g. food support to OVC, extended and foster families, institutions that take care of orphans, and food for training that targets OVC);
- Support training that promotes livelihood diversification (e.g. training focused on income-generating activities and vocational skills);
- Support education and prevention activities;
- Support home-based care (HBC);
- Promote adherence to treatment of individuals with tuberculosis (TB);
- Promote adherence to (and uptake of) paediatric anti-retroviral therapy (ART) of children living with HIV/AIDS, provide nutritional support (and adherence) to prevention of mother-to-child transmission (PMTCT) of pregnant and lactating mothers living with HIV/AIDS; and
- support adherence to and uptake of ART.

7.4.2 Linking CFSVA findings to programme targeting and design in high HIV/AIDS prevalence countries

Identifying areas that are food insecure is the first step in targeting areas that receive WFP assistance. However, targeting all WFP interventions is a complex exercise that takes place at different levels and different stages of the programme design process. This section helps clarify how findings from a food security and vulnerability assessment can inform targeting and design of WFP interventions in high-prevalence countries.

122.WFP, 2003, "Programming in the Era of AIDS: WFP's response to HIV/AIDS" (WFP/EB.1/2003/4-B). WFP, 2006, "Five years later: An update on WFP's response to HIV/AIDS" (WFP/EB.A/2006/5-D/1).

In the context of HIV/AIDS, WFP interventions can be categorized in two main groups:

1) Food assistance for universal access to care, treatment and support; and
2) Other interventions with food security objectives.

7.4.2.1 Food assistance for universal access to care, treatment, and support

Table 7.4 outlines the main activities supported by WFP in southern Africa for universal access to care, treatment, and support.

Table 7.4: Food support for universal access to care, treatment, and support: Main activities supported by WFP in southern Africa

Intervention type	Activity type	Population prioritized	Intended outcome
Food support for universal access to care, treatment, and support	Home-based care support	Households with a chronically-ill household member	Alleviation of the impact of HIV/AIDS-related illness on the household
	ART support	Take-home rations provided to food-insecure mothers at risk	Supporting adherence to and uptake of ART
	PMTCT support	Pregnant and lactating mothers living with HIV/AIDS	Prevention of parent-to-child transmission of HIV/AIDS
	Paediatric aids support	Children living with HIV/AIDS	Supporting adherence to and uptake of paediatric ART
	TB support	Individuals with TB	Adherence to DOTs

Source: WFP Southern Africa, 2007, "Social Protection and Human Security for Chronically Food-Insecure Populations in Countries with a High Prevalence of HIV and AIDS."

Implementing these kinds of interventions, which are strictly linked to services supporting PLHIV, does not depend on the findings from a food security and vulnerability analysis. Geographic and community targeting are still needed for targeting food assistance to therapeutic programmes and home-based care. WFP policy is to give first priority to the most insecure areas that also have high prevalence of HIV and second priority to areas that are generally food secure but have high prevalence rates (under the hypothesis that the disease will increase their food insecurity). Therefore, food security and vulnerability analysis can inform programme targeting by combining data on food security and HIV prevalence and by identifying food-insecure areas with high prevalence of HIV.

Food assistance for universal access to care, treatment, and support is meant to sustain medical treatment of people living with HIV/AIDS. Selection of beneficiaries is therefore determined by prioritizing the most food-insecure people who are enrolled in TB or ARV treatments, PMTCT or HBC programmes. WFP's implementing partners are encouraged to adopt a multi-dimensional approach based on clinical, social, and demographic criteria for identifying individuals eligible for food assistance.

7.4.2.2 Other interventions with food security objectives

In southern Africa, HIV/AIDS considerations are included in several programmes with food security objectives. The process of targeting programmes with food security objectives includes the following steps:

- Step 1 – Geographic targeting to identify areas with high food insecurity (first priority) and high HIV prevalence.
- Step 2 – Community targeting to identify districts and communities with greater food security needs (similar criteria as those in Step 1).
- Step 3 – Household targeting to identify food-insecure households and households vulnerable to food insecurity. Household targeting can be undertaken through community-based targeting or partnering with medical facilities. In the context of HIV/AIDS, targeting criteria should look at HIV indicators.

Table 7.5, on page 314, summarizes WFP programme interventions most relevant in southern Africa for providing assistance to at-risk populations. Intervention types, activities, populations, and intended outcomes have been extracted from the paper "Social Protection and Human Security for Chronically Food-Insecure Populations in Countries with a High Prevalence of HIV and AIDS." The last column has been added to suggest how vulnerability analysis and mapping (VAM) can contribute to each type of activity.

As outlined in the table, food security and vulnerability assessments can inform the design of these interventions in several ways:
- Identify food-insecure areas that have a high prevalence of HIV and a high concentration of OVC;
- Provide evidence on educational gaps for OVC;
- Provide evidence on the need to support skills development and create livelihood options and to prioritize households affected by HIV and AIDS.

Empirical evidence about the impact of HIV/AIDS on livelihoods and food security can also be used to refine beneficiaries' selection tools.

7.4.3 Key references: HIV/AIDS

- WFP. 2003. *Programming in the Era of AIDS: WFP's Response to HIV/AIDS*, WFP/EB.1/2003/4-B.
- ibid. 2006. *Five Years Later: An Update on WFP's Response to HIV/AIDS*, WFP/EB.A/2006/5-D/1.
- ibid. 2006. *Getting Started: HIV, AIDS and Gender in WFP Programme*. Rome: WFP.
- ibid. 2008. HIV/AIDS Analysis: *Integrating HIV/AIDS in Food Security and Vulnerability Analysis*, VAM Branch and HIV/AIDS Service, Rome, Italy.
- WFP Southern Africa. 2007. *Social Protection and Human Security for Chronically Food Insecure Populations in Countries with a High Prevalence of HIV and AIDS*. Johannesburg, South Africa.

7.5. KEY REFERENCES: CONCLUSIONS AND RECOMMENDATIONS

- TANGO International. 2004. *Development Relief Program Guidance - Part III, Analytical Framework, Methods, and Tools*. Office of Food for Peace, Bureau for Democracy, Conflict and Humanitarian Assistance.
- WFP. 2009. *Emergency Food Security Assessment Handbook*, second edition. Rome.

Table 7.5: WFP-supported interventions with food security objectives in southern Africa

Social Protection Interventions				
Intervention type	**Activity type**	**Population prioritized**	**Intended outcome**	**VAM contribution**
Food support to vulnerable households emerging from or at risk of shocks to their food security and well-being.	Vulnerable group feeding	Chronically food-insecure populations regularly unable to meet food needs	Alleviation of household food insecurity	Provide geographic and social targeting criteria
	School-based take-home rations	Orphaned and other vulnerable children (OVC)	Safeguarding orphaned and other vulnerable children's access to primary education	Provide information on reduced access to education for OVCs and reasons for this (e.g. child labour)
Nutritional support to groups particularly vulnerable to malnutrition.	Supple-mentary feeding	Pregnant, lactating, and child-rearing mothers, and children under 5 years of age (MCH)	Prevention of maternal and child malnutrition and improved pregnancy outcomes	Identify areas with high levels of maternal and child malnutrition and possible linkages to household food security
	School feeding	Children attending basic education programmes in formal and non-formal schools	Safeguarding the enrolment, attendance, and retention of orphaned vulnerable children in formal and informal schools	Identify areas with lower educational outcomes Identify areas where access to education is reduced for orphans and other vulnerable children
	Early childhood care and development	At-risk children under 5 years of age	Prevention of early childhood malnutrition	Provide evidence on levels of child malnutrition Support monitoring and evaluation for nutritional rehabilitation centres
Protection to agriculture-based livelihoods.	Conservation agriculture	Rural poor agriculture-based households in areas affected by declining yields	Sustaining agricultural yield and soil quality	Provide evidence on the link between HIV/AIDS and • reduced labour at household level • decline in quantity and quality of crops • fallow land returning to bush • reduction in soil fertility • decline in on-farm conservation and/or irrigation practices • reduced expenditure on agricultural inputs
	Food-for-work/assets/training	Communities in areas with generalized food insecurity that require development of the local agricultural infrastructure	Preservation of agricultural production with access to appropriate infrastructure and training	Identify areas with short-term food insecurity for FFW interventions in the lean season
	Junior Farmer Field and Life Schools	Orphaned and other vulnerable children who require education in basic farming practices	Safeguarding the capacity of children to engage in agriculture-based livelihoods	Identify areas with higher numbers of OVC Provide evidence on the drop-out rates of orphans and/or vulnerable children Provide evidence on the loss of agricultural knowledge, practices, and skills in the affected households and communities

Source: WFP Southern Africa, 2007, "Social Protection and Human Security for Chronically Food-Insecure Populations in Countries with a High Prevalence of HIV and AIDS."

CHAPTER 8

Report preparation and dissemination

The preparation of the report is often the longest task in the entire CFSVA process, as it involves many factors and actors. In order to provide the CFSVA writer(s) with concrete help, this chapter gives brief guidance on:

- Preparing and presenting a draft report;
- Preparing the final report; and
- Communicating the results of a CFSVA.

8.1 PRESENTATION OF RESULTS

The presentation of results, either in written, tabular, or graphic form, is just as important as generating the results themselves. If you cannot make others understand what the data are telling you, then the analysis will not have the intended impact.
SPSS offers a variety of options for generating tables and graphics. However, Excel offers better graphing options.

8.1.1 Data tables[123]

Overview of data tables
Tables are the most common type of display for communicating the results of data analysis. There are many options available for tables. However, creating effective tables is often an underestimated task. It takes time to generate attractive and easily understandable tables.

A table is an array of regularly spaced numerals or words. Tables can be divided into three major types:

- Sentences listing a few numbers in the text (best for 1-5 numbers, where all are values of the same variable);
- Text-tables, i.e. indented text lines (3-8 numbers, for one or two variables; often shown as a list of bulleted points such as this one); and
- A full table that can cope with 5-100 numbers.

The following few simple rules might help you design effective tables that look professional and are easy to read.

a) Use the rule of seven.
Try not to use more than six rows or columns in one table. Too many rows and columns inevitably compromises the size of the text and consequent readability. More columns and rows demand more space and decrease the white space available for margins and separation among items. The work will then look more crowded and will be difficult to understand.

Of course, there will be times when you will need more than 6–7 columns or rows, but be aware that it will be harder for the user to effectively interpret the table.

123. This section draws heavily from Caldwell, R., 2006.

b) Use Serif fonts.
Due to the fact that most presentation tables allow only for short, basic caption words or brief phrases, Serif fonts, such as Arial, Verdana, Helvetica, and Univers, are better in most situations. With its small size, Verdana works particularly well, as its design offers apparently much bigger-looking characters while using the same point size of other fonts (a small text written at 9 pts. will look much bigger and more readable in Verdana than in Arial).

c) Create "breathing" space among all rows and columns.
By utilizing white space or the background colour, always keep a sensibly sized empty margin between each row and column of your table to guide the eye in reading without need for gridlines and borders. Creating a table that can visually stand without the use of any of its initial default border lines is a major achievement in information design. In the best designed tables, graphic lines are used only to define the title space or the starting and ending area of the table, without becoming too redundant in spelling out the obvious separation of all rows and columns.

d) Format titles for immediacy.
In a table, column and row titles (also known as heads) are essential in that they offer initial and immediate explanation of the table contents. By formatting titles with emphasis (bold) and by providing titles that are short and easy to understand, one ensures that the first information component that will be read is optimized for immediacy and clarity.

e) Align with precision.
It is essential that all elements in a table are precisely aligned and that each column appears to the eye by virtue of its perfect alignment with its neighbouring elements and not by need of vertical lines showing to which column an item belongs. If your table cannot survive without gridlines and borders, something is wrong with it. Work at improving alignment, spacing, and margins until it looks good without any lines. Then add back a few formal lines where emphasis or area definition are needed.

f) Include size of the sample.
This is almost never done except in research findings, but it is a good idea to include the number of observations that means and other statistics are based on. When you interpret a table, the number of observations are very important.

Level of reporting
Only key results should be presented in the main report – if they are important to the "story" of food security in the country studied. All other results should be presented in annex tables.

How to aggregate?
Only the most relevant survey domain should be selected; this is usually the domain that shows the most significant differences for the indicator. However, for operational reasons, administrative units might be preferred in the report. A map or chart could suffice, since the data are reported in the annexes.

For example, the report could feature expenditures on food by livelihood group instead of by administrative zone, since livelihood has a typical effect on these expenditures. The experience of certain shocks can, on the other hand, be presented by administrative zone.

8.1.2 Graphs

Graphs (also known as charts) are visual representations of numerical or spatial information. Computer software has made it easy to make all types of graphs – both good and, all too often, bad.

A graph uses a spatial arrangement on the page (or screen) to convey numerical information. This has several advantages:
- Graphs can have very high information density, sometimes with no loss of data. By contrast, stating only the mean and standard deviation provides a summary that loses information about, for example, the number and position of outliers.
- Graphs allow rapid assimilation of the overall results.
- The same graph can be viewed at multiple levels of detail (e.g. overall impression, close-up, and exact location of several adjacent points).
- Graphs can clearly show complex relationships among multivariate data (in two, three, four, or even more dimensions).

However, graphs also have some disadvantages, especially if badly designed:
- Graphs take up a lot of space if showing only a few data points. Hence it is best not to use them if there are only a few numbers to present.
- A graph may misrepresent data, for example, by plotting regularly spaced bars for irregular data intervals.
- A line may suggest interpolation between data points where none applies.
- It can be hard to read exact numeric values, especially if badly chosen axis scales are used. If exact numeric values are required, a table is best.

It is important to understand how to make the best of graphs. Note that it may not be necessary to display all available data in your graph. The key requirement is that the graph honestly and accurately represents the data you collected or want to discuss.

The following tips will help you create graphs that are clear, legible, and easily understandable.[124]

- **Make the data stand out.** It is the most important part of the graph. Anything that distracts from data is undesirable.
- **Use clearly visible symbols**, which are more noticeable than any other text on the graph, such as axis labels.
- **Reduce clutter on the graph.** For example, use relatively few tick marks: four to six per axis is usually sufficient.

124. Borrowed from Dave Kelly, Jaap Jasperse, and Ian Westbrooke, *Designing science graphs for data analysis and presentation: The bad, the good and the better.* Dept. of Conservation, Technical Series 32. Wellington, New Zealand.

- **Labels on the graph should be clearly offset from the data** or even outside the axes, to ensure they are not confused with the data; appropriate abbreviations can help keep labels short.
- **Keep notes and explanations outside the data region, if possible.**
- **Overlapping symbols or lines must be visually separable.**
- **Allow for reduction and reproduction,** since most printed graphs will be reduced and photocopied at some stage: sometimes through several generations! If you can reduce a graph to 0.71 twice (i.e. reduce by 50 percent) and it is still readable, it will suit most presentation purposes.

8.1.3 Mapping of data and results

VAM produces many types of maps showing different kinds of data targeted at different audiences and through different kinds of media. The aim of this section is to define, within the context of the CFSVA, a framework and a roadmap for the standardization of these outputs, for a standardized methodology for map creation that allows for creating, updating, and re-using map elements, and for standardizing the look and feel of maps created using different media and portraying different messages.

8.1.3.1 Thematic Maps

The purpose of producing maps of survey data is to let the reader easily understand geographical patterns in the results. It is important, however, that the results display patterns that are real. The legend should clearly show the thematic ranges or classification and other important feature symbology (e.g. to distinguish administrative boundaries of different levels). This is explained further here.

- The boundaries between colours should reflect differences in underlying data, not just how administrative zones have been (arbitrarily) drawn. With a sampling frame set up by administrative borders, the maps might display “strange” results. For example, (let us say in a Sahel country) 95 percent of the households live in the very southernmost part of district A and are mostly farmers; the map of the entire area of district A will indicate the importance of farming, whereas in reality, in the less densely populated northern part of district A, pastoralists are much more numerous. Therefore, it is meaningful to add other information, for instance from livelihood zoning, to the information obtained. Ideally, small area mapping techniques should be employed.

It may be useful to show rivers, lakes, roads, etc., on the map, but if the symbology is simple and obvious, it is acceptable to exclude these from the legend, for the sake of simplicity.

It is advisable to keep the legend dynamically linked to the layers as much as is possible. Often graphical necessities require us to convert the legend into graphical elements, thus losing the dynamic link with the layers. In this case it is advisable to make a copy of the dynamic legend before conversion and paste it outside of the page layout.

Range values

The message the map gives is therefore determined by the range values (generally referred to as class breaks). How these class breaks are determined is therefore very important and should be considered with care. The following is a list of considerations to be taken into account when determining the class breaks:

1. Precision or margin of error of the indicators. Differences in colour should reflect differences in the underlying data. Regions with small differences (far from significant) should be displayed in the same colour. However, since we might be interested in "tendencies," these differences between regions should not necessarily be significant at the 95 percent level; we only want to be reasonably confident that real differences are reflected[125] in the map. As a rule of thumb, the class break should not be much smaller than the margins of confidence (at 95 percent) of the estimates. For example, if the average land area per farm is estimated in the various districts with a precision of around +/- 0.25 ha, a typical class break could be: <1 ha; 1-1.5 ha; 1.5-2 ha, 2-2.5 ha; >2.5 ha.
2. Values that represent proportions (between 0 and 100 percent). The estimated prevalence of an indicator will have a higher degree of confidence for extreme values (close to 0 percent or close to 100 percent) than for average values (around 50 percent). On this basis, the ranges for the extreme values can be reduced compared to those for average values. For example, if we have 200 households per reporting domain (and a design effect of 2), good class breaks could be: 0% – 10% – 25% – 50% – 75% – 90% – 100%
3. Distribution of the values. There is the need to show an indicator on a map only if areas of the map will be coloured differently, otherwise the map will not convey a message. If the margins of error of the various estimates are so large, and the mean values for the reporting domains so similar that no values are significantly different, a map should not be included.

Typical data to be mapped

Subsequently, the following variables can be compiled and explored at the sub-national level as *intermediary* candidates/indicators for the hazard and vulnerability geospatial layers.[126]

(a) National level related to hazards and other shocks:

- Frequency of extreme weather hazards
- Probability of below-threshold WRSI and low crop productivity
- Number of households exposed to flooding, droughts, etc.
- Percentage of population working in agriculture or other main livelihood strategies
- Percentage geographically sensitive land areas
- Assessment of double exposure (e.g. climate change and international trade interaction)
- Hazard specific and composite hazard exposure

125. This is an issue of the type I vs. type II error. Are we more satisfied that the few differences displayed are certainly significant? (Even if we have hidden a lot of differences because we were not sure of them at the 95 percent level, many of the hidden ones might actually have been real.) Or is it better to show more differences, reducing the number of missed differences (this means including more untrue differences, however). For maps, which give more of an overview of the situation, it is important to show tendencies. Hence, it is justified to display differences that do not meet the 95 percent confidence criterion (but still meet the 80 percent one, for instance).

126. ODI: Poverty and Climate Change: Assessing Impacts in Developing Countries and the Initiatives of the International Community,
http://www.odi.org.uk/iedg/publications/participation_in_negotiations/lse_report.pdf.

(b) Household food security and Vulnerability
- Percentage of households with an FCS below a threshold, with CSI above a threshold
- Percentage of households vulnerable to a particular shock
- Percentage of households in lowest wealth quintile
- Percentage of households without access to water and sanitation
- Percentage of households living in marginal areas
- Percentage belonging to vulnerable livelihood groups (subsistence farmers, pastoralists, etc.)
- Malnutrition rates
- Literacy rates, gender disparity, etc.
- Composite livelihood and composite vulnerability zones

(c) Composite risk
- Composite hazard
- Composite vulnerability
- Composite risk

8.1.3.2 Other types of cartographic output

Maps showing boundaries or locations
Maps showing boundaries or locations portray accessibility or divisions of territory and are therefore related to the demography or infrastructure. For this reason it's recommended that such maps contain a populated places layer and, if necessary, an infrastructure layer showing accessibility and the impact of the boundaries or locations to the population of interest. It would also be useful to include a Hillshade backdrop.

In general, it is necessary to indicate dimensions (scale bar, scale text, projection information) and it may be necessary to show a coordinate graticule on the map. It may be important to highlight when a boundary or location appears similar or close to an existing feature yet does not necessarily coincide with it (e.g. a river and an administrative boundary).

Raster (or matrix) layers
Data in raster format must be further distinguished into two categories:
- Image layers: non-thematic rasters where the pixel value is not an indicator value; it has no meaning when removed from its neighbouring pixels since it is only part of a picture; and
- Raster (or matrix) layers: a thematic layer where each pixel value refers to an indicator value relative to that theme (e.g. elevation or NDVI).

The most common data set used in this category is elevation data (see following section on "Elevation data"). Other raster data sets commonly used are derived from vegetation or meteorological data (e.g. NDVI, RFE or WRSI). To display these, a thematic legend must be created, either as a continuous shading based on the pixel value or as single colours assigned to discrete pixel value ranges.

Elevation data
Elevation is hard to portray. The best way to give the idea of the terrain's morphology is usually with the Hillshade product, although in this format the elevation information is lost, and thus by definition it becomes an image-type raster. In CFSVA maps, since elevation values are not required, it is advisable to simply use height ranges to separate the country into two to three areas of substantially different altitudes, a probable determinant of livelihood type.

Several layers are derived from elevation data and are implicitly related to it:
- Height: elevation raster where each pixel is the elevation value;
- Slope: raster where each pixel is the vertical angle with respect to the ground;
- Aspect: raster where each pixel is the horizontal angle with respect to north;
- Hillshade: image representing shadows of hills from an imaginary light source;
- Contour lines: vector lines representing linear continuums of equal elevation; and
- Height ranges: polygon vectors created by dissolving areas defined by contour lines.

The scientific way to express elevation data is to create contour lines through interpolation of the raster data. This functionality is available through the Spatial Analysis extension in ESRI products (or Vertical Mapper, in MapInfo).

Slope and aspect layers are used in analysis (e.g. for crop production and soil erosion) and fit into the raster data category and should be represented as continuous variables.

Image Layers
Image layers should be included as is. It is generally difficult to mask values or use transparencies since the pixel values do not have a specific meaning, and because compression technologies (such as .jpg) will create small disturbances in the pixel values.

8.1.3.3 Cartographic elements
Maps are used and published in a variety of formats and media. In CFSVA, maps are generally used to show how an indicator relates to the country's geographical layout, for example, with regard to urbanization, ecological zoning, elevation, presence of water bodies, soil and other farming issues, neighbouring countries. These chloropleth maps are known as "thematic maps" because the main layer is based on an indicator value.

Unlike maps that need to show dimensions or scale, thematic maps do not require precise measurements, and therefore there is no need for a scale indicator or the cartographic projection specification. A very simple scale bar may be useful for giving the idea of the size of a country. Only a very simple north arrow is needed if the map has been rotated to fit properly into a rectangular area.

As with all maps that will be inserted into reports, the legend, title, footnotes, logos, and other graphical elements should be inserted using the graphical capabilities of MS Word. See section 8.1.3.6 on "Exporting maps to the report."

8.1.3.4 VAM Symbology

VAM symbology is under creation as a set of styles and legends for ESRI ArcMap and ArcView 3. These will include the following:

- WFP VAM map layout template; and
- WFP logo as a legend item to be used for WFP office locations.

Maps created for reports will use the same report style used in images for maps, and should use standardized WFP office locations and any other similar symbols for map features.

VAM Spatial Information Environment (VAM SIE) as a style repository

All symbology and layouts should be stored in the VAM SIE repository, with an appropriate description using the metadata indexing.

- Spatial extent should refer to the area for which the template/style is relevant; generally this will be global coverage;
- Description; and
- Author name and contact information.

Other layout elements

All layout elements should be defined in a standardized form inside the layouts and styles. As well as legend symbology, this includes (but is not limited to) the following layout features:

- Titles
- Logos
- Frames
- Scalebar
- North arrow and other references
- Disclaimers
- Metadata

8.1.3.5 Consistency

To ensure correct portrayal of the indicator information, an indicator map could contain the indicator's numerical values as well as the shading or graduated symbols. These numerical values can be displayed as a map label, or as a table next to the map or in the report's text or annex. In the case of a set of maps showing the same theme for different geographic areas, or similar themes, the class breaks of both maps should be the same as well as the colour scheme. For example, two maps illustrating enrolment rates for boys and girls should have the same class breaks with the same colours, so that the reader, when comparing the two maps, intuitively understands the graphics.

A common mistake to avoid is the creation of similar maps (e.g. for different countries) with each using its own class breaks (determined by each country's data set) and then showing all the maps together using the same colour scheme.

8.1.3.6 Exporting maps to the report

The use of desktop publishing applications (DTP) or MS Word is fundamental in order to embed or copy thematic and other maps into report documents or to create high quality posters. These applications can also be used to improve the inclusion of other key elements on the map layout (e.g. logos, images, tables, text with styles) and can prove useful for exporting into graphic formats that are not supported by the mapping application.

When using DTP applications for including maps in reports, the following guidelines should be followed:

- The map should be created in a GIS application and exported as a map image without legend, title, or other elements.
- The legend should be exported as a separate image.
- Other necessary map components should be exported separately, for example the scalebar.
- The map and other graphical components should be exported as a WMF (windows metafile) or EMF (enhanced metafile). These formats can preserve raster, vector, and font types without degrading the quality of the image.
- The exported components should be assembled in the DTP application and given a title and other text elements following the style of the report.

8.1.3.7 Exporting bitmaps

Exporting bitmaps (such as .jpg, .gif, .tif, or .bmp) is not advised, since it will make the size of the document too large. The user must keep in mind that the human eye responds differently to images printed on paper compared to images seen on the screen. Therefore, results from compression become easily visible on paper when they were not visible on the screen. Also, resolution is roughly twice as high on the screen as on paper. Considering this, the recommendations are:

- Bitmaps should be avoided except if there is a strict need, for example, if font problems are incurred using metafile formats.
- Bitmaps should be exported at 300–600 dpi for photographic-quality print (glossy publication).
- JPEG format should be used only when satellite or aerial imagery is used in a map, and should be rendered at low compression (maximum: 50 percent).
- PNG (or GIF/TIFF/BMP) should be used for discrete colours. These formats can be reduced by limiting to 256, or fewer, colour palettes (8 bit) to drastically reduce file size.

8.1.3.8 Software applications

Although there is no declared standard GIS application for mapping within VAM, the following applications are currently in use or have been used in the past:

- ESRI ArcMap ArcInfo or ArcMap ArcView, currently the application used for map publication
- ESRI ArcView, a well-known lightweight application for map publication; currently the platform for the FIRM application (Food Insecurity Risk Mapping), used for COFSU reports
- ESRI ArcInfo Workstation, used for spatial analysis but not advisable for cartographic output. ArcMap and ArcView both access ArcInfo workstation data
- MapInfo, previously used in VAM for cartographic output but currently not in use
- Idrisi, used for raster analysis but not for cartographic output
- Erdas, used for raster analysis but not for cartographic output
- VAM SIE (currently GeoNetwork open source), the VAM spatial infrastructure architecture, and therefore the standard for GIS data catalogue and repository functionality (i.e. metadata sharing and file sharing). GeoNetwork incorporates Intermap as an online mapping system (which is currently not activated) and the future version 3 will include Geoserver and Openlayers as programmable client/server architecture for online mapping.

	ESRI ArcMap	ESRI ArcView3/FIRM
Ability to save layout templates	X	
Ability to save legend files	Legends can be exported as .LYR files and imported into projects; however, this is not dynamically linked (the next time the project is opened the legend will be in the project and the .LYR file will not be consulted for changes)	Legends can be exported as .AVL files and dynamically linked to layers
Ability to define custom symbology	X	
Ability to construct different maps from single project		X

8.1.4 Annex tables

Annexes to the CFSVA should give a complete list of indicators disaggregated by important reporting domains, and possibly other levels of disaggregation based on their importance in the local context. For example:

- Administrative unit
- Food security zone
- Livelihood group
- Food security class
- Sex of the household head
- Households with disabilities
- Households headed by minors

The following is a list of key indicators to be included and analysed in a CFSVA. Descriptions of the key indicators and indices and calculated scores are presented in the annexes.

- Household size
- Age dependency ratio
- Age of household head
- Percentage of dependents
- Crowding index
- Child-headed households
- Elderly-headed households

- Marital status of household head
- Physically challenged members of household
- Sex of household head
- Age of household head
- Food stocks
- Food consumption score
- Food consumption clusters
- Food sources
- Number of meals per day (adults and children)
- Stunting
- Wasting
- Underweight
- BMI of reproductive women
- Mortality rate (global and under-5 children)
- Caring practices
- Wealth index
- Years of schooling
- Level of education
- Net enrolment rate
- Gross enrolment rate
- Literacy
- Housing construction materials
- House ownership
- Sanitation
- Water sources
- Distance from water source
- Source of lighting; cooking fuel
- Assets (household, physical)
- Household shocks
- Coping strategies
- Access to agricultural land
- Area of agricultural land
- Major crops cultivated
- Agricultural production
- Ranked main and secondary income activities
- Contribution from different income strategies to household income
- Livelihood groups
- Household expenditures and access to credit

8.2 PREPARING AND PRESENTING A REPORT

8.2.1 Before starting the preparation of the report

In the majority of cases, the CFSVA process (as seen in Chapter 2 – Managing the implementation of a CFSVA), induces collaboration with many partners and colleagues. This is why it is essential during the first stage of preparation of the survey (and ideally in the terms of reference of the study approved by all partners) to include clear responsibilities about the writing of the report. Different cases can be observed:

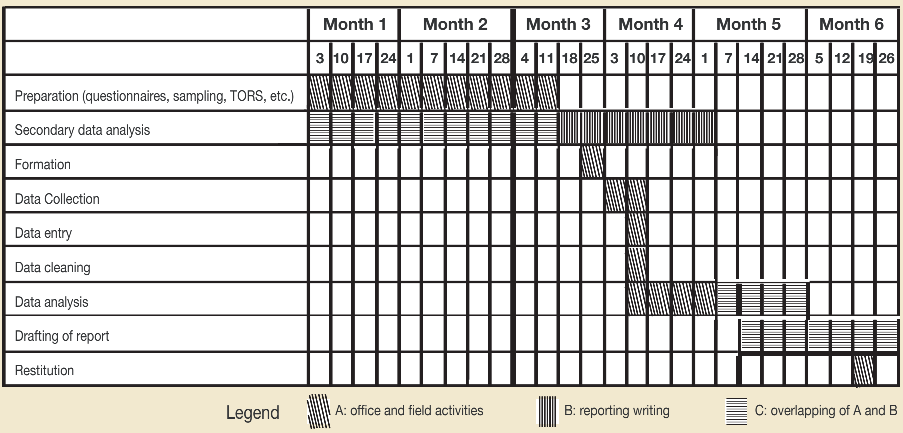
- The report is written by a single person: In this case, it could be the VAM officer or a consultant hired for the study. She/he will take care of all the sections and share it with the partners/colleagues in due time.
- The report is written by many, and then parts are combined to form the final draft. This option requires a more complex and well defined set of responsibilities.

It is also essential from the beginning to identify the person, group, or organization who will be responsible for the final clearance of the document. This is country-specific and depends on who is involved in the CFSVA, and how. Usually, the WFP country director or a ministry (or both) have the final responsibility but the role of the regional bureau (RB) or Headquarters in supplying support and guaranteeing quality is equally important. Once again, it is crucial that these responsibilities are clearly defined from the very start of the CFSVA.

8.2.2 Preparing a draft summary report

While it takes a great deal of time to complete a CFSVA, timeliness of reporting is a key factor. The CFSVA team/writer should attempt to produce reports within two to three months of the completion of data collection. This means that the data analysis and the writing of the report are often executed concurrently. It also means that most of the secondary data analysis and reporting is completed before the drafting of the results of primary data. The timeframe for writing the report could be as shown in Figure 8.1. (Blue represents the conducting of various office and field activities; writing of the report is represented in red.)[127]

Figure 8.1: Example of a timeframe for report writing



127. The example here is based on the fact that a CFSVA can take from four to eight months to complete.

The writing of the report can start as early as the preparation phase, when the objectives, sampling, and methodology of the study are clearly defined. The secondary data analysis can then be added. This first draft will be a good base for the insertion of the data analysis results. The merging of the two and the drafting of the rest of the report can then be completed. The drafting of the report continues after the restitution phase, as any final comments or last-minute contributions are inserted.

8.2.3 Sharing and presenting the draft report

Once the draft report is prepared, including results from the secondary data analysis and the primary data, it should be shared with all partners involved in the CFSVA (this first draft might not yet include recommendations). Normally, the draft report is first shared internally with partners, WFP colleagues at CO, RB, and Headquarters and among more "technical" experts from ministries and NGOs. Their comments can then be included in the draft report before it is shared more widely to less technical actors such as programme officers. At any time during this stage, the report should always clearly indicate that it is a draft, to avoid any confusion. It helps to put the date of the last revision on the document and to regularly update partners on the review process.

The review process needs to follow a clear and established timeline, and when the report is shared for the first time, a reasonable deadline for receipt of comments has to be made. The clearer the review process, the easier it will be for all reviewers and for the writer(s) to finalize the document. An example of a review process is described in the box on the right; the process varies by country. It cannot be emphasized enough that clear and regular communication between all actors at this stage is essential.

Review of draft report in Mali

For the CFSVA in Mali, the CO organized a two-day retreat, to gather all comments on the draft report, with all main actors (ministries, United Nations, NGOs, RB, and Headquarters) involved in the drafting of the report. This procedure allowed all actors to discuss issues, agree on comments and deadlines, and determine who would incorporate the comments. It also saved time by bringing all actors together.

Following the presentation of the draft report and the insertion and discussion of the comments received, the draft is ready to be cleared by the designated organizations and can be called final.

8.2.4 Preparing the final report

The final report of the CFSVA is the result of a long process involving many partners and a great amount of data. It should be the product of compromises and final agreement among partners (especially the government) but should be based on sound analysis and present the numbers as they appear. WFP often takes the lead in the survey and uses the final results for its PRROs and country programmes. As such, the final report should be as transparent as possible and all data and the methodology should be made available to readers. Do not hesitate to refer to other finalized CFSVAs.

Despite the amount of information used in the CFSVA, the report should remain concise and clear. The editing and formatting of the document should, if possible, be done professionally,[128] and the document should be translated into the local language when needed (sometimes just the executive summary and the recommendations need to be translated). If the document is written in one of the official languages of the United Nations[129] other than English, the summary or executive brief should be translated and distributed in English for certain audiences (such as donors).

Avoid language that could be ambiguous or misunderstood. Avoid jargon and the excessive use of acronyms.

Box 8.1: Common WFP acronyms and phrases (jargon) to avoid, or to explain, when communicating CFSVA findings to general audiences

- EMOP, PRRO, SO, VAM
- Asset depletion
- Vulnerable group
- Food security/access/availability/utilization
- Coping strategies
- Livelihood group/wealth index
- Food basket
- Wet ration, dry feeding/ration
- Difference between acute and severe malnutrition
- Stunting and wasting
- Supplementary and therapeutic feeding

After allowing time for all concerned to review and comment on the draft, finalize the report and send it to all concerned government entities, United Nations agencies, NGOs, donors, the WFP regional bureau, and WFP Headquarters.

8.2.5 Essential elements of a final CFSVA report

The detailed standard outline for a CFSVA is presented in Annex 7, but the following general elements should guide the writer in drafting the report:

- Essential in a CFSVA report is an Executive Summary and/or CFSVA Executive Brief, with substantive findings, graphical overview, and information for programme design for any recommended response (see 8.4.1). Often, the Executive Summary and the CFSVA Executive Brief are one and the same.
- Analysis with key results of statistical and spatial analysis will form the main body of the document. Include individual chapter summaries. Short summary presentations are a good way to organize the findings and present them in a coherent way for a larger audience.
- Use harmonized line graphs, scatter-plots, and bar charts to illustrate trends or changes in an easy-to-read fashion.
- Provide contextual information that influences the status of food and livelihood security (through secondary data analysis).
- Keep discussions about less significant findings brief. In some cases this information can be made into a footnote or placed in an annex.

128. It is important, for example, to keep figures, graphs, and maps legible and harmonized.
129. Arabic, Chinese, English, French, Russian, and Spanish.

- Place information about statistical tests, such as the type of test performed, the significance value (p-value) of the test, standard errors, and other technical details in footnotes, endnotes, or annexes.
- Provide annexes that contain the lists of frequencies and descriptive statistics, methods, figures, maps, and questionnaires (these should be put on a separate CD).
- The final document should include all logos or reference to the donors and contributors to the study (an acknowledgments section) and standard WFP credit section for the report and pictures (see annexes and other CFSVAs).

8.2.6 Recommendations

The recommendations of a CFSVA are of crucial importance, as they will be what WFP decision-makers and partners will use to design, adjust, and implement their food security programmes, and what other assessments will first look at when an emergency occurs.

Depending on the number of partners involved and the timing of the survey, the recommendations should be prepared with the widest audience possible, including experts in all sectors touched upon by the CFSVA (education, health, food security, migration, agriculture, etc.). Ideally, the partners should have carefully read the CFSVA and been invited to contribute to the drafting of the recommendations. Usually a culling of the results is organized and recommendations are agreed upon during a one- to two-day meeting. It is important that recommendations be presented in a clear and concise manner indicating: (i) the cause for the intervention, (ii) the type of intervention proposed, (iii) the primary target group(s) and the agencies/ministry in charge of the response (see example here). The recommendations can also be divided into different sections for more clarity, separating them by main sector (health, education, policy, etc.).

Example of recommendation matrix for a CFSVA

Cause of the intervention	Type of intervention	Primary target groups	Ministries/agencies
There are still many food-insecure villages without access to roads. This limits farmers' ability to sell and purchase commodities at markets.	Farm to market road construction through FFW, if food not consumed.	Remote food-insecure villages without road access.	Ministry of Labour and Social Welfare, WFP.

8.2.7 Evaluating the quality of the final report

Prior to the report's circulation, a thorough evaluation of the content is advised. Annex 7 provides a checklist for evaluating the quality of a CFSVA report.

8.3 HOW TO COMMUNICATE THE RESULTS OF A CFSVA

Sharing information is an essential component of humanitarian response. Food security analysis serves its purpose only if it leads to more efficient, appropriate, and targeted responses to a food crisis. It is intended to enable WFP managers to examine response options and recommendations, for partners to consider interventions in other complementary sectors, and for donors to allocate resources.

It is therefore crucial to ensure that key food security and vulnerability analysis findings, and food and non-food response options, are communicated in an effective way to the wider humanitarian community, in country as well as in the donor capitals.

WFP has a responsibility to communicate CFSVA findings to all stakeholders, humanitarian actors, and interested parties. This is in line with WFP's commitment to transparency. It also strengthens the links between assessment recommendations and decisions on programme design and funding allocations. Sharing information will also foster dialogue and contribute to building a common understanding of a problem.

Beyond being a programming tool, the CFSVA constitutes **an information base**. A CFSVA provides an in-depth understanding of food insecurity faced by communities and threatening livelihoods. It also provides data on the characteristics of the households (demographics, asset ownership, sources of income and labour, access to infrastructures such as markets and health posts, their shocks and coping strategies, etc.), their level of susceptibility to various risk factors (droughts, floods, pests, insecurity, etc.), and their ability to cope. The breadth of data and analysis can be used outside the food security sector, making it essential to share the information as widely as possible. In addition, in an emergency this information should be easily accessible, as it is a key reference for assessing the effect of a new crisis.

Communicating the results of a CFSVA is an important step, despite being the last, of the CFSVA process. Attention must be paid to this activity, despite fatigue at the end of the process. Without effective communication, it is unlikely that the data and analysis will be used outside WFP.

Remember, CFSVA findings are used for multiple purposes and should be communicated to a range of audiences, including the following:

WFP audience

- Managers, who require reliable and transparent information to make sound decisions about the scale and scope of a food insecurity and vulnerability;
- Programmers, who rely on CFSVA reports for designing interventions that are appropriate and operationally feasible; and
- Staff at all levels, who need timely and accurate information that can be communicated to government, donors, and other humanitarian actors for programming, resource mobilization, and advocacy purposes.

The wider humanitarian community

- The host government, which uses the results for its own programming and policy-making;
- United Nations partners, which will need to know about the food security situation of a given country in order to design and implement their own operations and for coordination purposes;
- NGO partners, which will use WFP findings to complement their own analysis and design and to implement their own interventions;
- Implementing partners, who will need to understand the food security situation to implement interventions in the most efficient and transparent manner;
- Donors, who need information on the magnitude and severity of a crisis in order to make funding allocation decisions; and
- Media, in order to report on the situation using balanced and accurate information.

The following sections discuss the channels and communication tools through which the findings of the CFSVA can be disseminated.

8.3.1 CFSVA executive brief

In addition to the main report, the CFSVA team leader should prepare an executive brief which presents the key findings in a concise and easily accessible way.

In theory the executive brief could be different from the executive summary. The executive summary summarizes the findings of the report. The executive brief provides information tailored to fit the information needs of the audience, information that would be useful for decision-makers. However, it is possible to combine the executive summary and executive brief into only one document. This brief should be distributed separately from the report.

Audience: It is intended for humanitarian decision-makers (WFP, donors, government representatives, United Nations partners) who are not interested in the technical aspects of the assessment and who need to have a clear overview of the food security situation and the appropriate responses. It is also intended for local authorities.

These actors need to know the following:
- Who is food insecure, or vulnerable to food insecurity?
- How many people are food insecure, or vulnerable?
- Where are they?
- Why are they food insecure or vulnerable?
- What is the likelihood that the situation could deteriorate?

The executive brief should present the various recommended responses, and their objective, content, target group, and duration.

When to prepare it?
The team leader should start preparing the executive brief when the findings, main conclusions, and recommendations are available and internally agreed upon. It should be a stand-alone document.

When to use it?
The brief can be used to start disseminating the information before the final report is available. For example, it can be used at a debriefing with partners. However, the brief must first be cleared by the country director.

Format (See format and guidance in Annex 7):
It should be ideally two to three pages. It should contain legible visual instruments, such as a map, tables, charts, or graphics to facilitate the reader's understanding. Names and e-mail addresses of the people responsible for the assessment, who can be contacted for additional information, need to be provided.

8.3.2 Official launch

For every CFSVA, an official launch should be organized to present the results and recommendations to the government, partners, and donors at a high administrative and political level. The presentation should be viewed as an opportunity to ensure a wide dissemination of the results and high visibility for donors. It is advised to work with the public information officer and WFP's Communications and Public Policy Strategy Division at Headquarters to invite journalists and prepare a press release.

CFSVAs typically include recommendations for both food and non-food responses. For WFP, the presentation to stakeholders is an opportunity to bring the non-food recommendations to the attention of partners so that they can consider them for their own decisions and planning.

The launch must be organized in advance to ensure good attendance. Both technical staff and managers or decision-makers from the following audience should be invited:
- United Nations partner agencies
- Government representatives from the relevant ministries
- Local and international NGOs
- Implementing partners
- Donors

8.3.3 Disseminating by e-mail and hard copy

The country director is responsible for disseminating the report. Once finalized, the report and the executive brief should be e-mailed to all relevant stakeholders, both inside WFP and outside WFP. This should be done in a systematic way. A proper list of recipients in the country but also at the regional level should be drawn up (in the case of WFP, the support of Headquarters and the regional bureau should also be sought).

Mailing or delivering hard copies should also be considered, as some decision-makers and managers will not have the time to open their e-mails, download the report, and print it out, and therefore may find it more convenient to receive a printed version. The official launch could be a good opportunity for disseminating hard copies to all partners.

8.3.4 Media work

The media are another important channel for disseminating information. It is important to ensure that the Public Information Officer (in the country office or regional bureau) is aware of the CFSVA from the start; he/she should be informed of the analysis progress, findings, and recommendations. The Public Information Officer acts as a spokesperson and deals with media inquiries and may have to answer journalists' questions on the survey, so it is important to keep him/her in the loop.

Public Information Officers have various tools or methods in communicating information: including press releases, media advisories, news briefings, press conferences, web stories.

To inform the media on the findings, a press release is likely to be the most appropriate tool. The purpose of a press release is to issue a straightforward message. It is for the country director to decide if a press release will be issued. Various factors will be taken into account when deciding to issue a press release, such as the importance of the findings, whether they are newsworthy, the political context, and the fund-raising needs.

The press release can be issued internationally, regionally, or in the country only (in which case, it will be a local press release).

The Public Information Officer is responsible for preparing the press release. Press releases are written following specific guidelines (e.g. with the most newsworthy information at the top), using non-technical language, and are tailored to fit journalists' needs.

8.3.5 Posting on websites

Once cleared by the country director, the report should be sent to the Food Security Analysis Service at WFP Headquarters, for posting to the WFP internal and public websites.

Posting on other websites should also be considered:
- Reliefweb, run by OCHA, disseminates humanitarian information by updating its website around the clock. Relevant reports can be e-mailed to submit@reliefweb.int.
- United Nations country team websites often include a list of documents and reports produced by United Nations agencies.
- Humanitarian Information Centres, managed by OCHA, when present in the country, are a good place to post reports.[130]

130. The Humanitarian Information Centre (HIC) is a service endorsed by the Inter-Agency Standing Committee (IASC). On behalf of the IASC, OCHA acts as steward of HICs, which aim to ensure that individuals and organizations involved in humanitarian operations have access to the benefits of information management tools to assess, plan, implement, and monitor humanitarian assistance. HICs are an inter-organizational resource whose products and services are available to the entire humanitarian community

8.3.6 Advocacy

Advocacy is a coherent strategy to influence policy formation, reform, and implementation. In the context of CFSVA, advocacy ranges from working with governments to inform understanding of food security conditions and recommending ways to alleviate poverty and hunger, to lobbying to donors, or raising general public awareness (WFP 2004).
Advocacy should address both food and non-food options for assisting needy populations. This is because food and nutrition insecurity is multifaceted, and in addition to food assistance, interventions may be required in such areas as education, hygiene promotion, water and sanitation, infrastructure development, and agriculture.

WFP may not be able to implement some of the response recommendations, and must therefore encourage the government or other organizations to do so. The presentation meeting is the starting point for discussions with possible collaborating partners. This should be followed up by more specific meetings with interested agencies. The purpose of such meetings is to:
- provide a detailed explanation of the reasons why particular interventions have been suggested;
- emphasize the complementary nature of different interventions, food and non-food; and
- seek commitments from other organizations regarding intervention in specific sectors.

It is usually necessary to organize a series of such meetings.

8.4 KEY REFERENCES: REPORT PREPARATION AND DISSEMINATION

- Caldwell, Richard. 2006. Using Tables and Graphs for Presenting Results: A session provided to the ADRA Monitoring and Evaluation Training. Bangkok, Thailand.
- FEWS NET/WFP. 2006. Rwanda Food Security Update: Poor Households Food Stock Already Gone. September.
- WFP. 2006. *Comprehensive Food Security and Vulnerability Analysis: An External Review of WFP Guidance and Practice.*
- ibid. 2007. CFSVA Methodology Workshop Report. Rome. April.

CHAPTER 9

Archiving

9.1 VAM-SIE

To facilitate data exchange and information-sharing, VAM has already developed and installed in all the regional bureaux, as well as in several countries, the Spatial Information Environment (SIE). The main goal of VAM SIE is to create a standardized but decentralized spatial information management environment that will enable WFP country offices and regional bureaux to access geo-referenced food security databases and cartographic products from a variety of sources. SIE includes tools and protocols for creating standards that could be employed by various users. This will support and strengthen the appropriate use of spatial information and facilitate the collaborative efforts to increase accessibility to relevant food security and vulnerability information.

All maps and data produced should be documented in VAM SIE. All maps should be published in an image format, and all data should be made accessible to the VAM team.

The metadata should describe the following:
- source of the data
- the reference year
- the spatial extent
- the data description
- the projection
- the precision (scale)

Vector data should be exported preferably into a shapefile format with projection information (PRJ file). If symbology is defined, share this as a LYR or AVL file.

Raster data should be exported preferably in a TIFF format, with projection information in a PRJ file and coordinates in a TFW (world) file. Again, if a legend is created, include it as a LYR or AVL file.

Tabular data should be included as an MS Access database or a DBF or CSV file, rather than MS Excel. An original-source data file or reference should be included, if possible.

With the SIE platform, practical applications that allow analysis of food security and monitoring trends, providing regular updates to guide humanitarian responses, will be developed, and data sharing mechanisms will be put in place.

Index

D

E

F

G

H

I

J
K

L

M

Q

R

S

T

U

Resources
on the Food Security Analysis Guidance DVD

The following documents, reference materials, handbooks and other resources are available on the Food Security Analysis Guidance DVD.

- **EFSA Handbooks (2009 & 2005 editions)** in English, French and Spanish
- **CFSVA Guidelines & Annexes**
- **VAM Thematic Guidelines** (some also available in French)
  - Initial assessments
  - Urban assessments
  - Market analysis tools
  - Nutrition
  - Livelihoods
  - Food consumption
  - Estimating population numbers & sampling
  - Qualitative analysis
  - HIV/AIDS
  - Participation and gender
  - National response capacity
  - Other
- **UN/Inter-Agency Assessment Guidelines**
  - WFP/UNHCR Joint Assessment Guidelines, March 2009
  - FAO/WFP Joint Guidelines for Crop and Food Security Assessment Missions (CFSAMs), January 2009
  - IASC Initial Rapid Assessment (IRA) Guidance notes
  - CAP Needs Analysis Framework, April 2005
- **Inter-Agency & Partners Food Security Analysis Guidelines**
- **Humanitarian Principles & Protection**
- **WFP Contingency Planning Guidelines**
- **WFP tools for EFSAs, CFSVAs & Inter-Agency Assessments**
  - Example terms of reference (TORs)
  - Example checklists & questionnaires
  - Example executive briefs & template
  - Example reports
  - Quality Monitoring Checklist templates

Field notes

Field notes

Field notes

Field notes

Credits

Photographs
Cover, from left to right:
WFP / Tom Greenwood
WFP / Peter Smerdon
WFP / Simon Derijke
Back cover:
WFP / Evelyn Espaillat
Page 21: WFP / Antonia Paradela
Page 37: WFP / Richard Lee
Page 47: WFP / Diego Fernandez
Page 63: WFP / Caroline Chaumont
Page 179: WFP / Stephen Wong
Page 207: WFP / Naho Asai
Page 301: WFP / Tom Greenwood
Page 315: WFP / Anne Poulsen
Page 337: WFP / Elmer Martinez

Editor
Jenna Dolan

Proof-reader
Valerie Guidi

Graphic design
Energylink di Fausta Masini

Printer
Arti Grafiche Tilligraf srl

Production
Caroline Chaumont
Mark Bennett

Comprehensive Food Security & Vulnerability Analysis Guidelines

January 2009 • first edition

World Food Programme
Via Cesare Giulio Viola, 68-70
Parco de' Medici - 00148 Rome (Italy)
www.wfp.org

WFP
World Food Programme

vam
wfp food security analysis