1. Objective

This document is a summary of the *technical brief on Secondary Data Review (SDR)*. The following pages provide a quick overview of how to undertake a SDR for needs assessment purposes and which issues should be taken into account. The objective of this document is to guide practitioners while undertaking a SDR in the immediate aftermath of a sudden onset disaster. To gain an in-depth understanding of the review of secondary data, please refer to the full technical brief.

2. The purpose of secondary data review

SDR is one of several existing methods for obtaining information during a needs assessment. Secondary data[1] plays a crucial role within assessments in emergencies, specifically in phase I and II when collecting data and information produced from outside the field assessment is necessary to provide a baseline with which to compare primary data. SDR complements and benefits the primary data collection.

The purpose of a SDR is:

- *To form a clearer, more detailed and up-to-date analysis of the situation at local level prior to the crisis. It will provide the background information about the affected area, groups of interest, risks and vulnerabilities as well as sectoral pre-disaster information*

- *To support the identification of what and where the problems might be (will provide a description of the character and plausible explanations of the nature and causes of the disaster impact as well as the related secondary threats)*

- *To provide a baseline with which to compare your primary data collection results.*

- *Identify information gaps and determine the most appropriate method to access this information (e.g. identify if a rapid assessment is necessary, information needs, etc.)*

- *To design subsequent primary data collection[2] phase and identify which sites to visit. This will provide additional information that can be used for sub-dividing the area into relatively homogeneous zones (food economy, rural/urban, coastal/mountainous, IDPs/Non IDPs, etc.)*

---

[1] Secondary data is data that has been collected, collated and analysed by other agencies, institution or bodies.
[2] Primary data is data collected directly through first-hand experience, for instance through the use of surveys, meetings, focus group discussions, interviews or other methods that involve direct contact with the respondents.
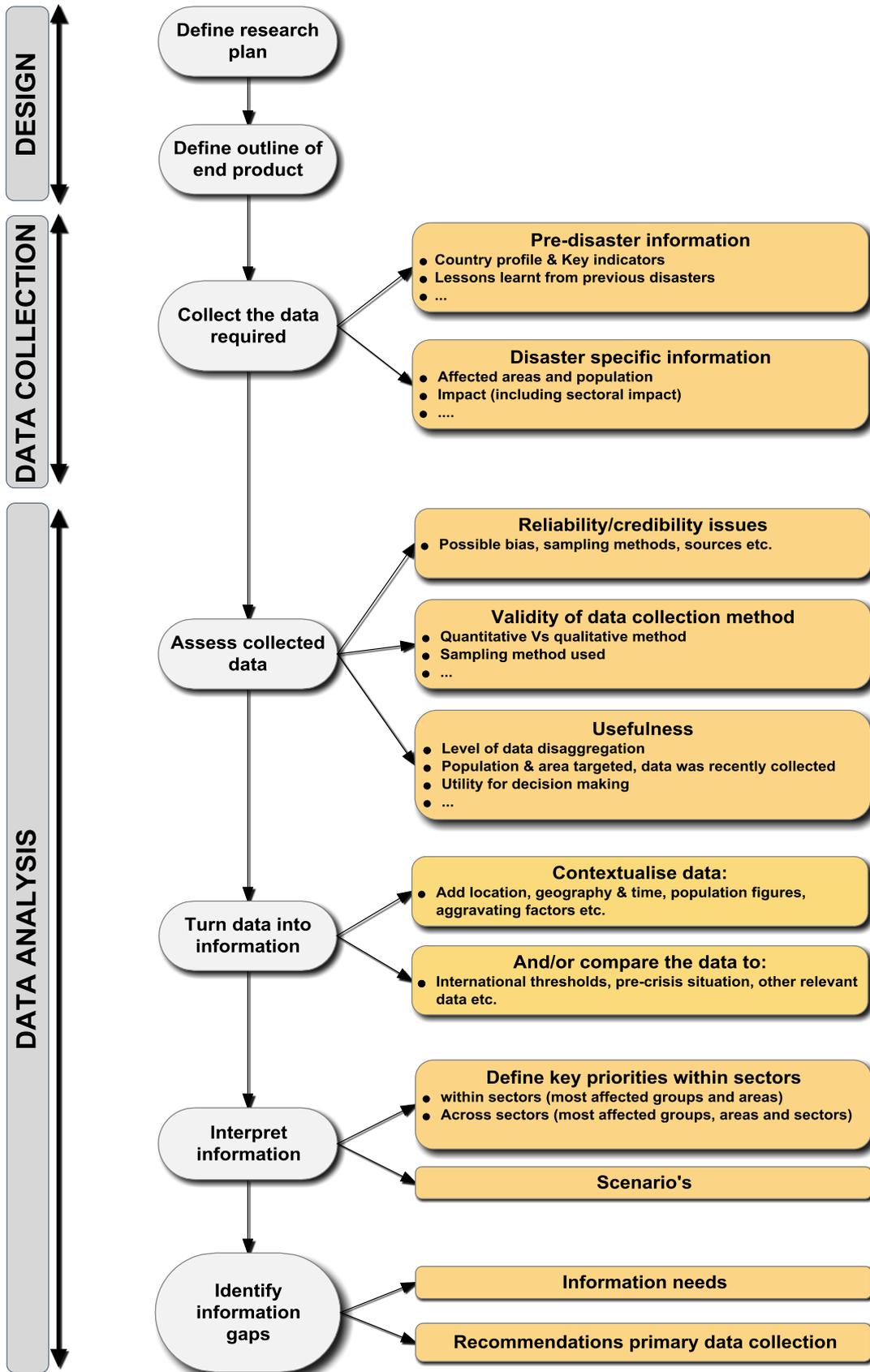
3. General principles

<u>Bringing together the right people at the right place:</u>

- The process of secondary data collection and analysis must involve people with different perspectives and competencies. Data collection may be undertaken by people with limited research training if the information needs, the process and the communication channels are well structured, while data analysis will require more experienced people.

- The three core competencies of staff required for secondary data analysis are people with **sectoral skills**, people with general **emergency programming skills and responsibilities** and people with **good local knowledge** of the geographic areas being discussed. As conducting a SDR is time consuming, make sure you have enough **dedicated resources** to capture and analyse the large volume of available data after humanitarian crises.

- If dedicated resources are not available at country level, SDR can and should **be undertaken remotely** by experienced people, ensuring a strong linkage between field offices and the SDR back up team. **Do not waste field resources on tasks that can be undertaken elsewhere, for instance on HQ or regional level.**

- Although analysis takes place throughout the data collection, it is important to integrate individuals in the final analysis who did not participate directly in the data collection to ensure a fresh perspective.

<u>Providing the right information at the right time:</u>

- **Timeliness:** provide information and analysis in time to inform key decisions about response (e.g. a Flash Appeal) and subsequent primary data collection.

- **Adequacy:** use information that is "good enough" - do not seek more detail or precision than needed

- **Relevance:** provide the information and analysis most relevant to decisions which have to be made.

- **Coverage:** collect data which is adequate to the scale of the problem.

- **Transparency:** be explicit about the assumptions made, methods used and information relied on to reach conclusions, as well as about the limits of accuracy of the data used.

- **Objectivity:** use a variety of sources when collecting and analyzing information so as to provide varied and balanced perspectives for addressing problems and recommending solutions.

4. Step by step approach for undertaking a secondary data review

DESIGN

- Define research plan
- Define outline of end product

DATA COLLECTION

- Collect the data required

Pre-disaster information
- Country profile & Key indicators
- Lessons learnt from previous disasters
- ...

Disaster specific information
- Affected areas and population
- Impact (including sectoral impact)
- ....

DATA ANALYSIS

- Assess collected data

Reliability/credibility issues
- Possible bias, sampling methods, sources etc.

Validity of data collection method
- Quantitative Vs qualitative method
- Sampling method used
- ...

Usefulness
- Level of data disaggregation
- Population & area targeted, data was recently collected
- Utility for decision making
- ...

- Turn data into information

Contextualise data:
- Add location, geography & time, population figures, aggravating factors etc.

And/or compare the data to:
- International thresholds, pre-crisis situation, other relevant data etc.

- Interpret information

Define key priorities within sectors
- within sectors (most affected groups and areas)
- Across sectors (most affected groups, areas and sectors)

Scenario's

- Identify information gaps

Information needs

Recommendations primary data collection

5. Secondary data collection

What information to look for?

Prior to the data collection, ensure the objectives of the secondary data review, the areas of interest (research plan) and the outline of your end product are clearly defined. Annex I provides an overview of the key information you need to look for during SDR for needs assessment purposes.

Proceed from general to specific and more disaggregated data. The following examples are the most common areas of focus while undertaking SDR for needs assessment purposes:

Focus	Content
Pre-post crisis	Pre crisis Vs post crisis data
Geographical	National key indicators Vs "affected area" key indicators
Group	Total population Vs specific sub-groups demographic data (refugees Vs residents)
Livelihood	Characteristic of different sub-set of socio-economic profiles (farmers Vs pastoralists)
Vulnerability	Characteristics of different vulnerable groups (disabled, food insecure, unemployed, etc.)
Catchment area	Characteristic of different livelihood zones (urban Vs rural, mountainous Vs riverine)
Gender and age	Characteristics of different categories of the population (Women Vs men, elders Vs youth)
Sector	Characteristics of different sectors (WASH, Health, Food security, etc)

Use proxy information when data is not available or too old to be relevant (e.g. coping mechanisms as a proxy to the severity of the crisis).

Refer to similar recent crises (same area, country or region) and likely impact to guide your data collection (See *Disaster Summary Sheets*[3])

How to proceed?

- Collect data first at **national level** (in depth reports available on the web or at country level) and afterwards look for disaggregated data for the population group or affected geographic area

- **Tag the data** collected according to the area of interest (pre or post crisis data; group, area and sector concerned; capacity, risk, need or response related information, etc) or the related heading within your final report to facilitate the information flow throughout the process.

- **Proceed to regular recap** of data collected so far and identify information gaps and "Known unknowns" that may guide further data collection. Updates on post crisis situation may also reveal new groups or geographical area of concern that will require additional area of research.

- Look for **important and relevant quantitative information** such as census, humanitarian profile, pre-disaster data sets, health statistics, demographic data, etc.

---

[3] Disaster Summary Sheets (DSS) provide a general profile of the potential impact of a disaster, based on lessons learned from previous similar disasters. 3 types of DSS (Earthquakes, Floods and Tropical Cyclones) are available.

Statistics may provide useful indications on the evolution of exogenous factors, patterns and trends.

- Get population data to lowest administrative level possible

- **Use snowball effects:** Use the references generally placed at the end of collected reports and documents to guide to more in depth research.

- **Use/build your information network:** Identify key resources (at local, national, regional and HQ level) that can support and contribute to the data collection. Reciprocity is of key importance; exchange information by providing short updates to your network regularly. Quote people/sources in your report (if not sensitive). When searching for secondary data or questioning the quality of a source that you have already collected, seek advice from sector specialists and other experts with local knowledge. For local level information and data, NGOs or local contacts might also have small libraries that provide additional information or local contact that can facilitate information and relevant data.

- **Customize your archiving procedures**: Standardized architecture should be used while archiving collected data in order to ensure easy retrieval of documentation or easy incorporation of new data collectors. For post disaster information, ensure data is stored in a way that simplifies daily updates of humanitarian profile/caseload and allows for visualization of trends (e.g. number of affected population, missing, injured, IDPs, etc.). Each document should be renamed as follow: Date/source/name of the document (e.g. 2008 WFP Food Security Assessment Armenia)

Where to find it?

Pre-disaster information	Disaster specific information
- National institutions (Ministries, research institutes, Universities, etc) - Large Survey (DHS, MICS, Census, etc..) - International development institutions (i.e. World bank) - Sector fact sheets, e.g. WHO country epidemiological profile - Common operational datasets (COD) - UN, Local and international NGOs survey reports - UN global data sets or Country portals - Geospatial data - Online databases (i.e. EM-DAT, prevention web) - Previous Flash appeals, CAP - ALNAP, evaluation reports, After Action reviews - DevInfo, world development indicators, MDGs	- National institutions (Ministries, LEMA, etc..) - Media reports - Assessment reports from local and international NGOs - Funding Appeals - Situation reports (OCHA, clusters, Gvt) - Humanitarian profile (CODs) - Geospatial data from UNISAT, Google Earth etc. - Satellite imagery, UNISAT or Private providers - Social media

Key principles for secondary data collection:

- *The more disaggregated the data, the more useful it is for identifying the most vulnerable people.*

- ***Importance of the data vs. the time needed to find it.*** *Some of the required data will not exist or will be difficult to find. Decide whether the importance of the data justifies the time required to find the data.*

- *Collect only what you know you can use. Know the question you are trying to answer*

and the data you are looking for.

- *Provide a clear timeframe for data collection and identify priorities. Ensure everyone is aware and regularly updated about groups and geographical areas of concerns.*

- *Let the data speak to you. Be prepared for redirecting your collection efforts accordingly.*

6. Data analysis

Assess your data:

- Evaluate the **reliability and the credibility** of the data (level of bias, source's credentials, data collection method, confidence intervals, etc.) and **validity** of data collection methods used. Also evaluate the **usefulness** of the data (information sufficiently recent and relevant to the SDR, level of disaggregation, etc.)

- Be prepared for the following commonly encountered problems while assessing secondary data:
  - Data is outdated, numbers are provided without specifying how (or when) it was collected, data is usually at national and at best provincial level, data is rarely disaggregated, and data may be season dependant.
  - **Inconsistent** information. Try to verify important information by comparing inputs from at least three different sources **(triangulation)**. Decide whether the **inconsistencies** will affect the assessment conclusions. Resolve important issues by estimating the confidence you have in each of the sources.

What to look for?

- Identify the **factors aggravating** the impact of the current crisis (existing vulnerabilities, exhausted coping mechanisms, exposure level of livelihoods to the hazard/shock, etc.). Pre-disaster information will provide very useful insights on how the crisis may have affected the livelihoods, systems and infrastructures in the considered area.

- **Identify the humanitarian profile of the crisis** (Overall number of affected population, dead, missing, injured, etc.). Define, quantify (best estimates) and categorize groups, sectors and areas affected directly and/or indirectly as a result of the crisis.

How to proceed?

- **Involve experts!** These can be individuals with knowledge of the local context sector specialist etc.

- **Proceed to a sectoral analysis** (life-saving sectors to start with) before to combine and consolidate findings into a cross sectoral analysis section.

- Look outside the collected data, **contextualise it**. Compare situation "before" and "after", compare to international standards/thresholds or to other relevant data (population figures, geography, time, etc). Use experience and lessons learnt from past similar situations to identify risks and likely evolution of the crisis.

- **Cross-analyse key data** and use additional information sources to understand or make reasonable sound inferences about unmeasured conditions or situations; this allows to

better understand not only what is happening and where it is happening but also why it is happening.

- Look at **what difference exists** between groups, sub-groups, sectors and places. Proceed through a "more or less" type of analysis, by using the following key questions: *Who are the most affected groups, what are the most affected areas, what are the sectors requiring immediate interventions, what are the key issues.* Prioritize areas, groups and interventions.

- Make a clear difference between the impact related to the crisis and pre-existing vulnerabilities that may exacerbate the impact.

- Identify **constraints**, **information gaps** and needs for further assessment phases. Always ask: *What's missing?*

- To overcome the 'known unknowns', use **assumptions, judgement** and **"educated guesses"**.

- **Articulate results.** Translate conclusions into easily understandable results. Focus on value added for target audience.

Key principles for secondary data analysis:

- ***Scrutinize information*** *and identify the underlying details of important facts, patterns, trends, significant differences or anomalies that are not always readily visible. Consider the details.*

- *Separate the matter into key parts and/or essential elements; break things down; identify causes/key factors or features/possible results.*

- ***Ensure there is enough time to turn data into information****. Often a great deal of time is spent collecting information, but too little time given to preparing for data collection, or analysing it.*

- ***Challenge your own assumptions and conclusions.*** *Discuss your findings with your colleagues and reach consensus on conclusions.*

- *Consider bias and reliability/credibility. Don't rely on one source only.*

- ***Be sceptical when dealing with comparisons.*** *Researchers like to do something called a "regression," a process that compares one thing to another to see if they are statistically related. They will call such a relationship a "correlation." Always remember that a correlation DOES NOT mean causation.*

- ***Be careful of the actual meaning of terms used.*** *Terms such as 'affected', 'household', or 'community' can mean different things in different areas. Definitions may change over time and where this is not recognised, erroneous conclusions may be drawn. Provide a definition for potentially confusing or sensitive terms.*

- *If you use **technical terms**, make sure you define them correctly. E.g., specify which type of malnutrition you are referring to (stunting, wasting, etc.)*

- ***Ensure the secondary data review is properly referenced.*** *A well-documented secondary data review and analysis allows for easier use of the material by other interested parties and allows for greater credibility of the product.*

- *Clearly define when information is based on assumptions instead of on facts or sufficiently verified information.*

- *Think about whether or not your findings make sense (Does it fit in with the history and context? Does it make sense to the people living there? Etc.)*

ANNEX I: What kind of information is needed?

CONTEXT AND VULNERABILITY (NATIONAL AND AFFECTED AREA LEVEL)

General information	• Outcome Indicators of:
	  ✓ Crude Mortality Rate as Deaths /10,000/day
	  ✓ < 5 Mortality Rate as Deaths /10,000/day
	  ✓ Maternal Mortality rate as Deaths/100000 live births
	  ✓ Malnutrition prevalence (6-59 months) as % <-2 Z-Scores below mean Weight for Height reference population
	• CODs: Population size and spatial distribution by administrative unit and locality (Census, Population projections), administrative boundaries, etc..
	• Age-sex structure of the population (Census, Population projections)
	• Socioeconomic and cultural characteristics of the population (literacy, economy, ethnic group, language, religion, minorities, marginalized).
	• Livelihood profile, income levels and basic indicators on access to basic services and goods
	• Administrative maps
	• Reference coordination bodies (ministries, NGOs, Cluster) and contact list
	• DRR, disaster management activities and bodies, contingency plans
	• Humanitarian system in place (clusters, coordination bodies, etc...) before the crisis, 4w's
	• Security context

	Food security	Health & Nutrition	WASH	Shelter & NFIs	Access / security
Sectoral information	• Relevant sectoral indicators, fact sheets, maps • Livelihood strategies of population groups. Source of income and expenditure profile • Available resource materials about agro-ecological and agro-economic context. Crop Calendar • Food supply and food security situation of the country • Food security indicators at national (e.g. Food Balance Sheets) and local level • Maps of chronic food insecurity areas and GIS data • Any recent and comprehensive report on food security and livelihood	• Relevant sectoral indicators, fact sheets, maps • Country epidemiological profile • Health system details • Health centre type and localization • Previous nutritional surveys at national and regional level. KAP surveys on the causes of malnutrition • National and regional statistics on health: infant feeding practices, mortality and morbidity for the main diseases, and CMR for total and < 5 population • Related maps and GIS data • Any recent and comprehensive report on Health status of the population	• Relevant sectoral indicators, fact sheets, maps • KAP survey and traditional behaviours information regarding excreta disposal, hygiene awareness and water use. • Statistics regarding national, regional and provincial waterborne diseases statistic, including diarrhoea incidence, Cholera endemic zone, Malaria incidence, etc. • Water and sanitation coverage indicators/data. Water source type (urban/rural) • Any recent and comprehensive report on Water and sanitation	• Relevant sectoral indicators, fact sheets, maps • Type of buildings and habitat per area • Type of fuel used for cooking and heating • Geography, geology and climate conditions in affected areas	• Relevant sectoral indicators, fact sheets, maps • Security incident statistics, human right violations, threat types and level • Targeted or marginalized population (women, children, minorities, ethnics, etc..) • Transport means, electricity, phone coverage, logistic hub and reference bodies

DISASTER SPECIFIC INFORMATION – IMPACT AND RISKS (AFFECTED AREA LEVEL)

General information	• Hazard maps and historic records of past similar shocks, impact and humanitarian profiles (severity and frequency of the hazards, number of affected people, dead, etc.) • Reports that indicate how exposure and vulnerability to such hazards may have changed as a result of recent environmental or DRR policy or activities. • Lessons learned and evaluation of past emergencies and relief/recovery interventions • NGOs presence and crisis related activities (3W). In-country capacities likely to support response to the crisis. Regional capacity • Humanitarian profile of the current crisis and Common Operational Datasets • Any Situation reports, press release, news, funding decisions, assessment reports, social media reports, describing impact on areas, groups or sectors (consider also previous flash appeals or CAP in same area or region).

	Food security	Health & Nutrition	WASH	Shelter & NFIs	Access / security
Sectoral information	**Look for data revealing effect on or existence of risk related to:** *Access to Food* • Market distribution systems and capacities • Population access to markets or shops (distance, security, time); • Population access to usual income earning opportunities • People's coping mechanisms and how they are/may have been disrupted/exacerbated by the shock; • Changes in daily food intake and dietary diversity *Availability of food* • Availability of food on local markets, shops, etc. • Crisis impact on local food production. Importation, exportation • Status of household food stocks and reserves • Capacity of the government and other organizations to supplement market mechanisms.	**Look for data revealing effect on or existence of risk related to:** *Health and Nutrition Status* • Health and Nutrition Profile of population • Incidence of communicable diseases *Access to Health and Nutrition Services* • Distance and time to clinics and hospitals • Access to primary health care • User Fees *Availability of Health and Nutrition Services* • Health and Nutrition infrastructure damage reported • Availability of drugs and food Supplies • Availability of qualified staff • Capacity of the government and other organizations for health and nutrition crisis response.	**Look for data revealing effect on or existence of risk related to:** *Access to Water and Sanitation* • Distance to water source • Existence of functioning latrines • User fee *Availability of water (supply)* • Sources of water • Quantity of water per person/day • Quality (safe, protected or not) • Latrines or defecation sites per population. Changes in the type of latrines used (HH/ community latrines) • Incidence of water borne diseases • Capacity of the government and other organizations to cope/ face the WASH emergency.	**Look for data revealing effect on or existence of risk related to:** *Access to Shelter* • Type of housing affected • Key community structures • Changes in shelter used (tents, school buildings etc.) • Access to cooking utensils and fuel, NFIs *Availability of Shelter* • Availability of shelter materials in market place, including plastic sheeting, tents, blankets, etc. • Availability of public buildings for sheltering • Capacity of the government and other organizations for shelter/NFI crisis response	**Look for data revealing effect on or existence of risk related to:** • Any security/risks constraints on movements and access (for affected population and staff) • Others INGOs/UN presence and logistical capacity to respond • Constraints on logistics/delivery systems (taking account of the impact of past events) • Capacity of the government and private companies to ensure safe working conditions and restore communications (roads and network)

Annex II: Reference documents

- IMAS, 2005. *Mine Risk Education, Best Practice Guidebook 2: Data collection and needs assessment*. http://www.mineaction.org/downloads/1/2%20-%20Data%20Collection%20and%20Needs%20Assessment.pdf

- FAO, 1997. *Marketing Research and Information Systems*. http://www.fao.org/docrep/w3241e/w3241e00.htm#Contents

- CARE, 2005. *Tips for Collecting, Reviewing, and Analyzing Secondary Data*, http://pqdl.care.org/Practice/DME%20%20Tips%20for%20Collecting,%20Reviewing%20and%20Analyzing%20Secondary%20Data.pdf

- Niles, R. 2010. *Data analysis*. http://www.robertniles.com/stats/dataanly.shtml

- UNDAC, 2006. *Field Handbook*. Chapter G: Disaster Assessment http://ochaonline.un.org/OCHAHome/AboutUs/Coordination/UNDACSystem/UNDACHandbook/tabid/6012/language/en-US/Default.aspx

- UNFPA, 2010 *Guidelines on data issues in humanitarian crisis situations.* http://www.unfpa.org/webdav/site/global/shared/documents/publications/2010/guidlines_dataissues.pdf

- WFP, 2009.*Comprehensive Food Security & Vulnerability Analysis Guidelines*. http://www.reliefweb.int/rw/lib.nsf/db900sid/PANA7DYCEL/$file/wfp_oct2003.pdf?openelement

- WFP, 2003. *Key Issues in Emergency Needs Assessment*. http://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp189624.pdf

- WFP, Monitoring and Evaluation Guidelines: Module choosing methods and tools for data collection. http://www.wfp.org/content/monitoring-and-evalutaion-guidelines

- IM Course, OCHA

- Garfield 2010, Panel session on principles and norms of data collection and analysis, NATF/ACAPS Coordinated Assessment Training Course, 31-10-2010, Revinge, Sweden.