

## **FOOD FOR PEACE INFORMATION BULLETIN**

### **MEMORANDUM FOR ALL FOOD FOR PEACE OFFICERS AND AWARDEES**

**TO:** USAID/W and Overseas Distribution Lists; Food for Peace Awardees

**FROM:** DCHA/Food for Peace/Director (Acting), John Brannaman

**SUBJECT:** Submission of Baseline/Final Evaluation Datasets by Food for Peace Awardees

**DATE:** August 11, 2011

#### **Food for Peace Information Bulletin 11-02**

### **I. Purpose and Background**

The U.S. Agency for International Development (USAID) Office of Food for Peace (FFP) requires Title II awardees to carry out a baseline survey in the first year of food aid program implementation and a final evaluation near the end of the food aid program. Food for Peace (FFPIB 09-07) specifies when reports from those surveys are due to FFP. This FFPIB further specifies that the datasets generated by Awardees to carry out their baseline/final evaluation surveys shall also be submitted in full to FFP/Washington and the USAID/ Development Experience Clearinghouse (DEC), along with the data dictionary and program files used to process the data.

In this FFPIB, FFP describes the general approach to use in preparing and submitting baseline/final evaluation datasets. This document should be made available to the person(s) and/or firms responsible for managing the evaluations—e.g., by attaching it to their Scope of Work so they can prepare and submit the required information as described here. Awardees are encouraged to coordinate with FFP throughout the life of the award to ensure they are fulfilling reporting requirements. Questions about these requirements should be directed to the food aid program's agreement officer's technical representative (AOTR).

### **II. Introduction**

The USAID Forward initiative aims to transform USAID from a traditional aid agency to a modern development enterprise. One area of focus is to use evidence-based research to determine future policy and program approaches. In line with this focus, FFP is compiling the large quantity of information collected by its multi-year development program Awardees as they carry out their quantitative baseline and final evaluation surveys into a single meta-database. FFP's intent is to use this excellent resource to advance its understanding of global programs' performance and to carry out special analyses that will help fine tune its programs and evaluation methodologies. Awardees' contributions to the meta-database will not be used to verify individual programs' reporting, and except in special circumstances (in which case interested

parties will be consulted), there will be no program specific analyses of the data. To ensure that the data contributed by Awardees is usable, this FFPIB provides instructions on the content and format in which the datasets should be submitted.

### **III. Preparing and submitting evaluation datasets**

#### ***What to submit***

Both the baseline and final evaluation datasets should be provided, along with the supporting documentation required to understand and interpret the data. Datasets and documents should be submitted in electronic format on a CD-ROM.

#### ***Where to submit***

Awardees shall submit their datasets and accompanying information on a CD-ROM to the AOTR at FFP/Washington and to USAID/ Development Experience Clearinghouse (DEC) by following the instructions found on the DEC website here: <http://dec.usaid.gov/index.cfm?p=docsubmit.aboutDoc&CFID=17345831&CFTOKEN=47374847>.

#### ***When to submit***

Evaluation datasets and supporting documentation are due at the same time the survey final reports are due, per FFPIB 09-07, i.e. no later than 3 months after the completion of data collection for the baseline study and upon submission of the new program proposal or upon expiration of the program, whichever comes first, for the final evaluation.

#### ***Format***

The CD-ROM should contain a separate folder for each of the following sections: Dataset, Survey Final Report, Data Dictionaries, Processing Steps, Questionnaires and Field Manuals, Sampling Frame, Description of the Sampling Design. Additional subdirectories may be created under the root directory or within each subdirectory, if this helps organize the information. For instance, separate folders may be created if the programmer used discrete processes to generate anthropometric data; the minimum adequate diet indicator; and the Household Hunger Scale.

In addition, a “Readme File” should be present in the root directory; then a set of subdirectories should be created that organize all the information in a way that facilitates the use of the data. The “Readme File” is an easily readable text file (MS Word, MS Notepad, etc.) that explains the format and organization used in preparing the CD-ROM. This text file should state: 1) the country, name of Awardee, and years during which the program was active; 2) the dates in which the baseline and final evaluations were carried out; 3) the native format of the data (what software/software version it uses); 4) a description of the subfolders’ content; and 5) the name of the person(s) or firms who were responsible for collecting the data and for processing it, with their full reference information.

## ***Datasets***

The datasets themselves should be in a matrix format, with rows corresponding to cases/units of sampling and columns corresponding to variables. The cleaned, raw dataset should always be provided, along with its dictionary (see below). Any additional datasets generated for the purpose of computing new variables (for instance, to compute anthropometric Z-scores) should also be provided, with clear indications of the processing steps used in preparing the data for analysis.

Before sending the data to FFP, Awardees should delete all information that could potentially allow one to identify the household or individual from whom data were collected (e.g. GPS location or address of household, names of individuals, etc). However, cluster identification should be maintained in order to facilitate future variance computations.

The datasets should be provided in one of the following software formats: Excel, EpiInfo; SPSS; SAS; or Stata. If the software used is different from the ones listed here, then it should be exported to one of these softwares or else to MS Excel. Note that while dedicated software like SPSS and EpiInfo already contain the data dictionary, Excel does not. Therefore data submitted using Excel should be accompanied by a separate dictionary file.

## ***Supporting Documentation***

To ensure appropriate handling of the data, a complete package of supporting documentation should accompany each dataset submitted. USAID/FFP requests that all supporting documentation be provided in English.

## ***Survey Final Report***

The final version of the baseline and of the final evaluation reports produced from the analysis of the data should be included in an easily readable format (PDF, Word, etc).

## ***Data Dictionaries***

A complete data dictionary should be submitted that specifies at a minimum for each variable: its name; label (if any); value labels (if any); format (alphabetic, numerical, currency; etc); type (dichotomous/categorical/continuous); and size (number of digits, including decimal points). When using dedicated software (e.g. SPSS, Epi-Info, SAS or STATA), the dictionary is already part of the data file so it is not necessary to create one separately; but the dictionary information should be provided in a text file.

## ***Processing steps***

The steps taken to read the data into the software, and then to transform the raw data into usable information should be carefully documented and referenced, so they can be easily replicated later by a third party. All the information required to replicate the steps should be “bug-free” and present on the CD-ROM. For instance, if a processing step calls for the merging of two different datasets, those two datasets must be present on the CD-ROM and properly referenced in the processing step.

### *Questionnaire(s) and Field Manuals*

The final questionnaire(s) used to collect the data must be included, preferably in PDF format. If the questionnaire was written for a PDA (handheld computer), the name of the software used should be specified, and the program developed to carry out the survey should be included. The Field Manual used by enumerators while carrying out the survey should also be included. It is expected that the Field Manual will contain an accurate description of all the questions covered in the questionnaire.

### *Sampling Frame*

Please provide the frame used to select the sample clusters, along with any associated size measures (such as number of households in each cluster on the frame). Explain the source of this list and the approximate date it was created.

### *Description of the Sampling Design*

Include a description of the sampling design used for the survey (i.e., multi-stage cluster sampling, with or without stratification). When a cluster sample was used for data collection, the document should also describe the sampling procedure used to select both clusters (e.g., PPS sampling) as well dwellings within each cluster (e.g., systematic random sampling, random walk, segmentation). If there was a specific treatment given to multi-household dwellings, this should be specified (e.g., sampling all such households within a selected dwelling). In addition, the approach used to sample children within households should be explained (i.e., all children within the target age range sampled in each household (recommended), or one child of the target age range sampled per household). An indication should be given whether or not sampling weights were calculated (recommended). If so, both the final sampling weights and the sampling weights at each stage of selection should be included in the dataset and data dictionary. Finally, the method for treating non-responding households should be documented. That is, the follow-up strategy that was implemented in the field should be described and any residual non-response should be reported as a percentage of the total expected sample size. In those cases where an adjustment to the final sampling weights was made to compensate for the residual non-response at the household level, this should be indicated as well.