Feed the Future

Survey Implementation

Document

In-Country Data Manager’s Manual

Zone of Influence Survey

[COUNTRY] [YEAR]

April 2018

This publication was prepared for review by the United States Agency for International Development. It was prepared by ICF for the United States Agency for International Development, USAID Contract Number GS-23F-8144H/AID-OAA-M-12-00006, Feed the Future PEEL.

**Recommended Citation**

Genevieve Dupuis, Jasbir Kaur, and Kiersten B. Johnson. 2018.

*Feed the Future Zone of Influence Surveys: In-Country Data Manager’s Manual.*

Washington, DC: Bureau for Food Security, USAID

**Contact Information**

Feed the Future

1300 Pennsylvania Ave, NW

Washington, DC 20004

[www.feedthefuture,gov](http://www.feedthefuture,gov)

Table of Contents

[Abbreviations iv](#_Toc523150649)

[1. Introduction 1](#_Toc523150650)

[1.1 Survey objectives 1](#_Toc523150651)

[1.2 Sample 1](#_Toc523150652)

[1.3 Survey modules 1](#_Toc523150653)

[1.4 Survey implementation 2](#_Toc523150654)

[1.5 Confidentiality 2](#_Toc523150655)

[2. Feed the Future Zone of Influence Surveys 3](#_Toc523150656)

[2.1 ZOI Survey data management 3](#_Toc523150657)

[3. ICDM and [CONTRACTOR] data manager roles 5](#_Toc523150658)

[3.1 Role of the ICDM 5](#_Toc523150659)

[3.2 Role of the [CONTRACTOR] data manager 6](#_Toc523150660)

[3.3 Communication between the ICDM and [CONTRACTOR] data manager 6](#_Toc523150661)

[3.4 Additional staffing recommendations 6](#_Toc523150662)

[4. The ICDM menu system 8](#_Toc523150663)

[4.1 Introduction 8](#_Toc523150664)

[4.2 File structure and programs 8](#_Toc523150665)

[4.3 ICDM menu options 9](#_Toc523150666)

[4.4 Accessing questionnaires 11](#_Toc523150667)

[5. Data quality review 12](#_Toc523150668)

[5.1 Data structure check 12](#_Toc523150669)

[5.2 Secondary editing reports 12](#_Toc523150670)

[5.3 Field check tables 13](#_Toc523150671)

[5.4 Cluster status reports 13](#_Toc523150672)

[Appendix A: Data structure messages and guidance 15](#_Toc523150673)

[Appendix B: Secondary editing guidelines 19](#_Toc523150674)

[Appendix C: Limits for length and weight of children 30](#_Toc523150675)

[Appendix D: Cluster status report 31](#_Toc523150676)

[Appendix E: Data issues log 32](#_Toc523150677)

[Appendix F: Data processing diagram 33](#_Toc523150678)

[Appendix G: Data processing communications diagram 34](#_Toc523150679)

[Appendix H: CSPro variable names and labels 35](#_Toc523150680)

List of figures

[Add figure numbers and captions. Check that text callouts precede each figure.]

[Figure 1: Screen shot of Central Office CSPro Menu shortcut 8](file:///C:/Users/21990/Documents/IHD%20files/PEEL/for%20formatting/formatted%20files/20180416%20FTF%20ZOI%20Survey%20Doc%20-%20In-Country%20Data%20Manager%27s%20Manual_formatted.docx#_Toc512246175)

[Figure 2: Screen shot of Central Office Main Menu 8](file:///C:/Users/21990/Documents/IHD%20files/PEEL/for%20formatting/formatted%20files/20180416%20FTF%20ZOI%20Survey%20Doc%20-%20In-Country%20Data%20Manager%27s%20Manual_formatted.docx#_Toc512246176)

[Figure 3: Screen shot of household case tree 10](file:///C:/Users/21990/Documents/IHD%20files/PEEL/for%20formatting/formatted%20files/20180416%20FTF%20ZOI%20Survey%20Doc%20-%20In-Country%20Data%20Manager%27s%20Manual_formatted.docx#_Toc512246177)

[Figure 4: Screen shot of Go To Field (F6) navigation box 10](file:///C:/Users/21990/Documents/IHD%20files/PEEL/for%20formatting/formatted%20files/20180416%20FTF%20ZOI%20Survey%20Doc%20-%20In-Country%20Data%20Manager%27s%20Manual_formatted.docx#_Toc512246178)

[Figure 5: Screen shot of secondary editing example 11](file:///C:/Users/21990/Documents/IHD%20files/PEEL/for%20formatting/formatted%20files/20180416%20FTF%20ZOI%20Survey%20Doc%20-%20In-Country%20Data%20Manager%27s%20Manual_formatted.docx#_Toc512246179)

List of tables

[Add table numbers and table titles. Check that text callouts precede each table.]

# Abbreviations

EA Enumeration Area

CAPI Computer-Assisted Personal Interviewing

CSPro Census and Survey Processing System

ICDM In-Country Data Manager

QCS Quality Control and Support

USAID United States Agency for International Development

ZOI Zone of Influence

# 

# 1. Introduction

This manual provides detailed instructions on how to implement a critical component of the United States Agency for International Development (USAID) Feed the Future Zone of Influence (ZOI) Surveys—data processing and management.

## 1.1 Survey objectives

Feed the Future is the U.S. Government’s global food security initiative that seeks to reduce poverty, hunger, and undernutrition among women and children and to increase, resilience, income, women’s empowerment, dietary diversity, and appropriate feeding practices. Feed the Future’s programmatic efforts are concentrated in ZOIs in a number of countries, including [COUNTRY].

Progress in achieving Feed the Future’s objectives is tracked using information collected through representative cluster sample surveys called ZOI Surveys. These surveys are designed to provide estimates for key Feed the Future indicators related to poverty, agriculture, nutrition, hunger and resilience, with a 95 percent confidence level.

## 1.2 Sample

This Feed the Future Phase ZOI Survey will be conducted among a sample of the entire population living in a Feed the Future ZOI. The study will follow a two-stage cluster sampling design. In the first stage, enumeration areas (EAs) will be selected; in the second stage, households will be selected in each EA. Before fieldwork begins, a complete household listing will be conducted in each EA, from which [XX] households will be selected randomly for interviews in each EA. During fieldwork, if more than one household is discovered in a single dwelling unit, all resident households will be interviewed for the survey.

The aim of the Feed the Future ZOI Survey is to produce single-point-in-time estimates of indicators, including their standard errors and confidence intervals, and to enable a statistical test of differences to detect changes in indicators over time.

## 1.3 Survey modules

The ZOI Survey has eight modules, identified as Modules 1–6, 6M, 7, and 8:

* Module 1, Household Identification, Informed Consent, and Household Roster and Demographics: One module will be administered in a household; each household roster contains basic data for all household members including age, gender, and educational attainment.
* Module 2, Dwelling Characteristics: One module will be administered in a household.
* Module 3, Food Security and Resilience: One module will be administered in a household.
* Module 4, Women’s Anthropometry and Dietary Diversity: One module will be administered to each woman age 15 to 49 years listed in the household roster.
* Module 5, Child Anthropometry and Infant and Young Child Feeding: One module will be administered for each child age 0 to 5 years listed in the household roster.
* Module 6, Empowerment in Agriculture Index–Primary Female Decisionmaker: One module will be administered in a household, but only in households with a primary adult (age 18 or older) female decisionmaker.
* Module 6M, Empowerment in Agriculture Index–Primary Male Decisionmaker: One module will be administered in a household, but only in households with a primary adult (age 18 or older) male decisionmaker.
* Module 7, Agricultural Technologies: One module will be administered to each household member responsible for cultivating the selected value chain commodities.
* Module 8, Household Consumption Expenditure: One module will be administered in a household.
* [COUNTRY-SPECIFIC MODULE].

Survey questionnaires will be translated into the major languages used in the ZOI in [COUNTRY].

## 1.4 Survey implementation

The ZOI Survey is funded by USAID, and the USAID-funded Feed the Future [SURVEY CONTRACTOR’S PROJECT NAME] project is providing technical assistance to the survey. [SURVEY CONTRACTOR] has subcontracted with [SUBCONTRACTOR] to conduct the ZOI Survey.

## 1.5 Confidentiality

All data collected by the interviewers for the survey will be completely confidential. Survey staff should not discuss these data with anyone other than the [SURVEY CONTRACTOR] data management staff, survey management personnel, and field supervisors. Data from this survey, including names of respondents, should not be shared.



# 2. Feed the Future Zone of Influence Surveys

Feed the Future Zone of Influence surveys are conducted in cluster samples. This means that out of all communities in the country, a number of communities is selected randomly to participate in the survey, and from all of the households in those selected communities, a sample of households is selected randomly to be interviewed. The sampled communities are referred to as “clusters.” A cluster may be the same as a village in rural areas, or may constitute a block in urban areas, with some exceptions. The clusters are generally Census EAs defined by the [COUNTRY] [national statistical agency]; information on the size of the population within each EA is based on the most recent census.

Clusters may vary in size, sometimes comprising several hundred households; however, the number of households selected from each cluster for participation in the survey is constant across all clusters. The number of households sampled in each cluster for this survey is [30]. Households that are selected are referred to as “sampled” or “selected.”

Interviewer teams are expected to complete interviews in all sampled households in a cluster before moving to the next cluster. While teams will transmit their data to the Central Office (via Dropbox) at least daily as backup, their final transmission of data to the Central Office will take place after all data in the cluster have been collected and checked in the field for completeness. This process of checking data for completeness and finalizing it before the final transmission to the Central Office is called “closing clusters.” Each cluster must be closed before the interviewer team leaves the area. In remote areas, access to the Internet may be unavailable. If a connection is not available in the cluster, the field supervisor must wait until the field team is in a location with Internet access. At that point, the field supervisor should transmit all un-transmitted, final household forms from the previous cluster from the interviewers’ tablets. The field supervisor should transmit files as soon as Internet access is available, even if this requires delaying fieldwork slightly. The field supervisor may need to make a special trip to the closest location with cellular or Wi-Fi access to transmit files in a timely fashion. Note that while teams may need to transmit data after the team has left the cluster due to Internet connectivity problems, the teams must still close each cluster before departing.

## 2.1 ZOI Survey data management

ZOI Survey data management requires collaboration between the in-country data manager (ICDM) and the [SURVEY CONTRACTOR] data manager. Contact information for the [SURVEY CONTRACTOR] data manager for [COUNTRY], the primary point of contact for the in-country data manager, is as follows:

[SURVEY CONTRACTOR data manager NAME]

[SURVEY CONTRACTOR data manager EMAIL]

[SURVEY CONTRACTOR data manager PHONE NUMBER]

Data management for these surveys begins after data from closed clusters have been transmitted from the tablet computers to the [SURVEY CONTRACTOR] server and [IMPLEMENTING AGENCY] using a secure Dropbox. Those data are then extracted for additional quality control review, editing, and analysis.

After downloading data from the field teams, the ICDM will first run a structure check (see Section 5.1) on the data for each closed cluster to ensure that all data sent by the interviewer teams have been received, with no missing data. The ICDM will then produce secondary editing reports and review them to either edit the data or clear the errors identified in the reports (see Section 5.2 and Appendix B). The ICDM will generate field check tables to assess the quality of work being done in the field and to identify any data collection issues (see Section 5.3). The ICDM will have access to the data using a Census and Survey Processing System (CSPro) data processing system. The [SURVEY CONTRACTOR] data manager will also examine the field check tables (see Section 5.3) and help with the secondary editing review (see Section 5.2), as needed. After the fieldwork ends and data from all clusters have been transmitted and received, the [SURVEY CONTRACTOR] data manager will review all data, including the complete secondary editing list and frequencies of the responses to each question in the survey. The data manager will review any issues with the ICDM to find the best solution.

Data received during fieldwork are assumed to be from closed clusters that the interview teams have already left; therefore, when the ICDM or data manager find issues or discrepancies in the data, two courses of action are to be taken:

* Data will be edited after having been received at the [IMPLEMENTING AGENCY] and the [SURVEY CONTRACTOR]. No data will be fabricated to fill missing data. The ICDM and [SURVEY CONTRACTOR] data manager will follow the guidelines (see Appendix B) provided by [SURVEY CONTRACTOR] to ensure proper and consistent editing procedures. At the conclusion of fieldwork, the [SURVEY CONTRACTOR] data manager will review all editing.
* Feedback on data quality will be provided to field supervisors either directly or through the [SURVEY DIRECTOR]. Fieldwork practices may need to be adjusted, which could include a recommendation to provide further training for interviewers, increase monitoring of field teams, or replace interviewers.

# 3. ICDM and [CONTRACTOR] data manager roles

## 3.1 Role of the ICDM

The ICDM has a large responsibility that necessitates frequent communication with the field supervisors, the survey manager, the Quality Control and Support (QCS) teams, and [SURVEY CONTRACTOR] about problems that are discovered during fieldwork. It is vital for data quality that the ICDM communicate any issues or concerns about data collection to the survey’s staff, and that survey staff take appropriate actions.

The data quality checks are implemented primarily through the use of household structure checks, field check tables, secondary editing reports, and cluster status reports, described later in Section 5. The ICDM is responsible for conveying the information from these sources to the QCS teams, field supervisors or other appropriate survey management. The responsibility for maintaining data quality does not end with feedback to the survey management; it requires following up with the field teams to ensure that the recommendations are carried out and that the issue with data collection is no longer a problem.

The ICDM will provide summaries of findings to the QCS teams, and will provide feedback and guidance for QCS team members to relay to the field teams. The in-country data manager will give the QCS team copies of the QC reports in hard copy if the QCS team visits headquarters or by email if the QCS team is in the field. The in-country data manager’s summaries will include positive feedback on areas where good performance has been observed and constructive criticism on areas where improvement is needed.

The ICDM also will be the lead contact for problems coming from the field concerning computer-assisted personal interviewing (CAPI) data collection, which includes repairing broken or malfunctioning tablets, managing field staff information in the CSPro system, making recommendations for further training on tablet procedures, and troubleshooting problems with data transmission. To support these functions, the ICDM should be fully knowledgeable of the CAPI data collection systems, tablet troubleshooting, and supporting data files for fieldwork. Supporting files include data files holding the identification information for all fieldworkers, households selected for interview, and the clusters included in the survey. These files are prepared by the CSPro programmer, who will modify them as needed.

The ICDM is also responsible for all post-field data processing; that is, the ICDM must track data as they are transmitted from the field teams through the sequential data processing steps. The ICDM will use the CSPro system for processing data in the central office. Steps include checking the data structure as it arrives to ensure that all data were transmitted successfully, carrying out secondary editing, and producing field check tables and status reports. Data from all clusters must be received, edited, and finalized and the ICDM must make sure this happens in a timely manner; data should be finalized no later than one week after it is received from the field.

See Appendix G for a visual representation of the communications among the ICDM, [SUBCONTRACTOR] Survey Manager, [CONTRACTOR] data manager, QC Teams, Field Supervisors, and Interviewers.

## 3.2 Role of the [CONTRACTOR] data manager

The ICDM, working in the Central Office in-country, will oversee the quality of data received from the field teams. The [SURVEY CONTRACTOR] data manager will review the reports that the ICDM produces and will be available to answer questions and support the ICDM. The [SURVEY CONTRACTOR] data manager will work with the ICDM to resolve all problems identified or reported from the field.

The [SURVEY CONTRACTOR] data manager will review all field check tables and track the data processing progress. If the data processing is progressing too slowly, the [SURVEY CONTRACTOR] data manager will identify any issues, locate the cause(s) for delays, and work with the ICDM to resolve them.

The [SURVEY CONTRACTOR] data manager will be responsible for CAPI system or CSPro programming problems. If an error is found during fieldwork, the [SURVEY CONTRACTOR] data manager is responsible for resolving it and ensuring that all interviewers or systems have received updated programs for data collection using CSPro.

After the fieldwork and data processing are completed, the [SURVEY CONTRACTOR] data manager will review the data quality and secondary editing, working with the ICDM to finalize the datasets.

## 3.3 Communication between the ICDM and [CONTRACTOR] data manager

The ICDM will review the findings of data quality reports and communicate with the QCS teams and field supervisors as needed, such as recommending additional training for interviewers or additional supervision. Unresolved issues with data transmission, structure checks, secondary editing, or field check tables must be recorded in the data issues log. If the ICDM cannot resolve the issue through application of standard guidelines, this should be noted in the data issues log and shared with the [SURVEY CONTRACTOR] data manager.

The ICDM and [SURVEY CONTRACTOR] data manager should communicate freely and openly about the survey data processing progress. Regular—at least weekly—phone calls should be planned to check in, although more frequent communication likely will be required.

## 3.4 Additional staffing recommendations

The ICDM is responsible for data processing activities in the Central Office, overseeing the processing of data arriving from the field through to finalization. In order to support teams in the field and manage data processing in the Central Office, the ICDM should have the support of a second data processing staff person and an IT specialist.

ICDM support staff will assist with responding to technical requests received from the field. The second data processing staff person will be trained to carry out secondary editing under the supervision of the ICDM and assist with requests from the field regarding use of the CAPI data collection system. The IT specialist will be available to assist in handling all IT requests, such as tablet repair or replacement, as well as IT requirements in the Central Office, such as networking and internet connectivity. Depending on the needs of the survey, these roles may be filled by the same full-time staff person that has attended all survey trainings and has the necessary skills to assist with both data processing and IT needs.

# 4. The ICDM menu system

## 4.1 Introduction

The ICDM will use a CSPro system menu to automatically track the data processing and produce data review reports on the quality of the fieldwork. Using a menu system simplifies the ICDM’s work and allows automatic tracking of the data processing status. Data from every closed cluster in the field will be sent to the central office by the field supervisors. The ICDM must use the menu to receive, edit, and finalize the data, and all data review reports will be run directly from the ICDM menu. Many of the processes performed by the ICDM menu automatically are described in this document, though they occur behind the scenes and require no action by the ICDM, other than to choose the options in the order they appear in the menu. The entire step-by-step process of data collection to data finalization is illustrated in Appendix F.

## 4.2 File structure and programs

All data and programs used to process Feed the Future ZOI Survey data are stored on the central office computer (and/or server), and managed by the ICDM. Files will be stored in the root directory of the computer in a folder, “C:\PEELcc,” where cc is the two-letter country code abbreviation. Subfolders in the C:\PEELcc folder will hold the data for processing:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | **Root Directory** | | |
|  | | C:\PEELcc | | |
|  |  | \CENTRAL |  | All programs the ICDM uses to process, review, and finalize data. |
|  |  | \CONTROL |  | Files used to control data processing and track cluster status. |
|  |  | \DATA |  | All survey data, in subfolders: |
|  |  |  | .\FIELD | Original unchanged data received from the field. |
|  |  |  | .\RAW  .\ACCEPT | All data being processed from the field.  Data that has passed basic structure checks and is ready for editing |
|  |  |  | .\FINAL | Final data after editing and finalization. |
|  |  | \DICTS |  | Dictionaries for all programs used to process data. |
|  |  | \DOCS |  | Documentation and manuals for all components of fieldwork and data processing. |
|  |  | \ENTRY |  | All data entry programs and menus used by interviewers. |
|  |  | \REPORTS |  | Results of all data review programs. |
|  |  | \SUPERVISOR |  | All programs and menus used by field supervisors to manage fieldwork. |

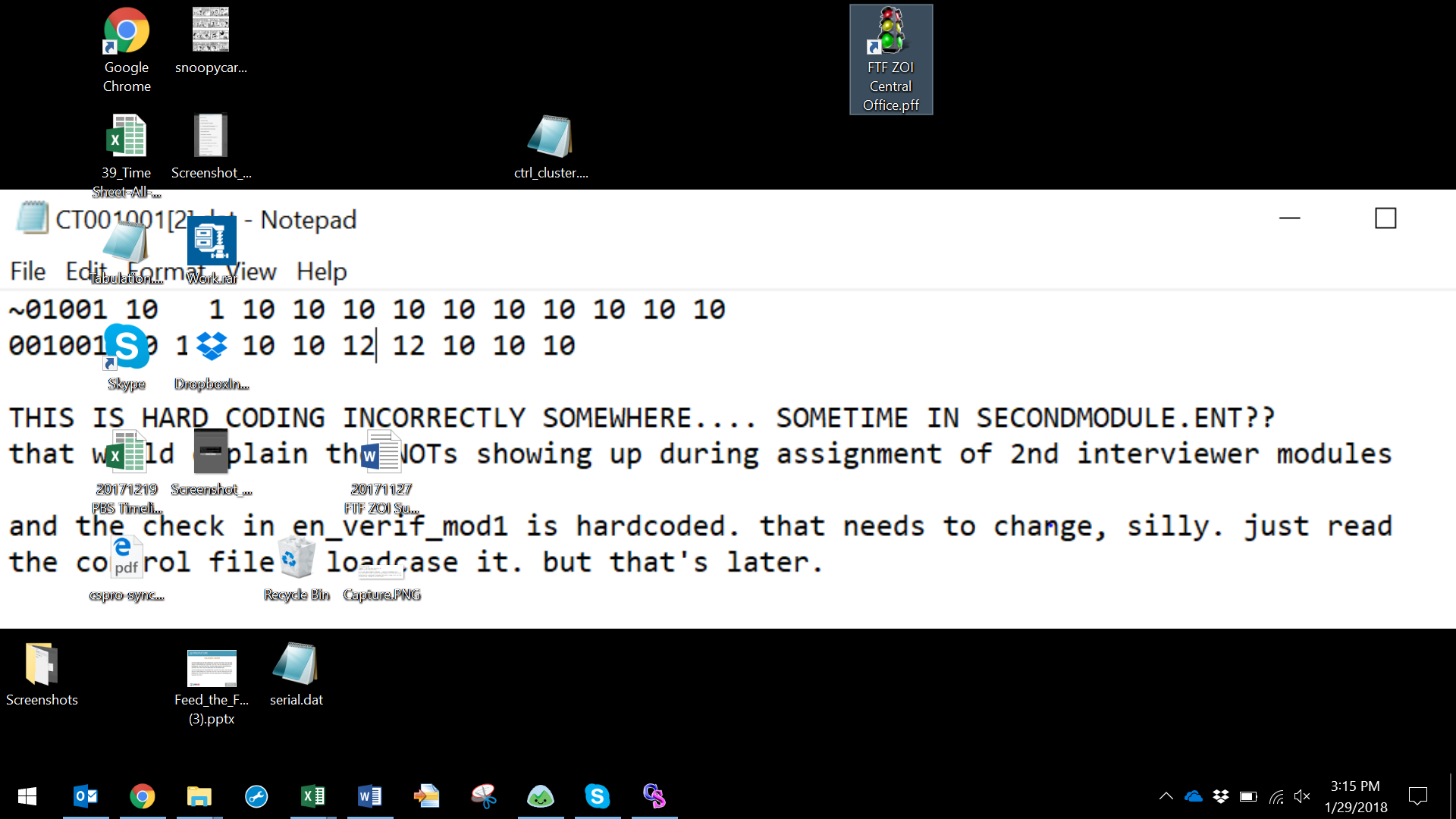
The ICDM will use the following program files to process data and review data quality:

|  |  |  |
| --- | --- | --- |
| \CENTRAL | | |
| CentralOffice.ent |  | Menu system for ICDM work. |
| \CONTROL |  |  |
| StructureCheck.bch |  | Checks for completeness of incoming data and result code consistency. |
| SecondaryEditing.bch |  | Checks data consistency and generates reports on outstanding data issues. |
| FieldCheckTables.bch |  | Generates field check tables for review of major trends in data collection. |
| StatusReport.bch |  | Produces a status report on all clusters and assesses progress toward data processing completion. |

## 4.3 ICDM menu options

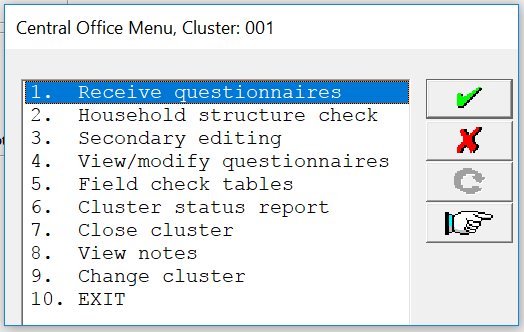
The ICDM menu has options for the steps required for processing data and monitoring data quality. An order of operations will be followed as data from each cluster are received in the central office. The ICDM must follow these steps and process all levels for each cluster before the data are considered final. As the cluster data files are received from the field, the in-country data manager uses the menu to automatically download the files and begin processing. Then the ICDM will access the files from the CSPro menu, shown in Figure 1.

Figure 1: Screen shot of Central Office CSPro Menu shortcut



A data processing control file is used to store dates that each stage of data processing is completed and to ensure the order of operations for the ICDM. When new data arrives to the central office, the ICDM menu first displays a report indicating new clusters have arrived in the central office and are ready for processing. After reviewing the list of clusters available for processing, the ICDM must enter the cluster number in which they would like to begin working. The data processing steps are ordered in the menu shown in Figure 2. Note that some menu options can be run at any time.

Figure 2: Screen shot of Central Office Main Menu



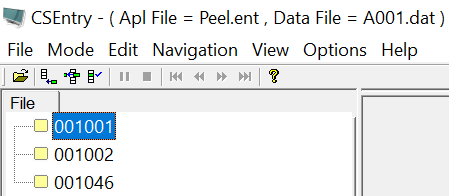
1. **Receive questionnaires.** As the data for each cluster are received from the field, this menu option will download the incoming data from a secure Dropbox folder to the subfolder, \RAW. This option will also simultaneously copy the original, unchanged data into the \FIELD subfolder, where it will be preserved as a backup. When received, the data processing control file is automatically updated to record the date of receipt.
2. **Household structure check.** This menu option will run a structure check on the data received from the field to ensure that all applicable survey modules were completed in all selected households in the cluster, and that no data are present that should not be included in the dataset. It will also ensure that all data were transferred successfully. The result will be a list of all households and modules completed, followed by any error messages to indicate problems. The control file for data processing is automatically updated to indicate the household structure check has been run and approved. This structure check report must be run with no errors before continuing with data processing.
3. **Secondary editing.** This menu option will list potential errors and inconsistencies in the data. Not all errors will require changes to the data; however, all error messages should be reviewed according to the standard guidelines for each message, provided in the secondary editing manual (see Appendix B). The ICDM or the data manager must run this report at least once before the cluster of data can be finalized. Each date that the secondary editing report is run is automatically recorded in the control file.
4. **View/modify questionnaires.** This menu option will allow the ICDM to access questionnaire data by cluster, either for review or for modification. To access the data, the ICDM must know the cluster and household number, indicated with each message in all reports on data quality.
5. **Field check tables.** This menu option will produce a series of tables of all data that have been received from the field and have passed the household structure check in Step 2. Only data with the proper structure will be included in the tables, which is why it is important for the ICDM to run the household structure check as soon as data are received from the field. Individual cluster data will be concatenated before running the field check tables program. These tables will be produced in an .RTF format, which can be viewed in Microsoft Word. The field check tables should be reviewed at least once a week, but can be run more often, as needed.
6. **Cluster status report.** This menu option tracks the status of data collection and processing at the cluster level. The results show the dates that data from each cluster have been received from the field and the dates as each step in the data processing are completed. This report can be produced anytime, but at least weekly, so that the [SURVEY CONTRACTOR] data manager can project data processing completion.
7. **Close cluster.** This menu option will close a cluster after the structure check and secondary editing steps are complete. Closing a cluster indicates that processing has finished for this cluster. Before a cluster can be closed, a program will check that previous processing stages were completed, and then will automatically update the data processing control file.
8. **View notes.** This menu option will load and display all notes that interviewers in the field have made for the cluster. Notes include anything that the interviewer felt should be noted for a particular question, module, or household.
9. **Change cluster.** This menu option allows the ICDM to switch between clusters in the data processing menu. To work in a different cluster, the user should first close all reports and programs, select this option, and enter the new cluster number.
10. **Exit.** This menu option will close the ICDM menu and should be used to leave the menu.

## 4.4 Accessing questionnaires

From the ICDM menu option “4 – View/modify questionnaires,” the ICDM can view and modify data from household questionnaires that have been received in the central office. The ICDM can view data any time; however, the ICDM should modify the data only during the household structure check and secondary editing steps. The ICDM should not create or change data unless an error message indicates that a modification is necessary and provides the specific guidelines for the change. The following steps describe how the ICDM can access data:

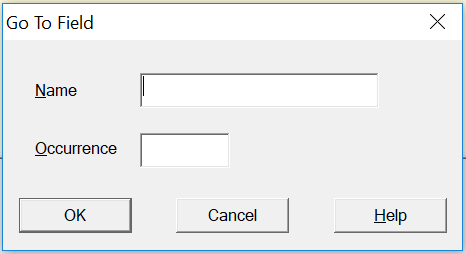
* From the ICDM menu, enter the cluster number of the questionnaire. Then choose option “4 - View/modify questionnaires”.

Figure 3: Screen shot of household case tree



* A list of households will appear on the left side of the screen, as shown in Figure 3, and double-clicking on the household will open the file to view or modify questionnaire data.
* Use one of the following options to navigate to a particular question or variable:

Figure 4: Screen shot of Go To Field (F6) navigation box



* Use the arrow keys to move forward and backward through the questionnaire.
* Press F6 to open the Go To Field box and navigate to a specific variable, such as “V201” or “V6301.” Enter the information in the dialog box, as shown in Figure 4, using the exact variable name. Appendix H lists the exact variable names.

To exit the questionnaire, press F10 to advance to the end of the questionnaire; or to the last question if the questionnaire is incomplete; or if a modification to a variable changed the flow of the questionnaire, requiring a response to a question that was not originally asked during fieldwork. If this happens, a code for “missing” must be entered if the real response to the data cannot be determined. It is important to press F10 to advance to the end of the questionnaire to exit. Closing the questionnaire from anywhere except the end or in any other way will flag the questionnaire as partially complete. Although data will not be lost, the system will require reopening the questionnaire and advancing to the end to exit properly.

# 5. Data quality review

Potential data quality issues or discrepancies in the data will be discerned by triangulating the results of a combination of four data quality assessments: data structure checks, secondary editing reports, field check tables, and cluster status reports. The ICDM will be responsible for producing daily data quality reports as data come in. The ICDM will review the reports and weekly send the results to the appropriate survey staff. All reports will be run directly from the ICDM menu to manage the survey data.

## 5.1 Data structure check

As clusters are closed in the field and data are transmitted to the central office, the ICDM will download the data. Using the household structure check menu option, the ICDM will check to ensure that all household questionnaires have a result code, all modules that should have been completed were completed, and all data have been transmitted and received properly. This household structure check step should be done regularly whenever data from the field is received. Falling behind in checking the household structure will cause backlogs in data processing and delay discovery of data quality problems, which will preclude timely corrections in the field. If any issues appear on this report, the data cannot be included in further data quality reports, and the ICDM will immediately contact the field supervisor to determine why the errors are showing up in the report. Appendix A shows a complete list of possible errors and guidance for resolving them. Using the error message number printed in the report, the in-country data manager should refer to Appendix A to read an explanation of the error, along with possible solutions that should be explored with the field supervisor. Most errors require the field supervisor to resolve an issue in the field and resend the data, while some errors may be resolved in the central office directly. Complete guidance is provided in Appendix A.

## 5.2 Secondary editing reports

After the ICDM completes the household structure check for a cluster, a secondary editing report will be run for each cluster. The secondary editing report, produced with menu option 3–Secondary Editing, will identify any inconsistencies within households in the data received in the central office. The ICDM will review all error messages to decide if action is needed, based on the secondary editing manual.

Each error message produced in the secondary editing report follows a specific format and should be addressed systematically. Figure 5 shows an example of an error message from a secondary editing report:



Figure 5: Screen shot of secondary editing example

The identification code for the household where the error occurred first appears – cluster number first, followed by the household number. The cluster number and household identification number are needed to review the data, or if required, to modify the data. After the identification code, the error message number appears, followed by a short description of the error. The ICDM should locate the error message number in the secondary editing guidelines (Appendix B), which provide an explanation of the error, what other variable or variables to check in the questionnaire to help resolve the error, and specific instructions on how to modify the data if necessary. Many of the secondary editing guidelines require no change to the data, and the in-country data manager should follow these instructions strictly.

If the secondary editing report error messages require changes to the data, the ICDM will use menu option 4 – View/modify questionnaire, described earlier in Section 4.4, to make changes using the data entry program that interviewers use in the field. To access the correct data, the in-country data manager can find the cluster and household identification directly from the secondary editing report and the error message. After making any required changes, the ICDM will run the secondary editing report again until all messages have been addressed. Note that this does *not* mean that all messages will disappear; many messages are warnings and do not require edits. Under no circumstances should data be edited solely to remove a message; guidelines for editing should be strictly followed, and no data should be fabricated. The secondary editing guidelines in Appendix B contain a complete list of secondary editing errors and guidance for handling them.

The [SURVEY CONTRACTOR] data manager will provide support for secondary editing to the ICDM, as needed. For any issues where the ICDM is not certain of how to handle an error message found in the secondary editing reports, the ICDM must contact the [SURVEY CONTRACTOR] data manager for guidance.

## 5.3 Field check tables

At least weekly the ICDM will produce field check tables to review key statistics from the received data to identify trends and inconsistencies, which may indicate quality problems in specific teams or with individual interviewers. Each table has guidelines and targets that field teams should meet during the fieldwork.

The ICDM can review the field check tables by team or by interviewer. It is recommended that the ICDM begin by reviewing tables by team, and subsequently by interviewer after data are available. It is important to review tables by both team and interviewer because some critical errors in collection by a single interviewer may not appear in team tables. After the reports are run, they are filed in the /REPORTS folder for the project files. The FTF ZOI Field Check Tables Manual provides field check table shells, guidance on the targets, and further explanation.

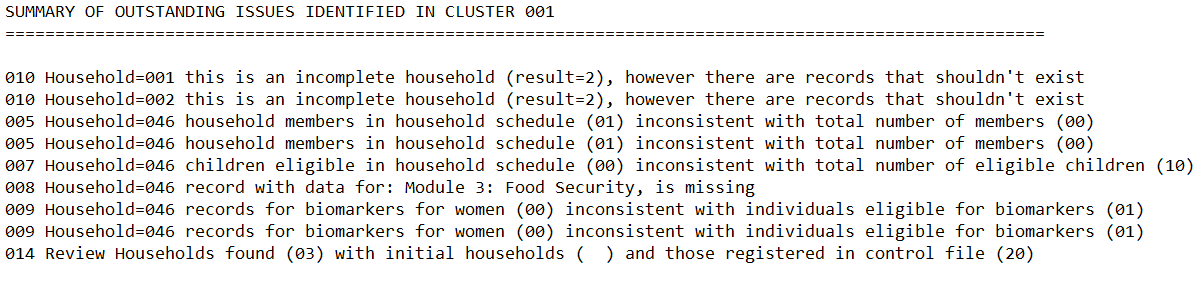
Each version of the field check tables produced will be provided to survey management, including the [SURVEY CONTRACTOR] data manager. The [SURVEY CONTRACTOR] data manager will review field check tables and provide feedback on data quality and progress to the ICDM, whose responsibility it is to communicate any findings to the field teams.

## 5.4 Cluster status reports

The ICDM will track the status of fieldwork and data processing, and report on it at the request of the [SURVEY CONTRACTOR] data manager and survey management. The ICDM will produce the cluster status report at least once a week, along with the field check tables. These reports will track the number of clusters that are in progress, closed, and at each stage of finalization in the central office. The reports show the data received in the central office by cluster and the dates that each processing stage has been completed, alongside response rates for each cluster. The report summarizes the number of clusters in each stage. The [SURVEY CONTRACTOR] data manager can project the completion of data processing using these reports. An example of these reports appears in Appendix D.

# Appendix A: Data structure messages and guidance

This appendix contains guidance for resolving errors encountered in checking the structure of data received from the field in the central office. Using the error code on the left of the screen, follow the guidance outlined in this Appendix A.



001 Control file for cluster n could not be loaded

This message can indicate that the current cluster is not part of the sample design or that a sample file has been corrupted. Please double check the cluster number. If the error persists, contact the [CONTRACTOR] Data Management Specialist.

002 Number of households to be collected in cluster n hasn't been assigned to the control file

This message indicates that the total number of expected households for this cluster is zero in the supervisor’s household assignment file. The ICDM should check that data from the field was properly downloaded and that there were no changes to the sample in the field (ex. added households during fieldwork). If the problem persists, contact the [CONTRACTOR] Data Management Specialist.

003 The household assignment file for cluster n could not be loaded

This message indicates that the current cluster household assignment is not part of the sample files or that the household assignment has been corrupted. Please double check the cluster number and confirm that the household assignment was properly downloaded from the field. If the error persists, contact the [CONTRACTOR] Data Management Specialist.

004 Household = n partially completed, structure will not be checked at this time

This error occurs when a household questionnaire has been exited in the middle of the data entry (during initial entry OR during a modification of the data) and the case was saved at its partially-completed stage. If the household has in fact been completed, it is likely that the interviewer re-opened the household to check the data or make a modification and did not advance to the end of the questionnaire before closing. If the household is complete, this error is resolved by opening the household interview and advancing to the end of the questionnaire (done quickest by using the “advance to end” option in the questionnaire menu). If the error appears after transmission from the field, the ICDM must contact the field supervisor to resolve the issue – either instructing them to complete the interview or to resend the data.

005 Household = n: number of household members in household schedule (x) inconsistent with total number of members (y)

This message indicates that the total number of people listed in the household schedule is different from a program calculated total. The ICDM should open the household questionnaire, review the household roster, and check if the data in household schedule have been corrupted. If all household members are properly listed, advance to the end of the questionnaire and rerun the program. If the error persists, the ICDM should contact the field supervisor to resolve the issue with the interviewers and resend the data.

006 Household = n women eligible in household schedule (x) inconsistent with total number of eligible women (y)

This message indicates that the total number of eligible women (age 15-49) listed in the household schedule is different from the number assigned to the variable AHWOMEN. The ICDM should open the household questionnaire, review the household roster and check if the data in household schedule contain accurate and eligible ages for all women. Once all household members are properly listed and ages of women correct, the ICDM should advance to the end of the questionnaire and rerun the program. If the error persists, the ICDM should contact the field supervisor to resolve the issue with the interviewers and resend the data.

007 Household = n children eligible in household schedule (x) inconsistent with total number of eligible children (y)

This message indicates that the total number of eligible children (age 0-5) listed in the household schedule is different from the number assigned to the variable AHKIDS. The ICDM should open the household questionnaire, review the household roster, and check if the data in household schedule contains accurate ages for all eligible children. Once all household members are properly listed and ages of children correct, the ICDM should advance to the end of the questionnaire and rerun the listing. If the error persists, the ICDM should contact the field supervisor to resolve the issue with the interviewers and resend the data.

008 Household = n record with data for *module-name*, is missing

This message indicates that the data for an expected module in the questionnaire do not exist. It may be because that the household questionnaire is not yet completed, although the interview result was entered as complete, or final data were not properly transmitted. The ICDM should open the questionnaire and advance to the module with missing data, confirm the data are present, and advance to the end before rerunning the program. If the error persists, the ICDM should contact the field supervisor to resolve the issue with the interviewers and resend the data.

009 Household = n record with data for male/female WEAI: *section-name*, is missing

This message indicates that the data for an expected section in the WEAI module does not exist. It may be because that the WEAI interview is not yet completed, although the interview result was entered as complete, or final data was not properly transmitted. The assigned interviewer should open the questionnaire and advance to the male/female WEAI module with missing data, confirm the data are present, and advance to the end before rerunning the program. If the error persists, the ICDM should contact the field supervisor to resolve the issue with the interviewers and resend the data.

010 Household = n records for anthropometry for women/children (x) inconsistent with individuals eligible for anthropometry (y)

This error occurs when women or children eligible for anthropometry collection have not yet been visited OR have been visited but the information collected has not yet been entered into the respective household by the interviewer that is assigned to Module 4 (women) and Module 5 (children). To resolve, the ICDM must confirm that the anthropometry questionnaire has been completed (including in cases of not present persons or refusals). If data are missing or the error persists, the ICDM should contact the field supervisor to resolve the issue with the interviewer assigned to Module 4 (women) or Module 5 (children) for the household and resend the data.

011 Household = n is an incomplete household (result = x), however there are records that shouldn't exist

This error is produced when an incomplete result code for the household has been entered (codes 2-96), yet data beyond the result code for the interview has been collected. To resolve, review the supervisor’s household assignment to determine the result code reported from the field. The ICDM should open the household interview – a message will appear stating this is a partially completed case and asking if the interviewer would like to go to the last question entered. Choose “No” to go to the beginning of the questionnaire. Reenter the valid incomplete code, and immediately close the questionnaire if it does not automatically do so. If data exist for some modules, the ICDM should contact the field supervisor to confirm the status of the interview and ensure that modules were not left incomplete in error.

012 Household = n assigned to interviewer xxx but collected by interviewer yyy. Household assignment will be updated

This message is a notification to the supervisor that one of the interviewers carried out a module interview that was originally assigned to another team member. The household assignment sheet will update automatically and no changes need to be made. The ICDM can rerun the program to remove the message.

013 Household = n not found in the household sample design file. It needs to be accepted by supervisor

This message shows that the household was not found in the supervisor’s assignment file. The ICDM should confirm any changes the household sample file with the field supervisor before accepting any changes to the sample. In addition, any households added to the sample should be communicated to the [CONTRACTOR] sampling statistician. If the error persists, contact the [CONTRACTOR] Data Management Specialist.

014 The number of households found and registered (x) is correct, but there are still errors that need to be resolved

This is a summary error—if any errors still exist, this message will appear. No specific action is required for this message.

015 Review households found (x) with initial households (y) and those registered in control file (z)

This message indicates that the total number of household questionnaires collected in this cluster is different from the expected defined in the supervisor’s household assignment file or the sample control file. The ICDM should confirm any changes the household sample file with the field supervisor that have not already been recorded during fieldwork. If the error persists, contact the [CONTRACTOR] Data Management Specialist.

016 Household = n one or more mandatory modules are missing

This message indicates that a household questionnaire with a result code of 1 is actually not completed yet, or data were sent from the field that were not finalized. Review the status of all modules in the household questionnaire and complete any missing modules. If everything is complete, advance to the end of the questionnaire and rerun the structure check program. If the error persists, the ICDM should contact the field supervisor to resolve the issue with the interviewers and resend the data.

# Appendix B: Secondary editing guidelines

**A. General guidelines for data editing**

The consistency editing guidelines include instructions on the steps to take to resolve inconsistencies detected during the editing process as well as the action to take if the inconsistencies cannot be resolved through an examination of the responses to other pertinent questions. While editing data, the In-Country Data Manager (ICDM) and secondary editors must review all the pertinent questions involved in an inconsistency before making a change. Changes should never be made hastily without a **thorough review** of all relevant responses.

During the secondary editing phase, the process should be organized to maximize consistency in the correcting process. Care should be taken to ensure that all procedures are followed in a standard manner. Each of the editing rules should be applied in accordance with the guidelines. All computer outputs specifying the errors detected at this stage should be retained.

Secondary editing should begin as soon as possible after the arrival of the questionnaires in the Central Office (via Dropbox) for a particular cluster. The editing process should then be repeated for the cluster until all errors have been addressed.

The rules outlined below for correcting inconsistencies should be applied uniformly to all questionnaires so that errors, which recur across questionnaires, are corrected in a standardized manner. Again, the primary rule of data editing must be observed: *UNDER NO CIRCUMSTANCES SHOULD AN ANSWER BE MADE UP*. Changes can be made only if there is evidence supporting a modification to the response. If an inconsistent value cannot be corrected and a modification is required, the value should be replaced with a code signifying inconsistency (7, 97, 997, etc.). Care should be taken not to confuse the 7 codes (used to replace an inconsistent response that cannot be corrected) with the 8 codes (98, 998, etc.), which are used by the interviewer to indicate that the respondent answered a question with the statement, “I don’t know.”

Finally, the editing phase necessitates close involvement of [SURVEY SUBCONTRACTOR] survey management, [SURVEY CONTRACTOR] survey management and data processing staff, and other senior project staff. Consequently, the data editing at this phase should be organized to encourage close interaction between the survey technical staff and the data processing personnel.

**B. Specific guidelines for data editing**

This section provides specific guidelines to apply in correcting inconsistencies that will be detected at the secondary editing stages. These general principles should be observed:

* When making corrections at this stage, the data editing staff must examine the questionnaires in which errors are detected and then check for the inconsistency in other parts of the questionnaire, not just by looking at the error listing.
* All pertinent questions must be examined before a decision is made regarding the manner in which inconsistent responses will be corrected. The data editing staff should never correct errors hastily, such as by entering a code 97, where it is possible to determine the correct response.
* The guidelines for correcting errors detected at this stage should be applied in a standardized manner. Uniform rules must be followed in making decisions to change responses that are inconsistent. Data editing decisions that are not covered by the following guidelines should be documented.
* The data editing staff should be careful to make changes only when there is evidence to support the new response. Again, the chief rule of data editing must apply: *UNDER NO CIRCUMSTANCES SHOULD AN ANSWER BE MADE UP.*
* Another basic data editing rule is to change the fewest possible responses. This is reasonable; the more data values that need to be changed to “fix” the error, the less likely it is that data editing staff are making the right decisions.
* Particularly in computer-assisted personal interviewing surveys, many errors cannot be resolved. Although every effort should be made to resolve error messages, it is important to recognize when there is not sufficient support to make a change to the data. Many errors cannot be resolved.

The task of resolving inconsistencies is frequently quite difficult. The data editing staff should consult with senior members of the survey staff for assistance in resolving difficult problems. Senior staff should also be sure to regularly review the work of the editing staff to ensure that they are observing the above principles.

**C. Data error messages and guidance**

Guidance for all possible secondary editing error messages is shown below. See Section 5.2 for further explanation. Error messages numbers are grouped according to module the error applies. This allows for additional checking and error messages customized to each survey. Therefore, not all error message numbers are represented in the manual that follows. For example, error messages 0010 – 0020 all apply to the household schedule, and messages 0021 – 0029 are available for additional editing messages.

**0001 Failed to update the CONTROL file for cluster n**

It was not possible to update the CONTROL file to indicate that a step of data processing had been performed on the cluster data file. This problem can occur if the CONTROL file is write protected or if the cluster number is missing from the CONTROL file. Ensure that the correct access rights have been given to this file and that the CONTROL file includes the cluster number. In network or multitasking environments, this message can also mean that the file is already in use by another program or in another window. If the problem persists, contact the [SURVEY CONTRACTOR] data management specialist for assistance in resolving the issue.

**0002 Count of questionnaires for cluster n (n:n:n) does not match CONTROL file (n:n:n)**

The CONTROL file contains the counts of questionnaires by result code based on the planned sample file. If this message is produced, the ICDM must check the field supervisor’s notes and data transmission to confirm the number of households found in the cluster. This error must be rectified before further processing can be attempted. Often this error occurs if a household included in the sample was found to be two or more households, resulting in added households to the sample.

**0003 Cannot open CONTROL file for cluster n**

See message 0001.

**0010 Only one person can be a primary decisionmaker with relationship code 01 (n were found)**

Although there are up to two primary decisionmakers in a household, only one can appear on the first line in the household schedule with code 01 for the relationship code (V103). The relationship of remaining household members is defined based on the male primary decisionmaker, or if no male primary decisionmaker, the female primary decisionmaker will be the reference person for relationship codes. Upon review of the entire schedule and the relationships between the other household members, the data editing staff should be able to identify who is the correct primary decisionmaker being used as a reference for the relationship codes and correct the other household members’ relationship code to the likely code (for example, the male and female primary decisionmakers could have both been given a relationship code of 01). If the likely relationship code of the second primary decisionmaker cannot be determined, change the relationship code (V103) to ‘98’ (don’t know).

**0011 No more than one spouse/parent/parent-in-law of each sex is allowed in the household**

In most households, there is one primary male decisionmaker, one primary female decisionmaker/spouse, one male parent of the head, one female parent, one male parent-in-law and one female parent-in-law. In some households, particularly polygynous households, this rule may not hold, and the secondary editing programs account for the potential of polygyny if it exists in the country.

Check the relationship codes of each member of the household and attempt to identify the correct relationships. If it is not possible to correct the problem or it is clear that the rule does not hold for this household, then no further action is necessary.

**0012 Primary decisionmaker (V101A(male) = x V101B(female) = y) sex/relationship information incorrect: V102(1) = n V103(1) = n V102(2) = n V103(2) = n**

The male primary decisionmaker (if one exists in the household) should be found on line 01. For CAPI surveys, if there is no male primary decisionmaker, then the female primary decisionmaker should be found on line 01; otherwise, in PAPI surveys and in all households where there does exist a male primary decisionmaker, the female primary decisionmaker will be on line 02. If either of the line numbers is incorrect according to the questions about decisionmakers prior to the household schedule, this error appears.

Review the responses to V101A and V101B, alongside the household roster. If there are no 18+ household members for each gender who qualify as primary decisionmakers, change V101A and/or V101B. Review Module 6 for respondents to clarify the presence of adult decisionmakers and confirm the roster and V101A/V101B. Ensure consistency among V101A, V101B, the household roster, and the WEAI modules. It is important to correctly identify the primary adult decisionmakers for the survey in each household.

**0013 Eligibility of member incorrect: Sex (V102 = x), Age (V104 = y), Eligibility (woman/child = z)**

To be eligible for the individual interview in the women’s nutrition and anthropometry module, a household member must be a woman 15 to 49 years of age. To be eligible for the children’s nutrition and anthropometry module, a child must be 0 to 5 years of age. If a household member is eligible, then the code entered for eligibility should be the same as the line number of the member (M1\_LINE)—otherwise, the eligibility code must be 00.

The eligibility of household members is checked at data entry to ensure that all eligible members and only those who are eligible are included in Modules 4 (women) and 5 (children). If this is a woman (V102 = 2), then ensure that her age (V104) is between 15 and 49 years inclusive. If this is a child, then ensure that his or her age (V104) is between zero and 5 years.

If all of these conditions are correct, then the household member is eligible and the eligibility code (V107/V108) must be the same as the member’s line number (M1\_LINE). If this member is eligible, the member must have an individual column in the appropriate module. If any of the conditions is not met, then the eligibility code must be 00 and no individual data should exist for this person in the corresponding modules. Check for the existence of an individual’s data in Modules 4 and 5 before determining which piece of information is incorrect.

If this error appears during editing, then not only should the household schedule be corrected, but the individual’s case may need to be added or deleted from the data file modules, depending on the source of the error.

**0014 Age (V104 = n) of primary decisionmaker or spouse is under x years**

The primary decisionmakers and/or the spouse of the primary decisionmakers are expected to be adults. A minimum age of 18 is used in checking the age of the primary decisionmakers, and age 12 is used as a minimum in checking the age of the spouse in a case in which the spouse is not also a primary decisionmaker.

During editing, if there is no obvious correction, then the data should be left as originally recorded.

**0015 Check primary decisionmaker’s age (line x, V104 = y) with his/her parent’s age (line x, V104 = y)**

It is expected that the parents of the primary decisionmakers would be at least 12 years older. Check other modules in the questionnaire for differences in ages. For example, if the parent is a woman who has been included in Module 4, see if the woman’s age matches the household roster. If not, then make no correction to the data.

**0016 Check spouse’s age (line x, V104 = y) with his/her parent's age (line x, V104 = y)**

It is expected that the parents of the primary decisionmaker’s spouse would be at least 12 years older. Check other modules in the questionnaire for differences in ages. For example, if the parent is a woman who has been included in Module 4, see if the woman’s age matches the household roster. If not, then make no correction to the data.

**0017 Sex (V102 = x) of spouse (line y) must be different from sex (V102 = x) of primary decisionmaker (line y)**

The sex of the spouse should be different from the sex of the primary decisionmaker. The first primary decisionmaker should be listed as the first member in the household.

The relationship code, sex, and age of each member should be checked. In many cases, a relationship code of 02 (spouse) has been recorded on the questionnaire instead of code 03 (child), and this can usually be clearly seen from the ages and line numbers of consenting parents for Module 5 (children’s nutrition and anthropometry). In other cases, the names of the members can often be used to check the sex of the members. Correct either the sex or the relationship of one of the members.

**0018 Number of eligible women/children incorrect: Expected x, found y**

The number of eligible respondents in the household schedule must equal the number of eligible respondents stated on the household questionnaire cover sheet. This error message primarily applies to PAPI surveys, but is checked in CAPI surveys as well to ensure the data has not been corrupted.

If these numbers are not in agreement, the household schedule should be reviewed. Each member in the household schedule should be checked regarding their sex, age, and marital status (if applicable). A column in the nutrition and anthropometry modules should exist for each eligible woman and child. To correct the mistake, either the total number of eligible women or children field on the cover sheet should be corrected, or the entries in the household schedule should be adjusted (for example, if the interviewer considered ineligible household members eligible for interview).

**0019 Education of member (V109 = x V111 = y) inconsistent with reading/writing ability (V112 = z)**

If the household respondent reports that a member went to secondary school or higher, it is unlikely that the member is unable to read or write. If the member can read or write, it is unlikely that the member did not attend school. Review ages to comply with school attendance and reading or writing ability. If no apparent mistake was made that can be resolved, make no change to the data.

**0020 No primary male/female decisionmaker found, but eligible member found. Review household roster.**

Identification of primary adult decisionmakers is very important to the survey results. Nearly all households will include at least one adult decisionmaker, and many will have both a male and a female decisionmaker. This message appears when there are members of the household, either male or female, who are 18 years of age or older but no primary male or female decisionmaker was identified. Review the household roster and the anthropometry module (for women) to determine whether household eligibility is correct. Because of the importance of this measure, if eligible primary decisionmakers are found but were not interviewed, immediately alert survey management of potential data collection issues.

If there is a Module 6 interview available, correct V101A/V101B accordingly. Otherwise, make no change to the data and note the error on the ICDM problem log.

**0030 Weight woman/child = n outside range expected for age = x, sex = y**

The weight of the woman or child is outside the expected range. For children, the range is dependent on the child’s age in months and the sex (the exact limits are provided in Appendix C). For women, the range is 20–150 kilograms.

Often this error occurs as a result of incorrect ordering of women or children in the anthropometry modules. Data for women should be entered in Module 4 and data for children should be entered in Module 5 in the order in which they appear in the household roster. If the data are not entered in this order, often the weights will be out of range. It is essential to enter the data in the correct order and therefore, reordering of the data may be necessary. To be sure that re-ordering is necessary, confirm the names listed with the line numbers in both the anthropometry modules and the household roster and make a change only if certain.

**0031 Height (length) woman/child = n outside range expected for age = x, sex = y**

The height or length of the respondent or child is outside the expected range. For children, the range is dependent on the child’s age in months and the sex (the exact limits are given in provided in Appendix C). For women, the range is 100–200 centimeters.

Often this error occurs as a result of incorrect ordering of women or children in the anthropometry modules. Data for women should be entered in Module 4 and data for children should be entered Module 5 in the order in which they appear in the household roster. If the data are not entered in this order, often the weights will be out of range. It is essential to enter the data in the correct order and therefore, reordering of the data may be necessary. To be sure that re-ordering is necessary, confirm the names listed with the line numbers in both the anthropometry modules and the household roster and make a change only if certain.

**0032 Second decimal for weight (AN518 = x) different from 0.5**

Weighing scales used in the survey are accurate to .0 or .5 on the final decimal. Check the data entry against the paper anthropometry forms, but if no error in keying was made, make no change to the data.

**0033 Children under 2 years are usually measured lying down, children 2+ measured standing up: Measured AN517 = x, age = y**

In measuring the height or length of children, the measurers are instructed to measure children less than 2 years lying down and children 2 years and over standing up. The code entered for this question does not agree with this rule.

Check the data entered from the paper anthropometry questionnaire, but make no correction unless a keying error was made.

**0040 Neither date of birth (V401=mm/yyyy) nor age (V402=n) specified**

The age of the respondent is one of the most important pieces of information in the questionnaire. For this reason, either the age or the year of birth are required for all respondents.

If no information has been recorded for either of these questions, then check other related pieces of information to see whether the age of the respondent can be deduced. If none of the related information gives any clear idea as to the correct date of birth or age of the respondent, use the respondent’s age as reported in the household schedule (V104) in V402.

**0041 Respondent’s age in individual questionnaire (V402=n) differs from age in household (V104(line = x) = y)**

The age of the respondent in the household schedule is not consistent with the response to V401/V402.

Check the paper anthropometry section and if no typing mistake has been made, then verify that the women are listed in their anthropometry section in the same order of listing as given in the household schedule. If not correctly ordered, the order must be corrected in the dataset.

If the ordering of respondents is correct, check the Module 6-WEAI if the respondent also is a primary decisionmaker and provided an age during that module. If the age given in Module 6-WEAI matches that of V104 or V402, make a correction to the incorrect piece of data. Otherwise, leave the ages unchanged.

**0042 Date of birth (V401=mm/yyyy) is out of range with respect to the interview date (mm/yyyy), range=mm/yyyy-mm/yyyy**

A respondent must be between 15 and 49 completed years to be eligible for anthropometry data collection. Depending on the date of interview, this translates into a minimum and maximum possible date of birth for the respondent. Occasionally a date of birth recorded is outside of this range.

Check that all information relating to the date of birth on the paper anthropometry questionnaire for the respondent is correct (i.e., no keying mistakes).

If the date of birth is 1 month before the minimum date of birth, and the age of the respondent is recorded as 49, then change the month of birth of the respondent to the month that follows what was recorded. This will ensure consistency of information without dropping the respondent from the sample.

If the respondent was clearly born outside the expected range, then the respondent should be dropped from the sample due to ineligibility. Correct the age and eligibility in the household schedule to reflect the correct age of the respondent.

**0050 Neither date of birth (V6101 = mm/yyyy) nor age (V6102 = n) specified**

The age of the respondent is one of the most important pieces of information in the individual data file. For this reason, either the age or the year of birth is required for all respondents.

If no information has been recorded for either of these questions, then check other related pieces of information to see whether the age of the respondent can be deduced. The information to check includes the following:

* Respondent’s age in the household schedule (V104)

If none of the related information gives any clear idea as to the correct date of birth or age of the respondent, then use the respondent’s age as reported in the household schedule (V104) in V6102.

**0051 Respondent’s age in individual questionnaire (V6102 = n) differs from age in household (V104 = n)**

The respondent’s age in the WEAI questionnaire (V6102) and the age given for her in the household schedule (V104) should be the same, but discrepancies can occur because the household schedule is often reported by a different member of the household.

If the age of the respondent has been recorded differently in the household schedule and the WEAI questionnaire, but both have been correctly entered, then the information should be left as recorded, **unless it affects her eligibility for interview.** Check the age of the respondent in the anthropometry section, if applicable, and try to determine the most likely age. If it is still not certain which of the three ages is correct, make no change.

Cases in which the age of the respondent is significantly different from the age recorded in the household schedule should be carefully reviewed. However, unless there has been a keying error, or the age of the woman in the household schedule makes her ineligible for interview, then the ages should not be changed.

**0052 Date of birth (V6101 = mm/yyyy) is out of range with respect to the interview date (mm/yyyy), range=mm/yyyy-mm/yyyy**

See message 0042.

**0053 Date of birth (V6101 = mm/yyyy, cmc=cmc-cmc) and age (V6102 = n) inconsistent with date of interview (mm/yyyy, cmc=cmc)**

The age calculated based on the month and year of birth of the respondent (V6101) must be equal to the respondent’s age in completed years (V6102).

In resolving inconsistencies in the calculated and reported ages, it may be necessary to consider other information in the questionnaire, including the following:

* The age recorded for the respondent in the household listing (V104)
* The age recorded for the anthropometry module (V402)

If, after reviewing all other relevant items of information, the inconsistency cannot be resolved, there are two ways of correcting the data. If the inconsistency is 1 year, correct the year of birth if the month is given; otherwise, correct the age. If the inconsistency is greater than 1 year, choose the age or date of birth, whichever appears more correct (i.e., is most consistent with other pieces of information). If the age is chosen, change the year of birth to 9997 and the month to 97. If the date of birth is chosen, change the age to 97.

Note that in many cases, the difference between the calculated and reported ages will not be large. Often the error occurs because the respondent reported her age at her next birthday (especially if the date is close to the interview date), rather than at her last birthday.

**0054 Month of birth (V6101M=n) inconsistent with other information: cmc=cmc-cmc [mm/yyyy-mm/yyyy]**

See message 0053.

**0060 Decisionmakers for WEAI (x) include partner/spouse, but respondent not married/living with partner (V6105 = y)**

The respondent reported in the specified question that the decisions included his or her spouse or partner but responded to V6105 as not married or living with a partner.

Check the similar questions in the same section of the module, but if no apparent pattern of responses exists suggesting an error in data entry, make no changes to the question.

**0070 Time at start of module (Module n = hh:mm) after time at end of module (Module n = hh:mm)**

The time recorded at the start of the interview for a particular module must be earlier than the time recorded at the end of the module.

If the time recorded at the start of the module is equal to or later than the time recorded at the end of the module, check the time recorded for the start and end of other modules conducted by the interviewer on the same date to see whether the inconsistency can be resolved. If the inconsistency cannot be corrected, change the values recorded for hours and minutes at the end of the interview to 97.

**0071 Activity duration (x.y hours) seems unlikely - please check**

Duration of activity exceeds expected length of time. Check the full time allocation to determine whether a data entry error was made. If no obvious error occurred, make no change to the data.

**0072 Sleeping and resting (A - hh:mm) with secondary activity. Review full time allocation.**

It is expected that no other activity can be recorded while sleeping. Check the time allocation surrounding the error and determine whether a keying error was made. If no obvious error occurred, remove the secondary activity if it is clear that sleeping was occurring (i.e., the surrounding time block is all sleeping). If it is the beginning or end of the sleeping time, make no change to the data.

**0073 Sleeping and resting (A) recorded as secondary activity in hour hh. Review full time allocation.**

See message 0072.

**0074 Travel for non-commuting reasons found before/after work or school activity**

There are two type of travel available to choose from for time allocation—one for standard commuting to and from work and one for all other travel. This error occurs when the “other” travel type is chosen immediately before or immediately after work or school activities. Check the full time allocation for further information, and if no other activity indicates “other” travel, change the travel code to a “commuting to and from work or school” travel code.

**0075 Typically daytime activities recorded during night hours at hh. Please check.**

Activities recorded during night-time hours seem unlikely. Check surrounding allocation, as well as full allocation, but if no obvious error was made, make no change to the data.

**0090 Minimum and maximum date of event n cross over: cmc=cmc-cmc [mm/yyyy-mm/yyyy]**

After all preparation for imputation, the minimum and maximum dates of an event have crossed over so that the minimum date of the event is now greater than the maximum date of the event. Check dates of birth, ages, and dates of interview along with accompanying messages. After resolving other messages, this message should clear. If it does not, contact the [SURVEY CONTRACTOR] data processing specialist.

**0091 Event n: Month of event (n) inconsistent with final range for event: cmc=cmc-cmc [mm/yyyy-mm/yyyy], imputed=cmc [mm/yyyy]**

See message 0053.

**1050 Neither date of birth (M6101 = mm/yyyy) nor age (M6102 = n) specified**

The age of the respondent is one of the most important pieces of information in the individual data file. For this reason, either the age or the year of birth is required for all respondents.

If no information has been recorded for either of these questions, then check other related pieces of information to see whether the age of the respondent can be deduced. The information to check includes the following:

* Age of the respondent in the household schedule (V104)

If none of the related information gives any clear idea as to the correct date of birth or age for the respondent, use the respondent’s age as reported in the household schedule (V104) in question M6102.

**1051 Respondent’s age in individual questionnaire (M6102 = n) differs from age in household (V104 = n)**

The respondent’s age in the individual questionnaire (M6102) and the respondent’s age in the household schedule (V104) should be the same, but discrepancies can occur because a different member of the household often reports in the household schedule.

If the age of the respondent has been recorded differently in the household schedule and the WEAI-M, but both have been correctly entered, then the information should be left as recorded, **unless it affects the man’s eligibility for the individual interview.**

Cases in which the age of the respondent is significantly different from the age recorded in the household schedule should be carefully reviewed. Unless there is a clear error that can be resolved, however, leave data as entered.

**1052 Date of birth (M6101 = mm/yyyy) is out of range with respect to the interview date (dd/yyyy)**

See message 0042.

**1053 Date of birth (M6101 = mm/yyyy, cmc=cmc-cmc) and age (M6102 = x) inconsistent with date of interview (mm/yyyy, cmc=cmc)**

See message 0053.

**1054 Month of birth (M6101M = mm) inconsistent with other information: cmc=cmc-cmc [mm/yyyy-mm/yyyy]**

See message 0053.

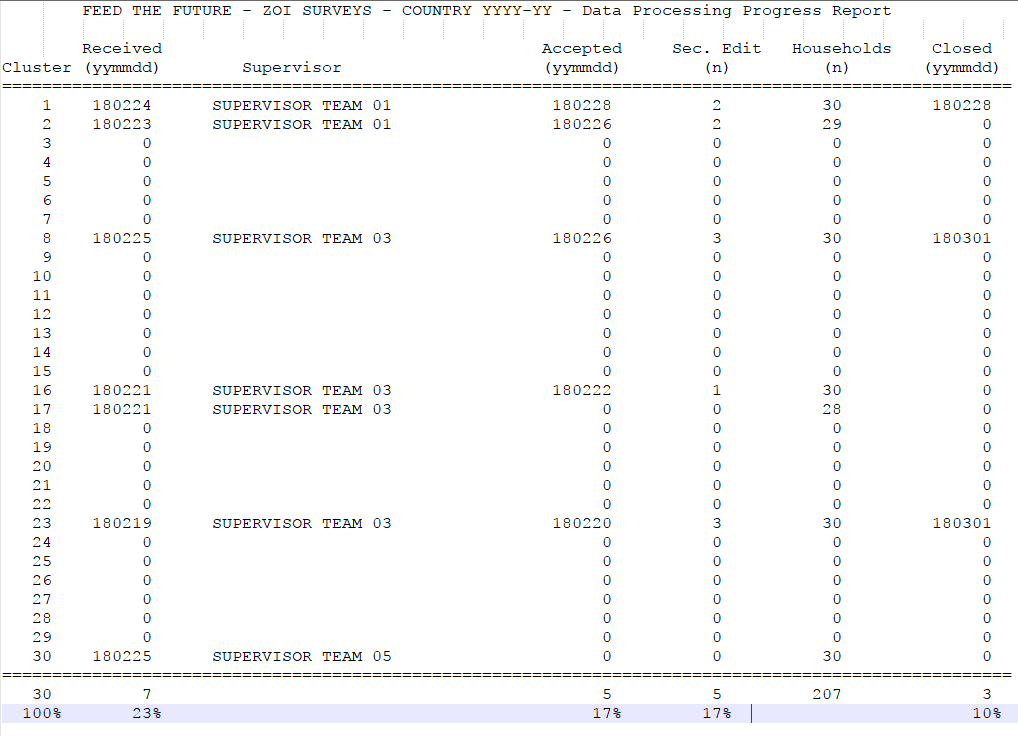
# Appendix C: Limits for length and weight of children

In editing the length and weight of children to ensure that no data entry errors are made, the following values are used as the minimum and maximum expected values. The ranges are dependent on the sex and age of the child and are given in centimeters for the length (height) of the child and kilograms for the weight of the child.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Age in months** | **Length (cm.)** | | | | **Weight (kg.)** | | | |
| **Males** | | **Females** | | **Males** | | **Females** | |
| **Minimum** | **Maximum** | **Minimum** | **Maximum** | **Minimum** | **Maximum** | **Minimum** | **Maximum** |
| 0–2 | 36.0 | 74.0 | 36.0 | 72.0 | 0.5 | 10.0 | 0.5 | 9.0 |
| 3–5 | 45.0 | 83.0 | 44.0 | 80.0 | 1.0 | 13.0 | 1.0 | 12.0 |
| 6–8 | 51.0 | 87.0 | 50.0 | 86.0 | 2.0 | 15.0 | 2.0 | 14.0 |
| 9–11 | 56.0 | 91.0 | 54.0 | 90.0 | 3.0 | 16.5 | 2.5 | 15.5 |
| 12–14 | 59.0 | 96.0 | 57.0 | 95.0 | 4.0 | 17.5 | 3.0 | 16.5 |
| 15–17 | 62.0 | 100.0 | 60.0 | 99.0 | 4.0 | 18.5 | 3.5 | 17.5 |
| 18–20 | 64.0 | 104.0 | 62.0 | 102.0 | 4.0 | 19.5 | 3.5 | 18.5 |
| 21–23 | 65.0 | 107.0 | 64.0 | 106.0 | 4.5 | 20.5 | 4.0 | 19.5 |
| 24–26 | 67.0 | 108.0 | 66.0 | 107.0 | 4.5 | 23.0 | 4.5 | 21.5 |
| 27–29 | 68.0 | 112.0 | 68.0 | 111.0 | 5.0 | 24.0 | 5.0 | 23.0 |
| 30–32 | 70.0 | 115.0 | 69.0 | 114.0 | 5.0 | 24.5 | 5.0 | 24.5 |
| 33–35 | 71.0 | 118.0 | 71.0 | 117.0 | 5.0 | 25.5 | 5.0 | 25.5 |
| 36–38 | 73.0 | 121.0 | 72.0 | 120.0 | 5.0 | 26.0 | 5.0 | 27.0 |
| 39–41 | 74.0 | 124.0 | 74.0 | 122.0 | 5.0 | 27.0 | 5.0 | 28.0 |
| 42–44 | 75.0 | 127.0 | 75.0 | 124.0 | 5.0 | 28.0 | 5.5 | 29.0 |
| 45–47 | 77.0 | 129.9 | 77.0 | 126.0 | 5.0 | 29.0 | 5.5 | 30.0 |
| 48–50 | 78.0 | 132.0 | 78.0 | 129.0 | 5.0 | 30.0 | 5.5 | 31.0 |
| 51–53 | 79.0 | 134.0 | 79.0 | 131.0 | 5.0 | 31.0 | 5.5 | 32.0 |
| 54–56 | 80.0 | 136.0 | 81.0 | 133.0 | 5.5 | 32.0 | 6.0 | 33.0 |
| 57–60 | 82.0 | 139.0 | 81.0 | 136.0 | 5.5 | 33.0 | 6.0 | 34.5 |
| 61–72 | 82.0 | 140.0 | 81.0 | 137.0 | 5.5 | 34.0 | 6.0 | 36.0 |

# Appendix D: Cluster status report

An example of a Cluster Status Report for a survey that has 30 clusters in total.

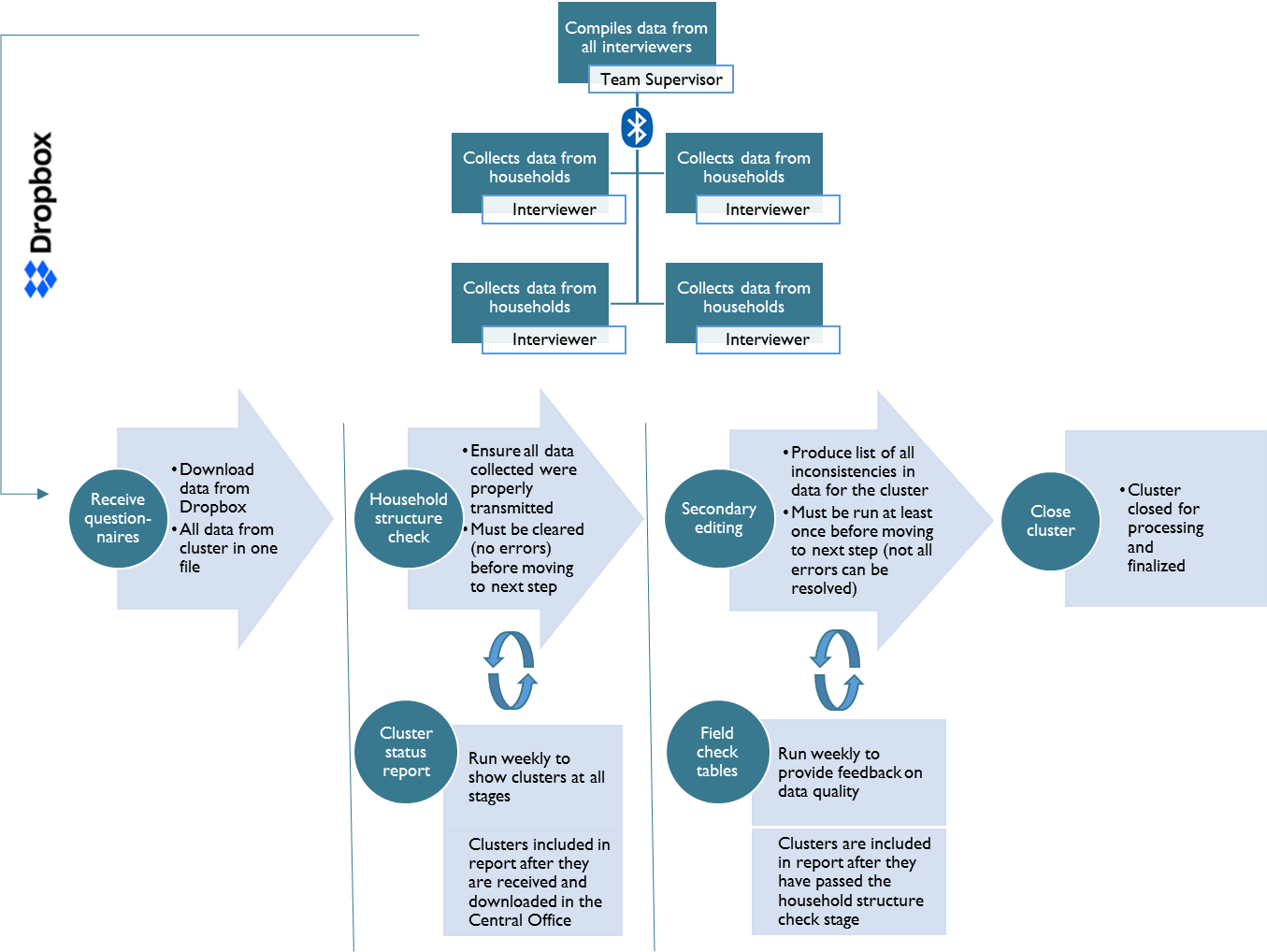


# Appendix E: Data issues log

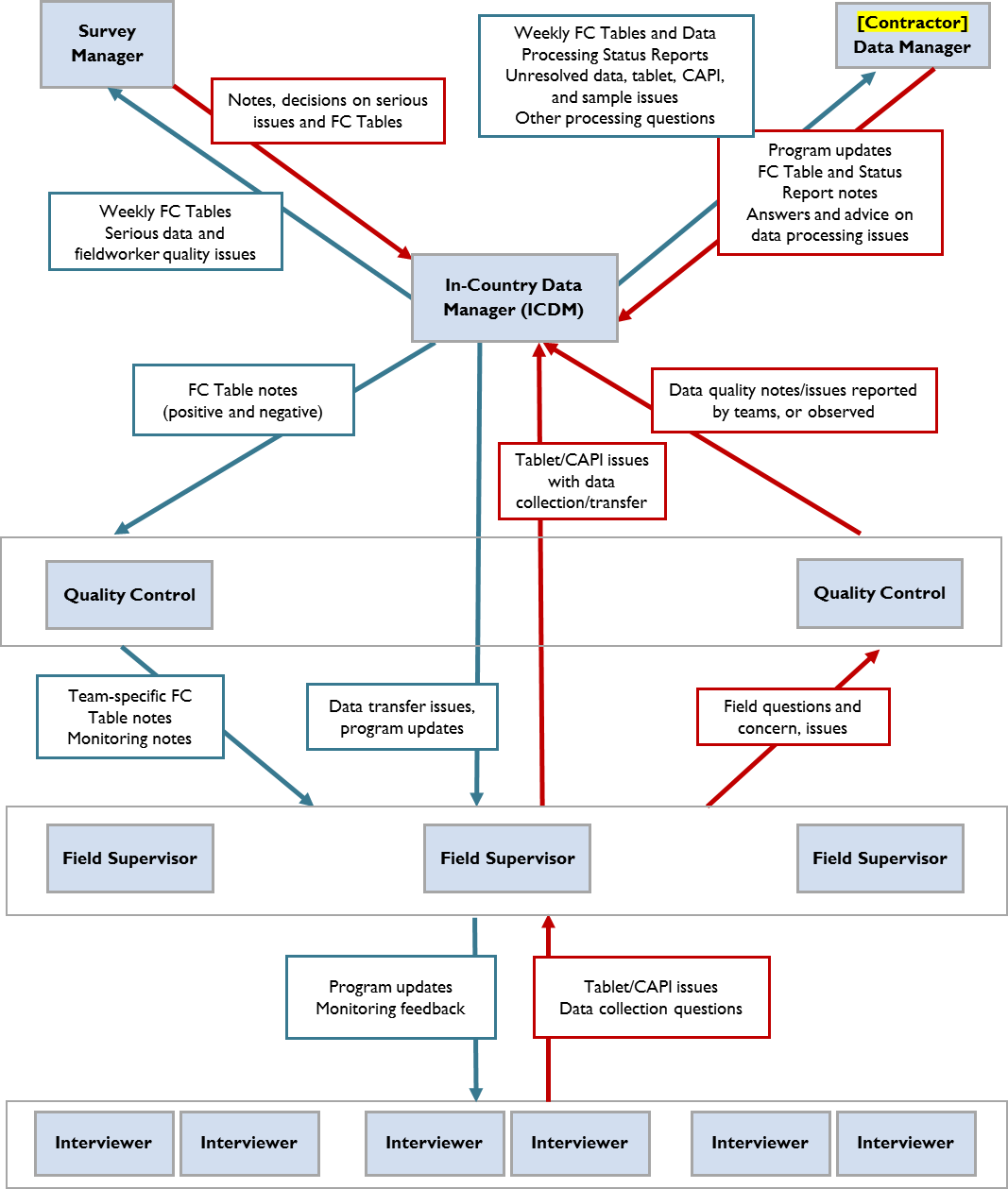
**Table E1: Data Issues Log**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cluster Name** | **Household** | **Question/ Variable** | **Description of Problem** | **Resolved** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

# Appendix F: Data processing diagram



# Appendix G: Data processing communications diagram



# Appendix H: CSPro variable names and labels

[INSERT CSPro VARIABLE NAMES AND LABELS]