Feedback from group discussion on status of data management status

6 October 2010

Ethiopia

Current status

• Data management plan and SOPs developed but not yet finalised. About one week’s work needed to finalise after taking into account results of pilot testing. Aspects to finalise include logistics and the use of the software.

• Full-time Data Manager is in place, with two other data managers on hand (from WHO/Federal Ministry of Health and from the Ethiopian Nutrition and Health Research Institute) to provide support when needed. Each field team also has one data checker.

• Software chosen for data entry and storage is CS-Pro ([http://www.census.gov/ipc/www/cspro/](http://www.census.gov/ipc/www/cspro/)). This will be used at the central data management unit. Chosen because of existing experience in National Statistical Authority (who have provided training and will provide help and support), and its adoption by the Ethiopian Nutrition and Health Research Institute.

• Data entry will be continuous during the survey. The software has been tested using test data, and will soon be tested further by using it to enter data collected during the pilot cluster (week of 3 – 9 October 2010).

• Confidentiality is primarily concerned with paper forms used for data capture where names and other personal identifiers are recorded. The electronic database will not contain personal identifiers, nor any free text, and will be based around survey participant identification numbers.

• Data quality assurance will be handled at different stages, for example:
  - in the field by a data checker
  - at data entry using logical and consistency checking programmed into data entry screens
  - after data entry by clerks checking forms against electronic screens
  - double-data entry for all identified TB cases

Difficulties/challenges

• How best to ensure against loss of forms as they are transported from the field to the central data management unit.

Possible solutions

• Photocopying forms before they are sent
• Scanning forms before they are sent
• Using serial numbers on each form and establishing a tracking mechanism to make sure all forms can be accounted for.

**Kenya**

*Current status*

• Protocol has not yet been finalized, so data management plan and SOPs have not been developed yet. Intending to start development in November 2010.
• Currently have joint statistician who is also acting as a data manager, but planning to appoint a full-time central data manager plus a data manager for each of the four field teams
• The data store will most likely be a central SQL-Server database, with each survey team having a local database. This system will build on existing in-country expertise in using such a system, plus support from CDC and KNCV. (??? Software for data entry screens and data transfer???)
• Data entry will be directly in the field, based on Personal Digital Assistants (PDAs) used by survey field team members. Data will be uploaded each day from PDAs to the field database and, depending on local conditions, from the field database to the central database. Piloting will be used to establish the frequency of uploads when daily uploads are not possible.
• Confidentiality: bar codes will be used for data capture in the field. Access to the database, and especially to the master list matching personal identification numbers encoded in bar codes to personal information such as name, address, etc, will be strictly controlled.
• The decision to use both PDAs and bar codes is based on extensive experience of using these technologies in other projects within the country.
• Data quality assurance will be handled at different stages, for example:
  o at data entry using logical and consistency checking programmed into the PDA data entry screens
  o at end of day by field data managers checking automated reports against manual tally of participants
  o running queries etc to identify data inconsistencies

Note: The team considered using scannable forms for data entry (for example the symptoms screening/interview), but settled instead on PDAs. Other countries may wish to consider scannable forms.

*Difficulties/challenges*

• Managing data transmission from field to central database
• Managing database stability and availability, eg using technologies such as database mirroring

*Possible solutions*

(None identified during discussions)
**Rwanda**

*Current status*

- Data management plan and SOPs not yet developed. Will start development in November, with assistance from KNCV.
- Candidates for position of data manager were interviewed last week with the aim of having a full-time data manager in place before work begins to develop the data management plan.
- The software to use for data entry and storage has not yet been chosen, but it will probably be either Epi-data, Epi-info or CS-Pro, as there is local expertise in using them.
- Planning to use paper-based forms in the field, with electronic data entry at the central data management unit. All processes will be tested and piloted, and data entry will be continuous as soon as forms start arriving at the data management unit.
- Confidentiality issues have not yet been addressed; approach is likely to be similar to that in Ethiopia (ie with no personal information entered into the electronic database).
- Quality assurance options: checking by team leader in the field; double-data entry of a fraction of forms to assess rate of data transcription errors; intermediate data analyses.

*Difficulties/challenges*

- Identifying and appointing a full-time experienced data manager.

*Possible solutions*

- Establishing partnership with the Institute of Public Health to get local help with data management.