10.5: Processing of census data

Most censuses involve the collection of substantial amounts of data. It is important to plan how these data will be processed, before the study is started. Usually, it will be desirable that the information is either entered electronically on collection into a PDA, tablet computer, or mobile phone or is entered into a computer shortly after it is collected, so that a large backlog of work does not accumulate. Rapid data entry and checks for transcription errors are especially important if the information collected at the census is to be used to produce forms for the recording of additional procedures to be performed on the trial participants shortly after the census. Furthermore, once the information is in a computer, consistency checks can be conducted, and errors or queries referred back to the relevant fieldworkers. Such feedback should occur as soon as possible after the original information has been collected.

In recent years, there have been major advances in computer systems and transmission of digital information, enabling the collection, processing, and checking of data virtually anywhere, eliminating many of the bottlenecks that, until quite recently, so slowed analysis of trial data in LMICs. Nevertheless, data management is a major task in all field trials and requires a well-defined data management strategy, as discussed in Chapter 20. The design of the recording system may need to allow for changes in the composition of the study population over time, due to in-and-out migration, and for movement between households within the population. It is usually desirable to seek help and guidance from an experienced statistician or data analyst for these aspects. This should be done at the start, rather than in the middle, of a study.