Chapter X
Quality assurance in surveys:
standards, guidelines and procedures

T. Bedirhan Üstun, Somnath Chatterji, Abdelhay Mechbal and Christopher J.L. Murray

On behalf of the World Health Survey (WHS) Collaborators [*]

World Health Organization
Evidence and Information for Policy
Geneva, Switzerland

Abstract

The quality of a survey is of prime importance for accurate, reliable and valid results. Survey teams should implement systematic quality assurance procedures to prevent unacceptable practices and to minimize errors in data collection. Establishment of effective and efficient strategies towards improvement of the quality of a survey will help achieve the timely collection of high-quality data and the validity of the results. "Quality assurance" may also be viewed as an organizing tool for implementation with pre-defined operational standards regarding the structure, process and outcome of the survey. Survey teams should adhere to explicit standards of quality and follow prescribed procedures to achieve such standards. The procedures should be transparent, systematically monitored and carefully reported as part of the general documentation of the survey implementation and results. It is also important that the survey be measured and summarized by quantifiable indicators, to the extent practicable.

The present chapter outlines a systematic approach to achieving quality assurance measures, going beyond simple control mechanisms. A large international survey -- the World Health Survey (WHS) implemented by multiple survey institutions in 71 different countries-- is used to illustrate the elaboration of the application of a total quality assurance programme. This survey was designed to gather comparable data to assess the different dimensions of health systems in participating countries using nationally representative samples. In accordance with the importance of the results of the WHS, rigorous quality assurance procedures were put in place utilizing international experts who were assembled to serve as an external peer review group and to support countries in achieving commonly agreed and feasible quality standards with regard to such matters as: sample selection methodology, achievement of acceptable response rates, treatment of missing data, calculation of measures of reliability and checks for comparability of the data across population subgroups and countries.

Key terms: quality assurance, quality indicators, World Health Survey, missing data, response rate, sampling, reliability, cross-population comparability, international comparisons.

---

[*] The WHS Collaborators are listed in full on the WHS web site: (http://www.who.int/whs/).

A. Introduction

1.   One of the basic features in respect of the design and implementation of a survey is the survey's "quality" (Lyberg and others, 1997). In every data-collection initiative, the results depend on the input; as the saying goes: garbage in-garbage out. In addition to the quality of the survey instruments and analytical techniques, the quality of the survey results depend mainly on the implementation of the survey including sound sampling methods and proper administration of the questionnaire.

2.   To achieve maximum quality, every survey team should adhere to a standard set of guidelines on survey implementation. These guidelines identify the following:

> (a)   Quality standards that need to be adhered to at each step of a survey;

> (b)   Quality assurance (QA) procedures that identify the explicit actions to be taken for monitoring the survey implementation in actual settings;

> (c)   Evaluation of the quality assurance process that measures the impact of quality assurance standards on the survey results and procedures towards improving the relevance and efficiency of the overall quality assurance process (Biemer and others, 1991).

3.   The overall aim of the guidelines is to provide support to improving quality rather than to audit the survey implementation. Since any survey is a large investment involving multiple parties with important results that have influence on the policies of a nation, it is essential that quality be a serious operational focus. Quality assurance is seen as an ongoing process throughout the survey from preparation and sampling through data collection and data analysis to report writing. The guidelines also aim to ensure a better understanding of the design of the survey among users. The purpose of establishing standard procedures is to help ensure that:

- The data collection is relevant and meaningful for the country's needs

- The data can be compared within a country and across countries to identify the similarities and differences across populations

- The practical implementation of the survey follows accepted protocols

- The errors in data collection are minimized

- The data-collection capability is improved over time

B. Quality standards and assurance procedures

4.   Quality assurance (Statistics Canada, 1998) is defined as any method or procedure for collecting, processing or analysing survey data that is aimed at maintaining or enhancing their
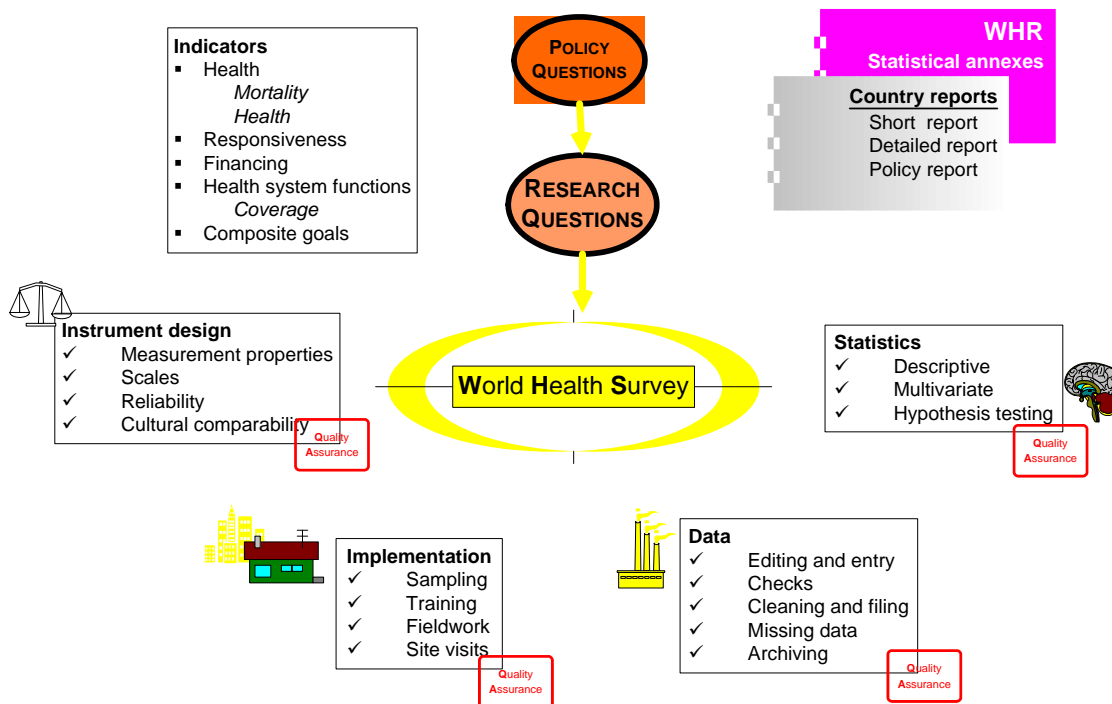
reliability or validity. Quality assurance could be understood as having similar yet differing meanings. In the present chapter, we utilize the total quality management paradigm that examines the survey process at each step and try to outline an approach not only to reducing sampling and non-sampling errors but also to improving the relevance and feasibility of the survey as well as the capacity of the country to implement surveys. To achieve this aim, yet remain practical, this chapter will make use of the World Health Survey (WHS) quality standards and assurance procedures (World Health Organization, 2002) referring to all the steps including:

- Selection of survey institutions
- Sampling
- Translation
- Training
- Survey implementation
- Data entry/data capturing
- Data analysis
- Indicators of quality
- Country reports
- Site visits

5.   Figure X.1 depicts the overall WHS life cycle indicating the above-mentioned steps in every phase of survey implementation. The quality assurance guidelines which were drafted by a large number of WHS participants as well as international experts, aim to identify best practices whose implementation, in order to achieve and monitor a good-quality survey, is feasible. Each step of survey implementation involves a certain examination of quality. For example, it is important that the survey instruments have good measurement properties, that the sampling be representative of the target population, and that the data be clean and complete.

6.   This set of procedures constitutes merely an example to demonstrate the "quality assurance approach" to survey design and implementation as a process and to improving the output of the survey in terms of its relevance, accuracy, coherence and comparability. Any survey team designing and implementing a survey could use a similar approach keeping in mind the specific aims of its own survey and the feasibility of the quality assurance standards proposed in this chapter. Most importantly, quality should be given distinct attention and should be guided and monitored within an operational context. The results of the quality assurance process should be reported both in quantitative terms using appropriate indicators where measurement is possible (for example, sampling ratios, response rates, missing data, test-retest reliability of the application) and in qualitative terms summarizing the structure, process and outcome of the survey.

Figure X.1. WHS quality assurance procedures

C. Practical implementation of quality assurance guidelines: example of World Health Surveys

7.   The overall quality assurance strategy described above has been implemented within the WHS to improve the quality of the surveys including in several developing countries in Asia and sub-Saharan Africa. The present section aims to make use of the quality assurance standards, procedures and reporting as a concrete guide. Other survey teams may use this example as it fits their purpose. To our knowledge this is the first-ever application of systematic application of quality assurance procedures in international surveys, and implementing agencies and collaborators have found them very useful in organizing and reporting their work. Initial data suggest that it was possible to detect errors early and prevent them, and increase completion, accuracy and efficiency of results.

8.   The World Health Organization (WHO) has initiated the World Health Survey (WHS) as a real-life data-collection platform for obtaining information on the health of populations and health systems in a continuous manner (Üstün and others, 2003a, 2003b; Valentine, de Silva and Murray, 2000; World Health Organization, 2000). WHS responds to the need of countries for a detailed and sustainable health information system and gathers data through surveys to measure essential population health parameters; and brings together standard survey procedures and instruments for general population surveys in order to present comparable data across WHO

member States. These methods and instruments are modular in structure and have been refined through scientific review of literature, extensive consultations with international experts and large-scale pilot tests conducted in more than 63 countries and 40 languages (Üstün and others, 2003a, 2003c; 2001). WHS is designed to evolve through its implementation by continuous input from collaborators including policy makers, survey institutions, scientists and other interested parties. The countries and WHO jointly own the data, and there is a commitment for long-term data collection, building local capacity and using the survey results to guide the development and implementation of health policy.

9.   This chapter systematically reviews each step of the survey process, except questionnaire design and testing, which is reviewed elsewhere (see Üstün and others, 2003b), and introduces the WHS quality assurance standards in each area. These are desirable standards though which to increase efficiency and prevent unacceptable practices. Greater attention to quality is needed now more than ever because of the increasing importance of the WHS data for WHO member States and their implications for health policies. WHS has therefore formulated general guidelines for survey practice in order to enhance the reliability and validity of WHS surveys by reducing possible preventable errors. Quality assurance guidelines as adopted will become primary organizing tools for WHS and also serve in the organization of survey work and the preparation and planning for implementation. This chapter therefore provides an overall guide to the critical aspects that need particular attention so as to ensure collection of good-quality data.

10.   These guidelines will also serve as an evaluation template for the survey managers and quality assurance advisers (a network of international experts with extensive survey experience who serve as peer reviewers of the whole process). They will make site visits to countries to support their efforts in implementing the WHS and undertake a structured and detailed assessment of the process, which will support countries in assessing quality in a systematic manner, and in identifying areas in survey activity that could be improved.

1. Selection of survey institutions

11.   Carrying out a national survey requires extensive knowledge, skills, resources and expertise. These requirements have resulted in the organization of survey activity in accordance with different styles and traditions in different countries and sectors. To ensure that a competent survey group in a given country carries out the WHS, it is important to establish the identification of good survey institutions and the specifying of standards as the contractual conditions. WHS usual practice is to consult with the ministries of health, regional offices and WHO country representatives or liaison officers to identify such institutions. Given the size and complexity of the survey, the feasibility should be demonstrated by a contractual bidding process as required by WHO regulations. This process starts with a call for competent survey institutions to make their bid for the WHS In accordance with the technical specifications of the sampling, interviewing and data collection. [Technical specifications for the WHS is available on the WHS web site (www.who.int/whs)]. These bids are compared according to a number of criteria before the final selection is made.

12.   Criteria for assessing performance standards of potential institutions include:

- Their previous track record (that is to say, their experience with at least five large national surveys in the recent past with sample sizes of 3,000 or more).

- Their capacity to carry out the whole survey process (namely, sampling, training, data collection and analysis).

- Their experience in different modes of data collection including face-to-face interviews (and other possible modes like telephone, mail, computer, etc.).

- Documentation on former surveys (including the survey metrics of sample representation, coverage of country population, quality of interviewing, cost and type of training, quality assurance and other survey procedures).

- Record of usual time lines for survey calendar and their ability to complete surveys within an established time frame.

- Their potential to develop and use a good infrastructure with regard to health information systems, working closely with the ministry of health, national statistical bodies and other agencies.

13.   The contractual bidding procedure is useful in identifying the best possible offer in terms of quality and costs, and allows for a comparative assessment of all possible providers in a country. In this way, WHO and the ministry of health can identify the best possible survey institution with a view to building capacity for further surveys and to incorporate WHS data into the health information system. The contractual process also allows for building in penalties for failure to deliver results and ensure adherence to quality. Consortium bids should be encouraged to ensure that relevant partners (for example, the ministry of health together with the national statistical office) work together to secure access to a good sampling frame.

14.   A careful review of the different proposals submitted using the list of criteria described above should be undertaken. This comparative analysis should be documented.

15.   In summary, it is important not only to identify a good agency that will meet the technical specifications of the desired survey in the country concerned but also to provide the agency with the necessary technical support in order to achieve the desired outcome. For large-scale national surveys, it is often necessary within a country to create a partnership of groups, institutions and persons that have the necessary expertise for design, training, implementation, data processing, analysis and report writing.

2. Sampling

16.   A survey is only as good as its sample. If either sample design/or implementation or both are faulty, there is little one can do to make up for the sample design's limited representativeness

or to fill in missing information. The survey results will then be biased in unknown ways and often of unquantifiable magnitude.

17.   Because there is a wide range of applications in the field, WHO and a group of international technical experts have identified a set of guidelines to secure a good sample for the WHS [WHS Sampling Guidelines for Participating Countries are available on the WHO web site (www.who.int/whs)].   Standards of scientific sampling are based on probability selection methods and are widely known and accepted (Üstün and others, 2001; Kish, 1995a). However, these are typically not followed because of poor operationalization, lack of supervision of the implementation of sampling procedures in the field and/or high costs of implementation in particular contexts and conditions.

18.   WHO guidelines emphasize the scientific principles of survey sampling as explicit standards for quality, give examples of good sampling plans, and identify quality assurance standards for countries to adhere to. WHO and technical advisers will provide technical support to countries when needed. Important aspects of WHS sampling are outlined below:

        (a)   The WHS sample should target the de facto population (that is to say, all people living in that country including guest workers, immigrants and refugees) and not the de jure population (the citizens of that country alone). It is important to create good representation as the "miniature" of the country's overall population. To this end, it is essential to represent all people living in the country and have full geographical coverage of the country;

        (b)   The size of the sample must be adequate to provide good (robust) estimates of the quantities of interest at national or subnational levels depending on the objectives of the survey; at the same time, survey managers must balance the need for larger sample sizes to achieve better estimates against the corresponding increase in survey costs. Large sample sizes do not make up for poor quality. For various purposes, it may be required to have adequate representation of minorities (for example, ethnic or other subgroups) which may require oversampling (that is to say, giving a higher probability of selection). If a subpopulation needs to be oversampled because of any scientific study question, then specifications for doing so must be clarified in detail. In case of oversampling, differential weighting at the data analysis stage should be applied to correct the distortion caused by oversampling;

        (c)   In the WHS, a sampling frame (that is to say, a list of the geographical areas, households or individuals from which the sample is selected, such as could be derived from a computerized population list, a recent census, electoral roll, etc.) with 90 per cent coverage of all key subgroups of interest is considered acceptable. Countries should use the most recent sampling frame available. If it is two or more years out of date, enumeration or listing of households to update the frame at the penultimate stage of selection is often necessary. Quick count methods may be used to update measures of size for the primary sampling units prior to selection; such methods include counting in selected tracks where an up-to-date frame is unavailable owing to obsolete cartography or other reasons. Besides quick counting approaches in the selected sampling areas, other sources such as postal addresses from local post offices, lists from water or electricity billing companies, etc. can be used to update the frame. It is essential that the population be scientifically weighted back to the most recent census;

205

(d)   The WHS sample targets all adult members of the general population aged 18 years or over as its sample.[22]   In most cases, it is based on the most recent census data as its sampling frame.   Households are selected using a multistage stratified cluster sampling procedure. One individual per household is then selected through a random selection procedure [for example, the Kish table method (Kish, 1995a), or alternative methods such as the last-birthday method, and the Trohdahl/Carter/Bryant method (Bryant, 1975)].   Random number tables could also be used at this stage provided that the selection numbers are carefully documented.   Whatever selection technique is used, all attempts should be made to reduce selection bias during actual implementation in the field. Countries should seek to design the simplest sample plan possible that meets the measurement objectives of the survey. With respect to an overly complex design, implementation may be difficult and errors may be out of control. Feasibility and having the data trail to monitor sampling design are key to the quality;

(e)   WHS uses the United Nations definition of household;[23] however, there may be variations in this definition owing to local circumstances. The possible impact of variations in the household definition on sampling should be elaborated in country reports. Should the countries use a sampling frame of households, it is suggested that they then use the same definition for a household in the survey as was used in the original frame;

(f)   WHS uses a scientific sampling strategy, which encompasses a known non-zero selection probability for any individual included in the survey. Use of strict probability methods at every stage of sampling is crucial, and makes it possible to extrapolate the sample data to the whole population. Otherwise, the survey results will not be representative and valid;

(g)   The inclusion of institutionalized populations in a general population survey is difficult because separate frames need to be developed. There are also many ethical implications in relation to interviewing in institutions (such as hospitals, nursing homes, army barracks and prisons). Given the wide ranges of differences in institutionalization in difference countries, a single solution cannot be found. As a possible solution, WHS attempts to include people who are institutionalized owing to their health condition if it is possible to interview them during the survey period.   We then use the institutional population rates from the census to check the concordance of the rates obtained in the survey. This is of specific concern to the WHS, since persons living in institutions such as nursing homes, long-stay hospitals, etc. are likely to be in worse health than those who are not in institutions and therefore need to be included in the sample to reduce the potential for underestimating health conditions;

(h)   WHS sampling guidelines clearly explain what is meant by unit non-response and calculation of non-response rates in terms of target and achieved samples. The sampling strategy of the WHS does not allow substitution of non-responses by another household or individual;

---

[22] Currently, the WHS only includes adults. Future work aims to develop a survey that will include children as well.
[23] The United Nations defines a household as a group of persons that live under the same roof and share cooking and eating facilities (in other words, eat from the same source). For the WHS, a person is usually considered part of the household if he/she is currently in an institution because of a health condition. Such institutionalized people must be included in the household roster.

(i)   Survey results on sampling should report the standard errors for the important survey variables so that users can see the measurement error in statistical terms;

(j)   Use of Geographical Information Systems (GIS) may prove useful in improving the quality of the results by verifying the field execution of the sampling plan; in other words, that the interviews have actually taken place in a certain location rather than so-called curbside or fictitious interviews (De Lepper, Scholten and Stern, 1995). GIS may also offer additional value to the data by linking information such as the distance to health-care facilities, water and other environmental resources to measured health parameters (such as health states, diseases, risk factors) in the survey. It may also demonstrate on a map the dispersion qualities of any parameter, thus indicating health inequalities. For this purpose, the WHS has been using Global Positioning System (GPS) devices and digitized maps to geo-code the data within certain guidelines (please refer to http://www3.who.int/whosis/gis). Certain legal measures have been taken to maintain the confidentiality of personal information because geo-coding information may violate data protection standards.

Evaluation of sampling

19.   The sampling strategy should be evaluated before the start of the survey to assess the appropriateness of the stratification, the adequacy of the representation of the population and the size and distribution of the clusters selected. The report should carefully document the exact procedures used in the field, also noting any departures from the design so that users can be better informed about the quality of the survey results.

20.   During data collection, implementation of the selection of households and individuals must be monitored carefully by the field and/or office supervisors for accuracy, in, for example, the use of the Kish tables and household roster completion.

21.   After data collection, the data analysis metrics (discussed further below) are used to assess the quality of the data by means of:

- A summary statistic, which we call the "sampling deviation index" (SDI)

- Test-retest reliability to indicate the "stability" of the instrument with respect to use by different interviewers

- Information about the degree of non-response and missing data

22.   These procedures are described in more detail in the section on data analysis. A detailed summary list for quality of sampling is given in table X.1.

Table X.1. Summary list for quality of sampling

- Overview of population composition (urban/rural, minorities, languages, oversampled groups)
- Sampling frame and number of stages of sampling:
  - Do(es) the sampling frame(s) cover all the target populations?
  - How recent is the sampling frame?
- Stratification within the sampling frame
- Sampling units at each stage: known selection probability
- Size of sampling units at each stage: ensure all sampling units have a measure of size that exceeds a predetermined minimum
- Checking of "on the ground" size of units and issues such as whether there is one or more households per selected "address", and how to select within these
- Size of sample selected
- Probability weight for household
- Probability weight for respondent
- Training in use of and proper implementation of Kish table (or alternative)
- Checking on procedure for selection of respondent in household
- Summary report on sampling on the actual implementation, deviations, weights, standard errors

3. Translation

23.   To make meaningful comparisons of data across cultures, one needs a relevant instrument that measures the same construct in different countries. The WHS instrument has been developed following scientific review of existing survey instruments, large-scale consultations with experts and systematic field-testing in a multi-country survey study (Üstün and others, 2003a). We have reported the survey instruments features, relevance and cultural applicability elsewhere (Üstün and others, 2003b). For any other survey, designers must aim to have the best instruments and measures and make certain that their instrument is fit for their purpose, has good measurement properties and has passed through pilot tests to assure its feasibility and stability.

24.   Once you have a good survey instrument, then translation is one of the key features of ensuring the equivalent versions of questions in different languages. Given the multicultural societies that we live in, it is essential that we have good translations that measure the same concepts in the survey.

25.   Often in one country, the instrument will be translated into multiple languages depending on the size of the different language groups within the country. It is suggested that any linguistic group that constitutes over 5 per cent of the population should be interviewed in its own language. For respondents who are interviewed in a language for which a formal translated version has not been produced, emphasis is placed on the understanding of key concepts. Interviewers work with one of the existing translations in the country to ask questions in the

language without translation, using the overall guidelines. A further challenge faced by a large multi-country survey exercise is that in many African and Asian countries languages are not written and no scripts are available. It is recommended, in such cases, that a standard translation still be prepared in keeping with the guidelines and transliteration with a script from another familiar language in the country be used to prepare the written version.

26.   Guidelines for the translation of the WHS instruments have arisen out of the extensive experience of WHO in developing and implementing international studies with multiple partners and linguistic experts. The WHS Translation Guidelines, which are available on the WHS web site (www.who.int/whs),emphasize the importance of maintaining the equivalence of concepts and ensure a procedure that identifies possible pitfalls and avoids distortion of the meaning. These guidelines stress that:

- Translation should aim to produce a locally understandable questionnaire

- The original intent of the questions should be translated with the best possible equivalent terms in the local language

- Question-by-question specifications should aim to convey the original meaning of the questions and pre-coded response options

- The questionnaire should first be translated by health and survey experts who have a basic understanding of the key concepts of the subject-matter content. A set of selected key terms and those that proved to be problematic during the first direct translation should be back-translated by linguistic experts who would then comment on all the possible interpretations of the terms and suggest alternatives. An editorial group under the supervision of the chief survey officer in that country should review the translation and the back-translation and report back to WHO about the quality of the translation.

- Focus groups and qualitative linguistic methods such as developing an inventory of local expressions, and comparing expressions with those in other languages, should be used to improve quality. WHO has already undertaken systematic studies of translation and cognitive interviewing in certain languages and incorporated the results of these studies in the current text of the WHS questionnaire. It is still recommended that "cognitive interviews" (that is to say, further exploratory studies of what subjects understood to be the meaning of questions) using the translated questionnaire be undertaken with local subjects. It is mandatory to translate all the WHS documents (namely, the WHS questionnaire, question-by-question specifications, the survey manual and training manuals) into the local language. The data entry program may remain in English. If, however, the country has translated the WHS questionnaire using the electronic media following WHO specifications, the data entry program can automatically be generated in the other languages.

- Each WHS country should submit a report on the quality of the translation work at the end of the pilot phase. For items that were found to be particularly difficult to

translate, specific linguistic evaluation forms should be requested that describe the nature of difficulty of translation.

- Quality assurance advisers for the country should pay special attention to the implementation steps in the translation process and should check the list of key terms with the chief survey officer in the country.

- In countries where there are many dialects and/or languages that are not available in written format, specific translation protocols should be discussed with WHO.

Evaluation of translation

27.   A full translation of the questionnaire should be submitted to WHO before the start of the pilot interviews in the WHS. This translation should be checked by relevant experts in the particular languages, and comments made to the country if required.

28.   The list of key terms back-translated together with a report on the translation process and issues arising therefrom should be reviewed. The linguistic evaluation sheets (Üstün and others, 2001) should be systematically examined by the Country Survey manager and later by WHO to spot particularly problematic items and to enable a common solution across languages wherever feasible.

29.   Discussions should be held with interviewers with respect to understanding the procedures employed in the field when a term, phrase or question is not understood. These discussions should review the extent to which interviewers are required to "explain" and "interpret" the questions to respondents.

Table X.2. Summary list for review of translation procedures

- Languages spoken in the country; coverage of major language groups
- Who was involved in the translation process?
- Were all the needed materials translated?
  - Questionnaire
  - Appendix
  - Guide to administration (only when the interviewers do not know English)
  - Survey manual (only when the interviewers do not know English)
  - Result codes
- What issues came up in the translation?
- What protocol was undertaken (for example, full translation sent to WHO or just list of key items)?
- Were linguistic evaluation forms completed?

D. Training

30.   Training of survey team is the key to quality. Training is an ongoing process that should be conducted before and during the data-collection process, and end with a detailed debriefing after the fieldwork period is completed.

31.   Training should be provided at all levels of the survey team involved in the survey, from interviewers to trainers and supervisors, as well as to the central team overseeing the process nationally. This will ensure that all involved persons are clear with regard to their role in ensuring good quality of data.

32.   The purpose of overall training is to:

- Ensure a uniform application of the survey materials
- Explain the rationale of the study and study protocol
- Motivate interviewers
- Provide practical suggestions
- Improve the overall quality of the data

33.   To fulfil part of the training purpose, WHO has organized WHS regional training workshops for principal investigators from all participating countries and produced various training materials, including a training video and an educational compact disk covering all training issues.

Selection of interviewers

34.   The use of experienced interviewers as well as people who are familiar with the topic of the survey is important.

35.   Interviewers should have at least completed the full period of schooling within their country and be fluent in the main language of the country. Individual countries must decide what further level of education is required as well as what formal assessments will be carried out prior to selection.

36.   The issue of whether the interviewers should be health workers or not is left to the individual countries to decide. The characteristics of the interviewers (age, sex, education, professional training, employment status, past survey experience, and so on) should be recorded on a separate database. This information can then be linked to the identification numbers of interviewers for each questionnaire completed and an analysis can be carried out of individual interviewer performance.

Length, methods and content of training

37.   Training should be long enough for the interviewers to become familiar with not only the techniques for successful interviewing, but also the content of the questionnaire to be used. For experienced interviewers, the training will be shorter than for less experienced ones.

38.   The recommended length of training for the WHS is from three to five days, with three days being appropriate for experienced interviewers requiring training on the questionnaire only. The longer period of training is recommended for all other interviewers.

39.   All the training should be carried out as far as possible by the same team to ensure a standard training either for all interviewers in one session or for different groups at different times and places. To cut down costs and provide for regional training, training may be decentralized and cascaded. However, these costing benefits are then outweighed by the disadvantages of a diluted or varying training.

40.   A booster session is strongly recommended if it can be accommodated at some point during the data-collection period. It should preferably be held sometime towards the middle of the WHS data-collection period. The booster session serves to review various aspects of data collection, focusing on those undertakings that are proving complex and difficult or those guidelines that are not being adhered to sufficiently by interviewers. This session could also provide feedback on how much has been achieved and the positive aspects, including feedback from the supervisors and central survey team to the interviewers, as well as from interviewers to the supervisors and survey team.

41.   The training methods should include as much role playing in interviews as possible (with a minimum of one per interviewer). This method provides assimilation of interviewing techniques more effectively. For role playing to be effective, different scripts must be prepared in advance of the training so that the different branching structures of the interview, the nature of explanations that are permitted, and anticipated problems during an interview with difficult respondents can be illustrated.

42.   In addition to role playing, there should be at least one opportunity, before starting the actual data collection, to conduct an interview with a real-life respondent outside of the interviewer group.   The practice interviews should be tape- or video-recorded as often as possible for review and feedback discussion during training sessions. WHS countries are encouraged to make a standard training video similar to the WHO video if this is possible. Feedback should be given after each role-play or practice interview.

43.   Training materials should be provided to all interviewers to use as reference material. Any material provided should be comprehensively reviewed during the training and, where relevant, should be translated into the languages used in the country.

44.   The content of training should include the following:

- Administrative issues
- Planning of fieldwork
- Review of all materials provided
- Contacting procedures, consent forms and confidentiality

Conducting an interview should encompass:

- Interview procedures in the field
- Supervision in field and reporting procedures
- Structure of the survey team and role of all members of the team

Evaluation of training

45.   Evaluation of training should occur at a number of levels. The interviewers must be evaluated in order to determine whether they are capable of interviewing effectively and what, if any, particular support they require. The interviewers may in turn evaluate the training provided and the trainers. There should be ongoing evaluation during the initial data-collection period and at the conclusion of the fieldwork.

46.   The supervisors must be similarly evaluated by the central survey team. It must be mentioned here that the nature of the training must be adapted to the tasks that the supervisors are expected to perform such as refusal conversions, cross-checking and verification of selected interviews and editing of interviews. Detailed protocols for these procedures must be drawn up and clearly explained during the training process.

47.   The interviewers can be given a formal assessment at the end of training and some form of certification provided to each successful interviewer. This must be decided and implemented by each country individually.

Table X.3. Summary list for review of training procedures

- Number of training sessions
- Number of days of training
- Who did the training and what was their expertise in training and in the area of health surveys?
- What documentation was used?
- Practical components: role playing observation in real context
- Problems experienced in training
- Evaluation of training

E. Survey implementation

48.   To plan and manage survey implementation is a complex task, logistically and otherwise. It requires much preparation, scheduling and moving around of forces in the field to obtain the desired sample. Strategically, survey implementation is a key element that determines whether survey data is of a good quality or not. It is therefore of great importance to pay careful attention

to the quality of implementation of the actual survey and monitor it in real time so that problems can be addressed while it is in progress.

49.   How a survey is actually carried out in the field is the quality-determining step in the overall process. Good and strong central organization of the survey in each country will help ensure quality. Each step (that is to say, printing questionnaires, making sample lists, enrolling subjects, sending out interviewer teams, carrying out daily supervision in the field, editing the questionnaires, and so on) should be planned and reviewed carefully for quality.   More specifically:

(a)   Each survey team should prepare a central survey implementation plan and a task calendar in which the details of the survey logistics are laid out clearly. This plan should identify how many interviewers are needed to cover an identified portion of the sample in a given region with a given number of calls (including callbacks) and success rate. Accordingly, it should take into account the anticipated non-response rate and incomplete interviews, and the survey team's presence in a location;

(b)   Each survey team should have a supervisor who oversees and coordinates the work of the interviewers, as well as provides on-site training and support. The ideal supervisor-interviewer ratio for the WHS varies between 1:5 and 1:10 depending on the country and the different locations;

(c)   Supervisors should set out the daily work at the beginning of the workday with the interviewers and review the results at the end of the day. In this review, interviewers will brief their supervisors about their interviews and results. Supervisors must examine the completed interviews to make sure that the interviewers' selection of the respondents in the household has been done correctly and that the questionnaire is both complete and accurately coded;

(d)   A daily logbook should be kept to monitor the progress of the survey work in every WHS country survey center. The elements to be recorded are:

- The number of respondents approached, interviews completed and incomplete interviews
- The response, refusal and non-contact rates
- The number of callbacks and outcomes of calls

Information must be maintained on each interviewer so that his/her work can be monitored by the supervisor on an ongoing basis. This interviewer base can then be used in order to give individual feedback and so that decisions with regard to future hiring can be made;

(e)   Each country should conduct a pilot survey at the beginning of the WHS survey period, which should last a week or two. The pilot should be used as a dress rehearsal for the main survey. Fifty per cent of the pilot sample would then be reinterviewed by another interviewer to demonstrate the stability of application of the interview. The pilot period should be evaluated critically and discussed with WHO. The data from the pilot should be rapidly

analysed to identify any particular implementation problems. Since the instrument to be used in the survey would already have undergone extensive pre-testing prior to the pilot, the intention of the pilot testing should be to identify minor linguistic and feasibility issues and enable better planning for the main phase. It would also be expected to identify some obvious particular mistakes in skip patterns, etc. in the survey. Feedback from the pilot will correct these errors and allow for minor adjustments to be made. After consultation with WHO, the main study should start;

(f)   The helpfulness of the printing and practical collation of questionnaires (for example, colour coding of sets of rotations, lamination of respondent cards) should be recognized. All countries should send WHO a copy of the printed documents;

(g)   Pursuant to WHS contract specifications, 10 per cent of the respondents should be randomly checked again by supervisors or other teams. This check can be done by phone or in person, and is structured to ensure that the initial interview has been conducted properly. The recheck interview should cover the basic demographic information and any information not collected at the initial interview;

(h)   Pursuant to WHS contract specifications, a randomly selected 10 per cent of the total sample of respondents should be given the whole interview again by another interviewer within seven days of first interview so that the reliability of the questionnaire can be assessed (the re-tested respondents should not be the same as the check-back respondents, as specified in (g) above);

(i)   Response rates should be monitored continuously and each centre should employ a combination of various strategies to increase participation in the survey and reduce non-response. For example, making public announcements in TV, radio, newspapers or local media channels, sending letters or cards to participants, asking assistance from local health workers, giving incentives for participation, negotiating with local traditional or other recognized authorities, etc. are all public relations techniques that may be used to maximize response. The use of particular methods is left to the individual centre;

(j)   Each survey should aim towards the highest attainable response rate.   WHS contract specifications require an overall response rate of at least 75 per cent. This threshold does not mean that 75 per cent should be a stop point in survey implementation. It simply denotes the minimum acceptable standard commonly agreed by WHS collaborators in view of the past surveys in many different countries. In many instances, WHS response rates have been higher. The response rate may vary across countries and has to be compared with that of other surveys in the same country. In calculating the response rate, the same definition of complete interview should be used in all countries. An algorithm is used during the data cleaning procedures to identify the completeness of an interview based on a set of key variables;

(k)   Callbacks: Pursuant to WHS contract specifications, survey teams should attempt up to 10 callbacks (including phone calls, leaving notes or cards indicating that the interviewer called). The average number of these callbacks depends on the response rate and each centre

should examine the gain in each additional callback and consult with WHO regarding the sufficient number for that particular country;

(l)   Survey implementation depends heavily on the resources at hand. Each survey should be evaluated within the context of the country. It is essential to compare with other comparable surveys in the same country. Local customs and traditions must be taken into account in the evaluation. The trade-off between having fewer interviewers do more interviews over a longer study duration versus having a larger number of interviewers do fewer interviews over a shorter study period needs to be considered in terms of impact on quality.

Table X.4. Summary list for review of survey implementation

Pilot survey
- Where was the pilot carried out?
- What training was provided for the pilot?
- Any data problems in data entry?
- Data analysis: see results; and what problems were experienced?
- Any changes in methodology arising from the pilot?
- Any changes in translation arising from the pilot?

Main survey
- Number of interviewers, supervisors and central coordinators:
    - How is supervision conducted? Feedback
- Logistic arrangements:
    - Travel: how easy was it to travel to the household? What sort of transport was used?
    - Team organization
- Contact procedures:
    - How easy was it to contact the respondent?
    - How many contact calls were made?
    - What was the refusal rate and what was the main reason for refusing to do the interview?
- Payment of interviewers
- Consent form signing and recording (as part of questionnaire or separate sheet)
- Checking procedures in field by supervisors
- Checking procedures centrally
- Return of questionnaires to central office and security
- Final check on questionnaire and procedure for correcting errors
- Checking procedures and supervision
    - **Weekly production status reports**:
    - To assess interviewing process
    - To review response, refusal and non-contact rates: ensure response rate
    - To monitor results and ensure that data collection is implemented

- **Verification of records**:
- Is the number of contacts (contact/contact attempt) recorded in detail?
  - Are at least 10 per cent of each interviewer's interviews verified to ensure that some answers remain constant (age, education, household composition) and that the interview has been conducted?
    - Check number of interviews already conducted and planning of interview schedule
    - Verify that final result codes for completed interviews and refusals have been assigned correctly
    - Check that informed consent forms are signed

All identifying information detached from questionnaires and data entry program.
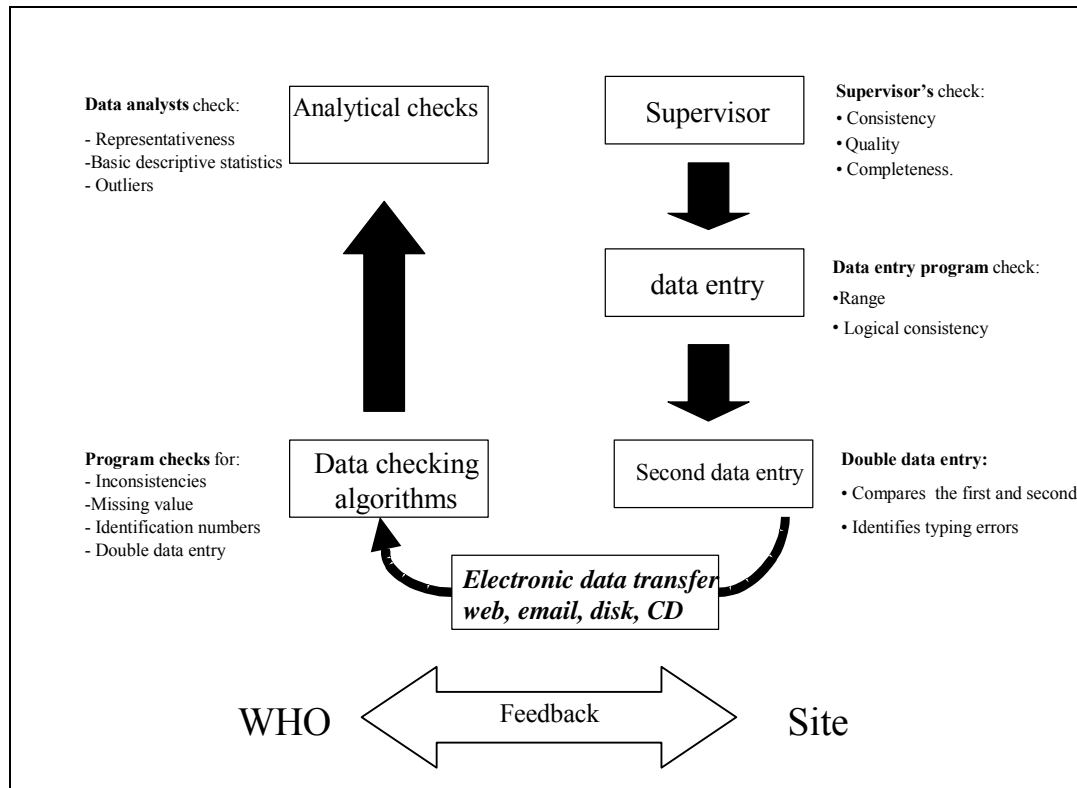
Draft report with recommendations for any action to be taken.

F. Data entry

50.   The everlasting output of the survey is the data. It is important to capture the data accurately and in a timely manner. The WHS data entry process is planned so that there is immediate local data entry and central coordination. It is essential that data be transferred to computer media as soon as possible after collection. In this way, standard routine checks can be easily conducted by use of local computers. Any errors found can then be dealt with while the survey is in progress in the field.

51.   Figure X.2 below describes the data flow in the WHS and the quality assurance steps that relate to this data flow. The tasks that are performed at the country level are presented on the right-hand side and the tasks that are performed at WHO are presented on the left-hand side.

Figure X.2. Data entry and quality monitoring process

52.   After the interview is administered, the following steps take place:

- Supervisor checks the questionnaire form before the data entry starts.

- Data entry (or data capture/registration) is performed by using the WHO data entry program. This program checks ranges (for example, the allowed response variable ranges) and checks to ensure logical consistency of related codes (for example, an illness cannot last longer than one's age, and men cannot have gynaecologic problems, etc.).

- Second data entry is performed for the purpose of identifying typing errors and accidentally skipped questions.

- Data are sent to WHO in batches using email, CD-ROM or diskette.

- Once the data are at WHO, programs check for inconsistencies, missing values, problems with identification numbers or test/re-test cases. These programs produce a report to be sent back to the countries. Also, any corrections received from the site countries are applied to the data.

- Data analysts check for representativeness, basic descriptive statistics and outliers. Representativeness is checked by comparing the age-sex distribution of the realized

sample with the expected population distribution. Basic descriptive statistics are used to determine the response distributions and identify any skewed distributions, odd results and outliers.

- WHO sends feedback to the countries. The countries will send, if needed, corrections and/or explanations in accordance with the feedback.

53.   Important quality issues concerning the data entry:

- Data entry should be carried out done using a data entry program, which provides quality check features. Use of other programs that do not include these features may therefore be disadvantageous.

- The completed interview forms should be checked by the supervisor before the data entry starts.

- The data entry program is accessible only to the responsible team members and to no one else. This is essential for the confidentiality of data.

- Double data entry is required so as to avoid data typing or editing errors. The data entry program identifies double data entry when the second entry is completed.

- The countries should be very careful in entering the identification (ID) number . A list of valid IDs is sent to the countries. The program has a checksum digit to make sure that the ID code is entered correctly. Using correct IDs is especially important for the re-test cases, since the ID is used to match the test cases with the re-test cases.

- Data must be submitted to WHO regularly, for example, on a daily or a weekly basis.

- Once WHO starts receiving data from the countries, it is checked and feedback is sent to the countries as the data collection continues.

- Certain rules are applied to maintain the integrity and accuracy of data involving, for example, checking to determine whether the same respondent is used twice and the extent of missing data.

54.   Identifying information will be detached from questionnaires and the data entry program will keep confidential information in a separate file if entered. It is the country's responsibility to maintain confidentiality. Security of data during transfer over the Internet is ensured through encryption.

Evaluation of data entry

55.   The following aspects should be carefully monitored and reviewed (see table X.5):

- The number of data entry personnel and their training

- The number of forms entered per day per person, including error rates

- Checking procedures and supervision of data entry

- Time period between completion of the interview in the field and data entry

- Number and regularity of completed interviews sent to WHO and problems encountered with respect to the sending of the data

56.   Though several problems with data entry can be minimized with computer-assisted interviews where the data are entered as the interview is in progress, these computer programs will require that checks be built in so as to ensure the correct application of the interview with all skip and branching rules and that consistent data within specified ranges are entered.

Table X.5. Summary list for the data entry process

- Who are the data entry personnel?
- What is the completion and error rate by data entry personnel? Are there data entry personnel who need retraining?
- Observe data entry process. What is the system used for keeping track of the number of questionnaires assigned to each interviewer?
- Discuss data analysis and calculation of data quality matrix, and need for further support
- **Questionnaires:**

Choose several completed questionnaires from each interviewer and check that:
  - Names are deleted from questionnaires
  - Coversheet has been detached from questionnaire
  - Household rosters have been randomized and completed appropriately
  - Handwriting is legible and neat
  - Options have been recorded appropriately (for example, options are circled, *not* ticked, underlined or crossed out)
  - Open-ended questions are answered when they need to be
  - Open-ended questions are recorded verbatim
  - Questions are skipped correctly
  - Questions to be answered by women are answered only by women
- Double data entry.
- Use of data entry program:
  - Verify confidentiality and security of data
  - Is data double-entered?
  - Check coding in database against hard copy
  - Check range, consistency, routing and other errors
  - Check extent of missing data

G. Data analysis

57.   In advance of substantive data analysis of the WHS data, there are a number of systematic checks of data quality. The compilation of these checks is called the "WHS survey metrics" and provides summary indicators of data quality.

58.   The components of survey metrics are:

- Completeness, which includes response rate (taking into account households whose eligibility status may be unknown, in which case an estimate must be made of the proportion of eligible households or, if such households are excluded from the calculation of response rates, a clear justification must be provided for the assumption that these households had no eligible respondents) and incomplete questionnaires or item non-response. Frequencies of missing data are calculated at the level of items across respondents and at the level of each respondent across all items. This helps identify problems of survey implementation, particularly problematic items in the questionnaire.

- Sample deviation index (SDI), which is a measure of the degree to which the sample deviates in representativeness from the target population. If this measure shows significant deviation then the analysis should be stratified. The SDI can be formally assessed using the chi-squared statistic. If some key subgroups have been intentionally oversampled, this should be taken into account so as to adjust the SDI by the intended oversampling factor.

- Reliability, which indicates replicability of results using the same measurement instrument on the same respondent at different times and with different interviewers. This analysis uses the data from the test/re-test protocol undertaken in 50 per cent of the pilot interviews and in 10 per cent of the whole sample.

- Comparison with external validators, that is to say, comparison with other survey results, such as the census, surveys and service data as well as private and public sector data.

59.   These metrics are further elaborated in the next section. Data processing is conducted at the country level, where the necessary capacity is available, as well as at WHO headquarters.

60.   Further country-level data analysis is seen as essential to ensure effective use of the results. WHO headquarters and regional offices will identify countries requiring support in the full analysis of the data and develop mechanisms for providing this support.

Evaluation of data analysis

61.   The evaluation of this aspect requires discussion on the availability of skills in the country to undertake the analysis and the level of support that is required or that can be provided by the country to other countries.

H. Indicators of quality

62.   It is useful to summarize the quality assurance by ways of indicators. These indicators may later be used to evaluate other contextual factors that affect the quality of the survey and the quality cycle is then completed. To our knowledge, there has not been a systematic set of indicators proposed to monitor and report the quality of a survey in summary measures. The WHS uses certain quantifiable indicators explained below as well as a structured qualitative assessment by a peer review process as a quality assurance report.

63.   In general, any household survey is subject to two kinds of errors: sampling error and non-sampling error. Sampling error occurs because a survey is carried out on a sample of the population rather than the entire population. It is affected by the sample size, the variability that occurs in the population for the quantities of interest and other aspects of the sample design such as stratification and clustering effects. Non-sampling errors, on the other hand, are affected by factors such as the nature of the subject-matter concepts, accuracy and degree of completeness of the sampling frame, fidelity of the actual selection procedures in the field vis-à-vis the intended sample design, and survey implementation errors. The last-mentioned factor entails such problems as poor design of the questionnaire, interviewer errors in asking the questions and respondent mistakes or misreporting in answering them, data entry and other processing errors, non-response and incorrect estimation techniques. Some of the non-sampling errors that lend themselves to measurement and quantification are illustrated below.

64.   In respect of monitoring the end result of survey data, the following standard indicators are currently being used to monitor the WHS data quality.

1. Sample deviation index

65.   Sample deviation index (SDI)[24] shows the proportion of age and sex strata in the sample compared with population data from an independent source, with the latter assumed to be the standard. The WHS has used, as the independent source, the United Nations population database, but any other more recent and reliable population data source may be used instead. The SDI is one indicator of the quality of the sample data in terms of their representativeness (that is
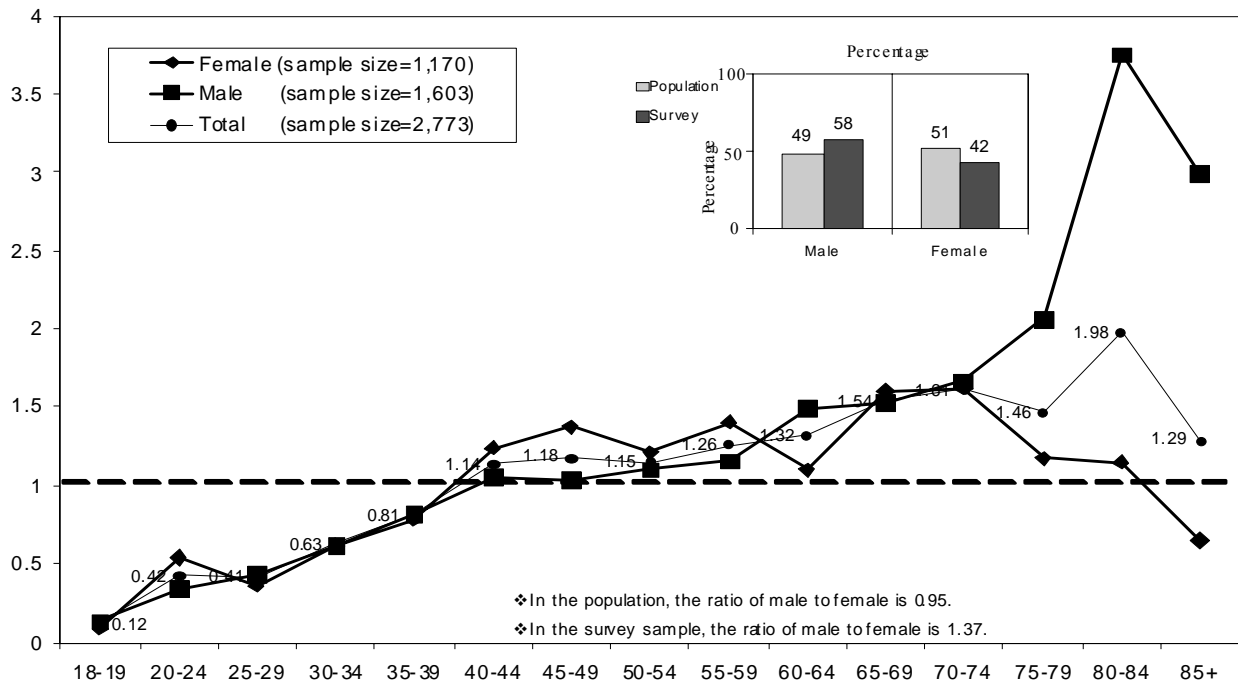
---

[24] $SDI = \sum_{a=1}^{a} \left| 1 - index_a \right|$, where a = age categories and the index is the ratio of the sample in the age category to the population in the age category from the UN population database or other updated source such as the country census. This index indicates the extent to which the sample represents the population in terms of age or sex distribution. The index can be tested by the chi-square or the pi-star tests for homogeneity.

to say, of how well the sample represents the overall population). A ratio of 1 shows that the survey sample matches the characteristics of the general population for an age or sex category, whereas deviations from 1 indicate oversampling or undersampling from that age or sex category.

66.   The expected value of 1 (ideal representativeness) is rarely observed in surveys because of sampling errors.   Figure X.3 presents the SDI for one of the surveys, showing underrepresentation at younger ages and overrepresentation at older ages, particularly for older men.

Figure X.3. Example of a sample deviation index

2. Response rate

67.   Response rate shows the completion rate of interviews in the selected sample, that is to say, the number of completed interviews among persons or households eligible for inclusion (a selected "household" that turns out to be a vacant dwelling, for example, is not eligible). This indicator shows how well the survey has performed with respect to achieving the ideal of 100 per cent response. A response rate of 60 per cent is generally regarded as the minimum acceptable, though the WHS requests a response rate of at least 75 per cent.

3. Rate of missing data

68.   The rate of missing data is defined as the proportion of missing items in a respondent's interview. The WHS measures the proportion of people failing to complete a

minimum acceptable range of items (for example, 10 per cent in the household face-to-face interviews) to determine the quality of the interviews. Problematic items with a high level of missing responses (over 5 per cent) across eligible respondents are also identified.

4. Reliability coefficients for test-retest interviews

69. Reliability coefficients for test-retest interviews show the stability of interview administration with respect to response variability on two separate occasions. These are calculated as chance-corrected concordance rates (that is to say, kappa statistics for categorical, and intra-class correlation coefficients for continuous variables). This indicator refers to how well a given item/question in the survey interview yields the same results in repeat administrations of the interview. Generally, a score greater than 0.4 is considered acceptable; a score greater than 0.6 is considered fair and a score greater than 0.8 is considered excellent (Cohen, 1960; Fleiss, 1981).

70. The main indicator of a survey's quality in terms of the error present in the data from the sampling component is the estimated standard error for each key statistic in the survey. It shows the estimated range of sampling error (for example, plus or minus 3 per cent) around a given estimate. A related measure, design effect coefficients for the multistage cluster samples of the WHS, are calculated when possible. This coefficient is the ratio of the variance from the actual sample to that of an assumed simple random sample of the same size. Since a true simple random sample is not practical in large-scale surveys owing to costs (including transportation costs), it is customary to calculate sampling variance (square of standard error) for comparison with a random sample (Kish, 1995b). A design effect of between 1 and 6 is generally considered to be acceptable for the indicators of interest to the WHS.

I. Country reports

71. An important feature of quality assurance relates to the final output in terms of reporting the data, because of the impact of the survey in terms of its added value to our knowledge base and the provision of further directions for policy. Proper reporting is obviously closely related to the relevance of the WHS to the country's needs. WHS results will be presented in a number of different types of reports, namely:

- (a)   Country reports for each individual WHS country:

  - (i)   Executive summary for policy makers and the public;
  - (ii)   Detailed report for researchers and other scientific users;

- (b)   Regional and international reports on specific issues.

72. The initial template for a country report [71(a) above] includes:

- •   Introduction encompassing (for example, the information to drive policy and available information on health systems).

- Discussion of survey implementation (encompassing, for example, the survey description, sampling methods, training, data collection and processing, quality assurance procedures, description of survey metrics).

- Overview of survey results and implications for policy (entailing, for example, the inputs to the health system, population and household characteristics, coverage of health interventions, health of the population, responsiveness of health systems; health expenditure).

- Conclusions: specific recommendations for health policy and monitoring the Millennium Developing Goals in the country.

73.   This template will be further developed in interactive collaboration with countries, regional offices and other interested parties.

74.   A dissemination strategy for the country report needs to be clearly developed through the media, workshops and other events. It is necessary to involve different stakeholders in the use of the information generated from the survey in policy debates.

75.   Countries themselves should be primarily responsible for generating their country reports. WHO will assist in providing the essential data and technical support and tools to prepare and discuss these country reports with production teams.

76.   The WHS is useful in obtaining information on different aspects on the health of populations and health systems. These elements include many components of the health system performance assessment framework. Moreover, the surveys provide detailed information on other aspects such as specific risk factors, functions of health systems, specific disease epidemiology and health services. It is therefore important to extract the best possible information value from the WHS data.

77.   Some countries may also wish to use WHS data for subnational analysis. In most cases, this may require larger sample sizes. In others, WHS data may be used together with other data sources such as the census and other surveys.

78.   In the long run, it is expected that the modular structure of the WHS will allow for integration of various surveys on health and health systems into a single survey.

Evaluation of country reports

79.   The analysis of the data and drafting of country reports is the culmination of the survey implementation. The quality of the reports and the manner in which the results are discussed will determine the way in which the future rounds of surveys are implemented as well as the impact the results will have on policy development and monitoring within the country.

J. Site visits

80.   WHS countries know in advance what is expected of them in terms of implementing the WHS and quality assurance procedures. It is important to document the fieldwork in this regard. To achieve this aim, WHO will contract independent quality assurance advisers who will make site visits in each country. These site visits will in effect constitute an external peer review of the survey implementation process and will independently record the adherence to QA standards. These site visits will also provide an opportunity to recognize any problems and solve them early in the process. The country team and the quality assurance adviser will then produce together a structured assessment of the overall survey quality along with the WHO guidelines.

81.   Quality assurance is a process, and is not reducible to the single event of a site visit. The relationship between QA advisers and the country teams can be seen as a long-term process in three phases: before, during and after the site visit.

82.   Before the site visit, countries and QA advisers should prepare a file for the visit, which will cover the basic format of the WHO QA guidelines as outlined in this document and include all aspects in the site visit checklist. Included in this file will be all background information available with regard to the site, survey institution, sampling design, local expertise, instruments and training package used locally, and template for the WHS country report. Information not available will be obtained during the site visit.

83.   Country officers at WHO headquarters and the QA advisers will be in direct communication with the principal investigator or chief survey officer within the country to make the QA process an integral part of the survey implementation process. This will help build a culture of quality assurance in surveys.   The aim of the QA process is not auditing or policing but achieving quality in the WHS through the provision of assistance and support.

84.   In order for the site visit to have the most impact, it should be scheduled towards the end of the training and the beginning of data collection. The site visit should focus on all aspects of the survey process, that is to say, diagnose problems, suggest remedies, be sensitive to local context and provide support and build an ongoing relationship.

85.   The role of the quality assurance advisers (QAAs) when visiting the countries, will be to diagnose the problems and note strengths within the survey implementation. Their main task is to examine the WHS implementation process used in the country and to identify any deviation from the expected QA standards. Their judgement as to whether this deviation is significant and how it could be remedied is essential. The QAA should also provide support directly through discussion with WHO headquarters or arrange for relevant support to be provided by another entity.

86.   The QAAs will perform their evaluation according to a structured checklist that will include the various steps in their order of importance. This evaluation should include the analysis of the "survey metrics" (as long as there are some data entered by the time the site visit occurs) which includes indicators for quality of data.

87.   The QA evaluation will be jointly discussed with the country survey team and WHO. Countries should know in advance what is expected of them in terms of quality assurance procedures.

88.   The site-visit report is succeeded by the WHS country report, which is the final product of the site visit and country support.   The site visit should start the process of drafting the country report and explore specific strategies for its production, including how to use the findings in policy development.

K. Conclusions

89.   Quality assurance is a core issue in survey implementation. It is necessary and possible to specify quality assurance mechanisms at each step of a survey. If these mechanisms are operationally defined, then they can be measured and an overall survey quality can be monitored.

90.   The establishment of quality assurance requires a change in the mindset of survey implementers, since examination and evaluation of each step become mandatory.

91.   The assessment of the quality indicators on an ongoing basis during the course of the entire survey is essential. The process should not be regarded merely as post hoc; it should also be used to make such midstream corrections as are warranted by detecting problems and intervening appropriately. This important continuous quality improvement or total quality management in the production process must be integrated into all surveys.

92.   The availability of computer tools now makes it possible to develop a survey management and tracking system that allows the continuous tracking of the survey process, which helps instil confidence in the data.

93.   It is important to document critical issues (for example, issues about survey implementation, training, etc.) in a systematic manner in terms of both qualitative reports and quantitative indicators (namely, the sample deviation index, response rates, missing data proportions, and test-re-test reliability) so as to give the users of data essential information about the quality of a survey.

94.   The desired outcome of the quality assurance process is to produce a survey that yields better-quality data. The results can then be documented as being valid, reliable and comparable.

95.   The continued implementation of these quality assurance procedures will set standards for acceptable international data-gathering exercises, and methods to monitor these standards will continue to evolve.

Acknowledgements

We would like to gratefully acknowledge the participation of the following survey experts from various countries and institutions in the production of WHS quality assurance guidelines:

Dr. Farid Abolhassani, Islamic Republic of Iran
Dr. Sergio Aguilar-Gaxiola, United States of America
Dr. Atalay Alem, Ethiopia
Dr. Lorna Bailie, Canada
Dr. Russell Blamey, Australia
Dr. Carlos Gomez-Restrepo, Colombia
Dr. Oye Gureje, Nigeria
Dr. Holub Jiri, Czech Republic
Mr. Mark Isserow, South Africa
Dr. Feng Jiang, China
Mr. Jean-Louis Lanoe, France
Professor Howard Meltzer, United Kingdom of Great Britain and Northern Ireland
Mr. Steve Motlatla, South Africa
Ms. Lipika Nanda, India
Dr. Kültegin Őgel, Turkey
Dr. Gustavo Olaiz Fernandez, Mexico
Dr. Mhamed Ouakrim, Morocco
Dr. Jorun Ramm, Norway
Dr. Wafa Salloum, Syrian Arab Republic
Dr. Shen Mingming, China
Dr. Benjamin Vicente, Chile

Sampling consultants

Professor Steve Heeringa, University of Michigan, Institute of Social Research, United States of America
Professor Nanjamma Chinnappa, India, ex-president of the International Association of Survey Statisticians

WHO regional advisers

Mrs. M. Mohale M., Regional Adviser for WHO Regional Office for Africa
Dr. Siddiqi Sameen, Regional Adviser for WHO Regional Office for the Eastern Mediterranean
Dr. Amina Elghamry, Regional Adviser for WHO Regional Office for the Eastern Mediterranean
Dr. Lars Moller, Regional Adviser for WHO Regional Office for Europe
Dr. Myint Htwe, Regional Adviser for WHO Regional Office for South-East Asia
Dr. Soe Nyunt-U, Regional Adviser for WHO Regional Office for the Western Pacific

References

Biemer, P.P., and others, eds. (1991). *Measurement Errors in Surveys*. New York: Wiley.

Bryant, B.E. (1975). Respondent selection in a time of changing household composition. *Journal of Marketing Research*, vol. 12, pp. 129-135.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement,* vol. 20, pp. 37-46.

DeLepper, M.H., H. Scholten and R. Stern, eds (1995). *The Added Value of Geographical Information Systems in Public and Environmental Health* Dordrecht, Netherlands: Kluwer Academic Publishers.

Fleiss, J.L. (1981). *Statistical Methods for Rates and Proportions*, 2nd ed. New York: John Wiley and Sons.

Kish, L. (1995a). *Survey Sampling*. New York: John Wiley and Sons.

_____ (1995b) Methods for design effects. *Journal of Official Statistics*, vol. 11, pp. 55-77.

Lyberg, L.E., and others, eds. (1997). *Survey Measurement and Process Quality*. New York: Wiley.

Statistics Canada (1998). *Quality Guidelines,* 3rd ed. Ottawa.

Üstün, T.B. and others (2001). *Disability and Culture; Universalism and Diversity*. Göttingen, Germany: Hogrefe Huber.

_____ (2003a). WHO Multi-country Survey Study on Health and Responsiveness 2000-2001. In *Health System Performance Assessment: Debates, Methods and Empiricim* (C.J.L. Murray and D.B. Evans, eds.). Geneva: WHO.

_____ (2003b). The World Health Surveys. In *Health System Performance Assessment: Debates, Methods and Empiricism* (C.J.L. Murray and D.B. Evans, eds.). Geneva: WHO.

_____ (2003c). *World Health Organization Disability Assessment Schedule II (WHO DAS II): Development and Psychometric Testing*. Geneva: WHO. In collaboration with WHO/National Institute of Health Joint Project Collaborators.

Valentine, N.B., A. de Silva and C.J.L. Murray (2000*). Estimating Responsiveness Level and Distribution for 191 Countries: Methods and Results*. Global Programme on Evidence Discussion Paper Series, No. 22. Geneva: WHO.

World Health Organization (2000). *World Health Report*. Geneva: WHO.

_____ (2002). *World Health Survey: Quality Assurance and Guidelines: Procedures for Quality Assurance Implementation by Country Survey Teams and Quality Assurance Advise.* Geneva: WHO.