Designing and Conducting Health Systems Research Projects

Volume I: Proposal Development and Fieldwork

Designing and Conducting Health Systems Research Projects

Volume I: Proposal Development and Fieldwork

Corlien M. Varkevisser
Indra Pathmanathan
Ann Brownlee

KIT Publishers, Amsterdam
International Development Research Centre

in association with
WHO Regional Office for Africa

Designing and conducting health systems research projects
Volume 1: Proposal development and fieldwork

Jointly published by KIT Publishers and the International Development Research Centre (IDRC), in association with the Africa Regional Office (AFRO) of the World Health Organization.

KIT Publishers
Mauritskade 63, 1090 HA Amsterdam, the Netherlands
publishers@kit.nl / www.kit.nl
ISBN 90 6832 148 X

International Development Research Centre
PO Box 8500, Ottawa, ON, Canada K1G 3H9
info@idrc.ca / www.idrc.ca
ISBN 1-55250-069-1 (Volume 1) / 1-55250-070-5 (Volume 2)

World Health Organization - Regional Office for Africa
Cite du Djoue, P.O.Box 06 Brazzaville, Congo
www.whoafr.org

Cover design: Mulder en van Meurs, Amsterdam
Production: Meester & de Jonge, Lochem
NUR 600

TABLE OF CONTENTS

FOREWORD

Health Systems Research (HSR) has proved to be a useful tool for health decision makers at all levels over the past 20 years, providing them with the necessary data for informed decision making.

The Joint HSR Project for the Southern African Region based in the WHO Office in Harare and supported by WHO Geneva, the Royal Tropical Institute (KIT) in Amsterdam and the Dutch Technical Development Co-operation (DGIS), has played a crucial role in the promotion of HSR in the African region since 1987. HSR was enthusiastically embraced by many Ministries of Health and universities. In 1996, the Regional WHO Office for Sub-Saharan Africa (AFRO) assumed full responsibility for implementing HSR. Following the recommendation of Health Ministers of the Region, WHO/AFRO in 1998 included HSR as a regular programme for all its 46 member states.

The present HSR training modules, developed by an interdisciplinary, international team of practical researchers, have been highly instrumental in raising the interest for HSR. Originally designed for health managers at different levels as a tool to develop problem solving research in the Southern African Region, the modules also proved useful in Malaysia and were further elaborated by staff of the School of Public Health The 1991 combined version, published by International Development Research Centre, Canada and WHO, Geneva,* was translated in French, Spanish and Portuguese, and sections of it appeared in Arabic, Vietnamese and Chinese. In different parts of the world the modules facilitated the development and implementation of hundreds of research protocols by health staff and researchers. The HSR modules are used in the Community Health and Social Science Departments of many African, Asian and Latin American universities to train students and prepare them for their fieldwork. They are also used by Masters of Public Health courses in Europe and the USA and by international research programmes interested in applied research.

This unanticipated application of the modules in academic as well as health management circles led to the rapid exhaustion of the 1991 edition and the several subsequent reprints. Various groups of users made many useful suggestions for changes and improvements. The HSR Unit in AFRO, with agreement from IDRC, therefore decided to organise a revision of the HSR modules. An interdisciplinary group of Southern African researchers reviewed and made revisions in two workshops in 1998 and 1999. Two of the three original editors finalised the present version. IDRC took on the final responsibility for the publication, which was financed by AFRO and IDRC and published by KIT.

It is hoped that this revised version of the modules will fulfil the same need as preceding ones have done. Certainly many new and persisting health problems urgently require operational research. How to support necessary health reforms and at the same time ensure equity in access to health care for high-risk groups remains a major challenge. HSR is one of the tools we have to obtain deeper insight in these challenges and optimally focus our resources.

Dr. Rufaro R. Chatora, Director of the Division of Health Systems and Services Development (DSD), WHO/AFRO, Harare

Dr. Christina Zarowsky, Senior Health Specialist, IDRC, Ottawa

Ms. Catherine Hodgkin, Head Health Department, KIT, Royal Tropical Institute, Amsterdam

* Corlien M. Varkevisser, Indra Pathmanathan and Ann Brownlee (1991) *Designing and conducting Health Systems Research projects. Part I: Proposal development and fieldword; Part II: Data analysis and report writing.* Ottawa: Health Sciences Division of the International Development Research Centre (IDRC) and Geneva:Programme on Health Systems Research and Development of the World Health Organisation.

PREFACE AND ACKNOWLEDGEMENTS

The present volume *'Designing and conducting Health Systems Research'* is a thorough revision of Volume 2 of the *Health Systems Training Series* which the International Development Research Centre (IDRC) in Canada and WHO HQ in Geneva published in 1991 and reprinted several times under the same name. It became necessary to revise the modules, because over the years inevitable shortcomings and gaps were detected which needed to be addressed.

Health managers, for example, stressed that implementation of the research findings and recommendations were somewhat underexposed in the modules. This point is now taken care of in Module 1 by adding a fourth, implementation phase to the Health Systems Research (HSR) training cycle which initially consisted of three phases: HSR proposal development (15 days), fieldwork (roughly 6 months) and data analysis and report writing (2 weeks). The implementation of research findings and recommendations is further elaborated in Module 33. Furthermore, health managers pleaded, understandably, for shorter courses. This wish has been taken care of by stressing more explicitly in Modules 1 and 3, as well as in the Course Guidelines (annexed to Part 1 of this volume) that the proposal development phase can be shortened by having research teams select their research topic in the field before the onset of the course, preferably under guidance of a facilitator. In addition, the WHO/AFRO HSR Programme based in Harare, is at present developing modules for participatory rapid action in health research at health centre and district levels which can be carried out and integrated in the day to day activities of staff and community members.

Research staff from Community Health, Social Science and other university departments/ research institutes in Sub-Saharan Africa or other parts of the world who are using the modules had other wishes. They advocated that, in addition to the already well-emphasized problem-solving, analytical research approaches, more weight should be given to descriptive research. A descriptive diagram has therefore been added to the problem analysis diagram in Module 4. In all subsequent research steps, if relevant, the distinction between analytic and descriptive studies has been elaborated. Qualitative research methods have also been given more weight and they were more thoroughly integrated with quantitative methods in the research methodology (Modules 8-14). This applies, for example, to Modules 10 (Data collection techniques) and 11 (Sampling techniques). Furthermore, two new modules have been added to Part 2 (Data analysis and report writing) of the volume: one on Measures of association based on risk (Module 25), which used parts of Module 30 in the 1991 version, and one on the difficult issue of Confounding variables (Module 26). This need for extension was also reflected in the most recent evaluation of HSR training (1997).*

Facilitators, finally, desired more elaborate examples of crucial research and data-analysis techniques. Therefore, Module 10B (Development of research instruments) has been elaborated with a section on interview techniques with interview exercises, and Module 10C (FGDs) now contains an example of a transcribed focus group discussion with codes in the margin. To Module 13 (Plan for data analysis), an example of a full-fledged questionnaire and of a master sheet have been added, and Module 23 (Analysis of qualitative data) now provides an example of a filled-in compilation sheet. Module 5 (Literature review), has been extended with an example of a literature review.

Apart from these additions, in all modules parts that had proven to be unclear or incomplete were rewritten, and many examples and references were replaced by more recent ones or elaborated.

The present revision was initiated in a workshop held from 2-11 November 1998 in Arusha by a group of interdisciplinary researchers and managers convened by the manager of the WHO/AFRO HSR Programme (since 1992 Gabriel Mwaluko). All participants had thorough experience with the modules and with HSR: Sambe Duale, Lawrence Gakuri, Pilate Khulumani, Steve Kinoti, Gabriel Mwaluko, Jude Padayachi, Brian Pazvakavambwa, Corlien Varkevisser and Godfrey Woelk. In August 1999 a group of three people (Alasford Ngwengwe, Corlien Varkevisser and Godfrey

---

* Corlien M. Varkevisser, Indra Pathmanathan and Ann Brownlee (1991) *Designing and conducting Health Systems Research projects. Part I: Proposal development and fieldword; Part II: Data analysis and report writing.* Ottawa: Health Sciences Division of the International Development Research Centre (IDRC) and Geneva:Programme on Health Systems Research and Development of the World Health Organisation.

Woelk) made further revisions and synchronised the different texts in the WHO/AFRO/HSR office in Harare, supported by staff of the HSR office (since 1999 headed by Isabel R. Aleta, with Makhamokha Mohale and Eric Naterop as APOs). Corlien Varkevisser and Ann Brownlee finalised and edited the modules, with the blessing of Indra Pathmanathan who this time could not participate. Deborah Karugonjo (Harare) and Merel Gallée (Amsterdam) provided highly valued assistance in the production of successive computerised versions. Funds for revising and publishing the HSR modules were made available by DGIS (Dutch Development Co-operation); SARA/AED, Washington; GTZ, Germany through the GTZ MCH/FP network for Health Systems Research in Southern Africa; WHO/AFRO and by WHO HQ, Geneva. IDRC, Canada assists in subsidised distribution of the modules.

A highly varied collection of people assisted in the production of earlier versions of the HSR modules. The cradle of the modules stood in Western Africa, where in the early eighties the Project for Strengthening Health Delivery Systems (SHDS), based in Boston University, USA, at the request of AFRO developed training materials in research protocol development. SHDS followed the step-by-step approach which till today is a major key to the success of HSR courses. Modules 1-17 in this volume are heavily adapted or new versions of the original SHDS modules.* The first adaptation took place in 1988, with 12 researchers from countries that participated in the Joint HSR Project (Omondi (Kenya), Sebatane and Makatjane (Lesotho), Chimimba and Msukwa (Malawi), Kitua and Savy (Seychelles), Tembo (Zambia) Munochiveyi, Taylor and Woelk (Zimbabwe) and Joint Project staff which also finalised the version (Corlien Varkevisser and Martien Borgdorff). These 'green modules'* found their way to Malaysia, where Indra Pathmanathan further developed them, with assistance from Maimunah Abdul Hamid, K. Mariappan and C. Sivagnanasundram (Sri Lanka), in the course of numerous protocol development workshops. The same occurred in Southern and Eastern Africa. At the initiative of Yvo Nuyens, who fathered the Joint HSR Project in WHO Geneva, and supported by IDRC (Annette Stark), the five volumes of the *Health Systems Research Training Series* emerged, of which *Designing and Conducting Health Systems Research Projects* formed Volume 2. These 'pink modules', published in 1991 in Ottawa by IDRC and WHO, form a thorough merge of the ever developing Southern African and Malaysian versions. They were integrated in Harare (Corlien Varkevisser and Leon Bijlmakers), in consultation with Indra Pathmanathan, and with thorough editing support from Ann Brownlee, one of the authors of the original SHDS modules. The present HSR modules are therefore a truly global production. It is even impossible to mention every contributor, because many HSR course facilitators and participants through their questions and critical remarks inspired further changes.

With such a colourful and interactive origin it seems highly unlikely that the present reprint will be the last one. Whenever the modules are used, they will be adapted. We hope, however, that in their present form they will last for some years and will be of use to health staff as well as university students.

Dr. Corlien M. Varkevisser, Royal Tropical Institute/University of Amsterdam

Dr. Ann Brownlee, University of California, San Diego

June 2003

---

* Regional Assessment of Health Systems Research Training in Eastern and Southern Africa. HSR Project and SARA/AED,Harare/Washington. SS Ndeki, 1997.

* Ann Brownlee, Thomas Nchinda and Yolanda Mousseau-Gershman (1983) *Health Services Research Course: How to develop proposals and design research to solve priority problems.* Boston: Boston University Health Policy Institute.

* Joint World Health Organisation/Royal Tropical Institute/Dutch Technical Development Co-operation Project on Health Systems Research for the Southern African Region (1988) *Health Systems Research Training Course: How to develop research proposals to solve priority health problems.* Geneva: World Health Organisation. WHO/SHS/HSR/88.3.

INTRODUCTION TO PART I: Proposal Development and Fieldwork

Since Module 1 of this volume provides a thorough introduction to the course and how it is organised, here we will only provide a brief overview of where the reader can find various types of information and how the course can be used.

Part I, Proposal Development and Fieldwork, contains modules 1-20, of which the first 18 will lead the course participants through all steps that the development of their proposal requires. Modules 19 and 20 guide them through the fieldwork period and preliminary data analysis.

Each module contains detailed instructions for group work on the successive steps in the development of the proposal. At the end of each module, facilitators will find Trainer's Notes, providing guidelines on how to present the modules and how to assist the groups in the writing of their research proposal.

After Module 20 an annex has been added with general guidelines for the planning and management of HSR workshops, the training methodology and the supervision of fieldwork. The annex includes an example of a course schedule and guidelines for budgeting an HSR course. Furthermore, an information circular for course participants and a course evaluation form have been added.

The course schedule presented applies to a full-time workshop for beginners, lasting just over two weeks. Depending on the level of the participants, the duration of the course can be shortened. The training materials can also be used in university settings, stretched out over a trimester or quarter with weekly sessions.

The most important guideline for the application of these HSR modules is BEING FLEXIBLE. The only general rule is that participants and facilitators should be conscious of the cyclical nature of the process of developing a research proposal. In many group work sessions, participants will therefore be referred back to earlier parts of the proposal they developed to make adjustments, it required. In Module 18 they are advised to review once again all sections of their proposal when writing the summary.

The managers of a course can adapt the time devoted to presentations and group work, as well as the sequence of modules, to the needs of their target groups. It is, for example, quite possible to combine Modules 4 (Statement of the problem) and 5 (Literature review). Also Modules 8 (Variables), 9 (Study type) and 10 (Data collection tools) are closely connected, and the group work for Modules 9 and 10 has already been combined. These modules form the core of any research proposal. Enough time will have to allotted to the group work for these modules and to field-testing of the data collection tools, as the quality of the tools will determine the quality of the findings.

For groups of participants that include health managers who already have extensive administrative experience it may be feasible to present modules 15 (Work plan), 16 (Budget) and 17 (Plan for project administration, monitoring and utilisation of results) as one block, entitled 'Management of a Research Project.'*

For Part II of this volume, Data Analysis and Report Writing containing modules 21-33, the same principles apply as outlined for Part I. We have included some specific guidelines on how to use these modules in the introduction to Part II.

It is advisable to distribute Part II during the first workshop. Participants who use qualitative research techniques, such as focus group discussions, would benefit from reading Module 23 (Analysis of qualitative data), when developing their plan for data analysis (Module 13). For those using quantitative techniques, Modules 22 (Description of variables) and 24 (Cross-tabulation of quantitative data) may be useful when variables are introduced (Module 8) or when data analysis is discussed (Module 13).

Designing and Conducting Health Systems Research Projects
Part I: Proposal Development and Fieldwork

Module 1

COURSE ORIENTATION

Module 1: COURSE ORIENTATION

COURSE OBJECTIVES

At the end of this course, you should be able to:

1. **Describe** what HSR is and understand the contribution it can make towards solving priority problems in health care within the local context.

2. **Prepare** a health systems research proposal by completing the following steps:

   - Identification, analysis and description of a research problem
   - Review of relevant literature and other available information
   - Formulation of research objectives
   - Development of an appropriate research methodology
   - Preparation of a work plan for the study
   - Identification of resources required and preparation of a budget
   - Development of a strategy for distribution and utilisation of research results

3. **Implement** this proposal in your own working situation during a period of 4-6 months.

4. **Analyse** and interpret the results.

5. **Prepare and present** a final report of the research findings, including recommendations for solving the problem and a plan of action for their implementation.

Whom is the health systems research course aimed at?

The health systems research (HSR) course has been developed for mid- and higher- level managers, health workers and health-related staff, as well as interested researchers.

What training method is used in the HSR course?

The training method applied is based on **learning by doing**. Course participants will themselves develop research proposals that they will actually carry out in the field.

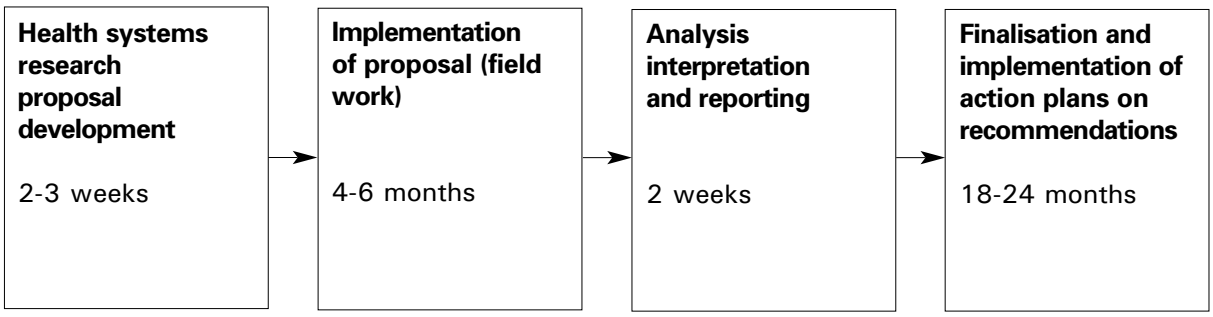
Each participant and trainer brings to this course his or her own experiences in applied research and in the management of health or health-related projects. Thus, the course should not be perceived as having a teacher-student orientation. It should rather provide a forum for sharing information where everyone can contribute the benefits of his or her own experience and knowledge. This sharing will add greatly to the richness and relevance of the course.

What type of projects will be developed?

Together with community leaders and other health decision-makers from the district, provincial or even national level, course participants will select priority problems in their own work situations that cannot be solved unless more information is collected. Preferably, the topics will have been selected before the training starts, although they may need more specification. In most cases, a **team of course participants** will carry out the planned research alongside their regular duties. Therefore, the project will have to be of modest size. For example, a maximum of 30 days for fieldwork and preliminary analysis per group member, and between 4,000 and 8,000 US$ per research project would be advisable.

How long is the course?

The course, which includes three main components, will cover a period of about 7 months, with an additional fourth component for the implementation of research results over 18-24 months. Thereafter the new activities resulting from the study are supposed to form part of the regular planning.



Component 1. HSR proposal development

The first 2-3 week workshop will provide an introduction to HSR**. If modules 1-4 are implemented in the field with guidance from a facilitator**, which we recommend, the proposal development workshop can be limited to two weeks or less. Participants will work in small groups and design research proposals, step by step, on the priority problems they have selected. As each new step is introduced, new concepts and research procedures will be presented. The participants will immediately apply these in the proposals they are developing. Modules 1-18 deal with proposal development.

Component 2. Implementation of the proposal

During the following 4-6 months, the same groups of participants will implement their proposals. It is therefore important that the groups are composed in such a way that they can easily cooperate during the fieldwork. Modules 19 and 20 give guidelines for the fieldwork and for writing a short fieldwork report - including preliminary results.

Component 3. Analysis of the data and report writing

After project implementation, participants will meet again for a 2-week workshop to further analyse and interpret the data. At the end of this workshop, a research report with recommendations for action will be prepared and presented to health policy makers, health staff and communities. Modules 21-33 pertain to data analysis, report writing, dissemination of research results and preparing implementation of recommendations based on the findings.

Component 4. Development and implementation of action plans

Together with the relevant stakeholders, (policy makers, managers, staff, community members), teams will draft action plans to implement the recommendations that are agreed upon. Because many of the participants are in direct positions of managerial responsibility, and because higher-level decision-makers and community members have been involved, it is expected that action plans can be implemented soon after the studies are completed. The proposed activities will normally be integrated in the district, provincial or national health plan and be subjected to regular monitoring and evaluation.

How will the research proposal be developed?

A number of basic steps have to be taken when developing a research proposal. These steps are presented in the flowchart below.

This flowchart appears on the back of each of the pages that mark the beginning of modules 3-18. The step in the proposal development process that the module addresses is indicated by *double lines* around the appropriate box in the flowchart.

Flowchart: Steps in the development of a health systems research proposal

Questions you must ask	Steps you will take	Important elements of each step
What is the problem and why should it be studied?	Selection, analysis and statement of the research problem	- problem identification - prioritising problems - analysis - justification
What information is available?	Literature review	- literature and other available information
Why do we want to carry out the research? What do we hope to achieve?	Formulation of research objectives	- general and specific objectives - hypotheses
What additional data do we need to meet our research objectives? How are we going to collect this information?	Research methodology	- variables - types of study - data collection techniques - sampling - plan for data collection - plan for data processing and analysis - ethical considerations - pre-test or pilot study
Who will do what, and when?	Work plan	- human resources - timetable
What resources do we need to carry out the study? What resources do we have?	Budget	- material support and equipment - money
How will the project be administered? How will utilisation of results be ensured?	Plan for project administration and utilisation of results	- administration - monitoring - identification of potential users
How will we present our proposal to relevant authorities, community and the funding agencies?	Proposal summary	- briefing sessions and lobbying

NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

It should be stressed that designing a research proposal is not a linear but a cyclical process. Throughout the course there will therefore be opportunities to review and, when the need arises, to revise parts of the proposal that have already been drafted. When developing the research methodology, for example, the teams may find that the objectives and even the statement of the problem need to be revised to be made more specific. When finalising the work plan and budget, the teams may determine that the research design, for financial reasons, may need to be revised so the project is more modest and thus less costly.

By the end of the first part of the course, each group will have developed a research proposal with the following chapters: (For details, see **Module 18**.)

EXECUTIVE SUMMARY

1. INTRODUCTION

1.1 Background information
1.2 Statement of the problem
1.3 Literature review

} Literature review may be partially or fully integrated in 1.1 and 1.2

2. OBJECTIVES

3. METHODOLOGY

3.1 Study type, variables, data collection techniques
3.2 Sampling
3.3 Plan for data collection
3.4 Plan for data processing and analysis
3.5 Ethical considerations
3.6 Pre-test

4. WORK PLAN (including description of project staff)

5. BUDGET (including explanatory note on major budget posts)

6. PLAN FOR ADMINISTRATION, MONITORING, AND UTILIZATION OF RESULTS

References

List of abbreviations (if applicable)

Data collection instruments (annexed)

In the second workshop for data analysis and report writing, a similar approach will be followed.

How may this set of modules be used?

The course has been organised in such a way that each module can be dealt with independently. A module includes:

- A **presentation** of the necessary theory and concepts to enable the participants to carry out this specific step in proposal development or data analysis and report writing. Presentations last between 30 minutes and an hour and include opportunity for questions and discussion.

- **Group work** during which groups, with assistance of their facilitator, utilise these concepts in the development of their proposal or in data analysis and report writing. The modules for proposal development, in particular, contain detailed instructions for group work. Group work may last from 1-4 hours per module, and sometimes longer.

- **Reporting** of the results of the group work in **plenary** by a member of each group, so that other groups and facilitators can comment. Plenaries are of crucial importance during the first workshop. During the data analysis workshop they are less frequent as not all modules are relevant for all groups. On the average, each group has 15 minutes for presentation and discussion, but for important topics this may be 30 minutes.

- Sometimes a module contains an **exercise**, either using examples provided during the presentation or using the group work results of other groups.

Depending on the level of the groups, it may be possible to combine certain modules and to shorten or lengthen the time allocated for presentations and group work and the total workshop time. For programme managers, for example, one week may be sufficient to prepare a first draft of a research proposal. Provincial and district level staff with some research experience may need 2 weeks, whereas novices to research will need the full 2 ½ weeks.

Note:

Participants are advised to read the course materials beforehand so that they can benefit, as much as possible, from the presentations and group work. It may be extremely useful for the participants to (re)read the course material after the presentation and group work as well, especially if they have had no previous research training or experience.

Trainer's Notes

Module 1: COURSE ORIENTATION

Timing and training methods

1-1½	hours	Personal introductions of participants and facilitators (if not completed the night before)
¾ -1	hour	Course orientation
½	hour	Administrative remarks
1-2½	hours	TOTAL TIME

Materials

- Name tags for participants and trainers
- Flipcharts and markers
- Course training materials for participants
- Overhead sheets for presentation

Personal introduction of participants and facilitators

If you were unable to do the mutual introduction of participants on the evening before the course begins, have all the participants (including the facilitators) introduce themselves. Make certain everyone indicates his or her profession, major activities and research experiences and interests. This may be done by having participants interview each other in pairs and then each introduces the person he or she interviewed. Names and a summary of the interview could be put on a flipchart and stuck to the wall.

The introduction may take 1 - 1½ hours.

Course orientation

- Present the major objectives of the course and stress its practical orientation. It should be clear to all participants that they will each work as part of a small group to develop a research proposal which they themselves will carry out. It should also be stressed that one important goal of the course is that the research findings will be used to help solve the problem the group has investigated. Therefore, decision-makers and users (government or non-government workers, and community members) should be involved in the choice of the topic, the review of the proposal, and the discussion of research findings and recommendations. Depending on the location of the course and the participants, each team should consider holding information sessions for interested persons on their return home. This will be discussed again later in the course.

- Emphasise the uniqueness of each participant's background and experience, pointing out how important it will be for everyone to contribute to the development of the proposal and to learn from each other.

- Distribute the course-training document to the participants. Describe how the course will be structured and how the training document will be used. Show the flowchart that appears at the beginning of each module. Explain that each session contains a presentation and group work during which each group will apply the concepts presented in the development of its proposal. Indicate that directions for group work are presented in boxes with double lines around them. Mention that some sessions also have exercises, which are presented in boxes with single lines. Discuss the fact that in some modules annexes provide more details on research methodology for those who are interested.

- Stress that the end product of the first workshop will be a research proposal that will be written, step by step, by the participants according to the plan presented on page 6.

Administrative issues

- Present any other information concerning the course and administrative arrangements that may be necessary and ask for final questions.

Designing and Conducting Health Systems Research Projects
Part I: Proposal Development and Fieldwork

Module 2

INTRODUCTION TO HEALTH SYSTEMS RESEARCH

Module 2: INTRODUCTION TO HEALTH SYSTEMS RESEARCH

OBJECTIVES

At the end of this session, you should be able to:

1. **Describe** the major characteristics of research.
2. **Describe** various components of the health system as a basis for understanding HSR.
3. **Describe** types of information for decision-making in the health system and the contribution various disciplines can make in providing such information.
4. **Describe** the purpose, scope and characteristics of HSR.4.

I. The development of health systems research

II. What is health systems research?

III. Participants in health systems research

IV. Guidelines for health systems research

I. THE DEVELOPMENT OF HEALTH SYSTEMS RESEARCH

Why did HSR develop?

By adopting of the philosophy and strategies for **Health For All**, politicians and health staff at all levels are committed to ensuring that *all* people will attain a level of health that enables them to participate actively in the social and economic life of the community in which they live.

Although research has made major contributions to health by providing knowledge of the causes of diseases and by developing the technology to cure and prevent disease and promote health, Health For All is far from being achieved.

Why is there still so much disease that could have been prevented or cured? Because health services by themselves cannot control all of the factors that influence health. Poverty and political systems which either widen or narrow the gap between rich and poor and which promote or neglect the education of girls, for example, influence the health of people. Drought and wars may bring malnutrition and disease with which the health services can hardly cope. While communicable diseases such as smallpox and, to some extent, leprosy may be gradually conquered due to improved environmental conditions and extra effort on the part of the health services, new diseases such as HIV/AIDS may appear which upset the whole health care system and society at large.

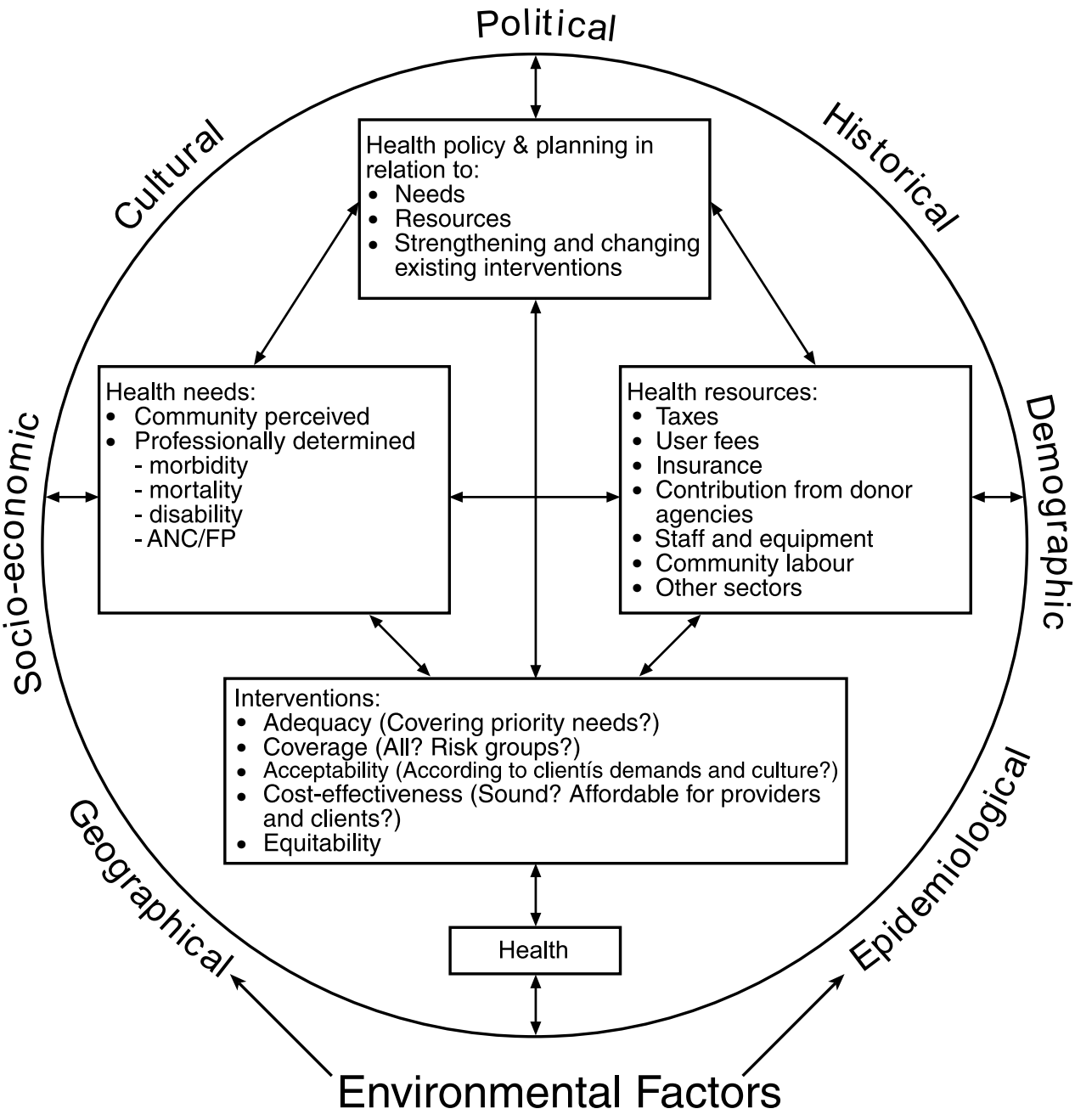
This **complex of environmental factors** – geographical, socio-economic, cultural, political, demographic, epidemiological – not only **influences the health of people**, it also affects the **health services**. Countries suffering from poor economics, wars and drought usually have poorly functioning health services as well.

Still, even within less favourable environments, some services function better than others. A very important factor is the **quality of information on which policy makers base their decisions**. Very often this information is vague or missing. Then decisions on interventions can be completely off track, which means that money is wasted. Basic questions which health policy makers need answered include, for example:

- What are the *health needs* of (different groups of) people, not only according to health professionals but also according to the people themselves? Can shared priorities be agreed upon?
- To what extent do the present *health interventions* cover these priority needs? Are the interventions acceptable to the people in terms of culture and cost, especially to the poor? Are they provided as cost-effectively as possible?
- Given the *resources* we have, could we cover more needs, or more people, in a more cost-effective way? Is it possible to introduce or expand cost-sharing through insurance, to reduce the risk of unexpected high costs, in particular for the economically vulnerable? Could co-operation with the private/NGO sector be improved? Could donor agencies help solve well-defined bottlenecks in the system?
- Is it possible to better *control the environmental factors* which influence health and health care? Can other sectors help (education, agriculture, public works/roads, etc.)?
  (See **Figure 2.1**.)

These questions cannot be answered without collecting more information through research. That is why, since the end of the 1970's, **Health Systems Research** (HSR) has been developed.

Figure 2.1: Environmental and health system factors influencing attainment of Health For All



II. WHAT IS HEALTH SYSTEMS RESEARCH?

What is research?

RESEARCH is the systematic collection, analysis and interpretation of data to answer a certain question or solve a problem.

Characteristics of research:

- It demands a clear statement of the problem.
- It requires clear objectives and a plan (it is not aimlessly looking for something in the hopes that you will come across a solution).
- It builds on existing data, using both positive and negative findings.
- New data should be *systematically* collected and analysed to answer the original research objectives.

Health research serves two major purposes:

First, **basic research** is necessary to generate new knowledge and technologies to deal with major unresolved health problems. Second, **applied research** is necessary to identify priority problems and to design and evaluate policies and programmes that will deliver the greatest health benefits, making optimal use of available resources.

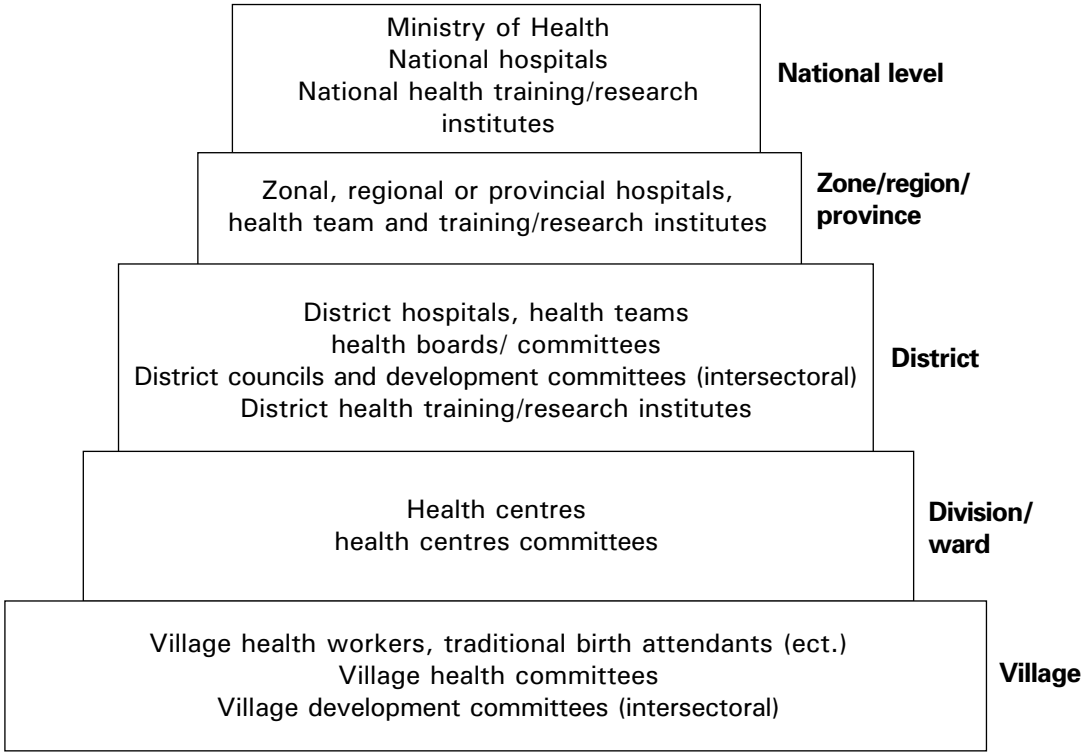
During the past two (or even three) decades there has been a rapid evolution of concepts and research approaches to support managerial aspects of health development. Many of these have been described by specific terms such as operations/operational research, health services research, health management research, applied research and decision-linked research. Each of these has made crucial contributions to the development of HSR (WHO 1990).

HEALTH SYSTEMS RESEARCH is ultimately concerned with improving the health of people and communities, by enhancing the efficiency and effectiveness of the **health system** as an integral part of the overall process of socio-economic development, with full involvement of all partners.

What is meant by a health system?

There are different interpretations of what a health system is. Some give a **narrow definition** and only consider the **different levels of the public health care services** as a health system (see **Figure 2.2**.)

Figure 2.2: Public health care system



The inclusion of the district council, district development committee and village development committee indicates, however, that some 25 years after Alma Ata* it has been widely recognised that local administration and other sectors than the health sector alone carry responsibility for the health of the people in a village, district or region.

Many HSR researchers have a wider perception of health systems. They also include the *private sector*. The private sector has many possible components:

- Non-governmental organization (NGO) care, provided by churches, Red Cross, local NGOs, etc.
- Medical practice by private doctors, nurses, or by quacks who provide injections and drugs without medical training.
- The pharmaceutical sector (licensed pharmacies or unlicensed sellers).
- The large 'non-biomedical' professionalised healing systems (Ayurvedic, Chinese, Unani, homeopathic, chiropractic, etc.)
- Traditional (or folk) medicine, with traditional birth attendants, herbalists and diviners, who may either identify natural or supernatural causes of disease (witchcraft, angry ancestors) and treat patients accordingly.

The **Primary Health Care (PHC) approach** has broadened the horizon of medical care providers considerably. PHC put individuals and communities in the centre of attention. Individuals providing s*elf-care* (what mothers and other relatives do to keep children and themselves healthy) and

* The WHO/UNICEF/World Council of Churches conference, which in 1978 laid the foundation for the worldwide Primary Health Care approach.

traditional/folk healers were accepted as important potential allies of health staff. So were personnel from *other sectors,* which could support health, for example, through the construction of roads, the improvement of education, water, sanitation, and through income generation.

Figure 2.3: A broadly defined health system

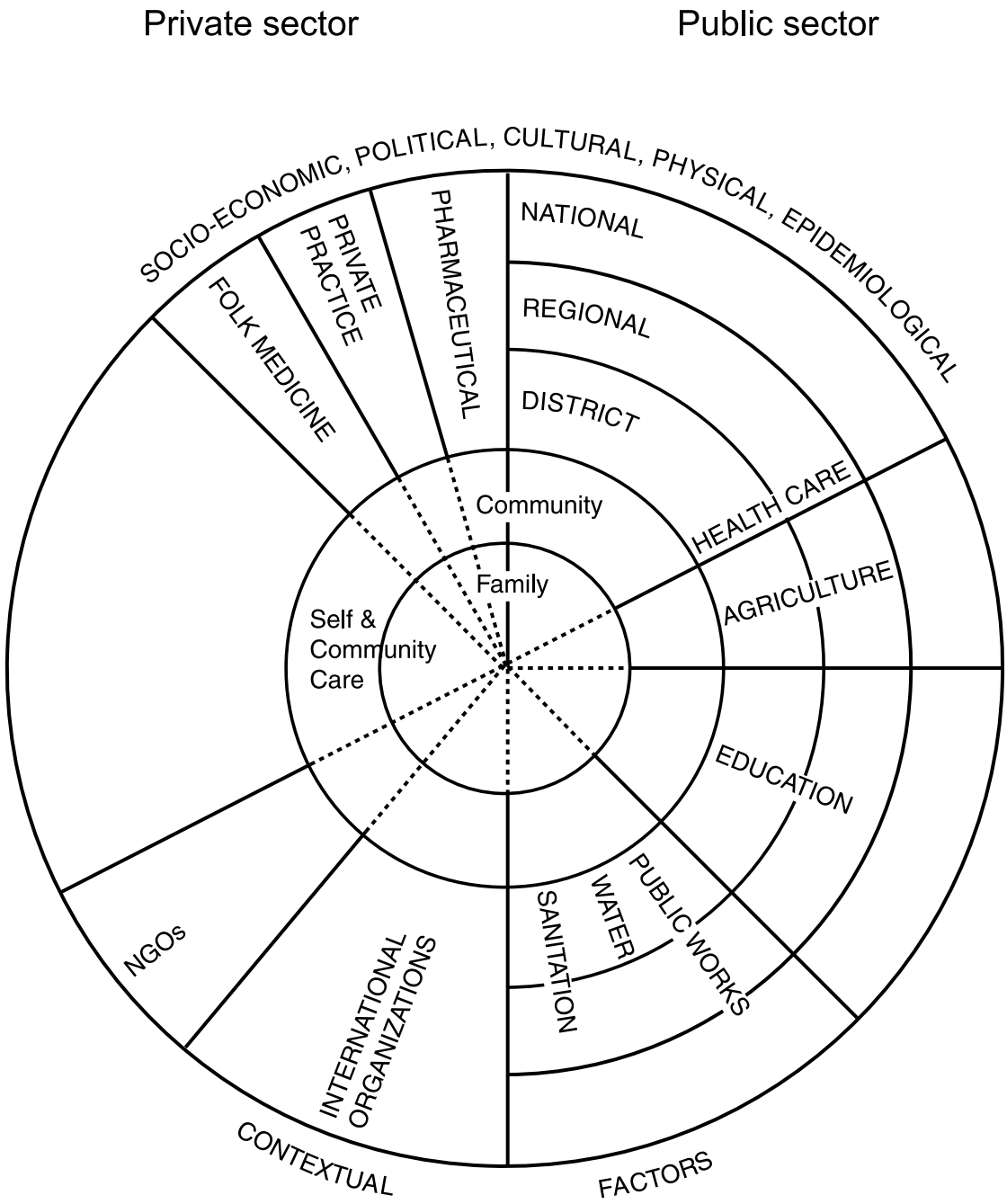


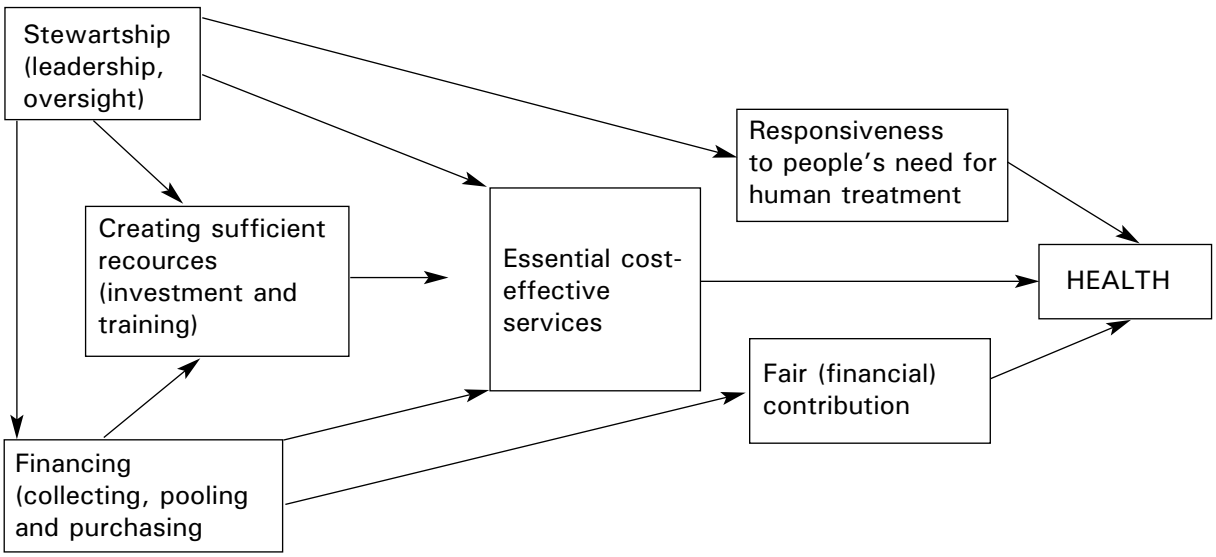
Figure 2.3 presents the **widest possible definition of a health system,** including *all public and private sectors/institutions which directly influence and support the health of people,* embedded in the wider environmental context that was described in **Figure 2.1**.

This figure would take different shapes in different societies, but everywhere **individuals** form part of a *network of family and community members* who are concerned about their health. This network prescribes or advises how to prevent illness and what to do in case of ill health. In many societies, mothers and grand mothers are key figures in early childcare. They determine nutritional and hygiene practices, alert children to dangers, provide care in case of disease, and teach children the basics of self-care.

At the other end of the spectrum, a **public authority** is responsible for the well being of all people inhabiting its territory. Nowadays governments of states organise public health care and, to some extent, regulate private health care initiatives. Through other social services (e.g., education, social welfare), through laws and taxes and police and army, governments are supposed to assure their citizens the resources to survive and live in peace. Since time immemorial this has been the duty of rulers, although each society has developed its own ways of ensuring 'health for all'.

When in the 1980s many countries were struck by chronic economic crises, the World Bank advocated **structural adjustment programmes** to reorganise the economies, which relied on market mechanisms rather than on state control with subsidies and protection. The health and educational sectors were inevitably affected and went through a series of reforms that hit the consumer hard. The World Health Organization recognised the need for **health reforms**, but under the condition that these would leave the goal of HEALTH FOR ALL in tact. It therefore focussed attention on fairness of the system, which should also be affordable to the poor, and at the same time stressed that the system should be responsive to the need of patients for human, respectful treatment (see **Figure 2.4**).

Figure 2.4: Objectives and functions of the health system



Health is expressed as life expectancy by the WHO, taking into account the time lived with a disability of any kind (also due to chronic disease and old age). In the highly industrialised countries of Western Europe, for example, the average life expectancy of men is 74 years, of which 6.5 years are with disability; for women it is 80.8 and 7 years, respectively. In Africa in areas most struck by AIDS, men live now on average only 45.6 years, of which 7.6 years are with disability; for women the respective means are 48 and 8 years. *Responsiveness* to patients' human needs would mean respecting the patient's dignity and autonomy and reducing the fear and shame that sickness brings with it. *Fairness* ideally means financial protection for everyone by payment according to financial capacity. This can best be assured by pre-payment through an insurance

system, with fees according to capacity. The insurance revenues are then pooled and costs of care paid from the pool, so that in fact the rich help to cover the treatment of the poor. Unfortunately, such a system is hard to organise in the least developed countries where rural areas harbour mainly poor, but WHO counts on international solidarity and donor agencies for contributions.

The health system comprises both public and private health services but, for the time being, no agricultural, educational or other sectors, however relevant. The first urgency is the *performance* of the health system, which should be as good as possible, given the available means. To reach that aim, WHO set some criteria. Ministries of Health should weigh the public health importance of proposed health actions, set priorities, and thoroughly investigate the cost-effectiveness of different possible interventions to select the highest value for the money. In terms of *resources,* they should strive for a balance between investments, the use made of these investments and their maintenance. For example, if staff members are highly trained but their knowledge is under-utilised, or if buildings, equipment and means of transport cannot be maintained, these investments are highly wasteful. Likewise the services don't function well if there is no money left for consumables such as essential medicines. The patients then have to buy medicines on the private market, out of their pockets and at unnecessary high costs, which the poor cannot afford. Good *oversight* is required to achieve an optimal balance among the different expenses, and it is one of the aims of HSR to provide the policy makers with the relevant data.

Good oversight and 'stewartship' is also required to develop a fair *financing system*. The Ministry of Health is usually the appropriate institution to collect money from taxes and donor agencies to finance the health care system. In the 1980's it became clear that even PHC services could never function adequately with the required coverage (health for ALL) without a contribution from the clients. User fees were introduced in countries that hitherto had provided care free of cost, but this appeared to hit the poor out of proportion despite exemption rules. Hence WHO proposes a more structural solution by introducing prepayment through insurance and pooling of resources, which is beneficial for the poor.

Although the MOH, in many developing countries, is still the principal provider of health care, if it is to achieve the most cost-effective care, it has to consider the use of the private sector and contract services out in cases where this would be cheaper. Consequently, the MOH has to set standards of care and control for deviation in the private sector as well as the public. To have oversight and control is one of the major present day challenges for Ministries of Health.

Specific questions for specific levels of service

HSR is not only of use to policy makers; at each level managers may have questions that require further research.

Health policy makers may, for example, want to know:

- What are the prospects for voluntary community-based insurance? What would acceptabl contributions for different income groups? Should the pooling of resources take place on a community or national basis?

- How can user-fees be used as an instrument to direct demands for care to the appropriate level?

Managers at district/provincial level may raise questions such as:

- Why is neonatal mortality in certain districts much higher than in other districts?

Hospital directors may ask:

- Why do we have such a high rate of complications during child birth? Are the first-line services available and adequate? Are our own services adequate? Are mothers coming late for delivery and, if so, why?

Managers at village level (village health committees) may want to know:

- How can we assist women with little or no education so that they can effectively recognise the symptoms of pneumonia and go in time to the health centre with their children?
- How much community labour will be required to manage the new water system?

(Please add your own examples.)

The **major objective of HSR** is to provide health managers at all levels, as well as community members, with the relevant information they need to make decisions on health-related problems they are facing.

We must be aware that problems at one level of the health system are usually connected with problems or deficiencies at other levels (see **Figure 2.2**). HSR should address problems from the differingperspectives of all those who are, directly or indirectly, involved. Otherwise we run the risk of coming up with results which only partly explain the problem and which are therefore insufficient to solve it.

III. PARTICIPANTS IN HSR

It is evident that many issues in health are interrelated and interact with issues in other sectors, such as production, education, the condition of wells or roads, and broader environmental factors. Research in health systems must recognise this. The research skills that are required may need to come from a variety of disciplines, e.g., public health/medicine, health economics, behavioural and social sciences, and agriculture. Therefore HSR is **multi-disciplinary** in nature.

Even simple research that is conducted at the operational level may require research skills from different disciplines to provide sufficient and relevant information to support decision-making. Therefore, training in HSR includes relevant aspects from various research disciplines.

Researchers who work in HSR will have to work in a **trans-disciplinary** way, which means working together as a **team** throughout all phases of the research. In the process, they need to acquire a basic understanding of the concepts and approaches as well as the potential and limitations of research techniques used in sister disciplines.

HSR, however, is not the concern of scientists alone.

Who should be involved in HSR?

The participatory nature of health systems research is one of its major characteristics. To ensure that the research is relevant and appropriate, everyone directly concerned with a particular health or health care problem should be involved in the research project(s) focused on it. This may include policymakers, managers from the health and other public services involved, health care providers and the community itself. Their involvement is critical if the research activities are to make a difference:

- If decision-makers are only involved after completion of the study, the report may just be shelved.
- If staff of health and other public services are only involved in data collection and not in the development of the proposal or in data analysis, they may not be motivated to collect accurate data or carry out the recommendations.

- If the community is only requested to respond to a questionnaire, the recommendations from the study may not be acceptable.

- If professional researchers are not involved in the implementation of recommendations, they may have little concern for the feasibility of the recommendations.

The roles that various types of participants will play in the research project will depend on the level and complexity of the particular study as well as its area of focus. Some projects are very complex and may need expertise from several levels, sectors and disciplines. Others may focus on simpler problems and require a more modest set-up. Health personnel may even play a major role in simple studies focusing on practical problems in their own working situations, although their projects may require assistance from researchers with skills in relevant disciplines.

Note:

Because of the **participatory nature** of HSR, in the modules that follow we will use the term **RESEARCHER** to mean anyone actively involved in planning and conducting the research.

IV. GUIDELINES FOR HSR

Bearing in mind that HSR is undertaken primarily to provide information to support decision-making that can improve the functioning of the health system, we summarise by suggesting some essential guidelines for success:

1. HSR should focus on **priority problems** in health care.

2. It should be **action-oriented**, i.e., aimed at developing solutions.

3. An integrated **multi-disciplinary** approach is required, i.e., research approaches from many disciplines are needed since health is affected by the broader context of socio-economic development.

4. The research should be **participatory** in nature, involving all parties concerned (from policymakers to community members) in all stages of the project.

5. Studies should be scheduled in such a way that results will be available when needed for key decisions; research must be **timely**. Otherwise, it loses its purpose.

6. Emphasis should be placed on comparatively **simple, short-term research designs** that are likely to yield practical results relatively quickly. Simple but effective research designs are difficult to develop but much more likely to yield useful results when needed.

7. The principle of **cost-effectiveness** is important in the selection of research projects. Program management and operational research should focus, to a large extent, on low-cost studies that can be undertaken by management and service personnel in the course of daily activities. (There is a need for larger studies as well, however, which may require outside funding and full-time research staff.)

8. Results should be presented in **formats most useful for administrators, decision-makers and the community**. Each report should include:

   - A clear presentation of results with a summary of the major findings adapted to the interests of the party being targeted by the research.

- Honest discussion of practical or methodological problems that could have affected the findings.

- Alternative courses of action that could follow from the results and the advantages and drawbacks of each, formulated with inputs from all parties concerned.

8. **Evaluation** of the research undertaken should concentrate on its ability to influence policy, improve services and ultimately lead to better health, rather than on the number of papers published.

Thus, an HSR project should not stop at finding answers to the questions posed, but include an assessment of what decisions and activities have evolved from the study.

Trainer's Notes

Module 2: INTRODUCTION TO HEALTH SYSTEMS RESEARCH

Timing and training methods
1 hour Introduction and discussion

Adapting the presentation to the participants

It is recommended that the content and focus of the module be adapted to the level and interests of the participants. For example:

1. Review the background of participants (e.g., primary health care, clinical medicine, research, policy-making, or community leadership).

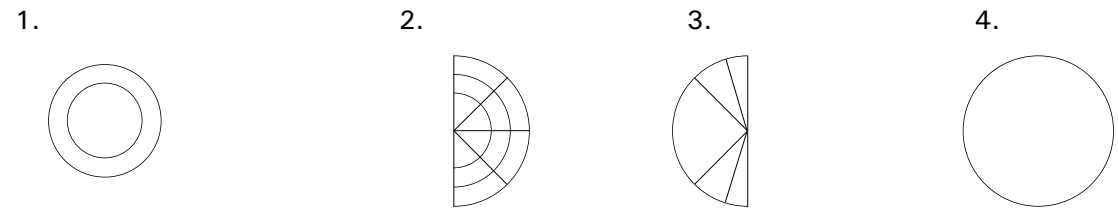
2. Based on this review, select suitable examples related to the background of participants to illustrate each concept.

   Remember that understanding of abstract concepts is facilitated if participants can relate them to their own experiences.

3. The focus and scope of this module can be varied in accordance with the expected future roles of participants in the research teams.

   If participants are fairly specialised personnel from a single discipline or from just a couple of disciplines, it would be useful to focus on the multi-disciplinary aspect of HSR and the types of information that disciplines other than those represented in the workshop can provide. For example, if workshop participants are hospital managers and clinicians, it may be most useful to illustrate the uses of research input from behavioural science for HSR; if participants are behavioural scientists such as health education officers and sociologists, it may be most useful to emphasize the importance of input from management sciences, health economics and clinical epidemiology.

4. When presenting figures, use different overhead sheets for the different components. For example, the health system (Figure 2.3) could be presented on four superimposed sheets.



5. Ask participants to give examples of topics suitable for HSR from their own working environment.

6. Ask whether they have participated in evaluations or other regular research activities. Demystify the concept of HSR. Identify in what stages of the research they have participated and whether the participation was optimal.

7. Try to draw out the points mentioned in the guidelines from the participants themselves. By the end of the introductory session they should be able to come up witsome of the points on their own.

Proposed pre-workshop reading (to be sent to participants before the workshop)

World Health Organisation (1990) *Health systems research: Background document at the World Health Assembly.* Geneva: WHO. A43/Technical Discussions/3.

Proposed additional reading (to be available in the course library)

Bobadilla JL (1998) *Searching for essential health services in low- and middle-income countries.* Washington DC: Inter-American Development Bank.

Joint Project on Health Systems Research for the Southern African Region (1994) *Summaries of Health Systems Research Reports 1988-1993.* Harare, WHO Sub-regional Office III (now AFRO).

Joint Project on Health Systems Research for the Southern African Region (1997) *Health Systems Research: Does it make a difference?* Update 1996 (3rd edition) WHO: Geneva.

Joint Project on Health Systems Research for the Southern African Region. *Series Health Systems Research: It can make a difference.*

Volume 1: Availability, provision and use of drugs (1994)
Volume 2: Factors associated with maternal mortality (1994)
Volume 3: Under utilisation of TB services in Southern Africa. Exemplary research protocol, research results and their implementation (1996)
Volume 4: Factors influencing the functioning of primary health care at village level (1996)

Murray CJL, Kreuser J, Whang W (1994) Cost-effectiveness analysis and policy choices investing in health systems. In: Murray CJL, Lopez A. *Global comparative assessments in the health sector: disease burden, expenditures and intervention packages.* Geneva: World Health Organization.

Scrimshaw SCM, Hurtado E (1987) *Rapid assessment procedures for nutrition and primary health care. Anthropological approaches to improving program effectiveness.* Tokyo: United Nations University and Los Angeles CA: Latin America Centre, University of California.

Taylor CE (1984) *The uses of health systems research*. World Health Organisation: Geneva. Public Health Paper 78.

World Health Organization (2000) *The World Health Report. Health systems: improving performance.* Geneva: WHO.

World Health Organisation (1988) *Health systems research in action: case studies from Botswana, Columbia, Indonesia, Malaysia, the Netherlands, Norway, and the United States of America*. Geneva: WHO.

World Health Organization, Programme on Health Systems Research and Development (1991) *From research to decision making.* Geneva: WHO.

Module 3

IDENTIFYING AND PRIORITISING TOPICS FOR RESEARCH

Flowchart: Steps in the development of a health systems research proposal

Questions you must ask	Steps you will take	Important elements of each step
What is the problem and why should it be studied?	Selection, analysis and statement of the research problem	- problem identification - prioritising problems - analysis - justification
What information is available?	Literature review	- literature and other available information
Why do we want to carry out the research? What do we hope to achieve?	Formulation of research objectives	- general and specific objectives - hypotheses
What additional data do we need to meet our research objectives? How are we going to collect this information?	Research methodology	- variables - types of study - data collection techniques - sampling - plan for data collection - plan for data processing and analysis - ethical considerations - pre-test or pilot study
Who will do what, and when?	Work plan	- human resources - timetable
What resources do we need to carry out the study? What resources do we have?	Budget	- material support and equipment - money
How will the project be administered? How will utilisation of results be ensured?	Plan for project administration and utilisation of results	- administration - monitoring - identification of potential users
How will we present our proposal to relevant authorities, community and the funding agencies?	Proposal summary	- briefing sessions and lobbying

NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 3: IDENTIFYING AND PRIORITISING TOPICS FOR RESEARCH

OBJECTIVES

At the end of this session you should be able to:

1. **Identify** criteria for selecting health-related problems to be given priority in research
2. Use **a group consensus technique to set priorities for research, applying the selected criteria on a number of research topics**
3. **Select** an appropriate subject for a research proposal that will be developed by your group during the course.

I. Identification of topic

II. Criteria for prioritising topics for research

III. Nominal group technique

Note:

If topics have been selected **before** the workshop, either by health managers who asked for the study or by the participants together with their health managers and community leaders, recapitulate **part I of Module 3**, then **go straight to Module 4**, even if the teams have already done some problem analysis in the field. If the participant teams need to re-examine the research topics they selected before the workshop, **section II of Module 3** can be used together with Module 4.

I. IDENTIFICATION OF TOPIC

In the previous module, a number of research questions were presented that may be posed at the various levels of the health system.

These questions can be placed in three broad categories, depending on the type of information sought:

1. **Description of the health situation, required for planning interventions**

   Planners need to know, for example, the magnitude and distribution of health needs in a population as well as of services; the risk factors for certain problems and people's awareness; the utilization patterns and cost-effectiveness of available and potential other interventions, in order to formulate adequate policies and adapt or plan interventions.

2. **Information required to evaluate ongoing interventions,** for example with respect to:

   - coverage of priority health needs
   - coverage of target group(s)
   - acceptability and quality
   - cost-effectiveness
   - impact on health,

   to assess progress and the need for adjustment on a routine basis.

3. **Information required to define problem situations in interventions in any of the fields mentioned under 2, and to analyse possible causes in order to find solutions**. These causes may include lack of or inequitable distribution of resources, vague policies, and any environmental factors affecting needs, interventions and resources (See **Figure 2.1**).

Although research in support of planning and evaluation (categories 1 and 2 mentioned above) is an important focus for HSR, the modules will emphasise problem-solving research, because health managers are frequently confronted with problems of this type. It is assumed, however, that research skills acquired in the present course will be of use in the broader field of planning and evaluation as well. In **Modele 4**, moreover, all three types will be treated.

Whether a problem situation requires research depends on three conditions:*

1. There should be a **perceived difference or discrepancy** between what exists and the ideal or planned situation;

2. The **reason(s)** for this difference should be **unclear** (so that it makes sense to develop research questions); and

3. There should be **more than one possible answer** to a question or more than one solution to the problem.

* This paragraph has been adapted from Fisher et al. (1983)

For example:

Problem situation:

In District X (pop. 145,000) sanitary conditions are poor (5% of households have latrines) and diseases connected with poor sanitation, such as hepatitis, gastro-enteritis and worms, are very common. The Ministry of Health has therefore initiated a sanitation project that aims at increasing the percentage of households with latrines by 15% each year. The project provides materials, and the population should provide labour. Two years later, less than half of the target has been reached.

Discrepancy:

35% of the households **should have** latrines, but only 15% **do have** them.

Research question:

What factors can explain this difference?

Possible answers:

1. **Service-related factors**, such as forgetting to adequately inform and involve the population, bottlenecks in the supply of materials, differences in training and effectiveness of sanitary staff, lack of co-operation between sectors.

2. **Population-related factors**, such as situations where community members lack understanding of the relationship between disease and sanitation or have other problems, for example due to poverty, which they consider more important.

3. **Physical factors/ecosystems**, such as hard soil, or land subjected to frequent flooding.

II. CRITERIA FOR PRIORITISING TOPICS FOR RESEARCH

Because HSR is intended to provide information for decision-making to improve health care, the selection and analysis of the research topic should involve those who are responsible for the health status of the community. This would include managers in the health- and health-related services, health care workers and community leaders, as well as researchers.

Each topic that is proposed for research has to be judged according to certain guidelines or criteria. There may be several ideas to choose from. Before deciding on a research topic, each proposed topic must be compared with all other options. The guidelines or criteria discussed on the following page can help in this process:

Criteria for selecting a research topic:

1. Relevance
2. Avoidance of duplication
3. Urgency of data needed (timeliness)
4. Political acceptability of study
5. Feasibility of study
6. Applicability of results
7. Ethical acceptability

1. Relevance

The topic you choose should be a priority problem. Questions to be asked include:

- How large or widespread is the problem?
- Who is affected?
- How severe is the problem?

Try to think of serious health problems that affect a great number of people or of the most serious problems that are faced by managers in the area of your work.

Also, consider the question of who perceives the problem as important. Health managers, health staff and community members may each look at the same problem from different perspectives. Community members, for example, may give a higher priority to economic concerns than to certain public health problems. To ensure full participation of all parties concerned, it is advisable to define the problem in such a way that all have an interest in solving it. Even **within** villages, opinions may differ on how important a problem is. It is therefore obligatory to discuss the problem with community leaders, as well as peripheral villagers, males as well as females, rich and poor, exploring their perceptions of the problem.

Note:

If you do not consider a topic relevant, it is not worthwhile to continue rating it. In that case you should **drop it from your list**.

2. Avoidance of duplication

Before you decide to carry out a study, it is important that you find out whether the suggested topic has been investigated before, either within the proposed study area or in another area with similar conditions. If the topic has been researched, the results should be reviewed to explore whether major questions that deserve further investigation remain unanswered. If not, another topic should be chosen.

Note:

Also, consider carefully whether you can find answers to the problem in already available, unpublished information or just by using your common sense. If so, you should **drop the topic from your list**.

3. Urgency of data needed (timeliness)

How urgently are the results needed for making a decision or developing interventions at various levels (from community to policy)? Consider which research should be done first and which can be done later.

4. Political acceptability

In general it is advisable to research a topic that has the interest and support of the local/national authorities. This will increase the chance that the results of the study will be implemented. Under certain circumstances, however, you may feel that a study is required to show that the government's policy needs adjustment. If so, you should make an extra effort to involve the policy-makers concerned at an early stage, in order to limit the chances for confrontation later.

5. Feasibility

Look at the project you are proposing and consider the complexity of the problem and the resources you will require carrying out your study. Thought should be given first to manpower, time, equipment and money that are locally available.

In situations where the local resources necessary to carry out the project are not sufficient, you might consider resources available at the national level; for example, in research units, research councils or local universities. Finally, explore the possibility of obtaining technical and financial assistance from external sources.

6. Applicability of possible results/recommendations

Is it likely that the recommendations from the study will be applied? This will depend not only on the management capability within the team and the blessing of the authorities but also on the availability of resources for implementing the recommendations. Likewise, the opinion of the potential clients and of responsible staff will influence the implementation of recommendations.

7. Ethical acceptability

We should always consider the possibility that we may inflict harm on others while carrying out research. Therefore, review the study you are proposing and consider important ethical issues such as:

- How acceptable is the research to those who will be studied? (Cultural sensitivity must be given careful consideration). Is the problem shared by target group and health staff/researchers?
- Can informed consent be obtained from the research subjects?
- Will the condition of the subjects be taken into account? For example, if individuals are identified during the study who require treatment, will this treatment be given? What if such treatment interferes with your study results?
- Will the results be shared with those who are being studied? Will the results be helpful in improving the lives or health of those studied?

These criteria can be measured by the following rating scales:

SCALES FOR RATING RESEARCH TOPICS

Relevance

1. = Not relevant
2. = Relevant
3. = Very relevant

Avoidance of duplication

1. = Sufficient information already available
2. = Some information available but major issues not covered
3. = No sound information available on which to base problem-solving

Urgency

1. = Information not urgently needed
2. = Information could be used right away but a delay of some months would be acceptable
3. = Data very urgently needed for decision-making

Political acceptability

1. = Topic not acceptable to high level policymakers
2. = Topic more or less acceptable
3. = Topic fully acceptable

Feasibility

1. = Study not feasible, considering available resources
2. = Study feasible, considering available resources
3. = Study very feasible, considering available resources

Applicability

1. = No chance of recommendations being implemented
2. = Some chance of recommendations being implemented
3. = Good chance of recommendations being implemented

Ethical acceptability

1. = Major ethical problems
2. = Minor ethical problems
3. = No ethical problems

In order to assist a group in selecting and rating different research topics, we will use the nominal group technique (NGT)

III. NOMINAL GROUP TECHNIQUE

The nominal group technique (NGT) is a group discussion technique that is useful when one wants to obtain a **consensus** from a group on a topic where decision-making can be usefully guided by the perceptions and opinions of the various group members. The sequence of the group discussion is usually as follows: individual expression, followed by 'voting', followed by discussion, and another round of 'voting' followed by discussion etc. The group discussion comes to an end when the results of the last vote are not appreciably different from the last-but-one vote.

Steps in applying the nominal group technique*

Participants (between 6-10, all familiar with the content area being explored) are assembled in a quiet room. They are seated in a U-shaped setting so that all participants can see the display (board, flipchart or overhead). The moderator is a non-participant (in our case a course facilitator) who explains and then guides the participants through the process. The steps of the NGT process are summarised below:

1. **Individual listing of ideas on paper.**

   In HSR courses, participants of a team each write one, at most two priority problems for research on a piece of paper. This is done in complete silence to prevent the group from becoming judgmental about the ideas too soon. The sheets are collected.

2. **Display of lists produced, followed by discussion.**

   The facilitator takes each sheet of paper and displays all sheets on the board so that all team members can see them. The leader requests the members to briefly explain their ideas and why they suggested them.

   No comments are made by the group at this time, but as the ideas for research topics are presented the rest of the group should study them and see whether they understand what the ideas are and why they are important. If clarification is needed, this is done after all ideas have been presented. Then the participants may attempt to combine overlapping ideas (for example, two almost similar research topics could be combined into one).

3. **Voting and ranking.**

   After the ideas have been clarified the facilitator asks the participants to select a certain number of ideas on the display (for example, five) that they consider most important, write them on a sheet of paper and rate them. The rating system used can vary but should be fixed in advance: for instance, 5 for the most important idea, 4 for the next most important, etc. The sheets of paper are then collected. For ranking the proposed research topics we would use the criteria and scales presented in the previous section.

4. **Summarising the results.**

   The facilitator writes each individual rating on the display, next to the idea. All scores are added, resulting in a total score for each idea. The ideas are then ranked, according to the score they received.

* Some of the ideas in this process are adapted from Williamson JW et al. (1981*) Health Accounting for Quality Assurance*. American Occupational Therapy Association.

5. **Discussion of the results.**

   The results of the first vote are discussed in plenary. All members are urged to contribute. The facilitator may wish to select two types of ideas for clarification: those with high votes and those with divergent votes (i.e. high as well as low weights). A few new ideas may be developed in this discussion. Also, it may be possible to identify a few 'sleeper' ideas among those given low votes. Sometimes such ideas may get a high vote when group members understand why the idea was brought forward.

6. **Second vote and re-ranking**.

   Participants are now asked to vote a second time and the whole process of ranking and discussion is repeated. Voting stops when the results from two consecutive votes do not yield a marked difference. The ranking of the revised final scores gives the order of importance of the ideas as perceived by the group.

Advantages of the NGT

- The discussion process is strictly separated from the voting process and voting may be done anonymously. This depersonalises the process and gives each member an equal vote, regardless of his verbal capacities.

- The results thus reflect input from all members of the group. The series of discussions and anonymous votes helps to minimise the chance that the results will be skewed toward the opinions of one or more dominant personalities.

- The voting process provides a useful means of aggregating individual judgements.

The NGT (or a modified version of the NGT) is particularly useful during the research process in health systems research to:

- assist a group of managers/researchers/community representatives in generating and prioritising lists of topics for which research information may be needed

- assist a research team in selecting a research topic from among alternatives proposed.

REFERENCES

Fisher A, Laing J and Stoeckel J (1983) *Handbook for Family Planning Operations Design.* NewYork: The Population Council.

For more reading about the nominal group technique the following literature is recommended:

Abramson JH (1990) *Survey Methods in Community Medicine. Epidemiological Studies, Programme Evaluation, Clinical Trials.* London: Churchill Livingstone/ Longman Group Ltd (4th ed.): 188-190.

Delbecq AL, Van de Ven AH, Gustafson DH (1975) *Techniques for program planning: A guide to nominal group and Delphi processes.* Glenview, Illinois: Scott, Foreman.

Van de Ven AH, Delbecq AL (1972) The Nominal Group Technique. *Am Jrnl of Public Health* ***62:***337.

EXERCISE: The Chobe district health team (DHT), selecting a research project

(To be carried out in plenary, ½ hour, if this is the first discussion of possible research topics)

Introduction to the exercise

The Chobe DHT, responsible for the health of a population of 525,000, including 313,000 in Chobe town, has to choose between two important study topics:

Possibility 1

The first possibility is a study that aims to contribute to the development of alternative health financing for low-income households. The study will **compare the effectiveness, feasibility and acceptability of two models of community health insurance in a rural community**.

In Chobe district, almost 90% of families depend on the informal sector for their livelihood. Day-to-day expenditure is met with great difficulty, and illness affects the household budget in two ways. Not only does the household have to spend on medical treatment, but illness also might reduce the ability of the household to earn during the period of illness. In Chobe, distance and travel time limit the access to public sector clinics and hospitals. Furthermore, public sector outpatient and hospital services suffer from poor quality and resource constraints. Hence most people use private sector health services. There is one private hospital in Chobe town. Nearly three fourth (70%) of the health care expenditure is met privately through out of pocket expenses by the patients. The poor avoid use of health services unless there is very serious illness. When such illness does occur, it frequently has a catastrophic effect on family finance. The concepts of insurance & risk sharing are still alien to the rural population. Health insurance, either social or private, has remained out of the reach of the low-income households due to unaffordable premiums. Recently, community-based initiatives have been designed aiming at providing alternate more equitable financing, particularly for hospitalization for the poor. However, there is limited experience in implementing such schemes, and several questions have arisen about the feasibility, acceptability, sustainability and effectiveness of such schemes.

It is proposed to test two models of community health financing in Chobe district, using the foundation of women's self help groups (SHG). During the past five years, the Department of Community Medicine has provided technical assistance for women's empowerment in Chobe by facilitating SHG. Village women are motivated to form small groups of 15-20 women and contribute a monthly fixed sum to a common fund. This fund is managed entirely by the group itself and women undertake income-generating activities with the help of loans from these funds. Most of the SHG now have reserve savings. It is proposed to introduce health insurance by motivating members to make a monthly contribution towards a common reserve fund for health insurance. SHG that have been in existence for more than three years would be invited to select one of two models of health insurance.

The first model would use the reserve fund of the SHG to subsidize cost of hospitalization in the Chobe Private Hospital. Each SHG will have to purchase from the Hospital three Health Insurance cards annually, at a predetermined cost. On producing the Health Insurance card at the time of hospitalization, a patient will have to pay only 25% of the total hospital bill. The group will reimburse this 25% from the common reserve fund of the group. The group itself will decide the reimbursement limit.
The second model would purchase health insurance policies for its members from a state sponsored insurance scheme. The State Scheme provides insurance policies with an annual individual premium and a family premium. The Self Help group would pay the premiums from the reserve fund and recover the cost from members in easy instalments. The insurance policy would cover reimbursement of hospitalization expenses in any public or private hospital or clinic for illness/disease or injury sustained up to a predetermined sum per person per year.

The Community Medicine Department would undertake a study to assess:

- Which model is more acceptable to the community in terms of cost to the family and services received;
- The numbers of beneficiaries and the types of illnesses covered in each model;
- Whether the self help groups would be able to sustain the scheme in future.

The study would take 18 months. The SHGs will be given the necessary support to develop guidelines to safeguard against the risks such as moral hazards, adverse selection, over-usage and fraud, which can derail the scheme. Concurrent with the study, a mass education programme of the community will be undertaken to render education and disseminate information regarding the benefits of health insurance. Base-line and end-line surveys together with focus group discussions will be conducted to obtain data on the research questions.

EXERCISE 1 (continued)

Possibility 2

The second possibility is to examine **the reasons for the increase of the number of reported suicides** in the district. This increase is alarming, not only in Chobe, but in the whole country. The victims are predominantly (85%) adolescents and young adults between 14-24 years and the number of reported suicides tripled between 1997 and 1999. 128 cases were reported in these two years nationally, of which 35% were in Chobe town (exploratory study carried out in 1999).

The number is highest in urban areas. The Minister of Health and Social Welfare expressed her concern and advocated for studies to get more insight in the reasons and possible remedies. From literature and first hand impressions, it appears that poverty and disintegration of families are major underlying factors to suicide. There is a high rate of migration of men to the mines in the neighbouring country, and women have to struggle to raise their children alone, often with minimal economic support from their husbands. There is also a significant rural-urban migration within the country (Census 1996). This migration erodes the extended family ties and also the cohesion within neighbourhoods. Yet in towns, social networks are still poorer, and for newly arriving youngsters often inexistent. Social services do not cope with the needs of the rapidly expanding urban population. Under these circumstances, young people may become easy victims of alcohol or drug use and unsafe sexual relationship, in an effort to please their peers or to stay alive. Adolescents may be suffering from all kinds of stress. In an effort to belong to and be accepted by their peers, there is extensive pressure on them to perform well in school, to obtain well-paying jobs, to take responsibility for their parents and younger siblings, and, for girls, to find solid partners. The reality of what they can expect to achieve is often very different, for boys as well as girls, which causes stress. For girls there is an additional risk of becoming pregnant against their will as they have little power to refuse men with more means and a higher status who approach them. Many feel forced to search for a (usually unsafe) abortion or risk the anger of their parents because of being expelled from schools. Then there is also the risk of STDs and HIV infection, of which youngsters are only vaguely aware, so that the shock – particularly for young girls – may be high when they become victims of such diseases.

Proposed Study: The district health team proposes to carry out a comparative study focusing on adolescents who attempted suicide and adolescents who did not (same age, sex, residence and workplace/school). Also close relatives of adolescents who committed or attempted suicide could be interviewed, as well as community members, teachers, youth leaders, and church leaders. The main objective of the study would be to get insight in the reasons for suicide and possibilities to prevent it and support adolescents in need.

Directions

Rate the two proposals in small groups, using the form on the following page, and prepare to defend your first choice in plenary. (When rating the topics on the criteria, you can either refer to the 'Scales for rating research topics' presented right before this exercise or use the summary scales at the bottom of the rating sheet).

EXERCISE (continued)

Proposed topic	1. Relevance	2. Avoidance of duplication	3. Urgency	4. Political acceptability	5. Feasibility	6. Applicability	7. Ethical acceptability	Total no. of points
1. Community health insurance								
2. Adolescent suicide								

Rating scale: 1 = low, 2 = medium, 3 = high

GROUP WORK (Approximately 2¼ hours if this is the first discussion)

Meet in your working groups to list and rank the research topics that you want to consider for the research proposal you will develop, as a team, during the course.

1. Choose a reporter who will present in plenary the topics you have considered and your final choice.

2a. **If this is the first discussion of possible research topics**, it is suggested that each group member write one or two topics on a piece of paper. Then all the topics can be listed on a flipchart and briefly discussed to eliminate duplications. Omit proposals that are obviously less relevant or too difficult to carry out. Ideally you should select no more than five to six topics for individual rating.

2b. **If a pre-selection has already been done in the field**, and/or different possibilities for research topics have emerged during problem analysis in Module 4, consider the two or three topics you have to choose from.

3. Each group member should then rate the selected proposals **individually**, using the scoring sheet on the following page. Then for each proposal the scores of the groups members for each criterion should be tallied on a flipchart and the total scores calculated. Discuss marked differences in individual ratings as these may be due to different interpretations of the criteria.

4. Then thoroughly review the (two) proposals that received the highest scores. At this point it is important to take into account, which **proposed study your group could most realistically carry out** within the coming 4-6 months. Ideally, all group members should be able to participate actively and benefit directly from the results.

5. Finally, select the topic for your upcoming research project and prepare a brief presentation for the other members of your course. Present the flipchart with the scores and provide reasons for your final choice.

6. Carefully document the arguments supporting your first choice and keep them for use in later sessions.

GROUP WORK (continued)

Criteria for selection of research topic								
Proposed topic	1. Relevance	2. Avoidance of duplication	3. Urgency	4. Political acceptability	5. Feasibility	6. Applicability	7. Ethical acceptability	Total no. of points
1.								
2.								
3.								
4.								
5.								

Rating scale: 1 = low, 2 = medium, 3 = high

Note:

If participants have been asked to make a preliminary selection of their research topics in the field, **Module 3** should be sent to participants and relevant managers **at least 6-8 weeks before the course starts**. Preferably a facilitator/trainer should be present as well to provide technical support during the selection process. Otherwise a set of guidelines could be prepared to assist the participants and their managers in the selection process. It is best to ask each group to come with at least two potential research problems, in case one of the topics, on further analysis, proves infeasible.

- If the preliminary selection of a research topic has been made before the workshop, you can proceed to. Part I of **Module 4** (Problem analysis). However, Part I of Module 3 should always be presented to all participants, as it is a basic introduction to Module 4.

- When problem analysis is completed, the second section in Module 4 asks the participant groups to reconsider their research problems. If the groups find their research problems involve the investigation of several sub-problems that cannot be combined in one study they may use section II of Module 3 to rank the sub-problems before making their final selections.

Trainer's Notes

Module 3: IDENTIFYING AND PRIORITISING TOPICS FOR RESEARCH

Timing and teaching methods

40 minutes	Introduction and discussion
50 minutes	Exercise: Chobe district including explanation of the nominal group technique (NTG) method
2 hours	Group work
1 hour	Group reporting (15 minutes per group)
4½ hours	TOTAL TIME

Materials

- Flip chart and markers
- Sticky stuff or clear tape, and
- Photocopies of the rating sheet for group work, if possible.

Introduction and discussion

Discuss the process of problem identification (Part I of the module), criteria for prioritising topics for research (Part II) and the Nominal Group Technique (Part III). Be sure you are thoroughly familiar with the concepts but let the criteria as well as the definitions come, as much as possible, from the group. (Before describing the criteria listed in the module, ask the participants to brainstorm, suggesting what criteria they think are most important to consider when selecting a research topic).

Exercise: The Chobe district health team (DHT) - selecting a research project

- Divide the participants in groups of 3-4 people so that they can do the exercise in plenary with minimal displacement.

- Ask the participants to carefully read both examples. Briefly explain how to use the rating sheet at the end of the exercise. Ask the groups to rate both examples, but let certain groups start with the first and others with the second. Give them 15-20 minutes at most to complete the rating process.

- Prepare a flip chart with the list of criteria and write down the ratings of all groups for both topics. Identify the criteria on which the rating differ most (for example, if one group rates a '3' for feasibility of the suicide study and another a '1', ask each group to explain why it decided on its score). The differences may be due to a different understanding of the criteria or to a different perception of the problems and the proposed methodology. Special attention should be given to uniformity in the interpretation of the criteria. After completing this exercise, participants should be able to see the importance of looking at all dimensions of a problem before moving ahead to select their own topics.

Note:

There are no right or wrong answers to the exercise. Either proposal may receive priority for different reasons.

Group work

- Review the logistics of supervising the group work with the other facilitators before the session starts. Choose meeting places for the four groups and be sure flip charts, sticky stuff or tape, and markers are available.

- When the selection process for choosing group topics is introduced, make sure that the participants realise they are involved in more than a 'hypothetical exercise'. Participants should be made aware that they will be developing the topics they select into research proposals during the course and that they will carry out these projects on their return home.

- Familiarise yourself thoroughly with the selection procedures as presented in the group work. These procedures are a simplified version of the **nominal group discussion technique**. Often one or two priority topics emerge after one round of discussions followed by individual rating and summarising the individual scores on a flip chart, so that a decision can be made during the second round.

- **If the total scores of two or three topics are very close, they may be discussed again. It may be useful, in particular, to re-examine criteria that were scored differently by the group members.** Special attention should be paid in this round to the question of whether the results of the study will be applicable and whether group members feel they can realistically carry out the research within the 4-6 months allocated.

- As facilitator, you may chair this first group work session but you should not dominate the discussion. You should make sure that the procedures run smoothly and that research topics which duplicate research already completed or which are not feasible are dropped before the rating starts. You should help insure, as well, that no important proposals or initiatives are dropped because the group is not yet familiar with handling the criteria.

- At the end of the group work for selecting projects, assist the reporter in editing and writing the list of topics debated by the group on a flip chart, along with the record of the combined group rating. Ask the secretarial staff to type out lists of topics considered and voting results for possible inclusion as an annex of the final course report.

Note:

During this first group work session it is not important which facilitator works with which group. Once the topics have been selected, final assignments of facilitators to specific groups may be made after considering the facilitators' familiarity with the topics chosen.

Module 4

ANALYSIS AND STATEMENT OF THE PROBLEM

Flowchart: Steps in the development of a health systems research proposal

Questions you must ask	Steps you will take	Important elements of each step
What is the problem and why should it be studied?	Selection, analysis and statement of the research problem	- problem identification - prioritising problems - analysis - justification
What information is available?	Literature review	- literature and other available information
Why do we want to carry out the research? What do we hope to achieve?	Formulation of research objectives	- general and specific objectives - hypotheses
What additional data do we need to meet our research objectives? How are we going to collect this information?	Research methodology	- variables - types of study - data collection techniques - sampling - plan for data collection - plan for data processing and analysis - ethical considerations - pre-test or pilot study
Who will do what, and when?	Work plan	- human resources - timetable
What resources do we need to carry out the study? What resources do we have?	Budget	- material support and equipment - money
How will the project be administered? How will utilisation of results be ensured?	Plan for project administration and utilisation of results	- administration - monitoring - identification of potential users
How will we present our proposal to relevant authorities, community and the funding agencies?	Proposal summary	- briefing sessions and lobbying

NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 4: ANALYSIS AND STATEMENT OF THE PROBLEM

OBJECTIVES

At the end of this session you should be able to:

1. **Analyse** a selected problem and the factors influencing it.
2. **Prepare** the statement of the problem for the research proposal you will be developing during the course.

I. Analysing the problem

II. Deciding on the scope and focus of the research

III. Formulating the problem statement

I. ANALYSING THE PROBLEM

In HSR the researcher is often required to do research on a problem with which (s)he is not very familiar. Health workers and managers or community members may be much more familiar with it. But even they may never have given critical attention to the various aspects of the problem.

A systematic **analysis of the problem**, completed jointly by the researchers, health workers, managers and community representatives, is a very crucial step in designing the research because it:

1. enables those concerned to pool their knowledge of the problem,
2. clarifies the problem and the possible factors that may be contributing to it, and
3. facilitates decisions concerning the focus and scope of the research.

Note:

In a workshop setting, it may be impossible to obtain input from **all** concerned. The opinion of people who cannot be consulted (e.g. local health staff or community members) should be solicited prior to and immediately after the workshop, before finalising the research proposal.

Steps in analysing the problem

Step 1: Clarify the viewpoints of managers, health care workers and researchers in relation to the problem.

Areas of concern within the health system are often expressed in broad or vague terms by managers and health care workers. For example,

'Complications of unsafe abortions need more attention'
'Teenage abortion is a problem'

During initial discussions with managers and health care workers who are involved in the problem area, clarify the issues by **listing all the problems** in the area of concern, as they **perceive** them.

Remember that a problem exists when there is a discrepancy between 'what is' and 'what should be' (see **Module 3**). Therefore, the perceived problems should be worded in such a way as to illustrate this discrepancy.

For example, health care managers and workers may determine that the general concern about the complications from unsafe abortions among teenagers includes the following problems:

- Increasing numbers of unsafe induced teenage abortions
- Poor health services management of complications of induced abortion
- Social stigma associated with premarital pregnancy
- Illegality of abortion
- Negative attitudes of health staff towards induced abortion
- Secrecy surrounding abortion

Step 2: Further specify and describe the core problem.

You should then try to identify the **core** problem and quantify it. Looking at the example discussed in Step 1, you may decide that the core problem is:

- *Increase in complications from unsafe abortions among teenagers*

You should now attempt to describe more elaborately:

- The **nature** of the problem; the discrepancy between 'what is' and what you prefer the situation to be, in terms of unsafe abortions and/or complications.
- The **distribution** of the problem - who is affected, when, and where and
- The **size** and **intensity** of the problem - is it widespread, how severe is it, what are its consequences (such as disability, death, waste of resources)?

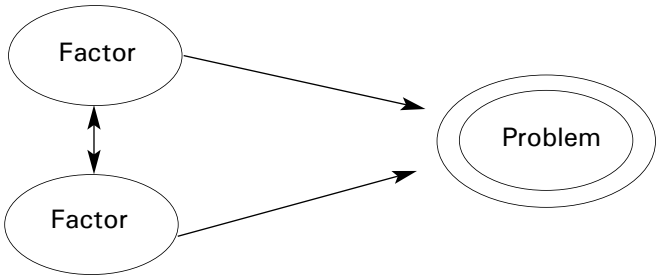
Step 3: Analyse the problem.

After identifying the core problem you should:

- Identify factors that may have contributed to the problem.
- Clarify the **relationship** between the problem and contributing factors.

It is helpful to visualise these interrelationships in the form of a **DIAGRAM.** The basic principles of constructing such a diagram are illustrated below:

Figure 4.1: Elements of a problem analysis diagram



The relationships between contributing factors and the problem can be indicated by arrows, either one-way (for cause-effect relationships) or two-way arrows (for mutual relationships). The core problem can be identified by drawing a double line around it.

Analysis of the problem involves several **sub-steps**.

Step 3.1: Write down the core problem(s) as defined in Step 2 in the centre of a blackboard or flipchart.

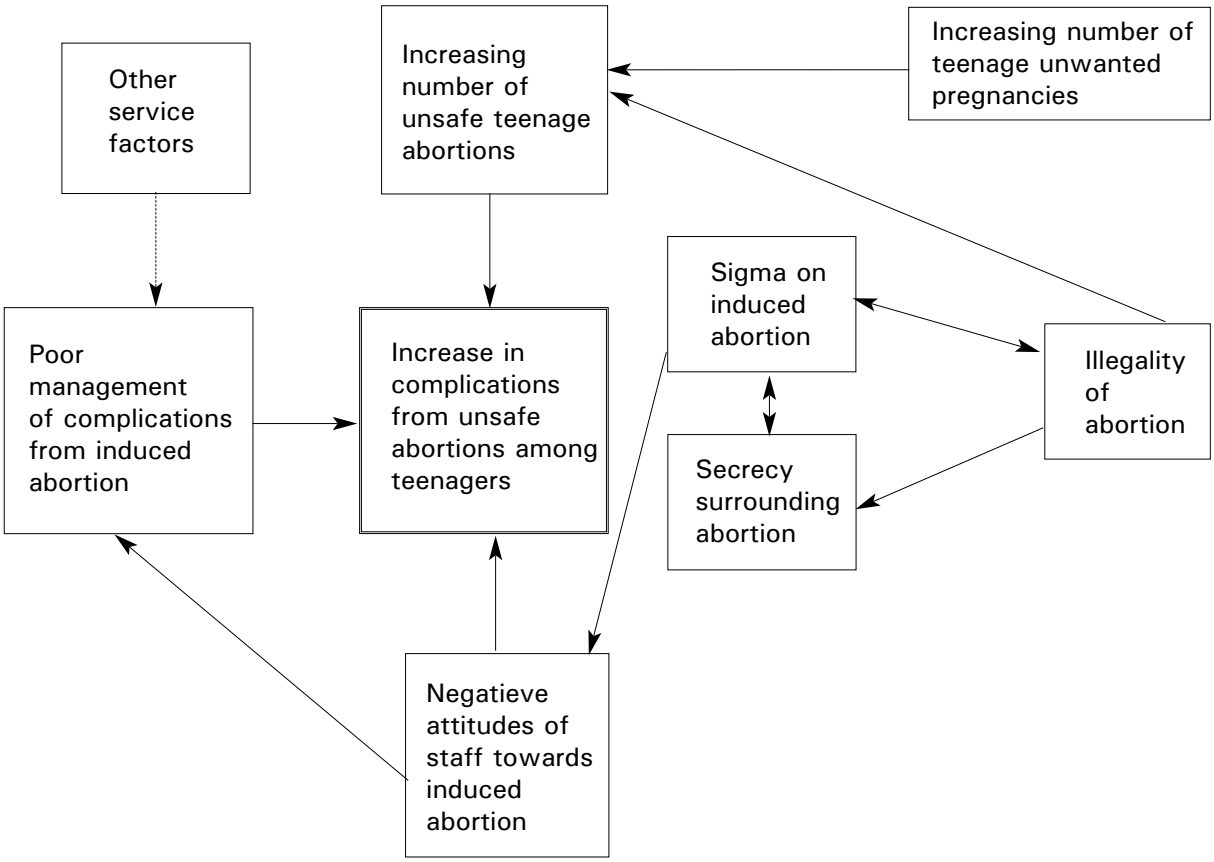
Step 3.2: Brainstorm on possible causes or factors contributing to the problem.

It is important that the viewpoints of managers, health care workers and/or researchers brought up during Step 1 are all included. Discuss the relationships between the different factors and the problem.

If desired, participants may use separate cards or pieces of paper on which to write possible contributing factors. The cards may be pinned or taped around the core problem on the board or flip chart and moved, revised, or eliminated as necessary, during the development of the diagram.

The initial diagram of the problem of complications from unsafe abortions among teenagers might look like this:

Figure 4.2: Initial problem diagram – Complications from unsafe abortion among teenagers



Note:

that many of the perceived problems mentioned in step 1 are related to each other in a **cause-effect relationship** (e.g. poor management of complications from abortion and high complication rate from abortions), or in a **mutual relationship** (stigma on induced abortion and secrecy surrounding induced abortion).

As you can see, this initial diagram suggests that further development of the analysis could proceed in at least two directions, i.e., analysis of factors related to:

- **Family and community**: increasing number of unsafe teenage abortions; secrecy sur rounding abortion.

- **Quality and accessibility of the services** provided (poor management of complications due to abortions negative attitudes of staff towards induced abortion).

These sets of community and service factors will appear in many HSR studies. In reality they usually prove to be **closely intertwined.** Poor management of abortion is, for example, influenced by the negative attitudes of staff towards abortion that, in turn, is influenced by community attitudes and perceptions.

Step 3.3: Identify further contributing factors.

Extend the problem analysis diagram further by identifying additional factors that could have contributed to or aggravated the problem. It may be possible to identify several 'generations' of predisposing factors, by asking 'but why'.*

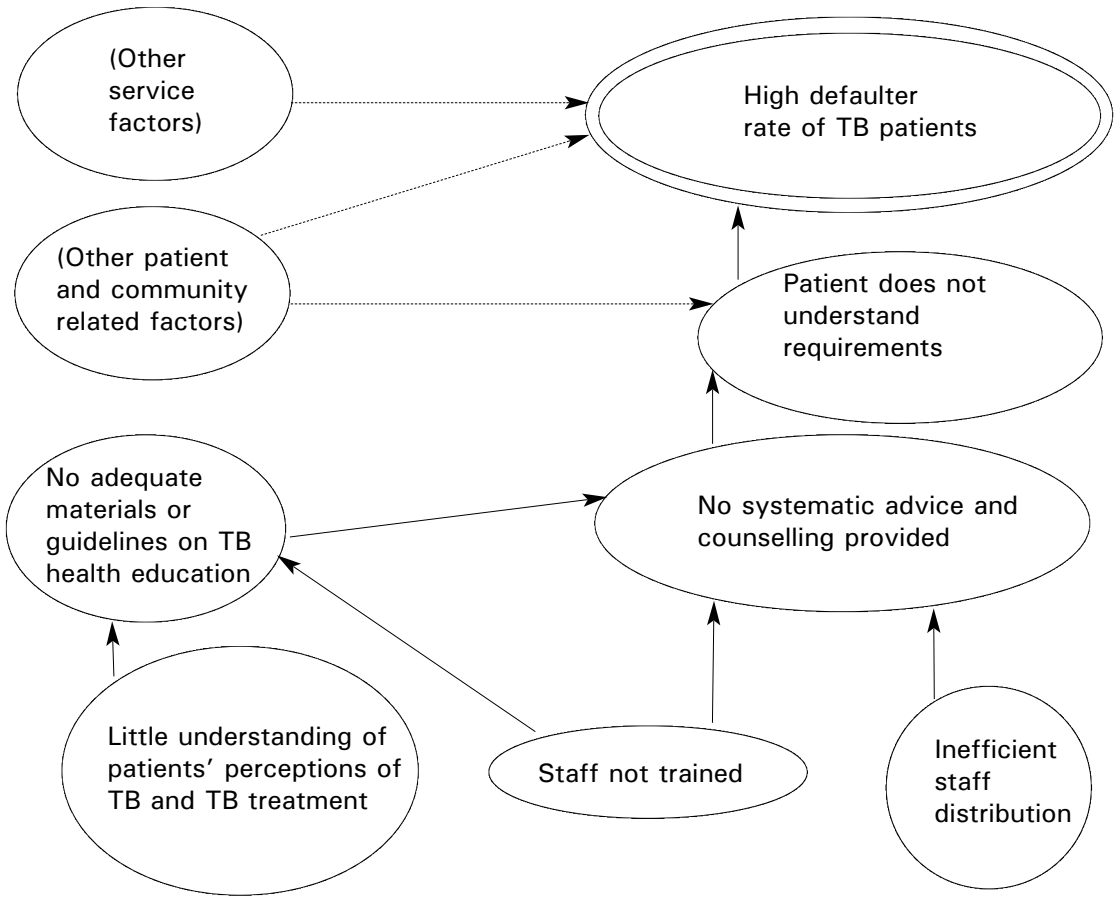
Let us take another example: High defaulter rate among tuberculosis (TB) patients (**Figure 4.3**).

One of the factors contributing to low patient compliance may be (and often is, see Joint HSR Project, 1996) low understanding by patients of the treatment requirements. Further asking 'but why' may lead to the assumption that patients are not systematically counselled on the 'what' and 'why' of treatment, which in turn may be influenced by a number of important weaknesses in the training of staff which need to be addressed.

It is desirable to continue identifying underlying contributing factors until you reach basic factors that need to be modified to solve the problem, and that **can be modified** within the existing context. This will facilitate the formulation of research projects that can **provide useful information for decision-making.** The process of continued analysis by asking 'but why' will necessitate several revisions or extensions of the initial analysis diagram. The final version should encompass all the critical factors that may be contributing to the problem which will be studied.

* See Barnett and Abbatt (1994) for further illustrations of the 'but why' method.

Figure 4.3: Identifying several 'generations' of predisposing factors causing high defaulter rate among TB patients



Step 3.4 Attempt to organise related factors together into larger categories, and develop your final draft of the diagram.

This final step in organising the diagram will help you not to overlook important factors and will make it easier to develop the data collection tools in a systematic way.

For example, the revised diagram focusing on the 'high defaulter rate' among tuberculosis patients may group contributing factors into three main categories:

- socio-cultural and economic factors
- service-related factors
- disease-related factors

For our TB example, we may categorise the contributing factors to defaulting into these three main groups (see **Figure 4.4**):

Socio-cultural factors, such as

- Personal factors such as a patient's age, sex, education, occupation, composition (and possible support) of the family

- Community determined factors such as:

- Poor or conflicting community knowledge of signs and causes of TB and of requirements for TB treatment
- Availability of other types of treatment in the community
- Preference for other types of treatment
- Poor understanding/support from employer

Service factors, such as:

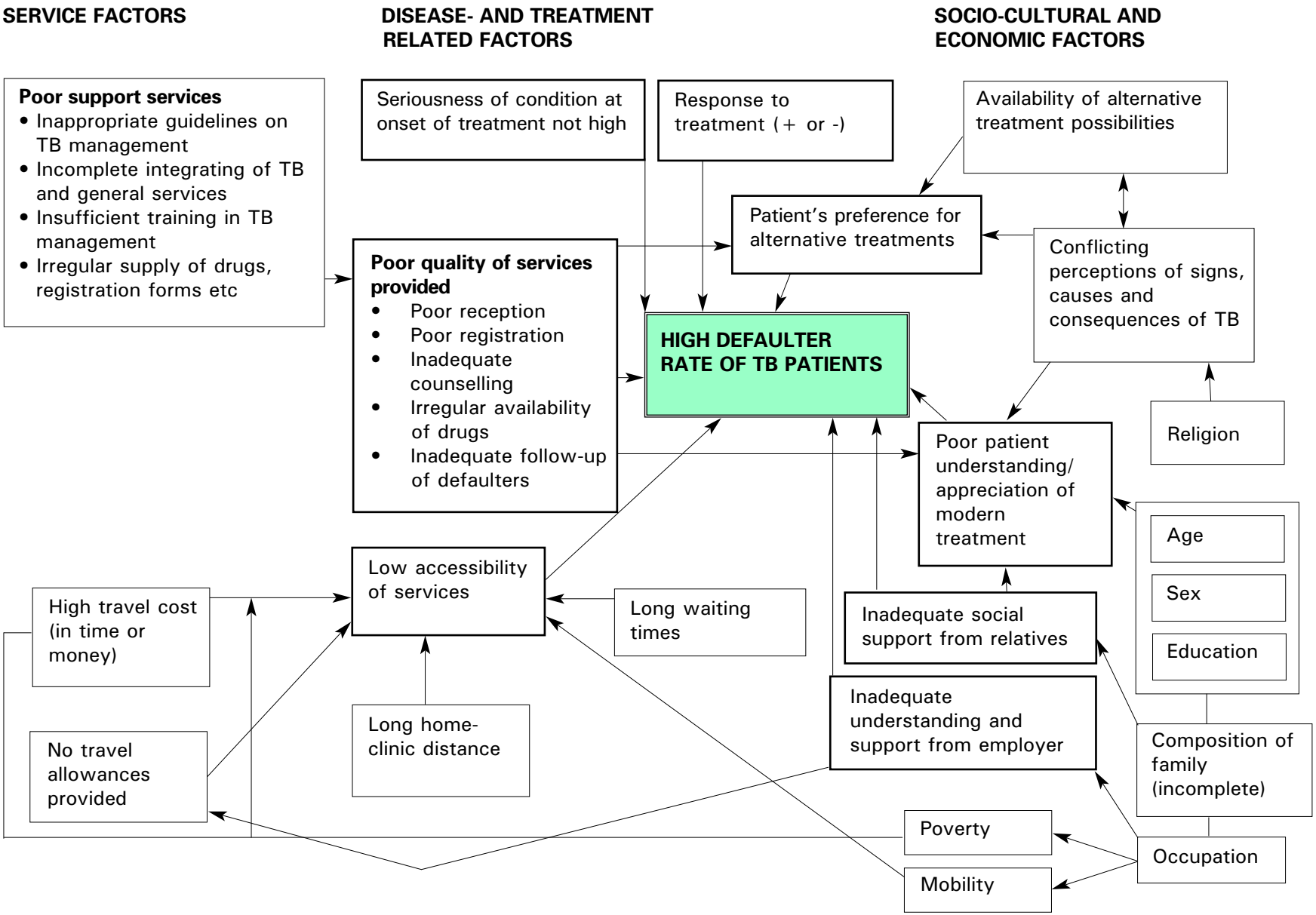
- Low availability and accessibility of services (including cost of treatment)
- Poor clinic management (poor reception of patients, inadequate counselling, no drugs)
- Poor support services (poor training, supervision, drug supply)

Disease-related factors, such as:

- Seriousness of the patient's condition at onset of treatment
- Physical response to the treatment (complications? or quick recovery?)

Note that diagrams are easier to read and work with if the factors that directly contribute to a problem are highlighted (see **Figure 4.4**).

Figure 4.4: Revised problem analysis diagram of factors contributing to high defaulter rate among TB patients

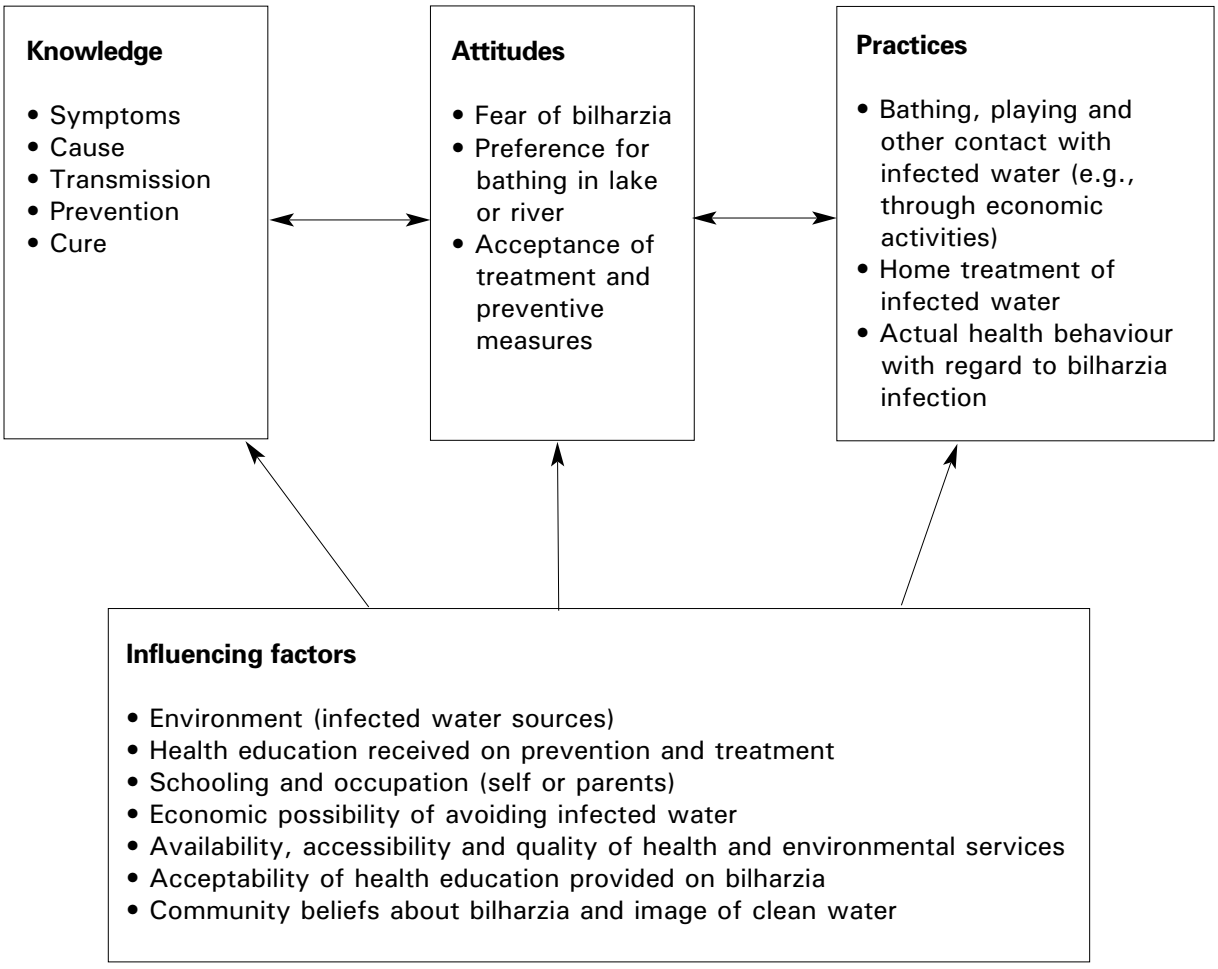


Note:

If the research will focus on a **description** of a situation or health problem (size, distribution) or an **evaluation** (see **Module 3**, p. 1), it may not be appropriate to make an analytical diagram looking for *causes* of a problem. Here the problem is lack of information.

For example, we may need information on knowledge, attitudes and practices (KAP) of teenagers with respect to bilharzia in order to develop adequate health-education materials for schools. This is a descriptive study, for which we can make a different diagram, listing knowledge, attitudes and practices that we want to explore in the study. We can, however, go one step further and also list the factors that may have contributed to the generation of the teenagers' KAP.

Figure 4.5: Knowledge, attitudes and practices with regard to bilharzia and its treatment and prevention (Example of a descriptive diagram)



If it is our aim to **evaluate** interventions or programmes, we evaluate in general against the objectives: to what extent have we reached them? Is the intervention the most cost-effective? Is it still appropriate (covering a priority need?) etc. (see **Module 3**).

Evaluation can also be applied to the health system as a whole. The WHO, for example, proposes that the health system should cover the whole population with fair (financially equitable) and the best attainable level of services, delivered in a consumer-friendly way. In that case we take diagram 2.4 in **Module 2** as our point of departure.

II. DECIDING ON THE FOCUS AND SCOPE OF THE RESEARCH

After detailed analysis of a problem, it is important to reconsider the focus and scope of the research. Several issues are particularly important to consider, including:

1. **Usefulness of the information**. Would the information that will be collected on this problem help improve health and health care? Who would use the findings related to the factors in the diagram that would be studied? How would the findings be used?

2. **Feasibility**. Is it feasible to analyse all the factors related to the problem in the 4-6 months available for research?

3. **Duplication**. Is some of the information related to factors in the diagram already available? What aspects of the problem need further research?

Reconsider your problem diagram with these issues in mind. If your problem is complex and has many possible contributing factors, identify and demarcate the boundaries of possible smaller research topics. If there is more than one possible topic, use the selection criteria and ranking method that were described in **Module 3** to assist you in your final decision concerning the focus and scope of your research.

Note of caution:

The dissection of the diagram into different parts and selection of one part for research is not advised if insufficient insight exists into the nature, relative weight and interrelationships of the various factors contributing to the problem. You would risk concentrating on marginal factors and coming up with marginal solutions. It is, for example, inadvisable to concentrate only on community factors or only on service factors to explain under-utilisation of services if you don't know how these factors are interrelated and where the main problem is.
An **exploratory study** would then be indicated, limited rather in the number of informants than the number of factors included in the study (see **Modules 9**, **10** and **11**).

III. FORMULATING THE PROBLEM STATEMENT

The first major section in a research proposal is the 'statement of the problem'.

Why is it important to state and define the problem well?

Because you will find that a clear statement of the problem:

- Is the foundation for the further development of the research proposal (research objectives, methodology, work plan, budget, etc.).
- Makes it easier to find information and reports of similar studies from which your own study design can benefit.
- Enables you to systematically point out why the proposed research on the problem should be undertaken and what you hope to achieve with the study results. This is important to highlight when you present your project to community members, health staff, relevant ministries and donor agencies who need to support your study or give their consent.

What information should be included in the statement of the problem?

1. A brief description of socio-economic and cultural characteristics and an overview of health status and the health-care system in the country/district *in as far as these are relevant to the problem*. Include a few illustrative statistics, if available, to help describe the **context** in which the problem occurs.

2. A concise description of the **nature of the problem** (the discrepancy between what is and what should be) and of the size, distribution and severity of the problem (who is affected, where, since when, and what are the consequences for those affected and for the services). For a descriptive or evaluation study you will elaborate the different **components** of the problem.

3. An analysis of the **major factors that may influence the problem** and a discussion of why certain factors need more investigation if the problem is to be fully understood.

4. A brief description of any solutions to the problem that have been tried in the past, how well they have worked, and why further research is needed (**justification** for your study).

5. A description of the type of information expected to result from the project and how this information will be used to help solve the problem.

6. If necessary, a short list of **definitions** of crucial concepts used in the statement of the problem.

A list of **abbreviations** may be annexed to the proposal, but each abbreviation also has to be written out in full when introduced in the text for the first time.

GROUP WORK

1. Select a reporter who will present the statement of the problem in plenary.

2. Discuss comments you received in the previous plenary session on the choice of your topic and revise your topic, if necessary.

3. Make an analysis diagram of the most important components of the problem or the most important factors that you think are influencing it. Use a blackboard or a flip chart and, if possible, separate cards for each factor. (See part 1 of this module for details on the steps in this process.) After making your initial diagram, try to rearrange the factors identified into broader categories.

4. Decide whether it is feasible to explore all the factors in your problem analysis diagram. In case of doubt, consider two possibilities:

   a. All factors seem important and interrelated; you could not easily split the diagram up in possible sub-studies. Just continue. We will come back to possible ways of increasing the feasibility of the study when discussing **Module 9** (Study type), **10** (Data-collection techniques) and **Module 11** (Sampling).

   b. The diagram is so complex that several studies would be necessary to cover it. If so, demarcate the boundaries of possible projects and use the criteria and ranking system in **Module 3** to select one of the sub-problems as the focus for your project.
   Note: Groups that selected their general problem areas before coming to the course may want to spend an hour or two at this point to systematically rank and choose between possible topics within their general problem areas, using the instructions in **Module 3**, part II.

5. Prepare a **first draft** of 2-3 pages of the statement of the problem for the topic you have selected in your group.

   - First prepare an outline covering items 2 through 5 in the list presented just before this group work session.

   - Then prepare one or two paragraphs of 'background information' (item 1) which places the problem in its context and will be used as the introduction to the statement of the problem.

   - Finally, define crucial terms and explain abbreviations, if necessary.

6. Identify information you need now, from the literature or from key informants, to help you focus your study and to further develop your statement of the problem. Ask course facilitators to give you assistance, if necessary. **Much, if not all, of your literature review (module 5) may be included in your statement of the problem.**

7. Present your flipcharts with the problem analysis diagram and outline of your problem statement in plenary. Justify the need for your study (¼ hour per group). See the guidelines for the discussion of the presentations on the following page.

8. Keep all materials presented, as well as your notes on the comments received in the plenary session, for use during further development of your proposal. Have the first draft of your statement of the problem computerised.

Guidelines for the plenary session following group work: presentation and discussion of problem statements

Each group should present its problem analysis diagram and problem statement in the plenary session. Facilitators and members of other groups should comment and provide suggestions for improvement.

Guidelines for presentation

- Each presentation should take about 10-15 minutes.
- Present the problem analysis using the diagram. Indicate, if necessary, the boundaries of the possible studies focusing on various aspects of the problem.
- Present the problem statement. It should be available to the audience either on flip chart or overhead projector or in written version.

Guidelines for discussion

As each team presents its work, consider the following issues for discussion:

- Is there too little, sufficient, or too much background information in the statement of the problem?
- Is the problem clearly described? (the nature, distribution, and magnitude/severity of the problem, and its different components (especially in descriptive studies).
- Are the possible contributory factors and their relationships clearly and logically described in the problem diagram?
- Have the boundaries of the project been clearly defined? Can the project be completed in 4-6 months, or should the focus of the project be narrowed further?
- Will the information collected be sufficiently specific to help solve the problem?
- Does the justification for selecting the research project appear to be rational?

Recommended reading:

Joint Project on Health Systems Research for Southern Africa, Harare, Zimbabwe, Series '*Health Systems Research - It can make a difference*':

Vol. 1: Availability, provision and use of drugs (1994)
Vol. 2: Factors associated with maternal mortality (1994)
Vol. 3: Under-utilisation of TB services in Southern Africa (1996)
Vol. 4: Factors influencing the functioning of Primary Health Care at village level (1996)

All these volumes contain exemplary research protocols with elaborate 'Statements of the Problem' and results of studies carried out in HSR Projects in Southern Africa.

Barnett, Liz and Fred Abbatt (1994) *District action research and education.*
London: The MacMillan Press Ltd.

Trainer's Notes

Module 4: ANALYSIS AND STATEMENT OF PROBLEM

Timing and teaching methods

1 hour	Introduction and discussion
3-5 hours	Group work
1 hour	Plenary
5-7 hours	TOTAL TIME (partly in the evening)

Materials

- Blackboard, flipcharts and tape or sticky stuff
- Several markers for each group
- Optional: Cards or sheets of A4 paper cut in two or three pieces (at least 20 'cards' per group)

Introduction

This module is designed to assist health staff, researchers, and managers to work together to analyse the problem and prepare a problem statement.

It would be beneficial to include health service managers and community leaders who are concerned with the problem and who are likely to use the findings as resource persons during the group work. Resource persons could include, for example, the district medical officer, the hospital manager, the co-ordinator of the Tuberculosis Programme, the mayor or traditional leaders, etc.

Presentation and discussion

It is extremely important that participants understand the principle of making a diagram, with a problem in the centre and contributing factors grouped around the problem. When you give examples, get them to participate actively, identifying what the core problem is and in what direction the arrows should point.

With respect to presenting the diagrams, you might advise the groups to work with different overhead sheets on top of each other.

The final TB diagram, for example, could consist of 4 sheets:
- the problem,
- socio-cultural and economic factors (contributing to the problem),
- service factors, and
- disease-related factors

Give due weight to descriptive studies which have **no** core problem or contributing factors, but simply consist of different, connected, components (e.g. knowledge, attitudes and practice studies).

Group work

This is the first time during the workshop that the participants will work in small groups. Therefore the group may experience the initial difficulties of establishing group-dynamics. Furthermore, many participants may never have had the experience of systematically analysing a problem before.

The topic for the first group work in the workshop is particularly complex. Adequate selection of a problem is crucial for subsequent development of the project. In order to ensure that the group work is productive, the **workshop facilitator could assume the role of chairperson** for this group work session.

In all subsequent modules, the participants should select their own group leader and reporter for each group work session.

Suggestions concerning various steps in the group work process are given below:

1. Listing viewpoints

- Each group member should silently and independently write a list of problems and contributory factors that are perceived to exist in the general problem area. Each factor should be listed on an individual card. (Remember that these statements are **perceptions** based on personal knowledge and experience. It is not necessary to produce supportive evidence at this stage.)

  Reassure participants that evidence can be gathered later during the literature search, screening of reports/interviews with key informants.

- Display **all** problems and contributory factors identified by group members on flip chart paper so that the entire list is visible to all members. (At this stage, **every** perceived problem and factor should be written down **without any discussion**. This will help insure that dominant group members do not override the perceptions of more timid members.)

2. Constructing a problem analysis diagram

- Place what appears to be the core problem in the centre of the flip chart or blackboard. Contributing factors should be written or 'pinned' around the problem by the participants. Relationships between these factors and the problem, and among different factors, should be indicated by arrows.

- If participants cannot yet decide at this point which is the core problem among a cluster of problems, let them select **any** of the problems that seem **major** to them and place them in the centre of the flip chart.

- Take out overlapping viewpoints and make sure that **every** viewpoint listed in Step 1 is represented in the diagram by the time the group finishes. Encourage analytical thinking by asking participants to determine which way the arrows should point between the various factors.

- **Aim to use only one sheet of flip chart paper** for the initial analysis. This forces participants to consider each other's viewpoints. If you use loose cards, each participant can pin his/her own 'contributing factor' on the flip chart around the problem in the centre.

- Now have the participants identify one (or two) core problem(s) to which all others contribute and rephrase those problems, if necessary.

- Then redraw the diagram, reorienting the position of the factors. Place the core problem(s) in the middle and add in the other factors already identified, encouraging the placement of factors in appropriate places.

- When regrouping the contributing factors, try to classify them (e.g., into service factors, socio-economic and cultural factors, factors related to the disease).

- If participants become 'locked to' a particular line of thought or are unable to think of other factors, ask:

  'What else could be causing this problem?'

  'Are there any other **service** or **community** factors that could be causing it?'

  Ask participants who have personal experience with the problem to recall and describe incidents that illustrate important aspects of the problem.

- Do not expect to produce a comprehensive diagram of the problem at the first or even the second attempt. It will be necessary for the participants to **revise the diagram** several times as they acquire a deeper understanding of the problem.

- If the topic chosen requires a **descriptive** or an **evaluation** study, let the group write different components of the topic on which more information is needed. They can also make a diagram.

3. Narrowing the focus of the project

- After the group has prepared its (problem) diagram, be sure that it spends some time seriously considering whether it is possible to give all factors the same weight during one project of 4 to 6 months (considering available staff time and resources). Keep the different possible ways to shorten data collection in mind (e.g., focus group discussions, or key informant interviews) if no full-fledged study on all aspects is possible.

  Make sure that participants have paid enough attention to possible linkages among factors in different categories.

4. Writing the statement of the problem

- Ask the group to list the major points it plans to include in the problem statement and rearrange them, if necessary, before preparing the written text for the problem statement.

- Encourage participants to use available reports to help specify their problem, and to search for information they do not have. Part of the literature review, if not all, may be included in the statement of the problem.

- Preparation of the written problem statement can be done in small groups or individually, but all group members should read each section and one person should be responsible for the final version.

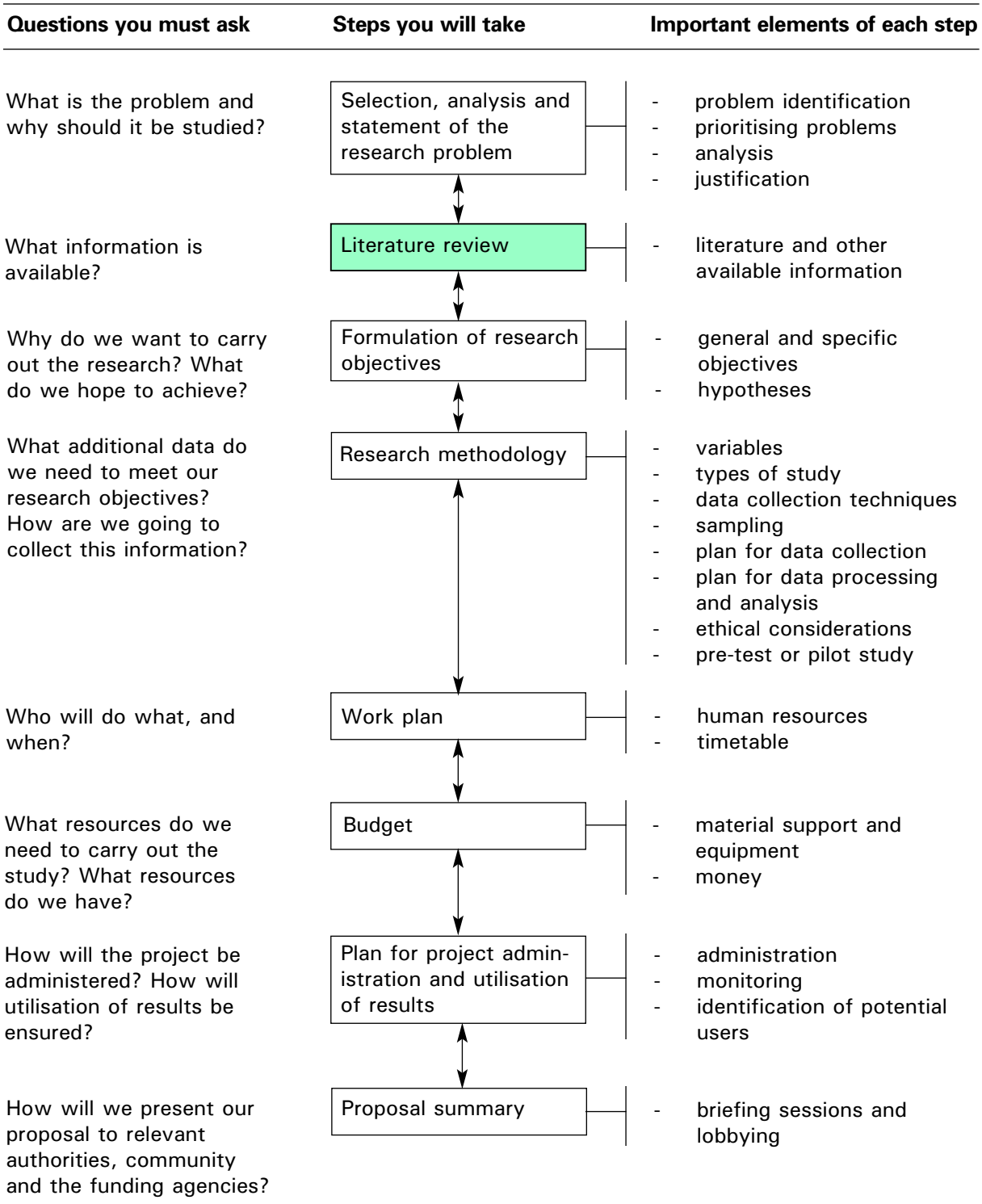
- The description of relevant socio-economic, cultural and health characteristics of the country and/or district in which the study takes place can be best done **at the end**, after a thorough description of the problem, to avoid irrelevant background data. It will be placed **at the beginning** of the problem statement, however.

Make sure that each group keeps a copy of the final version of the problem analysis diagram. It will serve as the basis for subsequently formulating specific objectives and constructing variables.

Module 5

REVIEW OF AVAILABLE LITERATURE AND INFORMATION

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 5: REVIEW OF AVAILABLE LITERATURE AND INFORMATION

OBJECTIVES

At the end of this session you should be able to:

1. **Describe** the reasons for reviewing available literature and other information during the preparation of a research protocol.

2. **Identify** the resources that are available for carrying out such a review.

3. **Prepare** index cards, computer entries or notes that summarise important information obtained from literature or interviews with key informants.

4. **Prepare** a review of literature and other information pertaining to your research topic that will adequately describe the context of your study and strengthen your statement of the problem.

Why is it important to review already available information when preparing a research proposal?

- It prevents you from duplicating work that has been done before.
- It helps you to find out what others have learned and reported on the problem you want to study. This may assist you in refining your statement of the problem.
- It helps you to become more familiar with the various research approaches that might be used in your study.
- It should provide you with convincing arguments for why your particular research project is needed.

What are the possible sources of information?

- Individuals, groups, and organisations;
- Published information (books, articles, indexes, abstract journals); and
- Unpublished information (other research proposals in related fields, reports, records, computer data bases)

Where can we find these different sources?

Different sources of information can be consulted and reviewed at various levels of the administrative system within your country and internationally.

ADMINISTRATIVE LEVELS	EXAMPLES OF RESOURCES
Community and district or provincial levels	• Opinions, beliefs of key informants (through interviews), in addition to written sources • Clinic- and hospital- based data from routine statistics, registers • Clinical observations, reports of critical incidents • Local surveys, annual service reports • Statistics issued at provincial and district levels • Newspapers, books, articles, mimeographed reports, etc.
National levels	• Articles from national journals, books identified during literature searches at university and other national libraries, WHO, UNICEF libraries, etc. • Special collections, e.g., newspaper clippings, archival records, library or congress collections • Documentation, reports and raw data from: - The Ministry of Health (e.g., 5-year plans) - central statistical offices - Non-governmental organisations

International levels

- Information from:
  - Bilateral and multilateral organisations (e.g., IDRC, USAID, UNICEF, WHO);
  - Computerised searches for international literature (from national library or international institutions).

You need to develop a strategy to gain access to each source and to obtain information in the most productive manner. Your strategy may vary according to where you work and the topic under study. It may include the following steps:

- Identifying a key person (researcher, decision maker or community member) who is knowledgeable on the topic and ask if he or she can give you a few good references or/and the names of other people whom you could contact for further information;
- Looking up the names of speakers on your topic at conferences that may be useful to contact;
- Contacting librarians in universities, research institutions, the Ministry of Health and newspaper offices and requesting relevant references;
- Examining the bibliographies and reference lists in key papers and books to identify relevant references;
- Looking for references in indexes (e.g. Index Medicus) and abstract journals (see **Annex 5.1**); which are available in libraries either as hard copies or in computerised form.
- Requesting a computerised literature search (e.g. Medline, see **Annex 5.2**).

Some agencies will assist with your literature search if requested by telephone or in writing. The request, however, should be very specific. Otherwise you will receive a long list of references, most of which will be not be relevant to your topic. If you are requesting a computerised search it is useful to suggest key words that can be used in locating the relevant references.

Note:

Facilitators should be able to provide specific information regarding national and international facilities to assist you with the search for literature.

References that are identified:

- Should first be skimmed or read.
- Then summaries of the important information in each of the references should be recorded on separate index cards (**Annex 5.3**) or as computer entries. These should then be classified so that the information can easily be retrieved.
- Finally the literature should be included in your protocol.

Information on an index card should be organised in such a way that you can easily find all data you will need for your report:

For an article the following information should be noted:

Author(s)' Surname followed by initials. Title of article. *Name of Journal.* Year, **Volume**, (number): page numbers of article.

Example:

Louria DB. Emerging- and re-emerging infections: The societal variables. *International Journal of Infectious Disease*. 1996, **1**(2):59-62.

For a book the following information should be noted:

Author(s)' Surname followed by initials. *Title of book*. Place: Publisher, year, Edition

Example:

Abramson JH. *Survey methods in community medicine*. Edinburgh: Churchill Livingstone, 1990, 4th ed.

For a chapter in a book, the reference can include:

Author(s) of chapter (Surname(s) followed by initials). Chapter title. In: Editor(s) of book, (Surname(s) followed by initials). (eds). *Title of book*. Place: Publisher, year: page numbers of chapter.

Example:

Todd J and Barongo L. Epidemiological methods. In: Ng'weshemi J, Boerma T, Bennett J and Schapink D (eds). *HIV prevention and AIDS care in Africa; A district level approach.* Amsterdam: KIT Press, 1997: 51-68.

The formats suggested above have been adopted as standard by over 300 biomedical journals and are referred to as the *Vancouver System*. In other journals and books it is common to put the year, between brackets, straight after the name of the author(s). This is called the *Harvard System.* There are more systems in use for referencing to literature. Always carefully look what system is used in the journal you are submitting an article to and **follow it systematically**.

At present many journals use as few punctuation marks as possible. We therefore have minimised punctuation marks in the examples above. In Harvard style, this looks as follows:

Abramson JH (1990) 4th ed. *Survey methods in community medicine*. Edinburgh: Churchill Livingstone.

Further, the **index card** or computer entry (one for each reference) could contain quotations and information such as:

- Key words;
- A summary of the contents of the book or the article, concentrating on information relevant to your study; and
- A brief analysis of the content, with comments such as:

  - Appropriateness of the methodology; possible weaknesses/comments in literature review
  - Important aspects of the study; and
  - How information from the study can be used in your research. (See **Annex 5.3**.)

Note:

Index cards or computer entries can also be used to summarise information obtained from other sources, such as informal discussions, reports of local health statistics, and internal reports. **If you don't use cards or computers, write the information retrieved from your literature in a systematic way on A4 sheets of paper.**

How do you write a review of literature?

There are a number of steps you should take when preparing a review of available literature and information:

- Take your problem analysis diagram as a framework
- Organise your index cards or notes in groups of related statements according to which aspect of the problem they touch upon, e.g., community factors, service factors. Use your problem analysis diagram as a framework for writing (and adapt the diagram in turn as you find more literature).
- Then, decide in which order you want to discuss the various issues. If you discover you have not yet found literature or information on some aspects of your problem that you suspect are important, make a special effort to find this literature. If there is no literature, this supports your justification for conducting the study.

Where do you put which information?

Clearly, you will use some literature when describing the local context (country, region) or your problem. Note that **all facts you mention need a source**, except some general and well-known statements.

Also, for the description of the selected problem, you will use all available raw, grey or published literature you can obtain, well quoted (see below). You may use literature from other countries or regions to illustrate your point. If these sources are many you could have a separate section on international literature (see **Annex 5.4**). More complex studies using theoretical models should have a separate section or chapter discussing these models, which could come after the 'Statement of the Problem' section.

Note:

When drafting your 'Background' section or the 'Statement of the Problem', you will usually not describe your sources one by one. Instead, you will write a **coherent discussion in your own words**, using all relevant literature linked to each other. It is possible to cite several sources for one statement you make. (See **Annex 5.4**)

Referencing

You always need to reference all the literature that you refer to in your review. When you use the Vancouver system, you will use consecutive numbers in the text to indicate your references. At the end of your paper or chapter (of a book) you will then list your references in that order, using the format described above. In your research proposal the references will come before the annexes (see **Modules 1** or **18**).

Alternatively, you can use the Harvard system and refer to the references more fully in the text, putting the surname of the author, year of publication and number(s) of page(s) referred to between brackets, e.g., (Shiva 1998:15-17). If this system of citation is used, the references at the end of the proposal should be listed in **alphabetical** order (see **Annex 5.5**).

The Harvard author/date system of referencing seems easier, as you can change the order of paragraphs without consequences for your referral system. However at present, computers have programmes that change the numbers of your references automatically if you reshuffle the text while using the Vancouver system.

Possible bias

Bias in the literature or in a review of the literature is a distortion of the available information in such a way that it reflects opinions or conclusions, which do not represent the real situation.

It is useful to be aware of various types of bias. This will help you to be critical of the existing literature. If you have reservations about certain references or if you find conflicting opinions in the literature, discuss these openly and critically. Such a critical attitude may also help you avoid biases in your own study. Common types of bias in literature include:

- Playing down controversies and differences in one's own study results;
- Restricting references to those that support the point of view of the author; and
- Drawing far-reaching conclusions from preliminary or shaky research results or making sweeping generalisations from just one case or small study.

Ethical considerations

The types of bias mentioned above would put the **scientific integrity** of the responsible researcher in question. Moreover, careless presentation and interpretation of data may put readers who want to use the study's findings on the wrong track. This may have serious consequences, in terms of time and money spent on HSR and it may even lead to wrong decisions affecting people's health. A similarly serious act, for which a researcher can be taken to court, is the presentation of research results or scientific publications from other writers without quoting the author. Therefore, appropriate referencing procedures should always be followed in research proposals as well as in research reports.

Introduction to group work

For this group work session you will choose a group chairperson. In the sessions that follow you will always have a group chairperson as well as a recorder.

The **functions of a chairperson** are to:

- Make sure that all parts of the group work assignment are understood and completed by the group as a whole
- Take care that all group members have a chance to contribute. (A chairperson should **not** dominate the discussions or always present the results of the group work in plenary sessions.)
- Make sure that tasks are distributed between group members, if they are many, but that the group as a whole has a chance to discuss the different contributions before they are presented in plenary
- Take care that 10 minutes before a plenary session is due to begin, flip charts or overhead sheets are prepared for presentation

- Keep flip charts and other group products together carefully for further use, or delegate this task to a group member

- Organise and co-ordinate the typing of various sections of the research report and carefully store the drafts or delegate this task to a group member

The major **function of the recorder is,** to take care that the flip charts or overhead sheets to be presented in plenary:

- Meet the requirements of the group work assignment

- Contain the main elements of the discussion

- Are **clearly written** and **readable** at a distance

Recorders may change each session, but the leadership role should remain with one group member, for efficiency's sake. This is especially important during the last week, when the final draft of the research proposal is being prepared.

GROUP WORK (2 hours)

1. Select a chairperson and recorder. Read their functions aloud and discuss whether you agree with these functions.

2. Outline the topics for which you need information that will be included in the 'Background' or 'Statement of the Problem' sections of your proposal, making use of your problem analysis diagram.

3. Search through the documents (books, articles, and bibliographies) available in the course library. List the most useful references you can find on your topic. Brainstorm on where to find additional literature.

4. Summarise the most important information from the references. Place this information on index cards, A4 sheets of paper or in computer entries. Divide this work among group members (and make sure it continues after having drafted the research proposal.)

5. Decide whether, apart from the background and statement of the problem sections based on literature, you need a separate section on international literature as introduction to or following the statement of the problem. You may have a section on relevant theory. Write coherent narratives based on the information you collected; analyse and comment on the contributions from various sources, rather than simply reporting on their content. A list of the references used should be presented straight after the text of your research proposal.

REFERENCES

Gibaldi J (1995) *MLA Handbook for Writers of Research Papers*. New York: Modern Language Association of America.

Jen Tsi Yang et al. (1996) *An outline of Scientific Writing: For Researchers with English As a Foreign Language.* Singapore: World Scientific Publishing. www.amazon.com/exec/obidos (through internet, Sept. 2000).

Lindsay D (1996) *Guide to Scientific Writing*. Australia: Addison & Wesley. (paperback)

Annex 5.1: Samples from an abstract journal

CURRENT HEALTH INFORMATION ZIMBABWE

Volume 11 Number 2 **April-June 1997**

Produced by the
UNIVERSITY OF ZIMBABWE MEDICAL LIBRARY
in association with the
MINISTRY OF HEALTH

CURRENT HEALTH INFORMATION ZIMBABWE is produced with the aim of providing information primarily to rural and other health professionals who have little access to current publications.

Levine C.
Orphans of the HIV epidemic: unmet needs in six US cities.
AIDS Care. 7 Suppl 1:S57-62, 1995.
*** Most issues available in U.Z. Medical Library**
In the United States, an estimated 72,000-125,000 children and adolescents will lose their mothers to AIDS by the year 2000. Six cities have been particularly hard hit: New York City, Newark, Miami, San Juan, Los Angeles, and Washington, DC. The most urgent unmet needs for children, their families and new guardians are for mental health services, including bereavement counselling; transitional services to help overcome the loss of AIDS-related benefits following the parent's death; legal services; housing supports, and appropriate evaluations and referrals by juvenile justice and school staff to community-based services. Professional staff need additional training and support. Public policies and legal standards should stress a preference for maintaining children in their extended families, broadly defined, whenever possible. Much more needs to be done to improve the lives and futures of these youth.

Pitts M. McMaster J. Hartmann T. Mausezahl D.
Lay beliefs about diarrhoeal diseases: their role in health education in a developing country.
Social Science & Medicine. 43(8):1223-8, 1996 Oct.
This study examines the beliefs and understandings concerning diarrhoea among 2 groups of Zimbabwean women. Mothers with formal education are compared to those with less formal education. Differences and commonalities of beliefs are examined. The findings show that traditional explanations of an illness such as diarrhoea can inhibit health education campaigns against this disease which kills many children every year.

Annex 5.2: Example of output from a computerised literature search on women and AIDS, through Medline

PubMed medline query

PubMed PubMed QUERY

Other Formats: Citation MEDLINE
Links: Related Articles Go to publisher site

Health Hum Rights 1998;3(1):20-36

Human Rights Approaches to an Expanded Response to Address Women's Vulnerability to HIV/AIDS.

Whelan D

International Center for Research on Women, 1717 Massachusetts Avenue, NW, Suite 302, Washington, DC 20036, USA.

[Record supplied by publisher]

Research from around the world has revealed how gender-related sociocultural norms and economic realities contribute to women's vulnerability to HIV infection, and how gender-related discrimination contributes to their vulnerability to the impact of AIDS. As the global response to the epidemic enters its second decade, the need for an expanded response to address the societal determinants of women's vulnerability to HIV/AIDS is widely accepted. However, public health has been ill-equipped to address the broader context of vulnerability. This paper analyzes the research on gender and vulnerability, including five key policy and programmatic responses that have emerged from the research, through the lens of human rights. Each recommendation will be presented in terms of the promotion and protection of enumerated rights under four human rights treaties, the realization of which can support the objectives of an expanded response to reduce women's vulnerability to HIV and the impact of AIDS.

PMID: 10347373

[redacted] the above report in Macintosh Text format
[redacted] documents on this page through Loansome Doc

http://www.ncbi.nlm.nih.gov 28-7-99

Annex 5.3: Example of a reference summarised on an index card

vd Geest S, Whyte SR eds. (1991) ***The context of Medicines in Developing Countries. Studies in Pharmaceutical Anthropology.*** **Amsterdam: Het Spinhuis Publishers. (2nd edition)**

- The study consists of two parts:

  (1) A description of the *transaction of medicine* (production, selling, consumption). The roles of drug company salesmen, pharmacists, street vendors and 'traditional' practitioners in selling commercial drugs are examined.

  (2) A description of the *meaning of medicines* to its users, e.g., perceived efficacy related to cost, colour, taste, packaging.

- The methodologies used are a combination of survey techniques (to identify where people obtain what drugs, how often etc.) and qualitative techniques (participant observation and in-depth interviews about what people actually use when, how and why).

- Studies cover a wide range of countries in Latin America, Africa, Southern and SE Asia, and even Ancient Europe and Medieval China.

Reverse side of index card:

- Points emphasised in the publication:

  - Self-care with 'western' pharmaceuticals and locally produced and marketed drugs and herbs has as yet received (too) little attention of public health authorities (no control!) and researchers. Still, 70-80% of people worldwide apply self-care.

  - Chinese, Ayurvedic, Unani medicines are increasingly commercialising and finding their way to the global market. Terms like 'indigenous' and 'modern' medicine are therefore loosing their distinctive value.

  - Western pharmaceuticals are often considered as fast, good for acute diseases; herbal medicine are considered as slower and better suited for chronic and recurrent conditions.

- Observations

  - New, interesting field; good mix of methodologies

  - For own HSR study, use pp 131-149,199-216,199-326

Annex 5.4: Example of literature review with references: Vancouver referral system

Part of the literature review in support of the study ***'Factors contributing to defaulting from out-patient treatment among tuberculosis patients who registered in Masvingo Province, Zimbabwe'.*** *

It is interesting that studies quantifying problems in case-finding and case-holding in TB are far more numerous than studies identifying contributing factors, and further, that knowledge, attitudes and practices of patients often receive more attention than those of staff. In areas as different as India and Honduras the same factors appeared responsible for reluctance of patients to come forward for treatment: lack of knowledge of early symptoms, fear of stigma attached to the disease, or a combination thereof.[20] Yet, a study in Japan showed that he average 'doctor's delay' (time between first visit to a doctor and actual diagnosis) always surpassed the 'patient's delay' (time between reported onset of symptoms and first visit to a doctor). For patients who reported early with complaints, within two months after onset of symptoms, the delays in diagnosis were relatively most extensive.[21] Aluoch[22] found the same for Kenya.

Also with respect to patients compliance with treatment we have to consider the multiple contribution of patients, community as well as services. In particular in turbulent times, for example when the TB services are abruptly integrated in the general health system, as happened in Botswana in the end of the 1970s, defaulter rates as high as 75% have been reported.[23]

Other studies in South Africa,[24] India,[25] Papua New Guinea[26] and Malaysia[27] as well as the Botswana study23 concentrate more on patient and community factors, sometimes in interaction with service factors. Poverty, mobility (migrant labour), poor accessibility of TB services, lack of support from relatives, peers or employers, socio-cultural factors (conflicting perceptions of causes and treatment preferences; fear of stigma) and illness factors (low severity of disease at diagnosis, duration of symptoms) emerge as good predictors of poor patient compliance. Very interesting is, however, the shared emphasis on the poor information patients receive on their disease. Roy found in his study in Malaysia[27] that even among patients who were hospitalised for two or three months before starting out-patient treatment, 70% did not know they were suffering from an infectious disease and 80% did not know how long they had to stay in hospital or how long the estimated total duration of their treatment would be. He states that most of the medical and paramedical staff have neither the training nor the interest to give health education, a task for which they have not been primarily employed.

In the field of health education, encouraging experiments have been carried out in some of the countries mentioned above. Papua New Guinea, where patients defaulted en mass both during in- and out-patients treatment, health education was turned into a participatory exercise. Patients not only were explained the basic facts of TB in understandable terms but also listened through the stethoscope to the crackles in their chests and were shown X-rays of affected lungs compared to healthy lungs. As in Malaysia, more experienced patients were involved in health education, besides the health staff; they became the peer educators of newly diagnosed victims of TB and their relatives. The percentage of non-complying out-patients dropped from 50 to 20 within half a year.[26]

The literature underlines the urgency of the planned research in Masvingo Province where the AIDS and TB epidemics seem to fuel each other. It also emphasises the necessity to study the problem of low patient compliance with treatment from different sides: that of the services, the patient and the community.

---

* In: Joint Project on Health Systems Research (1996) *HSR: It can make a difference*. Vol. 3: Under-utilisation of tuberculosis services in Southern Africa. Harare: WHO/AFRO.

References and notes

20. Westaway MS. Knowledge and attitudes about tuberculosis of black hospitalised TB patients. *Tubercle,* 1990,**71**:55-59.
21. Aoki M, Mori T, Shimao T. Studies on factors influencing patient's, doctor's and total delay of tuberculosis case-detection in Japan. *Bulletin of the International Union Against Tuberculosis*, 1990, **60**(3-4):128-130.
22. Aluoch JA. A survey to assess passive case-finding in pulmonary tuberculosis in Kenya. *East African Medical Journal*, 1983, **60**(6):360-365.
23. Varkevisser CM. *TB Control in Botswana; Problem identification*. Amsterdam: Royal Tropical Institute, 1977:42-53; 24-41.
24. Bell J, Jach D. Tuberculosis patient compliance in the Western Cape, 1984. *South African Medical Journal*, 1988, **73**:31-33.
25. Barnhoorn F, Adriaanse H. Insearch of factors responsible for non-compliance among tuberculosis patients in Wardha District, India. *Social Science and Medicine*, 1992, **34**(3):291-306.
26. Garner P, Hill G. Brainwashing in Tuberculosis management. *Papua New Guinea Medical Journal*, 1985, **28**:291-293.
27. Roy RN. Systematic health education of tuberculosis patients and of the population. *Bulletin of the International Union Against Tuberculosis*, 1990, **60**(3-4):33-34.

Annex 5.5: Same part of literature review as 5.4, now using the Harvard referral system

It is interesting that studies quantifying problems in case-finding and case-holding in TB are far more numerous than studies identifying contributing factors, and further, that knowledge, attitudes and practices of patients often receive more attention than those of staff. In areas as different as India and Honduras the same factors appeared responsible for reluctance of patients to come forward for treatment: lack of knowledge of early symptoms, fear of stigma attached to the disease, or a combination thereof (Westaway 1990). Yet, a study in Japan showed that he average 'doctor's delay' (time between first visit to a doctor and actual diagnosis) always surpassed the 'patient's delay' (time between reported onset of symptoms and first visit to a doctor). For patients who reported early with complaints, within two months after onset of symptoms, the delays in diagnosis were relatively most extensive (Aoki et al 1990). Aluoch (1983) found the same for Kenya.

Also with respect to patients compliance with treatment we have to consider the multiple contribution of patients, community as well as services. In particular in turbulent times, for example when the TB services are abruptly integrated in the general health system, as happened in Botswana in the end of the 1970s, defaulter rates as high as 75% have been reported (Varkevisser 1977).

Other studies in South Africa, (Bell and Jach 1984) India (Barnhorn and Adriaanse 1992), Papua New Guinea (Garner and Hill 1985) and Malaysia (Roy 1985) as well as the Botswana study (Varkevisser 1977) concentrate more on patient and community factors, sometimes in interaction with service factors. Poverty, mobility (migrant labour), poor accessibility of TB services, lack of support from relatives, peers or employers, socio-cultural factors (conflicting perceptions of causes and treatment preferences; fear of stigma) and illness factors (low severity of disease at diagnosis, duration of symptoms) emerge as good predictors of poor patient compliance. (etc., etc.)

REFERENCES

Aluoch JA (1983) A survey to assess passive case-finding in pulmonary tuberculosis in Kenya. *East African Medical Journal* **60**(6):360-365.

Aoki M, Mori T, Shimao T (1990) Studies on factors influencing patient's, doctor's and total delay of tuberculosis case-detection in Japan. *Bulletin of the International Union Against Tuberculosis* **60**(3-4):128-130.

Barnhoorn F, Adriaanse H (1992) In search of factors responsible for non-compliance among tuberculosis patients in Wardha District, India. *Social Science and Medicine* **34**(3):291-306.

Bell J, Jach D (1988) Tuberculosis patient compliance in the Western Cape, 1984. *South African Medical Journal* **73**:31-33.

Garner P, Hill G (1985) Brainwashing in Tuberculosis management. *Papua New Guinea Medical Journal* **28**:291-293.

Roy RN (1990) Systematic health education of tuberculosis patients and of the population. *Bulletin of the International Union Against Tuberculosis* **60**(3-4):33-34.

Varkevisser CM (1977) *TB Control in Botswana; Problem identification*. Amsterdam: Royal Tropical Institute:42-53, 24-41.

Westaway MS (1990) Knowledge and attitudes about tuberculosis of black hospitalised TB patients. *Tubercle* **71**:55-59.

Trainer's Notes

Module 5: REVIEW OF AVAILABLE LITERATURE AND INFORMATION

Timing and teaching methods

¾ hour	Introduction and discussion
3 + hours +	Group work (but literature will be read till the proposal is in its final version, and then till the study has been implemented and the report is finalised)
3¾ hours	TOTAL TIME

Materials

- Examples of:

  - Abstract journals
  - Index Card
  - Computer printouts
  - Literature reviews and reference lists

(You can use copies of **Annexes 5.1-5.5**, and supply your own examples.)

- Some blank index cards or blank sheets of paper for each participant

Be sure that a course library is ready for use. Prior to the workshop, facilitators should look through their own resources to find any relevant articles they may have for each research topic.

Introduction and discussion

Discuss why and how to do a review of the literature. Have participants suggest answers to the questions, but provide additional information when necessary.

- Refer to the annexes for examples of tools that can be used to find information relevant to a specific research topic and an example of a literature review
- It may be useful to have the assistance of a librarian in this session.
- Provide information on national library facilities that may be available during or after the course.
- Stress the importance of developing libraries at all management levels in organisations and ministries concerned with solving health problems.
- Present the points concerning preparation and use of index cards or computer entries, or at least small summaries on separate sheets of paper.
- Insist that the participants should be 'writing', not merely 'citing' when using the sources in different parts of their proposal, and provide different options for 'putting what where.'

- Discuss possible biases in documents and literature review. Stress the researcher's responsibility for presenting his findings honestly so that he does not put readers who want to implement the findings on the wrong track.

- Ask for comments or questions concerning the review of the literature and problems participants are likely to face. Determine how you can help the participants to overcome these problems.

Group work

- Ask the group to begin reviewing relevant literature and information departing from their problem analysis diagram. If possible, try to obtain relevant papers and reports from various sources for use even during the course.

- As a first step, each participant should review at least two articles, reports, or books, using index cards or blank sheets of paper to make summaries.

- Then the information should be put together in the 'Background' and 'Statement of the Problem' sections, and a decision will have to be made where to put international literature. **Make sure that references are made in a consistent way**, by using the Vancouver or the Harvard system (or another system of their choice).

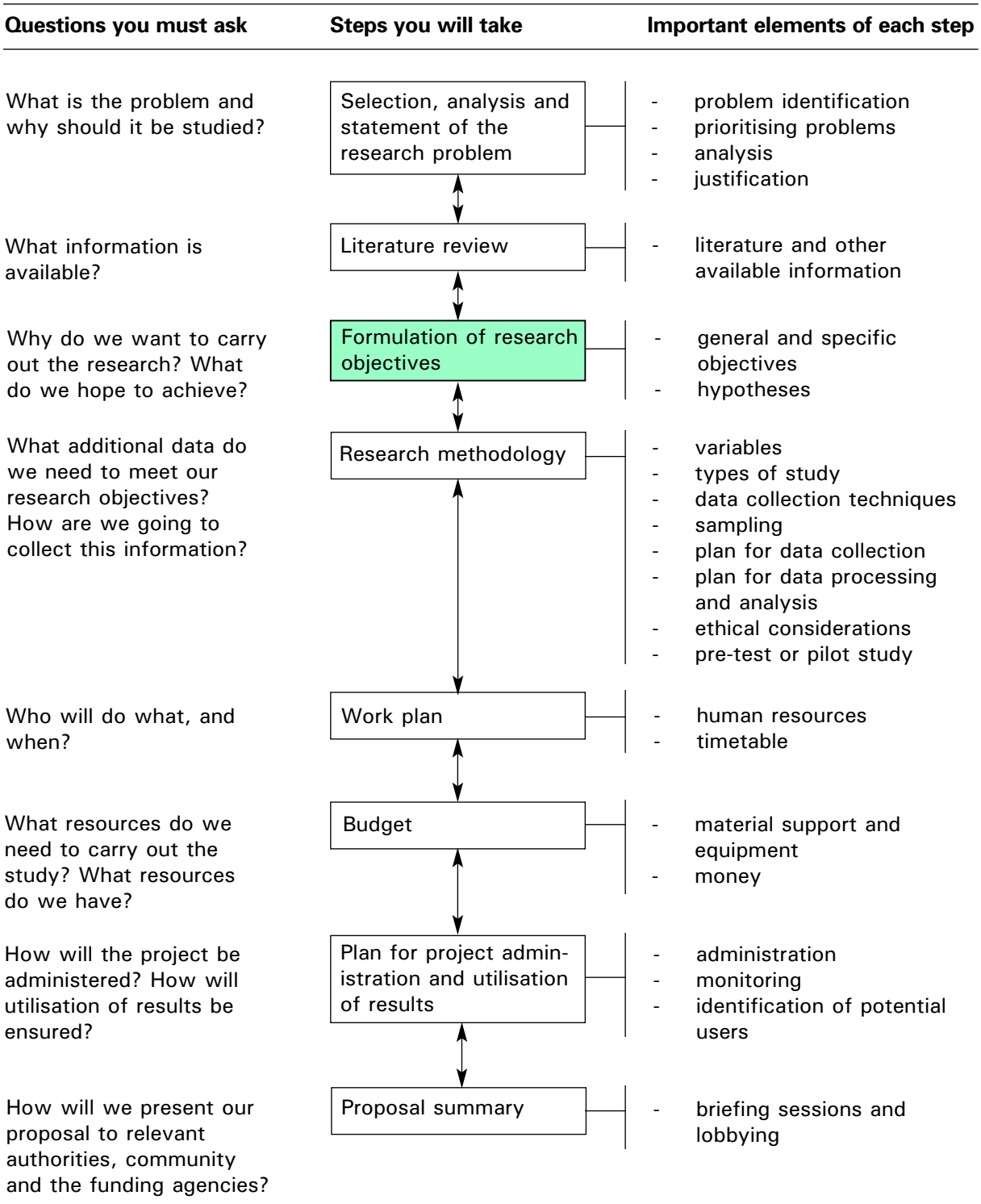
- Emphasise that the review of literature should be thorough and critical. Only references that relate directly to the proposed research should be discussed. Irrelevant literature should **not** be mentioned.

Designing and Conducting Health Systems Research Projects
Part I: Proposal Development and Fieldwork

Module 6

FORMULATION OF RESEARCH OBJECTIVES

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 6: FORMULATION OF RESEARCH OBJECTIVES

OBJECTIVES

At the end of this session you should be able to:

1. **State** the reasons for writing objectives for your research project.
2. **Define** and describe the difference between general and specific objectives.
3. **Define** the characteristics of research objectives.
4. **Prepare** research objectives in an appropriate format for the project you are developing.
5. **Develop** further research questions, and research hypotheses, if appropriate for your study.

I. Objectives

II. Hypotheses

III.Title of the study

I. RESEARCH OBJECTIVES

The **OBJECTIVES** of a research project summarise what is to be achieved by the study.

Objectives should be closely related to the statement of the problem. For example, if the problem identified is low utilisation of child welfare clinics, the general objective of the study could be to identify the reasons for this low utilisation, in order to find solutions.

The **general objective** of a study states what researchers expect to achieve by the study in general terms.

It is possible (and advisable) to break down a general objective into smaller, logically connected parts. These are normally referred to as **specific objectives.**

Specific objectives should systematically address the various aspects of the problem as defined under 'Statement of the Problem' (**Module 4**) and the key factors that are assumed to influence or cause the problem. They should specify **what** you will do in your study, **where** and **for what purpose**.

A study into the cost and quality of home-based care for HIV/AIDS patients and their communities in Zimbabwe, developed at an HSR workshop, for example, had as its general objective:

> To explore to what extent community home-based care (CHBC) projects in Zimbabwe provide adequate, affordable and sustainable care of good quality to people with HIV/AIDS, and to identify ways in which these services can be improved.
>
> It was split up in the following specific objectives:
>
> 1. To identify the full range of economic, psychosocial, health/nursing care and other needs of patients and their families affected by AIDS.
> 2. To determine the extent to which formal and informal support systems address these needs from the viewpoint of service providers as well as patients.
> 3. To determine the economic costs of CHBC to the patient and family as well as to the formal CHBC programmes themselves.
> 4. To relate the calculated costs to the quality of care provided to the patient by the family and to the family/patient by the CHBC programme.
> 5. To determine how improved CHBC and informal support networks can contribute to the needs of persons with AIDS and other chronically and terminally ill patients.
> 6. To use the findings to make recommendations on the improvement of CHBC to home care providers, donors and other concerned organisations, including government.

The first specific objective usually focuses on quantifying or specifying the problem.

This is necessary in many studies, especially when a problem has been defined (but not quantified) for which subsequently the major causes have to be identified. Often use can be made of available statistics or of the health information system. In the study on the high defaulter rate of TB patients, this rate should first be established, using the records, and only then would the contributing factors to defaulting be analysed.

In the example given, the needs of AIDS patients and their relatives for care and support have been defined in the first objective. The objectives which follow concentrate on adequacy, cost and quality of care provided whereas the last two objectives specify possible improvements with respect to CHBC, and to whom the results and recommendations of the study will be fed back.

Note:

It may be helpful to use the diagram as a point of departure and check whether the **problem** and all **major, directly contributing factors** (analytic study) or **major components** (descriptive or evaluation study) have been covered by the objectives.
An objective indicating **how the results will be used** should be included in every operational study, either as part of the general objective or as a specific objective.

Why should research objectives be developed?

The formulation of objectives will help you to:

- **Focus** the study (narrowing it down to essentials);
- **Avoid** the collection of data which are not strictly necessary for understanding and solving the problem you have identified; and
- **Organise** the study in clearly defined parts or phases.

Properly formulated, specific objectives will facilitate the development of your research methodology and will help to orient the collection, analysis, interpretation and utilisation of data.

How should you state your objectives?

Take care that the objectives of your study:

- Cover the different aspects of the problem and its contributing factors in a **coherent** way and in a **logical sequence**;
- Are **clearly phrased** in **operational terms,** specifying exactly what you are going to do, where, and for what purpose;
- Are **realistic** considering local conditions; and
- Use **action verbs** that are specific enough to be evaluated.

  Examples of action verbs are: to determine, to compare, to verify, to calculate, to describe, and to establish.
  Avoid the use of vague non-action verbs such as: to appreciate, to understand, or to study.

Keep in mind that when the project is evaluated, the results will be compared to the objectives. If the objectives have not been spelled out clearly, the project cannot be evaluated.

Using the previous example on cost and quality of CHBC, we may develop more specific **research questions** for the different objectives, such as:

- Do rural and urban CHBC projects differ with respect to the adequacy, quality, affordability and sustainability of HBC provided?
- How satisfied are AIDS patients, relatives and service providers with the care provided? Are there differences in perceptions between those groups?
- Is the stigma attached to being HIV+ the same strong for women as for men? Or are there gender differences in stigma?
- What impact does the care provided to AIDS patients have on the economy of the homestead? Is there competition with other basic needs (e.g. schooling of children, purchases of food)?

II. HYPOTHESES

Based on your experience with the study problem, it might be possible to develop explanations for the problem, which can then be tested. If so, you can formulate hypotheses in addition to the study objectives.

A HYPOTHESIS is a prediction of a relationship between one or more factors and the problem under study that can be tested.

In our example concerning the cost and quality of HBC in Zimbabwe it would have been possible to formulate and test the following hypotheses:

1. The role of first-line relatives in the provision of care to AIDS patients is more substantial in rural than in urban areas.
2. The silence and stigma surrounding AIDS makes the formation of self-help groups of AIDS patients and their relatives next to impossible, which in turn maintains the high level of stigma on HIV/AIDS.

Note:

Policy makers and field staff usually feel the need for research because they do **NOT** have enough insight into the causes of a certain problem. Therefore, most HSR proposals present the specific objectives in the form of **open statements** (as given in the examples earlier) instead of focusing the study on a limited number of hypotheses.

III. TITLE OF THE STUDY

Now you can finalise the title of your study. The title should be in line with your general objective. Make sure that it is specific enough to tell the reader what your study is about and where it will be calculated.

NOT: 'A study on community home-based care'

BUT: 'A study on cost and quality of community home-based care for HIV/AIDS patients and their communities in Zimbabwe'

You might also consider fancier titles:

> 'Do We Care? A study on cost and quality of CHBC for HIV/AIDS patients in Zimbabwe'*

Another example could be:

> 'WORKSHOPS: Blessings or Burdens? A study of the workshops held in 1999 in Province Y - Their utility and consequences for daily working activities of health staff'

* The study with this title, used as an example in the present module, was carried out by G Woelk, H Jackson, R Kerkhoven, K Hansen, N Manjonjori, P Maramba, J Mutambirwa, E Ndimande and E Vera. It was published in December 1997 by the Department of Community Medicine, University of Zimbabwe, the Southern African AIDS Information Dissemination Service (SAFAIDS) and the National AIDS Control Programme, Ministry of Health, Harare, Zimbabwe.

GROUP WORK (2 hours)

1. Choose a chairperson and a recorder.

2. Hang up the flip charts that you used to present your statement of the problem so they are visible to all group members. Incorporate useful suggestions for changes that were made when you presented them in plenary. Then, use the analysis diagram as a starting point for formulating objectives, focusing, for example, on:

   - Further quantifying and specifying the problem, if required;
   - Exploring the key factors or major groups of factors that, in your opinion, might influence or cause the problem; and/or
   - Any other major research activities you propose.

3. Prepare a general objective and specific objectives for the research proposal you are developing.

4. After formulating your objectives ask yourself the following questions:

   - Do the objectives deal with all aspects of the research problem in a logical and coherent way?
   - Are the objectives clearly phrased?
   - Are the objectives defined in operational terms that can be measured? Are they realistic?
   - Do they indicate where the study will be conducted?
   - Do they include the development of recommendations for how the research results will be used to solve the problem?

5. Prepare a flip chart with your objectives for use in the **exercise** and in the plenary discussion. Add on the title of your study and revise it, if necessary, to match the objectives

EXERCISE: Assessing the objectives of another group (½ hour)

Assess the research objectives formulated by another team using the criteria mentioned above. Compare them with the group's statement of the problem and the title of the study.

Trainer's Notes

Module 6: FORMULATION OF RESEARCH OBJECTIVES

Timing and teaching methods

½ hour	Introduction and discussion
2 hours	Group work
½ hour	Exercise: Assessing the objectives of another group
1 hour	Presentation by each group, followed by comments by the group that did the exercise and general discussion
1 hour	Adjustments
5 hours	TOTAL TIME

Introduction and discussion

- Emphasise that the formulation of clear and comprehensive objectives is critical to the development of all the other components of a research design, as well as to subsequent data analysis and report writing.

- Formulation of good objectives is a skill with which many participants have difficulty. Two types of problems come up quite often:

  - Difficulties with developing concise, operational objectives that focus clearly on what the study hopes to accomplish and cover all parts of the study in a logical order;
  - Difficulties in understanding the difference between programme objectives and research objectives. For example, many participants may not, in the beginning, see the distinction between a programme objective, such as, 'Make sure that Health Posts in District X are supplied monthly with sufficient drugs' and a research objective, such as 'To compare two methods of supplying drugs to Health Posts in District X'.

  Reference to the analysis diagram that groups developed in **Module 4** will help solve these problems. It should be stressed that they should first consider whether they need more data to specify their problem. Then they should systematically write objectives to cover the different categories of factors they have identified.

- Stress that it is not necessary to develop an objective for every single contributing factor they included in the diagram. The participants should try to limit their objectives to two or three for each major category in their diagram, including several factors in each objective, when possible.

Group work

Be sure to provide sufficient time for the groups to formulate good objectives for their chosen projects. As groups work from their analysis diagram, they may discover that changes are necessary (additions, regrouping, or dropping of factors). It is recommended that the diagram be displayed on a flipchart rather than on an overhead sheet with photocopies for individual group members, so it will be easier to focus the group's attention on it. The flip chart with the diagram can also be used in **Module 8** (Variables).

EXERCISE: Assessing the objectives of another group

Hold an exercise in which groups evaluate the objectives prepared by another group, using the criteria set out on the exercise sheet.

Plenary session

Have each group present their analysis diagram and the objectives they have developed. Immediately following each presentation, ask the group that analysed the objectives during the exercise to comment. Then open up the discussion to the rest of the class. (Allow 15 minutes per topic.)

Each group should also present the title of its research project.

It is important that each group receive clear feedback on the quality of the objectives they have developed, as well as practical suggestions for improvement. When providing feedback, ask yourself:

1. Do the objectives cover all parts of the analysis diagram, in a logical order?
2. Do the objectives really measure what the group wants them to?
3. If the objectives were met, would the study provide the results needed to solve the problem posed in the statement of the problem?
4. Are the objectives feasible? If too ambitious, could the scope of the study be reduced?
5. Is the title specific enough and does it cover the objectives?

Adjustments

Facilitators in past courses have found it useful to provide a second group work session in which participants can finalise their objectives, analysis diagram and title of the research project, after they have received feedback during the plenary session.

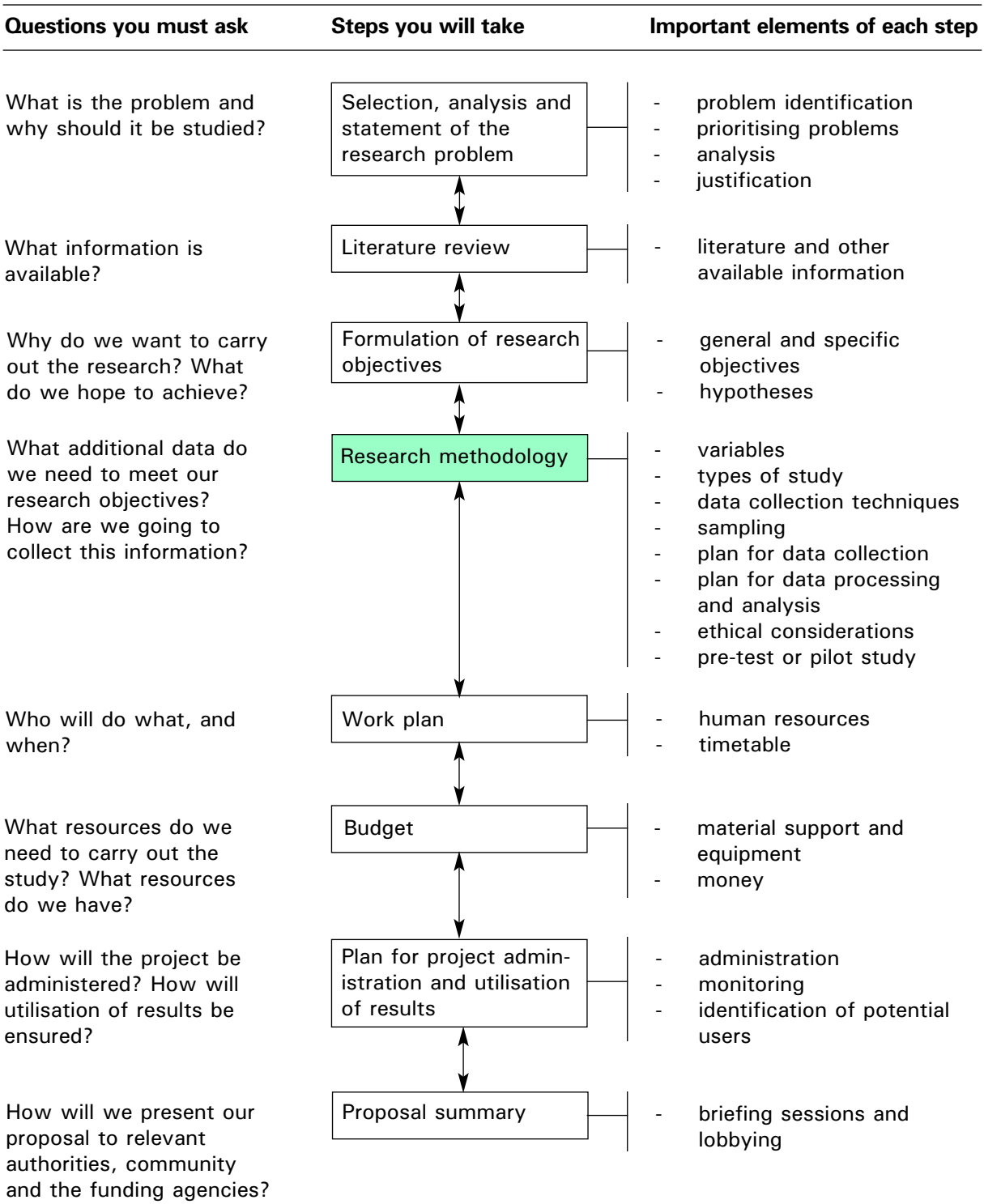
Designing and Conducting Health Systems Research Projects

Part I: Proposal Development and Fieldwork

Module 7

INTRODUCTION TO HEALTH SYSTEMS RESEARCH METHODOLOGY

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 7: INTRODUCTION TO HEALTH SYSTEMS RESEARCH METHODOLOGY

OBJECTIVES

At the end of the session you should be able to:

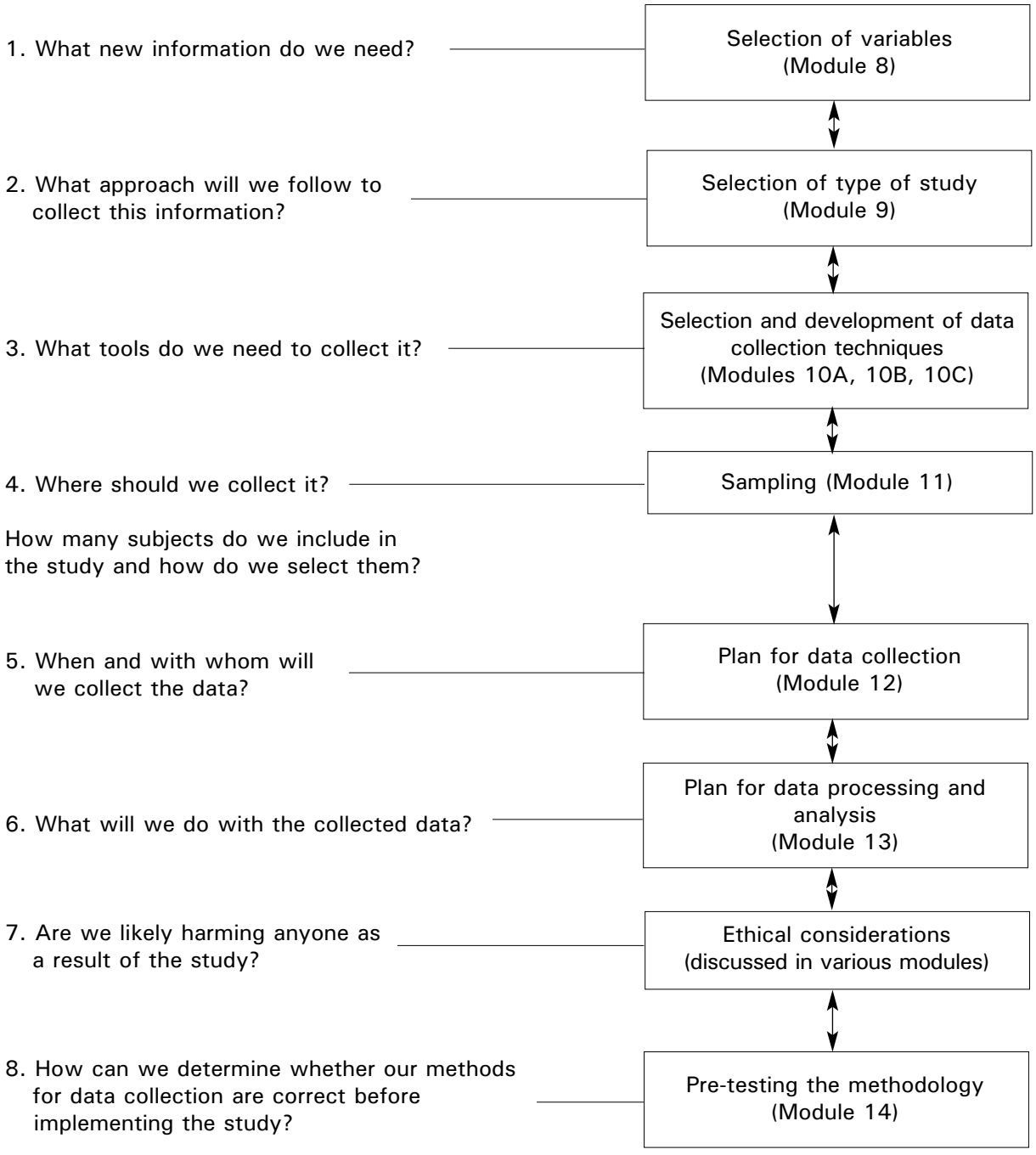
1. **Identify** the pertinent questions to consider when developing the methodology for your research proposal.
2. **Describe** the components that should be dealt with in the methodology section of your research proposal.

In the previous modules, you:

- Selected a research topic;
- Prepared a brief description of the problem and its importance;
- Conducted a literature and information review to determine what was already known about the problem; and
- Developed objectives that clearly state the purpose of the study, what study results are expected, and how the results will be used.

Now you must decide exactly how you are going to achieve your stated objectives: i.e., what new data you need in order to shed light on the problem you have selected and how you are going to collect and process this data. The questions in the flow chart on the next page cover the major issues that must be examined as you develop your **research design**. These issues will be dealt with in **Modules 8**, **9**, **10**, **11**, **12**, **13** and **14**.

RESEARCH DESIGN



Note: The steps are interrelated. The process is usually cyclical in nature. After completing a step, you should review previous steps to ensure consistency in your proposal.

Trainer's Notes

Module 7: INTRODUCTION TO HEALTH SYSTEMS RESEARCH METHODOLOGY

Timing and teaching methods

15 min.	Introduction and discussion
15 min.	TOTAL TIME

Guidelines for trainers

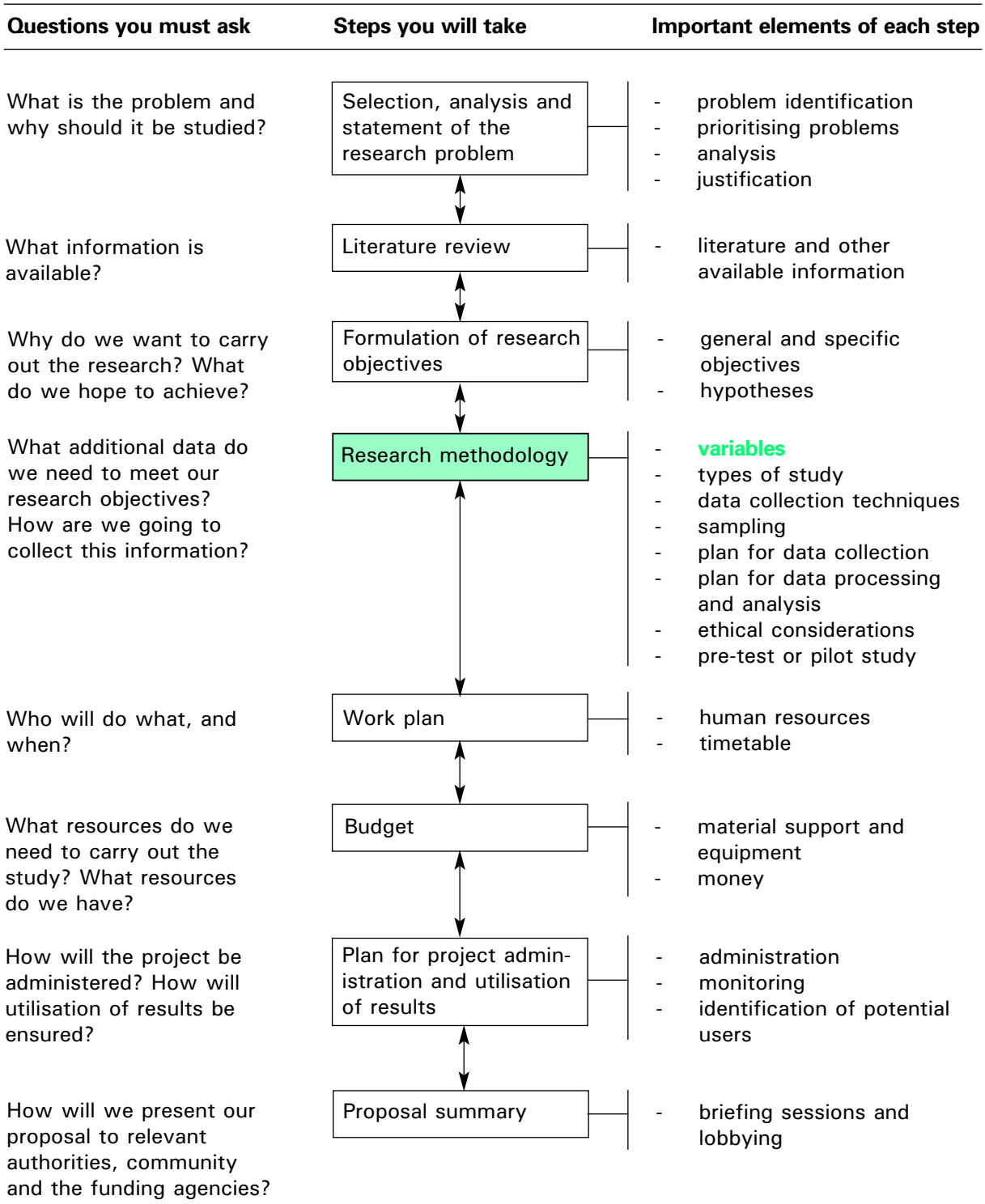
- List and explain the components of a good research design as outlined in the module.
- Stress the cyclical nature of the different steps in designing the methodology.

Designing and Conducting Health Systems Research Projects
Part I: Proposal Development and Fieldwork

Module 8

VARIABLES

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 8: VARIABLES

OBJECTIVES

At the end of the session you should be able to:

1. **Define** what variables are and describe why their selection is important in research.
2. **State** the difference between numerical and categorical variables and define the types of scales of measurement.
3. **Discuss** the difference between dependent and independent variables and how they are used in research designs.
4. **Identify** the variables that will be measured in the research project you are designing and develop operational definitions with indicators for those variables that cannot be measured directly.
5. **List** the variables that you hope to identify and describe during your planned study but that cannot be measured at this time (qualitative data).

I. Introduction

II. Formulating variables

III. Identifying indicators in qualitative studies

IV. Causes and associations; confounders

I. INTRODUCTION

In **Module 4** we analysed the problem we wanted to investigate. The problem itself and all the factors that might influence it were presented in a diagram, which then served as the basis for the formulation of research objectives. Now we have come to a stage where we must ask ourselves the question:

'What information are we going to collect in our study **to meet our objectives?'**

- In most studies, we must first describe the problem itself more precisely.

  For example, in a study that is investigating why so many tuberculosis (TB) patients default from out-patient treatment, we first want to know how high the defaulter rate is: is it 10%, 30%, 50%? To obtain the defaulter rate we need a clear definition of what we mean by defaulting (how many times treatment was missed).

- We also want to know **whether** certain factors do indeed influence the problem, and **to what extent.** If we know the extent to which a certain factor influences the problem, we are much more likely to be able to convince ourselves (and relevant others) to take action.

  For example, if we find that becoming a dropout of TB treatment is strongly associated with the following factors, we have clues that will help us to solve the problem:

  - The patient's lack of knowledge concerning the actual duration of treatment and the danger of relapse or death when the full course is not completed;
  - Living more than 8 km away from the clinic where the drugs have to be collected monthly; and
  - Being between 15 and 30 years of age.

To find these associations between problems and contributing factors, it is essential that we carefully define the problem itself, as well as each of the factors identified when analysing the problem in **Module 4**. We do this by formulating variables.

II. FORMULATING VARIABLES

What is a variable?

A **VARIABLE** is a characteristic of a person, object or phenomenon which can take on different values. These may be in the form of numbers (e.g., age) or non-numerical characteristics (e.g., sex).

A simple example of a variable in the form of numbers is 'a person's age'. The variable 'age' can take on different values since a person can be 20 years old, 35 years old and so on. Other examples of variables are:

- weight (expressed in kilograms or in pounds);
- home - clinic distance (expressed in kilometres or in minutes walking distance);
- monthly income (expressed in dollars, rupees, or kwachas); and
- number of children (1, 2, etc.).

Because the values of all these variables are expressed in numbers, we call them NUMERICAL VARIABLES.

Some variables may also be expressed in categories. For example, the variable sex has two districts categories, groups, male and female. Other examples are:

Table 8.1: Examples of categorical variables

Variables	Categories
Colour	• red • blue • green, etc.
Outcome of disease	• recovery • chronic illness • death
Main type of staple food eaten	• maize • millet • rice • cassava, etc

Since these variables are expressed in categories, we call them CATEGORICAL VARIABLES.

Further breakdown of numerical and categorical variables (optional)

Numerical variables can either be continuous or discrete.

i. ***Continuous***. With this type of data, one can develop more and more accurate measurements depending on the instrument used, e.g.:

- height in centimetres (2.5 cm or 2.546 cm or 2.543216 cm)
- temperature in degrees Celsius (37.2°C or 37.19999°C etc.)

ii. ***Discrete***. These are variables in which numbers can only have full values, e.g.:

- number of visits to a clinic (0, 1, 2, 3, 4, etc).
- number of sexual partners (0, 1, 2, 3, 4, 5, etc.)

Categorical variables, on the other hand, can either be ordinal or nominal.

i. ***Ordinal variables.*** These are grouped variables that are ordered or ranked in increasing or decreasing order:

For example: High income (above $300 per month);
Middle income ($100-$300 per month); and
Low income (less than $100 per month).

Other examples are:

Disability:	no disability, partial disability, serious or total disability
Seriousness of a disease:	severe, moderate, mild
Agreement with a statement:	fully agree, partially agree, fully disagree
Fear of leprosy:	will not share food with a patient; will not enter the house of a patient; will not allow patient to live in the community.

Note:

Fear of leprosy is an attitude, and attitudes are often scaled (you make them into ordinal variables).

It is obvious that the definition of what we would call high (income) or far (distance) will vary from country to country and from region to region. If a researcher has little idea about the distribution of a certain variable in a population (**for example,** if you don't know whether 30%, 50%, or 95% are below the poverty line of $100 per month), it is advisable to categorise numerical data only after the pre-test, or even **after** data collection (see **Module 13**).

ii. ***Nominal variables***. The groups in these variables do not have an order or ranking in them.

For example:

Sex:	male, female
Main food crops:	maize, millet, rice, etc.
Religion:	Christian, Moslem, Hindu, Buddhism, etc.

For examples of scales of measurement, see **Annex 8.1**. We will come back to these distinctions in **Module 22**, as continuous, discrete, ordinal and nominal data require different statistical tests.

EXERCISE 1:

Look at your problem analysis diagram and give examples of numerical (continuous and discrete) and categorical (ordinal and nominal) variables.

Factors rephrased as variables

When looking at your problem analysis diagram you will notice that most of what we called 'factors' are in fact variables which have negative values. We phrased the contributing factors negatively on purpose (e.g., lack of knowledge) as it is much easier to visualise these factors in the negative. However, in reality not everyone with good knowledge of TB treatment is a regular attender and not everyone with poor knowledge absconds from treatment. As we conduct our study we will try to determine **to what extent** these contributing factors play a role. Therefore we have to formulate them in a neutral way, so that they can take on positive as well as negative values. The table below presents examples of negatively phrased 'factors' and how they can be rephrased as neutral 'variables'.

Table 8.2: Factors rephrased as variables

Factors as presented in the Analysis Diagram	Variables
• Long waiting time • Absence of drugs • Lack of supervision • Poor knowledge of the signs, causes and consequences of TB	• Waiting time • Availability of drugs • Frequency of supervisory visit • Knowledge of the signs, causes and consequences of TB

Operationalising variables by choosing appropriate indicators

Note that the different values of many of the variables presented up to now can easily be determined. However, for some variables it is sometimes not possible to find meaningful categories unless the variables are made operational with one or more precise **INDICATORS**. Operationalising variables means that you make them 'measurable':

For example:

- In many HSR studies, you want to determine the **level of knowledge** concerning a specific issue in order to find out to what extent the factor 'poor knowledge' influences the problem under study (**for example** low utilisation of pre-natal care by pregnant women).

  The variable 'level of knowledge' cannot be measured as such. You would need to develop a series of questions to assess a woman's knowledge, for example on pre-natal care and risk factors related to pregnancy. The answers to these questions form an **indicator** of someone's knowledge on this issue, which can then be categorised. If 10 questions were asked, you might decide that the knowledge of those with:

  - 0 to 3 correct answers is poor,
  - 4 to 6 correct answers is reasonable, and
  - 7 to 10 correct answers is good.
- Nutritional status of under-5 year olds is another example of a variable that cannot be measured directly and for which you would need to choose appropriate indicators. Widely used indicators for nutritional status include:

  - Weight in relation to age (W/A)
  - Weight in relation to height (W/H)
  - Height in relation to age (H/A)
  - Upper-arm circumference (UAC)

  For the classification of nutritional status, internationally accepted categories already exist, which are based on so-called standard growth curves. For the indicator 'Weight/Age', for example, children are:

  - *well-nourished* if they are above 80% of the standard,
  - *moderately malnourished* if they are between 60% and 80%, and
  - *severely malnourished* if they are below 60%.

Note:

When defining variables on the basis of the problem analysis diagram, it is important to realise which variables are measurable as such and which ones need indicators. Once appropriate indicators have been identified we know exactly what information we are looking for. This makes the collection of data as well as the analysis more focused and efficient.

Defining variables and indicators of variables

To ensure that everyone (the researcher, the data collectors, and eventually, the reader of the research report) understands exactly what has been measured and to ensure that there will be consistency in the measurement, it is necessary to clearly define the variables (and indicators of variables). **For example**, to define the indicator 'waiting time' it is necessary to decide what will be considered the starting point of the 'waiting period' e.g., is it when the patient enters the front door, or when he has been registered and obtained his card?

Annex 8.2 gives examples of common variables with different possible choices for indicators.

III. IDENTIFYING INDICATORS IN QUALITATIVE STUDIES

Certain variables cannot be defined with indicators before the study, because the information to do this is lacking. The purpose of the study may be to find this information.

For example, policy makers in Nepal would like to eliminate leprosy. They have noticed that fewer women report for leprosy treatment than men and would like to know whether stigma keeps women from reporting for treatment and/or whether the services have to be more sensitive to the needs of women for privacy at diagnosis.

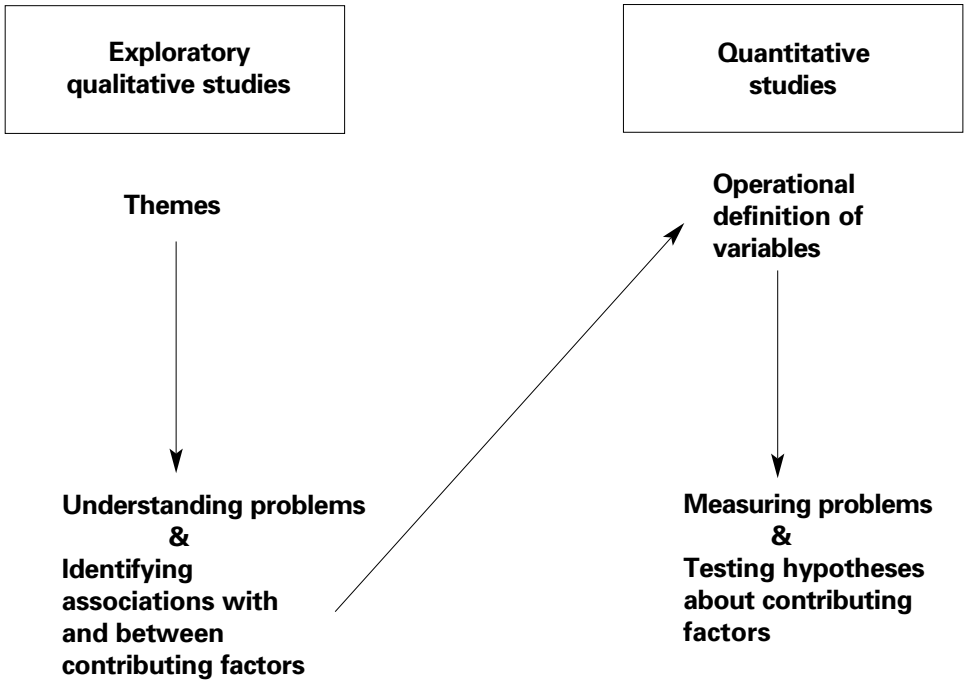
We define stigma as an *undesirable* differentness that disqualifies a person from full social acceptance (Goffman: 1963). However, we cannot fill in more precisely in what way men and women are discriminated against, as that has still to be studied. Some indicators for stigma could be the divorce rate of male and female patients, or the degree of isolation of the patient by the healthy spouse or by the community, but how the severity of this isolation should be measured is still unknown. Possibilities includd, for example, whether patients and spouses still share a house, share food, share one bed? Do community members still accept leprosy patients as village leaders, do they welcome patients to attend village meetings, and, if so, do they still drink beer or eat together, and do they ask patients to bring their own cups?

Note: that in many qualitative studies the researcher is not primarily interested in measuring variables, but rather in **identifying** variables or clusters of variables that help explain a problem or reasons for success. In that case, the researcher will often try to find indicators that make the variables measurable.

One could state that in exploratory, qualitative studies we study themes, such as stigma, to understand better how patients suffer from stigma and how they cope with it. We also discover contributing factors to stigma: in some societies women are more vulnerable to stigma than men; adolescents are more vulnerable than adults who have settled economically and socially; patients with deformities are always more vulnerable to stigma than those without visible signs.

By better understanding the problem of stigma we can now give an operational definition of the strength of stigma on a scale. This enables us to measure through a quantitative study the degree of stigma male and female patients suffer from, and the most important contributing factors to stigma. (See **Figure 8.1**)

Figure 8.1: Relationship between qualitative and quantitative studies in understanding and measuring problems



IV. CAUSES AND ASSOCIATIONS; CONFOUNDING

Dependent and independent variables

Because in health systems research you often look for causal explanations, it is important to make a distinction between **dependent** and **independent variables**.

> The variable that is used to describe or measure the problem under study is called the DEPENDENT variable.
>
> The variables that are used to describe or measure the factors that are assumed to cause or at least to influence the problem are called the INDEPENDENT variables.

For example, in a study of the relationship between smoking and lung cancer, 'suffering from lung cancer' (with the values yes, no) would be the dependent variable and 'smoking' (varying from not smoking to smoking more than three packets a day) the independent variable.

Whether a variable is dependent or independent is determined by the statement of the problem and the objectives of the study. It is therefore important when designing an analytical study to clearly state which variable is the dependent and which are the independent ones.
Note that if a researcher investigates why people smoke, 'smoking' is the dependent variable, and 'pressure from peers to smoke' could be an independent variable. In the lung cancer study ' smoking' was the independent variable.

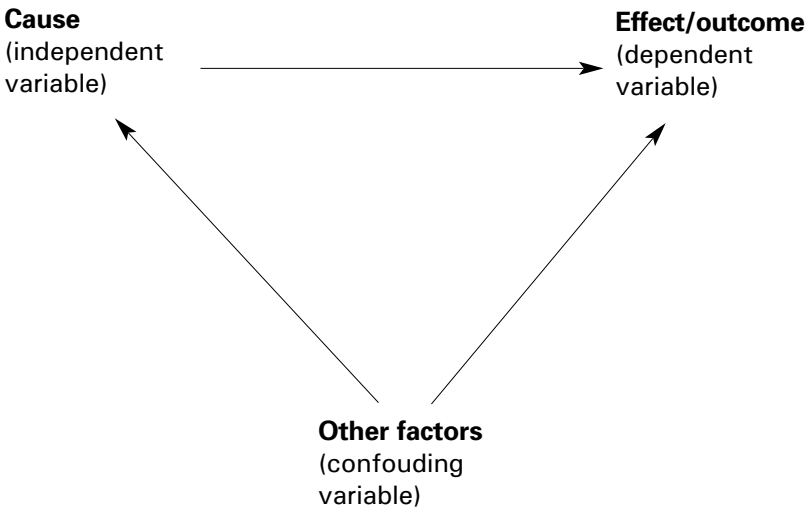
EXERCISE 2:

Look at your analysis diagram and see if you can give an example of a dependent variable and one or two independent variables in your own study.

Although in everyday language we may speak of possible **CAUSES** of problems, in scientific language we prefer to speak of **ASSOCIATIONS** between variables, unless a causal relationship can be proven. If we find an association between smoking and cancer, we can conclude that smoking **causes** cancer only if we can both demonstrate that the cancer was developed **after** the patient started smoking and that there are no other factors that could have caused both the cancer and the habit of smoking. Nervous people, for example, may both smoke more and suffer more from cancer than persons who are not nervous.

A variable that is associated with the problem **and** with a possible cause of the problem is a potential **CONFOUNDING VARIABLE**.

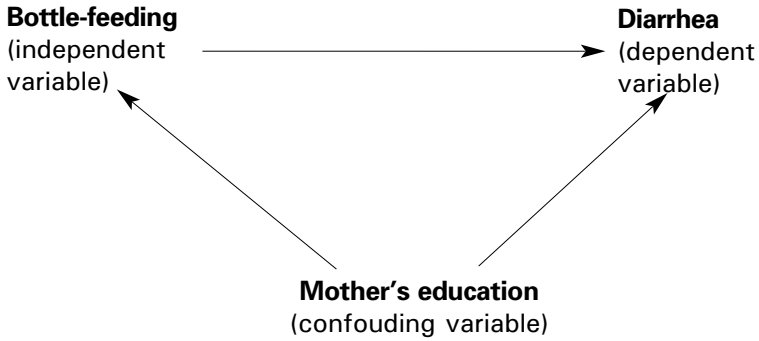
A confounding variable may either strengthen or weaken the apparent relationship between the problem and a possible cause.



Therefore, in order to give a true picture of cause and effect, possible confounding variables must be considered, either at planning stage or while doing data analysis.

For example:

A relationship is shown between bottle-feeding and diarrhea in under-twos. However, mother's education may be related to bottle-feeding as well as to diarrhea.



Mother's education is therefore a potential confounding variable. In order to give a true picture of the relationship between bottle-feeding and diarrhea of under-twos, the influence of mother's education should be controlled. This could either be addressed in the research *design*, e.g., by selecting only mothers with a specific level of education, or it could be taken into account during the *analysis* of the findings by analysing the relation between bottle-feeding and diarrhea separately for mothers with different levels of education.

Background variables

In almost every study, **BACKGROUND VARIABLES**, such as age, sex, educational level, socio-economic status, marital status and religion, should be considered. These background variables are often related to a number of independent variables, so that they influence the problem indirectly (hence they are called background variables). Only background variables important to the study should be measured. Background variables are notorious 'confounders'.

Note 1:

If you do a purely **descriptive** study, for example an inventory of knowledge, attitudes and practices related to bilharzia (schistosomiasis) or AIDS, you do not need to differentiate between dependent and independent variables, as there are no causal relationships between variables. In this type of study you may simply concentrate on variables and give operational definitions, with indicators if needed, to measure knowledge, attitudes and practices (see **Module 4** figure 4.5).

Note 2:

In **evaluation studies**, however, it is particularly important that we prepare good operational definitions, because here we want to compare and measure results at the beginning of the project phase and in the middle or at the end.

According to the WHO definition of health as an outcome of the health system (see **Module 2**, Figure 2.4) we can, for example, measure the improvement in the *health* of a population by comparing the estimated life expectance at birth and time lived with a disability over the past ten years (provided the epidemiological and other environmental factors did not change). Increased *fairness* of the health system could be measured by the percentage out-of-pocket spending on health by the poor (living on 1 US$ or less a day) of the total health expenditure, comparing, say, the past ten years.

Responsiveness to patients' need for human treatment is more difficult to measure, but a number of indicators could be developed, using the concepts: *respect* for patients (not humiliating or demeaning them); *confidentiality* with regard to a patient's diagnosis and treatment, providing patients with *essential information*, so that they can participate in choices about their own health and treatment, and *client-orientedness* in the services offered (prompt attention, clean premises) (WHO 2000: 32). It is interesting that one can not only make comparisons within one country over time, but also between countries.

Note 3:

When you select the variables for your study, it is important to review your objectives, as well as your problem analysis diagram. When you review your objectives you may find that you need to consider some new factors not originally included in your problem analysis diagram. On the other hand, you may discover that your objectives are too vague and can be revised and clarified, now that you have identified your variables

You should continue to adjust your problem analysis diagram, variables and objectives until they are all in line with each other.

REFERENCES

Abramson JH (1990, 4th ed.) *Survey Methods in Community Medicine.* London: Churchill-Livingstone.
(In particular Chapters 9 and 10)

Moser CA, Kalton G (1979) *Survey Methods in Social Investigation*. Hants, UK: Gower Publishing Company: 220-224.

World Health Organization (2000) *The WORLD HEALTH REPORT 2000. Health Systems: Improving Performance.* Geneva: WHO.

EXERCISE 3: Identification of variables in research

(to be carried out in plenary, ½ hour)

Look at the following descriptions of research problems and then answer the questions that follow.

Problem 1

A health researcher believes that in a certain region anaemia, malaria and malnutrition are serious problems among adult males and, in particular, among farmers. He therefore wishes to study the prevalence of these diseases among adult males of various ages, family size, occupations and educational backgrounds in order to determine how serious a problem these diseases are for this population.

Questions:

- What are the dependent and independent variables in the study?
- Which of these are categorical (ordinal and nominal) and which are numerical (continuous and discrete) variables?

Problem 2

A district medical officer (DMO) receives a complaint from the community that village health workers (VHWs) often run out of chloroquine. In preliminary investigations this shortage of chloroquine is confirmed. VHWs get their drugs at monthly meetings at the health centre. The DMO decides to investigate why the supply of drugs to VHWs is unsatisfactory.

Questions:

- What is the dependent variable in the study
- What would be a meaningful indicator for the dependent variable?
- How would you define 'short of chloroquine'?
- Can you think of some independent variables?
- Which independent variables are 'measurable' as they are and which ones need indicators?

Problem 3

Occasionally, a research project is carried out without considering some of the important variables. This may result in deceptive findings or an unclear relationship between independent and dependent variables.

In a study concerning prevalence of bilharzia (schistosomiasis) in the adult population of a village community, a researcher found that being a farmer was a risk factor for developing bilharzia. He was however not convinced that it was being a farmer that made these people more likely to develop bilharzia.

Question:

Are there any variables whose inclusion in the study might ensure that the researcher could show how much being a farmer actually contributed to a person developing schistosomiasis? Are there farmers who did *not* get bilharzia? Which variables might help explain why some farmers got bilharzia and others did not?

GROUP WORK (2½ hours)

1. Using the diagram of factors that possibly influence the problem you are studying (the diagram that you prepared for the statement of the problem), identify for each of these factors the variables that will be included in your study:

   - What is/are your dependent variable(s)? (List them.)
   - What are your independent variables? (List them.)
   - Which of the variables can be 'measured' as they are?

   } NB: This distinction cannot be made in a purely descriptive study.

   - Choose appropriate indicators for the variables that are not measurable as they are and/or formulate appropriate definitions for these variables/indicators.
   - State whether you have identified themes that need further exploration during your study before you can define the concepts adequately.

Use the table below for your work.

	Variable	Indicators (if needed) or further definition	Further exploration needed?	Which objective covered?
Dependent variable(s)	1. 2.			
Independent variables	1. 2. 3. 4. 5. etc.			

2. In the table we have included a column to state **which objective** is covered by your variables. You may discover that some objectives are not well covered by your variables (probably because your analysis diagram and objectives are not yet completely in line with each other). In that case, you need to rethink whether the objectives are indeed important for your study, and, if so, develop variables to measure them. You may discover that your objectives are too vague when compared to the type of data (or to the variables) you would like to collect. If so, you should make your objectives more specific.

   Before you finish you should review, as a group, your problem analysis diagram, objectives and variables, and make any adjustments needed so they are all in line with each other.

Annex 8.1: Example of a framework for defining variables

Conceptual definition of variable	**Operational definition i.e., indicator**	**Scale of measurement**
Age	Age at last birthday	Continuous: in months
Family size	Number of family members	Discrete
Use of clinic	Number of visits to clinic	Discrete
Haemoglobin	Haemoglobin concentration in capillary blood, measured by haemoglobinometer	Continuous: e.g., grams per 100 ml., rounded off to nearest gram
Nutritional status	Weight in relation to age compared to a standard growth curve	Ordinal: e.g., 1. well nourished = >80% of standard 2. moderately malnourished = 60% to 80% of standard 3. severely malnourished = <60% of standard
Patient's satisfaction	Response to a specific question about his/her satisfaction with services obtained, put to patients on discharge	Ordinal: e.g., 1. very satisfied 2. somewhat satisfied 3. somewhat dissatisfied 4. very dissatisfied
Immunisation coverage	Percentage of children immunised in a particular age group	Continuous: e.g., percentages; or ordinal, e.g., • high > 80% • medium 60% - 80% • low < 60%
Religion	As reported by informants	Nominal: Christian, Moslem, Hindu, Buddhist, etc.
Main source of carbohydrate in the diet	Main type of staple food eaten	Nominal: e.g., maize, millet, rice, cassava, etc.

Annex 8.2: Examples of variables with different options for indicators*

Occupation	• Occupation for which subject was trained (profession or trade), or work actually performed? If retired or unemployed, will previous occupation be used? Will women be classified by their own or by their husbands' occupation or both?
Education	• Number of years of education, or last grade attained, or type of educational institution last attended?
Income	• Personal income, family income, or average family income per member?
Crowding	• (Mean number of persons per room in housing unit) Which rooms are excluded from index (bathrooms, showers, toilets, kitchens, storerooms, rooms used for business purposes, entrance halls)?
Social status	• Based on occupation, education, crowding index, income, neighbourhood or residence, home amenities, or subject's self perception? Based on one of these, or a combination?
Marital status	• Expressed in terms of legal status (single, married, widowed, divorced), or in terms of stability (e.g., stable union, casual union)?
Parity	• Total number of previous pregnancies, or total number of children delivered?
Date of onset of	• Date when first symptoms were noticed, date when first diagnosed, or disease date of notification?
Presence of chronic disease	• Based on duration since onset? If so, what duration makes it chronic: 3 months, 6 months, a year? Or is 'chronic' defined based on the presence of certain diseases? Are some diseases defined as chronic whatever their duration? If so, what diseases? What about conditions that come and go (e.g., recurrent sore throats)?
Hospitalisation	• Is hospitalisation for childbirth included or not? Is the hospital stay of a well newborn baby included? Is overnight stay essential? Is overnight stay in a casualty or emergency ward included?

* Adapted from Abramson (1990)

Trainer's Notes

Module 8: VARIABLES

Timing and teaching methods

1 hour	Introduction and discussion (including first exercise)
½ hour	Exercise: Identification of variables in research (and discussion of answers)
2 hours	Group work
1 hour	Plenary
4½ hours	TOTAL TIME

Introduction and discussion

- Stress that it is important to define the problem as well as the factors influencing the problem in measurable terms.
- Let participants give some examples of numerical variables and discuss what different values these variables may have.
- Let participants give examples of categorical variables, after you have provided one or two examples. Make sure they understand that, once you have clear categories, you can 'measure' those variables, which means that you can determine their different values.
- Make sure that the participants understand that certain variables can be 'measured' directly and that others need indicators before they can be measured.

  Note: We use quote marks to indicate that 'measuring' of categorical variables such as sex or mode of transport means 'determining their values'.
- Discuss the relationship of the concept of dependent and independent variables to causality and stress that descriptive studies (see **Module 4**) do not have dependent and independent variables.

Exercise: Examples of dependent and independent variables

- Let the groups give examples from their own studies.
- Explain the difference between association and cause.
- Explain clearly that dependent variables factors with values such as low, medium, and high or sick and well need operational definitions to explain just what these values mean.
- Stress that sometimes measuring variables is not our concern, but rather identifying and describing them (if we know very little about possible causes of a problem).
- Stress the fact that when participants are working to list variables they have identified in their analysis diagram, they should also go back to their objectives to ensure that each objective is adequately covered. Since certain variables may need to be measured for several objectives it would be more complicated to start identifying variables by looking at the objectives rather than the analysis diagram.

Exercise: Identification of variables in research

- Conduct the exercise on 'Identification of variables in research' during the plenary session. Ask the participants to read and respond to the questions posed for each problem in the exercise individually or in small groups of two or three people. Give 4-5 minutes for each of the three problems, immediately followed by a group discussion. (Suggested answers are on the following 2 pages.)

Group work

- Ask the participants to meet in their working groups to select the variables that will be involved in the study being designed.

- Each group should then prepare a list of the selected variables for presentation and discussion in plenary and for inclusion in the methodology section of its research proposal. The groups should also indicate which variables would have to be further defined in the field.

ANSWER SHEET FOR EXERCISE 3: Identification of variables in research

(The following answers are by no means exhaustive)

Problem 1:

Dependent variables:

• % of males with malaria	- numerical	(continuous)
• % of males with anaemia or haemoglobin level in the blood	- numerical	(continuous)
• % of males moderately or severely malnourished	- numerical	(continuous)

Independent variables:

• age	- numerical	(continuous, but often expressed in number of months or years (discrete); if expressed in categories, it is ordinal)
• economic status	- categorical	(ordinal if expressed as high, moderate, low)
• educational background:		
- years of schooling	- numerical	(discrete)
- type of school	- categorical	(nominal)

Problem 2:

Dependent variable:

- Availability of chloroquine for village health workers

Indicator for availability of chloroquine:

'Short of chloroquine' should be defined in relation to the time since the date of the last drug supply and ideally also in relation to the size of the population.

For example:

If the number of tablets in stock is measured for all VHWs two weeks after the date of the last meeting at the health centre where drugs were supplied, one could say that any VHW who does not have enough drugs to treat 1% of the population for malaria is short of drugs. Since an adult needs 10 tablets for a full course, this would mean that a VHW should have at least 50 tablets available, if the village has a population of 500.

An alternative definition could be having **no** tablets in stock two weeks after the last date of supply.

Independent variables:

- Availability of drugs at the Health Centre (influenced by frequency of ordering and frequency of supply)
- Amount of drugs monthly supplied to VHWs
- Number of weeks since the VHW last came in for his supply of chloroquine
- Number of patients treated since date of last supply

Problem 3:

Important **independent variables** that could be taken into account include:

- Age
- Location in the village
- Contact with water
- Type of farming activities
- Division of labour
- Season

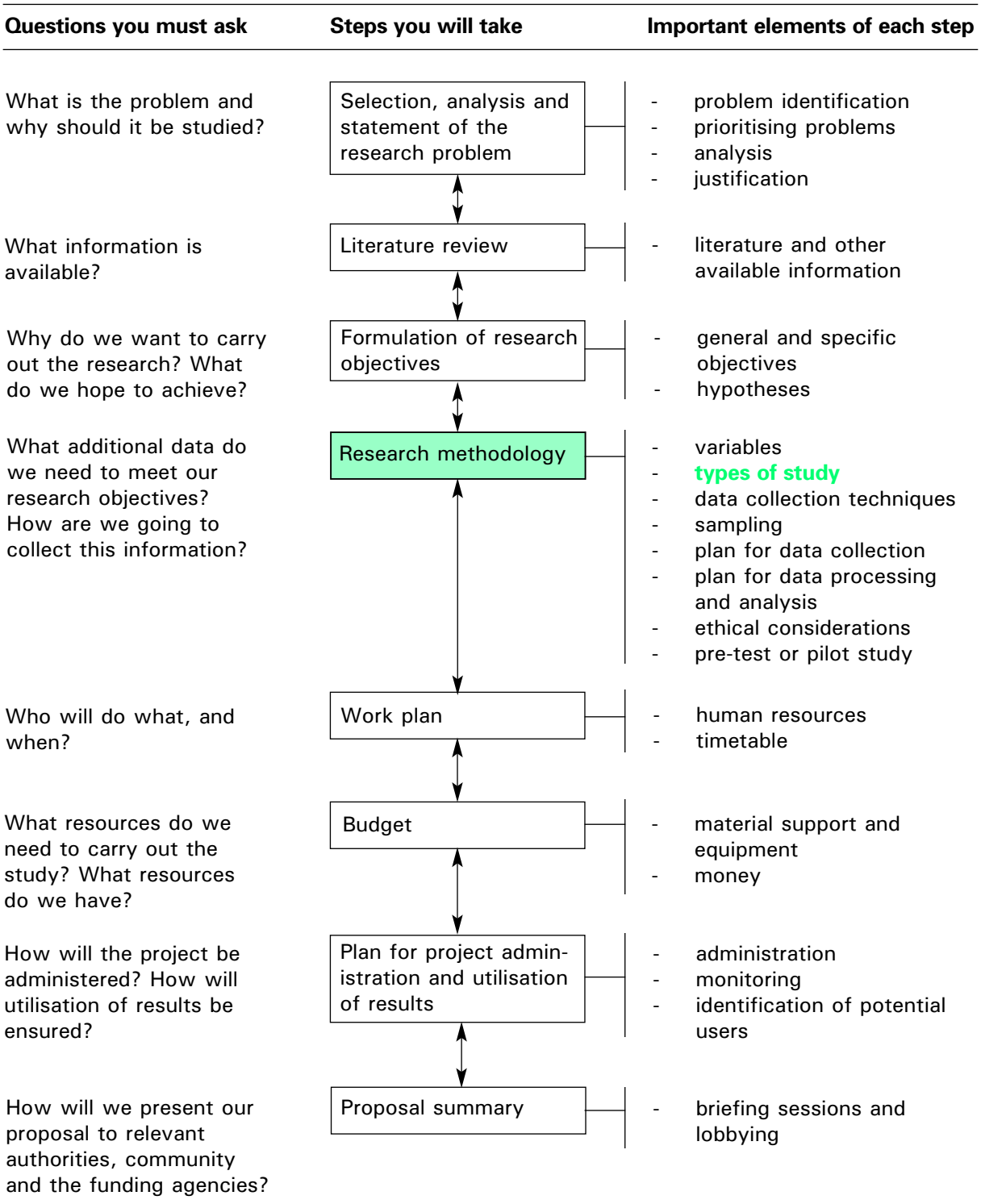
Closer study revealed that schistosomiasis was present among 70% of the young farmers between 20-25 years of age, while it was almost entirely absent in farmers older than 50 years of age. It turned out that younger farmers tended to have farms much further away from the village where the land was more fertile, and they had to cross a river where they bathed on their way home in the evening. The older farmers, on the other hand, had always had their farms close to the village and obtained water from wells.

Designing and Conducting Health Systems Research Projects
Part I: Proposal Development and Fieldwork

Module 9

STUDY TYPES

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 9: STUDY TYPES

OBJECTIVES

At the end of the session, you should be able to:

1. **Describe** the study types most used in HSR.

2. **Define** the uses and limitations of each study type.

3. **Describe** how the study design can influence the validity and reliability of the study results.

4. **Identify** the most appropriate study design for the research proposal you are developing.

I. Introduction

II. Overview of study types

III. Deriving valid and reliable conclusions

I. INTRODUCTION

Depending on the existing state of knowledge about a problem that is being studied, different types of questions may be asked which require different study designs. Some examples are given in the following table:

Table 9.1: Research questions and study types

STATE OF KNOWLEDGE OF THE PROBLEM	TYPE OF RESEARCH QUESTIONS	TYPE OF STUDY
Knowing that a problem exists but knowing little about its characteristics or possible causes.	• What is the nature/magnitude of the problem? • Who is affected? How do the affected people behave? What do they know, believe, think about the problem and its causes?	*Exploratory studies*, or *Descriptive studies:* • Descriptive case studies • Cross-sectional surveys
Suspecting that certain factors contribute to the problem.	• Are certain factors indeed associated with the problem? (e.g., Is lack of pre-school education related to low school performance? Is low fibre diet related to carcinoma of the large intestine?)	*Analytical (comparative) studies:* • Cross-sectional comparative studies • Case-control studies • Cohort studies
Having established that certain factors are associated with the problem: desiring to establish the extent to which a particular factor causes or contributes to the problem.	• What is the cause of the problem? • Will the removal of a particular factor prevent or reduce the problem? (e.g., stopping smoking, providing safe water)	• Cohort studies *Experimental or quasi-experimental studies*
Having sufficient knowledge about cause(s) to develop and assess an intervention that would prevent, control or solve the problem.	• What is the effect of a particular intervention/strategy? (e.g., treating with a particular drug; being exposed to a certain type of health education) • Which of two alternate strategies gives better results? Which strategy is most cost-effective?	Experimental or quasi-experimental studies

The type of study chosen depends on:

- the type of problem;
- the knowledge already available about the problem; and
- the resources available for the study.

When investigating health management problems, such as overcrowding in a hospital out-patient department or shortage of drugs at PHC level, a good description of the problem and identification of major contributing factors often provides enough information to take action.

When exploring more complicated management or health problems, we usually want to go further and determine the extent to which one or several independent variables contribute to the problem (for example, the contribution of low-fibre diet to cancer of the large intestine). For these types of problems more rigorous analytical or experimental studies will have to be conducted before we decide on appropriate interventions.

II. OVERVIEW OF STUDY TYPES

Several classifications of study types are possible, depending on what research strategies are used. The table below categorises studies, based on the combination of research strategies they use, including:

I. **Non-intervention studies** in which the researcher just observes and analyses researchable objects or situations but does not intervene; and

II. **Intervention studies** in which the researcher manipulates objects or situations and measures the outcome of his manipulations (e.g., by implementing intensive health education and measuring the improvement in immunisation rates.)

NON-INTERVENTION STUDIES

We will first concentrate on non-intervention studies and their use in health systems research. We will discuss:

- Exploratory studies
- Descriptive studies
- Comparative (analytical) studies

1. Exploratory studies

An EXPLORATORY STUDY is a small-scale study of relatively short duration, which is carried out when little is known about a situation or a problem*. It may include description as well as comparison.*

For example:

A national Acquired Immunodeficiency Syndrome (AIDS) Control Programme wishes to establish counselling services for Human Immunodeficiency Virus (HIV) positive and AIDS patients, but lacks information on specific needs patients have for support. To explore these needs, a number of in-depth interviews are held with various categories of patients (males, females, married, single) and with some counsellors working on a programme that is already under way.

When doing exploratory studies we *describe* the needs of various categories of patients and the possibilities for action. We may want to go further and try to explain the differences we observe (e.g., in the needs of male and female AIDS patients) or to identify causes of problems. Then we will need to *compare* groups.

Note:

Comparison is a fundamental research strategy to identify variables that help explain why one group of persons or objects differs from another.

In HSR, **small-scale studies that compare extreme groups** are very useful for detecting management problems. We could, for example, compare:

- Two district health teams (DHT) that are functioning well and two that do not function satisfactorily, in order to detect the possible reasons for bottlenecks in the functioning of the district health teams;*
- One community with high and another with low participation in health activities, in order to identify factors that contribute to community participation;
- 20 mothers who delivered in a maternity and 20 who delivered at home, in order to identify possible reasons for the low percentage of supervised deliveries.

Exploratory studies gain in explanatory value if we **approach the problem from different angles at the same time**. This is called **triangulation**. In a study that is looking for causes of the low percentage of supervised deliveries, it may be very useful to include observations and interviews with health staff in the maternity centres that should serve the mothers in question and interviews with their supervisors, as well as with the mothers themselves. In this manner, **information from different independent sources can be cross-checked**.

For some management problems such a 'rapid appraisal' may provide sufficient information to take action. Otherwise, a larger, more rigorous comparative study will have to be developed to test differences between groups with respect to various independent variables.

Note:

If the problem and its contributing factors are not well defined (see **Module 8** group work) it is **always advisable** to do an **exploratory study** before embarking on a large-scale descriptive or comparative study.

* Such small-scale studies may be called **exploratory** case studies if they lead to plausible assumptions about the causes of the problem and **explanatory** case studies if they provide sufficient explanations to take action (Yin, 1984).

2. Descriptive studies

A DESCRIPTIVE STUDY involves describing the characteristics of a particular situation, event or case.

Descriptive studies can be carried out on a small or larger scale.

(1) Small scale, descriptive case studies

Descriptive case studies describe in-depth the characteristics of one or a limited number of 'cases'. A case may be, for example, a patient, a health centre, or a village. Such a study can provide quite useful insight into a problem. Case studies are common in social sciences, management sciences, and clinical medicine. For example, in clinical medicine the characteristics of a hitherto unrecognised illness may be documented as a case study. This is often the first step toward building up a clinical picture of that illness.

However, if one wishes to test whether the findings pertain to a larger population, a more extensive, cross-sectional survey has to be designed.

(2) Large scale, cross-sectional surveys

Cross-sectional surveys aim at describing and quantifying the distribution of certain variables in a study population at one point of time. They may cover, for example:

- *Physical characteristics* of people, materials or the environment, as in
  - prevalence surveys (of bilharzia, leprosy, HIV), or
  - evaluation of coverage (of immunisation, latrines, etc.),
- *Socio-economic characteristics* of people such as their age, education, marital status, number of children and income,
- The *behaviour* or practices of people and the *knowledge, attitudes, beliefs, opinions* which may help to explain that behaviour (KAP studies), or
- *Events* that occurred in the population.

Cross-sectional surveys cover a selected sample of the population. If a cross-sectional study covers the total population it is called a **census**.

A cross-sectional survey may be repeated in order to measure changes over time in the characteristics that were studied. The surveys may be very **large**, with hundreds or even thousands of study units. In these cases only a **limited number of variables** will usually be included, in order to avoid problems with analysis and report writing. If cross-sectional surveys are **smaller** they can be **more complex**. They may include all the elements just mentioned. Small surveys can reveal interesting associations between certain variables, such as between having tuberculosis and socio-economic status, sex, and ways of coping.

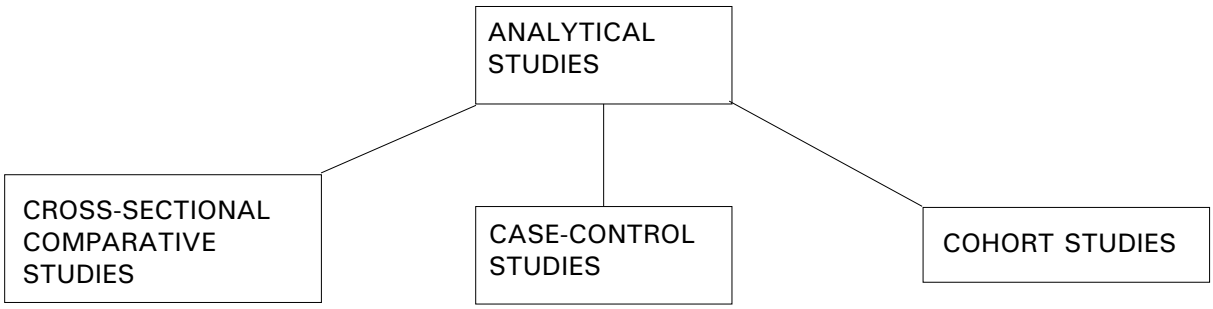
Researchers often go further and will combine a *description* of the study population with a *comparison* of a number of groups within that population (see below). Such combinations are very common, and thus the distinctions between descriptive and comparative studies are sometimes quite fuzzy.

3. Comparative or analytical studies

An ANALYTICAL STUDY attempts to establish **causes or risk factors** for certain problems. This is done by comparing two or more groups, some of which have or develop the problem and some of which have not.

Three commonly used types of analytical studies will be discussed here:

Figure 9.1: Types of analytical studies



(1) Cross-sectional comparative studies

Many cross-sectional surveys focus on describing **as well as** comparing groups.

For example, a survey on malnutrition may wish to establish:

- The percentage of malnourished children in a certain population;
- Socio-economic, physical, political variables that influence the availability of food;
- Feeding practices; and
- The knowledge, beliefs, opinions that influence these practices.

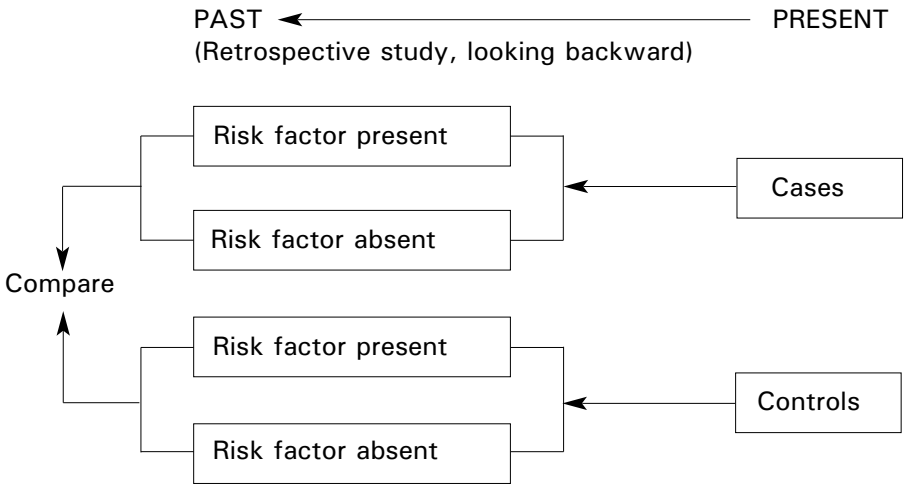
The researcher will not only describe these variables but, by comparing malnourished and well-nourished children, he will try to determine which socio-economic, behavioural and other independent variables may have contributed to malnutrition.

In any comparative study, one has to watch out for **CONFOUNDING** or **INTERVENING** variables. (Please look at **Module 8** for examples and discussion, as well as to the next page, and **Module 26**).

(2) Case-control studies

In a CASE-CONTROL STUDY the investigator compares one group among whom the problem that he wishes to investigate is present (e.g. malnutrition) and another group called a control or comparison group, where the problem is absent, in order to find out what factors have contributed to the problem.

Figure 9.2: Diagram of a case-control study*



For example, in a study of the causes of neonatal death, the investigator will first select the 'cases' (children who died within the first month of life) and 'controls' (children who survived their first month of life). (S)he then interviews their mothers to compare the history of these two groups of children, to determine whether certain risk factors are more prevalent among the children who died than among those who survived.

Note:

Controls should come from the same **'source'** population. For example, in a hospital case-control study where cases are being sought in the hospital, cases should normally be selected from patients attending at the same hospital. If controls are selected from another hospital, they might not be from the same source population because the referral pathways may be different, and therefore they would not really be comparable to the cases.

As with a cross-sectional comparative study, the researcher has to control for CONFOUNDING VARIABLES. In case-control studies, this may be done to some extent beforehand, by MATCHING the groups for expected confounding variables. Matching means taking care that the cases and controls are similar with respect to the distribution of one or more potentially confounding variables. However, we cannot then look at the effect of the matched variable as a risk factor because we have made the cases and controls exactly the same with respect to that variable.

For example, in a study on causes of malnutrition in children-3 years you may match the well- and the malnourished on age, because this factor may influence many other variables influencing your problem (e.g. time of weaning, time of teething, which are both related to diarrhoea and consequently to malnutrition). But you will *not* match them on economic status of parents, as you do want to know whether poverty influences malnutrition.

In general you will only match for strong confounders (such as age) that you cannot properly control by stratification during data analysis, unless you double or treble the number of informants. You also match for potentially confounding variables such as location/source of origin, as these can influence many other potential confounders (e.g. ethnic group, religion, economic status), some of which you may not even expect beforehand.

* Adapted from WW Holland et al., eds. (1985) *Oxford Textbook of Public Health, Volume 3: Investigative Methods in Public Health*. Oxford: Oxford University Press.

(3) Cohort studies

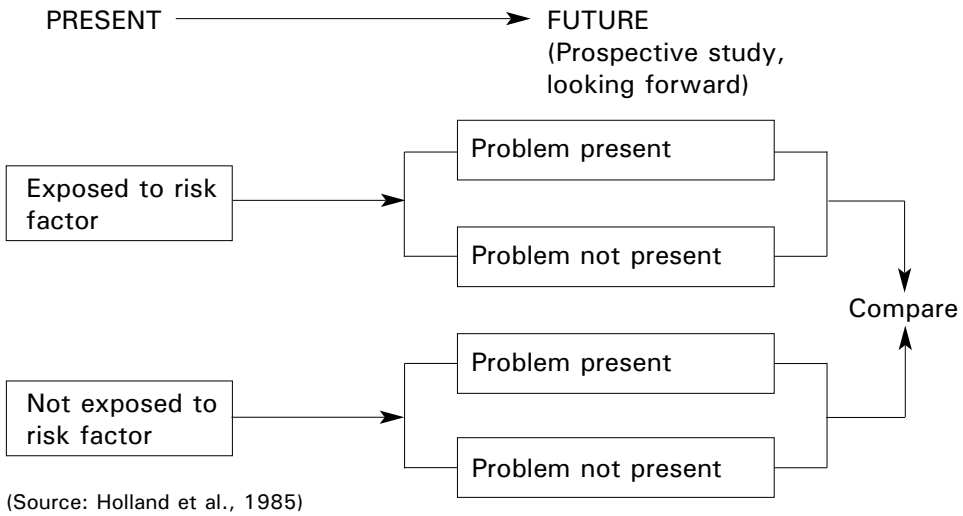
In a COHORT STUDY, a group of individuals that is exposed to a risk factor (study group) is compared to a group of individuals not exposed to the risk factor (control group). The researcher follows both groups over time and compares the occurrence of the problem that he expects to be related to the risk factor in the two groups to determine whether a greater proportion of those with the risk factor are indeed affected.

A well-known **example** of a cohort study is the study by Doll and Hill (1950) of smokers and non-smokers that was conducted among doctors to determine the importance of smoking as a risk factor for developing lung-cancer.

A study may start with one large cohort. After the study starts, the researchers determine who is exposed to the risk factor (e.g., smoking) and who is not, and follow the two groups over time to determine whether the study group (of smokers) develops a higher prevalence of lung cancer than the control group. If it is not advisable to select one cohort (for example, because only few people are affected by the risk factor, which necessitates a very large sample), two cohorts may be chosen, one in which the risk factor is present (study group) and one in which it is absent (control group). In all other respects the two groups should be as alike as possible.

The control group should be selected at the same time as the study group, and both should be followed with the same intensity.

Figure 9.3: Diagram of a cohort study



(Source: Holland et al., 1985)

Uses and limitations of different types of analytical studies

You may use any of the three types of analytical studies (cross-sectional comparison, case-control or cohort) to investigate possible causes of a problem.

For example, if you assume there is a causal relationship between the use of a certain water source and the incidence of diarrhea among children under five in a village with different water sources:

- You can select a group of children under five years and check at regular intervals (e.g., every two weeks) whether the children have had diarrhoea and how serious it was. Children using the suspected water source and those using other sources of water supply will be compared with regard to the incidence of diarrhea (**cohort study**).

- You can also conduct **a case-control study**. For example, you may compare children who present themselves at a health centre with diarrhea (cases) during a particular period of time with children presenting themselves with other complaints of roughly the same severity, for example acute respiratory infections (controls) during the same time, and determine which source of drinking water they had used.

- In a **cross-sectional comparative study**, you could interview mothers to determine how often their children have had diarrhea during, for example, the past month, obtain information on their source of drinking water, and compare the source of drinking water of children who did and did not have diarrhoea.

Cross-sectional comparative studies and case-control studies are usually preferred to cohort studies for financial and practical reasons. However, cohort studies are stronger in establishing causal relationships because confounding variables are to a large extent eliminated. If the study is well designed, the 'confounders' are equally distributed among the cases and controls. Experimental studies have the same advantage as cohort studies.

Cross-sectional comparative studies and **case-control studies** are relatively quick and inexpensive to undertake. With cross-sectional comparative studies, however, the number of stratifications one can make is limited by the size of the study. The problem with case-control studies is sometimes the difficulty of making a precise selection of a control group which is comparable to the study group on one or two specific variables (e.g., well- and malnourished children of the same sex and age, in months).

Cohort studies are a relatively sure way to establish causal relationships. However, they take longer than case-control studies and are **labour intensive**, and therefore **expensive**. The major problems are usually related to the identification of all cases in a study population, especially if the problem has a low incidence. Further, the following up all persons included in the study over a number of years may be impossible because of population movement.

INTERVENTION STUDIES

In intervention studies, the researcher manipulates a situation and measures the effects of this manipulation. Usually (but not always) two groups are compared, one group in which the intervention takes place (e.g. treatment with a certain drug) and another group that remains 'untouched' (e.g. treatment with a placebo).

The two categories of intervention studies are:

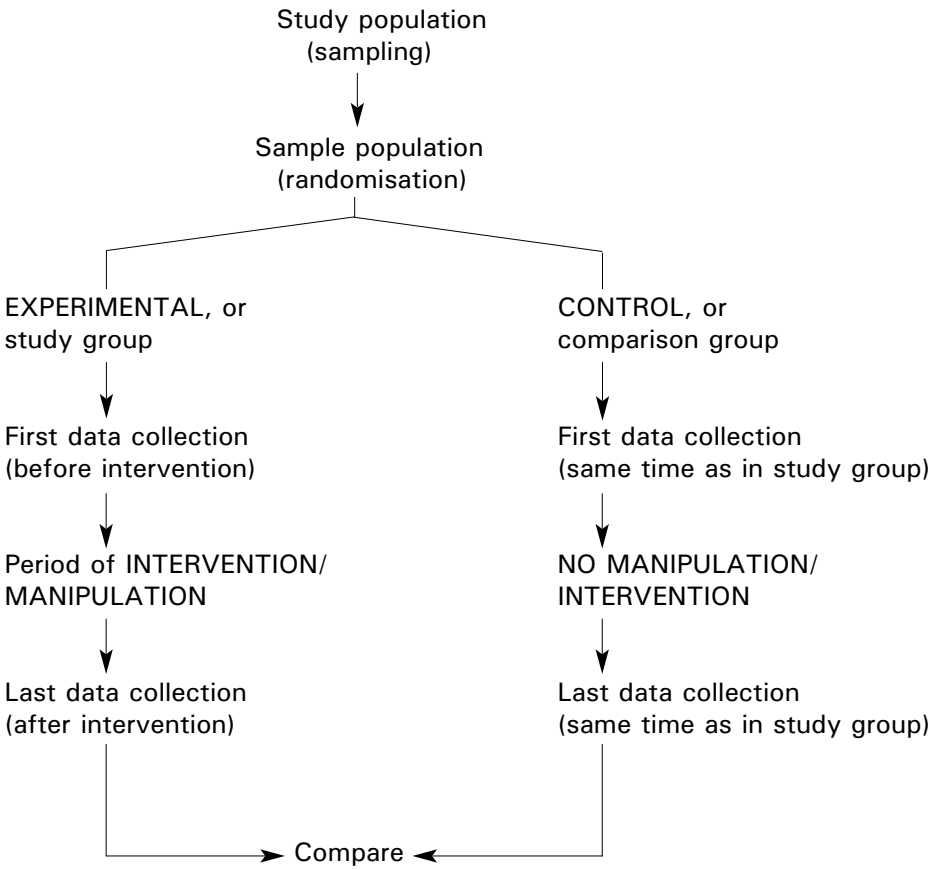
- experimental studies and
- quasi-experimental studies.

1. Experimental studies

An experimental design is a study design that gives the most reliable **proof for causation**.

In an EXPERIMENTAL STUDY**,** individuals are randomly allocated to at least two groups. One group is subject to an intervention, or experiment, while the other group(s) is not. The outcome of the intervention (effect of the intervention on the dependent variable/problem) is obtained by comparing the two groups.

Figure 9.4: Diagram of an experimental study



Note:

The strength of experimental studies is that by randomisation the researcher eliminates the effect of confounding variables through the equal distribution of confounders (both known and unknown) in the experimental and control groups.

A number of experimental study designs have been developed. These are widely used in laboratory settings and in clinical settings. For ethical reasons, the opportunities for experiments involving human subjects are restricted. However, randomised control trials of new drugs are common.

> **For example**, a researcher plans to study the effect of a new drug. (The drug has already been tested extensively on animals and has been approved for trial use.) He plans to include 300 patients in the study who are currently receiving the standard treatment for the same condition for which the new drug has been designed. He explains the study to the patients asking their consent to be divided into two groups on a random basis. One group will receive the experimental drug while the other group will continue to receive the standard treatment. He makes sure that the medications are disguised and labelled in such a manner that neither the research assistant administering them nor the patient know which drug is used. (This is called a 'double blind' experiment.)

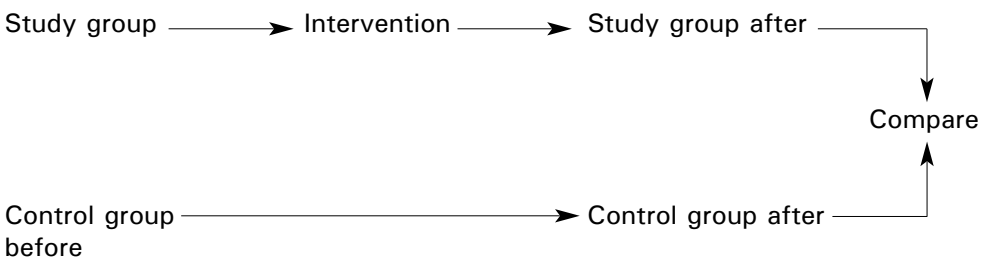
At community level, where HSR is frequently undertaken, we experience not only ethical but also practical problems in carrying out experimental studies. In real life settings, it is often impossible to assign persons at random to two groups, or to maintain a control group. Therefore, experimental research designs may have to be replaced by quasi-experimental designs.

2. Quasi-experimental studies*

In a QUASI-EXPERIMENTAL STUDY, one characteristic of a true experiment is missing, either randomisation or the use of a separate control group. A quasi-experimental study, however, always includes the manipulation of an independent variable which is the intervention.

One of the most common quasi-experimental designs uses two (or more) groups, one of which serves as a control group in which no intervention takes place. Both groups are observed before as well as after the intervention, to test if the intervention has made any difference. (This quasi-experimental design is called the 'non-equivalent control group design' because the subjects in the two groups (study and control groups) have not been randomly assigned.)

Figure 9.5: Diagram of a quasi-experimental design with two groups



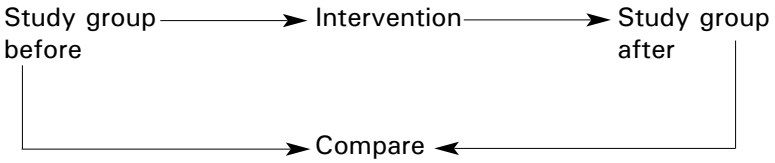
* For a more detailed explanation of experimental and quasi-experimental designs and their advantages and disadvantages, one excellent reference is Campbell and Stanley's *Experimental and Quasi-Experimental Designs for Research (1963).*

Example of a quasi-experimental study:

A researcher plans to study the effects of health education on the level of participation of a village population in an immunisation campaign. She decides to select one village in which health education sessions on immunisation will be given and another village which will not receive health education and serves as a control. The immunisation campaign will be carried out in the same manner in both villages. A survey will then be undertaken to determine if the immunisation coverage in the village where health education was introduced before the campaign is significantly different from the coverage in the 'control' village which did not receive health education. (Note: The study is quasi-experimental because the subjects were not assigned to the control or experimental groups on a random basis).

Another type of design that is often chosen because it is quite easy to set up uses only **one group** in which an intervention is carried out. The situation is analysed before and after the intervention to test if there is any difference in the observed problem. This is called a 'BEFORE-AFTER' study. This design is considered a 'pre-experimental' design rather than a 'quasi-experimental' design because it involves neither randomisation nor the use of a control group.

Figure 9.6: Diagram of a before-after study



Example of a 'before-after', pre-experimental study:

The out-patient clinic of hospital X is extremely crowded. Waiting times of over 5 hours for patients before they are attended to are not uncommon. The hospital management has a study carried out to analyse the bottlenecks and implements most of the recommendations made. Three months later, another study is done to check to what extent the bottlenecks have been solved and where further action is necessary.

This design is often used for management problems that pertain to one single unit (hospital, school, village). However, if the problems occur at a larger scale or if they might be influenced by other factors **apart from the intervention** during the trial, it is highly recommended that the design include both a study *and* a control group.

In the trial with health education on immunisation, **for example,** it would have been quite risky to work without a control group. Outside events (such as a health education campaign on immunisation by radio or other mass media) might have led to improved knowledge on immunisation in both the study group and the control group. (NB: The immunisation campaign by radio provides a so-called *'rival explanation'* for your results.) If you had had just a study group and no control you might have concluded erroneously that all of the increase was due to your own intervention.

III. DERIVING VALID AND RELIABLE CONCLUSIONS

Whatever research design is selected, a primary concern is that the conclusions of the study be **VALID** and **RELIABLE.**

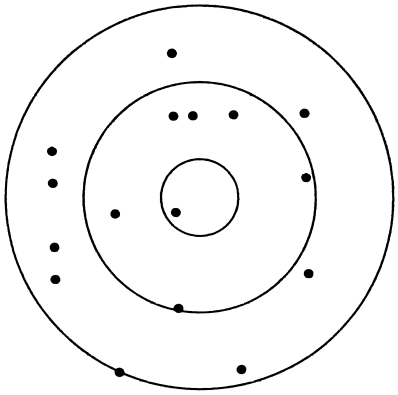
What are validity and reliability in research findings?

Validity means that your scientific observations actually measure what they intend to measure (your conclusions are true).

Reliability means that someone else using the same method in the same circumstances should be able to obtain the same findings (your findings are repeatable).

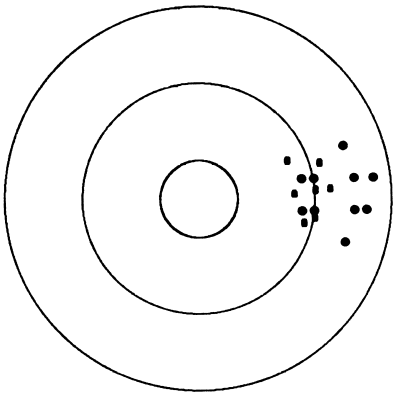
Reliability (repeatability) refers to the possibility to *replicate* (repeat) the observations and is related to the precision of the instrument used for scientific observations. Validity refers to the *soundness* of the observations and to the accurateness of the data collected by the research method/instrument.

Figure 9.7: Validity and reliability; graphic presentation of possible combinations



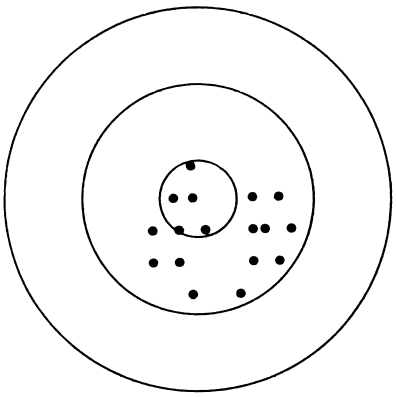
Neither valid nor reliable

The research methods do not hit the heart of the research aim (not 'valid') and repeated attempts are unfocussed



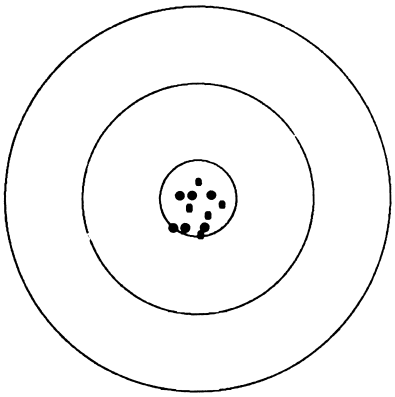
Reliable but not valid

The research methods do not hit the heart of the research aim, but repeated attempts get almost the same (but wrong) results



Fairly valid but not very reliable

The research methods hit the aim of the study fairly closely, but repeated attempts have very scattered results (not reliable)



Valid and reliable

The research methods hit the heart of the research aim, and repeated attempts all hit in the heart (similar results)

For example:

Four different teams of researchers set out to determine the body weights of three children whose true body weights were 10 kg, 15 kg and 20 kg respectively and obtained following four sets of results.

Team 1

Child	True body weight	First set of results	Neither
A	10 kg	8 kg	valid
B	15 kg	18 kg	nor
C	20 kg	19 kg	reliable

The first set of results is not valid because the results do not represent the true body weights. They are not reliable because they are sometimes too high and sometimes too low, and the relative difference from the true body weight varies from child to child.

Team 2

Child	True body weight	Second set of results	Reliable
A	10 kg	11 kg	but
B	15 kg	16.5 kg	not
C	20 kg	22 kg	valid

The second set of results is not valid because the results again do not represent the true body weights. However, they are reliable because the results are too high by the same proportion (10%) for every child.

Team 3

Child	True body weight	Third set of results	Fairly
A	10 kg	10.15 kg	valid but
B	15 kg	14.85 kg	not
C	20 kg	20.33 kg	reliable

The third set of results is fairly valid because the results are almost representing the true body weight. They are not reliable because two weights are too high and one is too low and the proportion by which they differ from the true body weight is different for each child.

Team 4

Child	True body weight	First set of results	
A	10 kg	10 kg	Valid
B	15 kg	15 kg	and
C	20 kg	20 kg	reliable

The fourth set of results is both valid and reliable because the results are the same as the true body weights, and these results have been obtained for every child.

Note: that it is possible to implement a research instrument with precision and yet obtain invalid responses! For example, a door-to-door survey on sexual behaviour of informants may give the same type of answers throughout and therefore seem reliable. But the chance that people are concealing their true sexual behaviour is high, so that the validity may be low.

How to deal with threats to validity and reliability

At various stages of the research validity and reliability could be threatened:

- At the moment of the **selection** of the study type and design. You should not start with a big survey when your knowledge of the situation and problem(s) is superficial, but always first do an exploratory study. Otherwise validity and reliability will be limited. Distortion may also occur during sampling or due to selectivity in assigning different subjects into various groups. (See **Modules 9** and **11** part III.)

- At the level of the **data collection** (related to the instrument): the instrument itself may be unreliable; bias (distortion) may occur at various stages of data collection. (See **Modules 10A** part III, **10B** part V.)

- At the level of the **analysis** of the data collected: confounding variables or events that disturbed your study design, and unnoticed weaknesses in study type and data collection may lead to misleading conclusions. (See **Modules 9**, **26**.)

Examples of threats to validity:

1. Confounding factors

Example:

You might find that children who have had pre-school education subsequently perform better in primary school. Can you conclude that pre-school education leads to better school performance?

Rival or alternate explanations include:

- Educational and income level of parents may be contributing to both pre-school education and school performance; and

- Educational and income level of parents may, through availability of education toys in the home, television etc., influence learning performance in pre-school as well as primary school.

Education and income are therefore **confounding factors**.

2. History

Unexpected factors beyond your control might have produced the same effect as the intervention you were studying, thereby making it impossible for you to know whether it was your intervention that produced the impact.

Example:

A well-known example is when a certain agency that had designed a health education program for early detection of breast cancer designed a study to test the effectiveness of the program by studying the increase in the proportion of women who reported doing self-examination of breasts. However, while the study was in progress the President's wife developed breast cancer and she appeared widely on mass media to advise women on early detection of breast cancer.

3. Differential subject loss in various groups

The type of subjects who drop out of your study or control groups may be related to some of the characteristics you are studying.

Example:

You are studying the effectiveness of a 'weight watchers' program by comparing the average weight loss in the 'weight watchers' group with that of a control group. However, a number of women in the 'weight watchers' group found the program too demanding and have dropped out.

4. Selectivity (or bias) in assigning subjects to various groups

Example:

You intend to study whether a programme on 'how to stop smoking' will be effective in helping the smokers in your hypertension clinic. Therefore you invite those who would like to attend to register themselves. You plan to compare the percentage who stop smoking among those who attend the programme with those who do not. However, it is likely that those who register themselves are those who are strongly motivated to stop smoking while those who are not motivated do not join the program (Also see **Module 11)**.

Strategies to deal with threats to validity

1. **Triangulation**. Approaching a research problem from different angles (e.g., by selecting complementary study populations or using different research techniques at the same time) (see **Modules 9** and **10**).

2. **Control group**. Observing a control group who is not exposed to the risk factor or intervention reduces threats due to unexpected and confounding factors.

3. **Appropriate sampling procedures and assignment of subjects to research groups.** This reduces threats due to selectivity (see **Module 11**).

4. **Before and after measurements**. This allows us to assess whether there has been selectivity as well as differential loss of subjects. If there has been an inevitable loss of subjects, it may enable assessment of the dropouts to determine whether they had peculiar characteristics that distinguished them from those who did not drop out (see this module).

5. **Unobtrusive methods** of data collection and allowing adaptation time for subjects to get used to being observed or interviewed (see **Module 10B**).

6. **Careful design and pre-testing** of instruments, stressing the participation of health managers, staff and community members, reduce bias due to instrumentation (see **Modules 10** and **14**). Training of interviewers and standardisation of interview techniques and tools such as questionnaires are also important in reducing this bias.

7. **Knowledge of the environment events** enables the researcher to be sensitive to external events that could affect validity (i.e., history). In case of an expatriate researcher, local key informants can contribute a lot to the validity of the study.

8. **Stratification and matching for confounding variables** during the analysis of the results (see **Module 26** on confounding).

Selection of study design

In selecting the design of the study, you have to consider the type of information you want to obtain and devise strategies to enable you to obtain that information.

The selection of an appropriate research design depends on:

- the state of knowledge about the problem
- the nature of the problem and its environment
- the resources available for the research
- the ingenuity and creativity of the researcher

REFERENCES

Abramson JH (1990, 4th ed.) *Survey Methods in Community Medicine.* London: Churchill-Livingstone.

Beaglehole R, Bonita R, Kjellström T (1993) *Basic Epidemiology*. Geneva: World Health Organization.

Campbell DT, Stanley JC (1963) *Experimental and Quasi-experimental Designs for Research*. Chicago: Rand McNally.

Doll R and Hill AB (1950) Smoking and carcinoma of the lung: Preliminary report. *Br. Med. J.* **2**:739-748.

Gordis L (1996) *Epidemiology.* Philadelphia USA: W.B. Saunders Company.

Holland W, Detels R, and Knox G (ed.) (1984-85) *Oxford Textbook of Public Health, Volumes 1-4.* Oxford: Oxford University Press.

Katzenellenbogen JM, Joubert G, Abdool Karim SS (1997) *Epidemiology; A manual for South Africa.* Capetown: Oxford University Press.

Kidder LH, Judd CM (1987) *Research Methods in Social Relations*.
Hong Kong: CBS Publishing Japan Ltd.

Moser CA, Kalton G (1989, 2nd ed.) *Survey Methods in Social Investigation*. Hants, UK: Gower Publishing Company.

Patton MQ (1990, 2nd ed.) *Qualitative Evaluation and Research Methods.* Newbury Park, USA: Sage Publications.

Rose J, Barber DJP (1989) *Epidemiologie for the uninitiated.* Plymouth GB: Latimer Trend & Co/ British Medical Association.

Yin RK (1984*) Case Study Research; Design and Methods*. London, UK: Sage Publications.

Vaughan JP, Morrow RH (1989) *Manual of Epidemiology for District Health Management*. Geneva: World Health Organization.

Trainer's Notes

Module 9: STUDY TYPE

Timing and teaching methods

1-1½ hour	Introduction and discussion
1-1½ hours	TOTAL TIME

Introduction and discussion

It is helpful if participants read this module the evening before the presentation, so they are more familiar with the topic.

The goal of the module is to give participants an understanding of the major issues involved in choosing different research strategies rather than having them learn the various possible study types by heart.

Present Table 9.1 at the beginning of the module to illustrate basic questions that lead to the choice of different study types **without getting into the details concerning each type.** Then proceed with the detailed discussion of each study type. Repeat the presentation of Table 9.1 at the end of the module, summarising the different study types that are possible.

It should be stressed that unless all variables to be investigated are clearly defined, small-scale studies are preferable to large-scale studies. A combination of study types (triangulation) can be considered if some variables still have to be explored (e.g., by open-ended questions) whereas other well-defined variables need to be measured on a larger scale (e.g., degree of utilisation of services).

Note:

Try to give examples of different study types in the fields in which the participants are interested. **Shorten** the presentation, especially of part III, if participants will not be engaged in analytical or quasi-experimental studies.

It is advisable to proceed with the presentation of **Module 10A** (Overview of data-collection techniques) before participants do their group work to choose the type(s) of study they will use for their research projects. After **Module 10A**, the participants can be asked to do an exercise which involves selecting a study type as well as data collection techniques for certain problems.

Note:

The group work session on selection of study types is combined with group work on selection of data collection techniques. It comes at the end of **Module 10A**.

Module 10

DATA COLLECTION TECHNIQUES

A: OVERVIEW OF DATA COLLECTION TECHNIQUES
B: DESIGN OF RESEARCH INSTRUMENTS;
INTERVIEW GUIDES AND INTERVIEW SKILLS
C: FOCUS GROUP DISCUSSIONS

Flowchart: Steps in the development of a health systems research proposal

Questions you must ask	Steps you will take	Important elements of each step
What is the problem and why should it be studied?	Selection, analysis and statement of the research problem	- problem identification - prioritising problems - analysis - justification
What information is available?	Literature review	- literature and other available information
Why do we want to carry out the research? What do we hope to achieve?	Formulation of research objectives	- general and specific objectives - hypotheses
What additional data do we need to meet our research objectives? How are we going to collect this information?	Research methodology	- variables - types of study - **data collection techniques** - sampling - plan for data collection - plan for data processing and analysis - ethical considerations - pre-test or pilot study
Who will do what, and when?	Work plan	- human resources - timetable
What resources do we need to carry out the study? What resources do we have?	Budget	- material support and equipment - money
How will the project be administered? How will utilisation of results be ensured?	Plan for project administration and utilisation of results	- administration - monitoring - identification of potential users
How will we present our proposal to relevant authorities, community and the funding agencies?	Proposal summary	- briefing sessions and lobbying

NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Introduction

Data collection is a crucial stage in the planning and implementation of a study. If the data collection has been superficial, biased or incomplete, data analysis becomes difficult, and the research report will be of poor quality.

Therefore, we should concentrate all possible efforts on developing appropriate tools, and should test them several times.

Depending on the type of study, different data-collection techniques may be used. In HSR studies we usually combine a number of different techniques and look at problems from different perspectives (**triangulation**).

To give full weight to data collection we have split the module into three parts:

- An overview of different data collection techniques
- Guidelines on how to develop interview guides or questionnaires, including guidelines on how to interview
- An introduction on how to conduct Focus Group Discussions with an exercise in FGD, which should be carried out in any protocol development workshop

All parts are highly relevant for young researchers who still need to gain experience in conducting research.

Designing and Conducting Health Systems Research Projects
Part 1: Proposal Development and Fieldwork

Module 10A

OVERVIEW OF DATA COLLECTION TECHNIQUES

Module 10A: OVERVIEW OF DATA COLLECTION TECHNIQUES

OBJECTIVES

At the end of this session you should be able to:

1. **Describe** various data collection techniques and state their uses and limitations.
2. Advantageously **use a combination** of different data collection techniques.
3. **Identify** various sources of bias in data collection and ways of preventing bias.
4. **Identify** ethical issues involved in the implementation of research and ways of ensuring that your research informants or subjects are not harmed by your study.

I. Overview of data collection techniques

II. The importance of combining different data collection techniques

III. Bias in information collection

IV. Ethical considerations

I. OVERVIEW OF DATA COLLECTION TECHNIQUES

Data-collection techniques allow us to **systematically** collect information about our objects of study (people, objects, phenomena) and about the settings in which they occur.

In the collection of data we have to be systematic. If data are collected haphazardly, it will be difficult to answer our research questions in a conclusive way.

Example:

During a nutrition survey three different weighing scales were used in three villages. The researchers did not record which scales were used in which village. After completion of the survey it was discovered that the scales were not standardised and indicated different weights when weighing the same child. It was therefore impossible to conclude in which village malnutrition was most prevalent.

Various data collection techniques can be used such as:

- Using available information
- Observing
- Interviewing (face-to-face)
- Administering written questionnaires
- Focus group discussions
- Projective techniques, mapping, scaling

1. Using available information

Usually there is a large amount of data that has already been collected by others, although it may not necessarily have been analysed or published. Locating these sources and retrieving the information is a good starting point in any data collection effort.

For example, analysis of the information routinely collected by health facilities can be very useful for identifying problems in certain interventions or in flows of drug supply, or for identifying increases in the incidence of certain diseases.

Analysis of health information system data, census data, unpublished reports and publications in archives and libraries or in offices at the various levels of health and health-related services, may be a study in itself. Usually, however, it forms part of a study in which other data collection techniques are also used.

The use of **key informants** is another important technique to gain access to available information. Key informants could be knowledgeable community leaders or health staff at various levels and one or two informative members of the target group (e.g., adolescents on their sexual behaviour). They can be involved in various stages of the research, from the statement of the problem to analysis of the data and development of recommendations. Other sources of available data are **newspapers** and published **case histories**, e.g., patients suffering from serious diseases, or their relatives, telling their experiences and how they cope.*

* For example, Noerine Kaleeba (1991) *We miss you all: AIDS in the family.* Harare: Women and AIDS Support Network.

Note:

In order to retrieve the data from available sources, the researcher will have to design an instrument such as a checklist or compilation sheet. In designing such instruments, it is important to inspect the layout of the source documents from which the data is to be extracted. For health information system (HIS) data, for example, the data compilation sheet should be designed in such a way that the items of data can be transferred in the order in which the items appear in the source document. This will save time and reduce error.

The advantage of using existing data is that collection is inexpensive. However, it is sometimes difficult to gain access to the records or reports required, and the data may not always be complete and precise enough, or too disorganised.

2. Observing

OBSERVATION is a technique that involves systematically selecting, watching and recording behaviour and characteristics of living beings, objects or phenomena.

Observation of human behaviour is a much-used data collection technique. It can be undertaken in different ways:

- **Participant observation:** The observer takes part in the situation he or she observes. (For example, a doctor hospitalised with a broken hip, who now observes hospital procedures 'from within'.)

- **Non-participant observation:** The observer watches the situation, openly or concealed, but does not participate.

Observations can be **open** (e.g., 'shadowing' a health worker with his/her permission during routine activities) or **concealed** (e.g., 'mystery clients' trying to obtain antibiotics without medical prescription). They may serve different purposes. Observations can give additional, more accurate information on behaviour of people than interviews or questionnaires. They can also check on the information collected through interviews especially on sensitive topics such as alcohol or drug use, or stigmatising diseases. For example, whether community members share drinks or food with patients suffering from feared diseases (leprosy, TB, AIDS) are essential observations in a study on stigma.

Observations of human behaviour can form part of any type of study, but as they are time consuming they are most often used in small-scale studies.

Observations can also be made on **objects**. For example, the presence or absence of a latrine and its state of cleanliness may be observed. Here observation would be the major research technique.

If observations are made using a defined scale they may be called **measurements**. Measurements usually require additional tools. For example, in nutritional surveillance we measure weight and height by using weighing scales and a measuring board. We use thermometers for measuring body temperature.

3. Interviewing

An INTERVIEW is a data-collection technique that involves oral questioning of respondents, either individually or as a group.

Answers to the questions posed during an interview can be recorded by writing them down (either during the interview itself or immediately after the interview) or by tape-recording the responses, or by a combination of both.

Interviews can be conducted with varying degrees of flexibility. The two extremes, high and low degree of flexibility, are described below:

- **High degree of flexibility:**

  For example:
  When studying sensitive issues such as teenage pregnancy and abortions, the investigator may use a list of topics rather than fixed questions. These may, e.g., include how teenagers started sexual intercourse, the responsibility girls and their partners take to prevent pregnancy (if at all), and the actions they take in the event of unwanted pregnancies. The investigator should have an additional list of topics ready when the respondent falls silent, (e.g., when asked about abortion methods used, who made the decision and who paid). The sequence of topics should be determined by the flow of discussion. It is often possible to come back to a topic discussed earlier in a later stage of the interview.

The unstructured or loosely structured method of asking questions can be used for interviewing individuals as well as groups of key informants. (For details concerning focus group discussions (FGDs), see **Module 10C**.)

A flexible method of interviewing is useful if a researcher has as yet little understanding of the problem or situation he is investigating, or if the topic is sensitive. It is frequently applied in exploratory studies. The instrument used may be called an **interview guide** or interview schedule.*

- **Low degree of flexibility:**

Less flexible methods of interviewing are useful when the researcher is relatively knowledgeable about expected answers or when the number of respondents being interviewed is relatively large. Then **questionnaires** may be used with a fixed list of questions in a standard sequence, which have mainly fixed or pre-categorised answers.

For example:
After a number of observations on the (hygienic) behaviour of women drawing water at a well and some key informant interviews on the use and maintenance of the wells, one may conduct a larger survey on water use and satisfaction with the quantity and quality of the water.

* Though in principle one may speak of loosely structured questionnaires, in practice the term questionnaire appears to be so hooked to tools with pre-categorised answers that we have decided to use the term interview guide for loosely structured tools. However, in reality there is often a mixture of open and pre-categorised answers (see Module 10B). In that case we will still use the term questionnaire.

4. Administering written questionnaires

A WRITTEN QUESTIONNAIRE (also referred to as self-administered questionnaire) is a data collection tool in which written questions are presented that are to be answered by the respondents in written form.

A written questionnaire can be administered in different ways, such as by:

- Sending questionnaires by mail with clear instructions on how to answer the questions and asking for mailed responses;
- Gathering all or part of the respondents in one place at one time, giving oral or written instructions, and letting the respondents fill out the questionnaires; or
- Hand-delivering questionnaires to respondents and collecting them later.

The questions can be either open-ended or closed (with pre-categorised answers).
(See **Module 10B** for details concerning design of interview guides and questionnaires.)

5. Focus group discussions (FGD)

A focus group discussion allows a group of 8 - 12 informants to freely discuss a certain subject with the guidance of a facilitator or reporter. (See **Module 10C** for a discussion of this technique.)

6. Projective techniques

When a researcher uses projective techniques, (s)he asks an informant to react to some kind of visual or verbal stimulus.

For example: An informant may be provided with a rough outline of the body and be asked to draw her or his perception of the conception or onset of an illness.

Another example of a projective technique is the presentation of a hypothetical question or an incomplete sentence or case/study to an informant ('story with a gap'). A researcher may ask the informant to complete in writing sentences such as:

- If I were to discover that my neighbour had TB, I would ...;
- If my wife were to propose that I use condoms, I would ...

Or (s)he may ask the informant: Suppose your child suffered from diarrhoea, what would you do?

Such techniques can easily be combined with semi-structured interviews or written questionnaires. They are also very useful in FGDs to get people's opinion on sensitive issues.

7. Mapping and scaling

Mapping is a valuable technique for visually displaying relationships and resources.

In a water supply project, **for example**, mapping is invaluable. It can be used to present the placement of wells, distance of the homes from the wells, other water systems, etc. It gives researchers a good overview of the physical situation and may help to highlight relationships hitherto unrecognised.

Mapping a community is also very useful and often indispensable as a pre-stage to sampling.

Scaling is a technique that allows researchers through their respondents to categorise certain variables that they would not be able to rank themselves.

> **For example**, they may ask their informant(s) to bring certain types of herbal medicine and ask them to arrange these into piles according to their usefulness. The informants would then be asked to explain the logic of their ranking.

Mapping and scaling may be used as participatory techniques in rapid appraisals or situation analyses. In a separate volume on participatory action research, more such techniques will be presented. (Also see the literature list at end of this module.)

Rapid appraisal techniques and participatory research are approaches often used in health systems research.

Differentiation between data collection techniques and data collection tools

To avoid confusion in the use of terms, the following table points out the distinction between *techniques* and *tools* applied in data collection.

Table 10A.1: Data collection techniques and tools

Data collection techniques	Data collection tools
Using available information	Checklist; data compilation forms
Observing	Eyes and other senses, pen/paper, watch, scales, microscope, etc.
Interviewing	Interview guide, checklist, questionnaire, tape recorder
Administering written questionnaires	Questionnaire

Advantages and disadvantages of various data collection techniques

Table 10A.2 summarises the advantages and disadvantages of various data collection techniques.

Table 10A.2: Advantages and disadvantages of various data collection techniques

Technique	Advantages	Possible constraints
Using available information	Is inexpensive, because data is already there. Permits examination of trends over the past.	Data is not always easily accessible. Ethical issues concerning confidentiality may arise. Information may be imprecise or incomplete.
Observing	Gives more detailed and context-related information. Permits collection of information on facts not mentioned in an interview. Permits tests of reliability of responses to questionnaires.	Ethical issues concerning confidentiality or privacy may arise. Observer bias may occur. (Observer may only notice what interests him or her.) The presence of the data collector can influence the situation observed. Thorough training of research assistants is required.
Interviewing	Is suitable for use with both literates and illiterates. Permits clarification of questions. Has higher response rate than written questionnaires.	The presence of the interviewer can influence responses. Reports of events may be less complete than information gained through observations.
Small scale flexible interview	Permits collection of in-depth information and exploration of spontaneous remarks by respondents.	The interviewer may inadvertently influence the respondents. Analysis of open-ended data is more difficult and time-consuming.
Larger scale fixed interview	Is easy to analyse.	Important information may be missed because spontaneous remarks by respondents are usually not recorded or explored.
Administering written questionnaires	Is less expensive. Permits anonymity and may result in more honest responses. Does not require research assistants. Eliminates bias due to phrasing questions differently with different respondents.	Cannot be used with illiterate respondents. There is often a low rate of response. Questions may be misunderstood.
Participatory and projective methods	Provide rich data and may have positive spin offs for knowledge and skills by researchers and informants.	Require some extra training of researchers.

II. IMPORTANCE OF COMBINING DIFFERENT DATA COLLECTION TECHNIQUES

When discussing different data collection techniques and their advantages and disadvantages, it becomes clear that they can complement each other. A skilful use of a combination of different techniques can reduce the chance of bias (see below) and will give a more comprehensive understanding of the topic under study.

Researchers often use a combination of flexible and less flexible research techniques.

Flexible techniques, such as

- loosely structured interviews using open-ended questions,
- focus group discussions, and
- participant observation

are also called **QUALITATIVE** research techniques. They produce qualitative data that is often recorded in narrative form.

QUALITATIVE RESEARCH TECHNIQUES involve the identification and exploration of a number of often mutually related variables that give INSIGHT in human behaviour (motivations, opinions, attitudes), in the nature and causes of certain problems and in the consequences of the problems for those affected. 'Why', 'What' and 'How' are important questions.

Structured questionnaires that enable the researcher to quantify pre- or post-categorised answers to questions are an example of **QUANTITATIVE** research techniques. The answers to questions can be counted and expressed numerically.

QUANTITATIVE RESEARCH TECHNIQUES are used to QUANTIFY the size, distribution, and association of certain variables in a study population. 'How many?' 'How often?' and 'How significant?' are important questions.

Both qualitative and quantitative research techniques are often used within a single study.

For example:

It has been observed in country X that children between 1 and 2-1/2 years, who have already started to eat independently, have unsatisfactory food intake once they fall ill. A study could be designed to address this problem, containing the following stages:

- Focus group discussions (FGDs) with 2 to 5 groups of mothers or in-depth interviews with 10 - 20 mothers, to find out whether they change the feeding practices for children in this age group when they suffer from (various) illnesses and how mothers deal with children who have no appetite when they are sick (exploratory study);

- A cross-sectional survey, testing the relevant findings of the exploratory study on a larger scale; and

- FGDs with women in the study area to discuss findings and possible questions arising from the survey and to develop possible solutions for problems detected.

In this example, the first, qualitative part of the study would be used to focus the survey on the most relevant issues (mothers' feeding behaviours and reasons for these behaviours) and to help phrase the questions in an optimal way in order to obtain the information that is needed.

The second, quantitative part of the study would be used to find out what proportion of the mothers follow various practices and the reasons for their behaviours and whether certain categories of children (e.g., the younger ones or children from specific socio-economic categories) are more at risk than others.

The third, qualitative part of the study would provide feedback on the major findings of the survey. Do the conclusions make sense to women in the study area? Have certain aspects been overlooked when interpreting the data? What remedial action is feasible to improve practices related to feeding sick children?

It is also common to collect qualitative and quantitative data in a single questionnaire. Researchers collecting both types of data have to take care that they:

- do not include too many open-ended questions in large-scale surveys, making data analysis more complicated; and
- do not use inappropriate statistical tests on quantitative data generated by small-scale studies.

III. BIAS IN INFORMATION COLLECTION

BIAS in information collection is a distortion in the collected data so that it does not represent reality.

Possible sources of bias during data collection:

1. Defective instruments, such as:

- Questionnaires with:
  - fixed or closed questions on topics about which little is known (often asking the 'wrong things');
  - open-ended questions without guidelines on how to ask (or to answer) them;
  - vaguely phrased questions;
  - 'leading questions' that cause the respondent to believe one answer would be preferred over another; or
  - questions placed in an illogical order.
- Weighing scales or other measuring equipment that are not standardised (see section 1).

These sources of bias can be prevented by **carefully planning the data collection process** and by **pre-testing the data collection tools**.

2. Observer bias:

Observer bias can easily occur when conducting observations or utilising loosely structured group- or individual interviews. There is a risk that the data collector will only see or hear things in which (s)he is interested or will miss information that is critical to the research.

Observation protocols and **guidelines for conducting loosely structured interviews** should be prepared, and training and practice should be provided to data collectors in using both these tools. Moreover it is highly recommended that data collectors **work in pairs** when using flexible research techniques and discuss and interpret the data immediately after collecting it. Another possibility - commonly used by anthropologists - is using a tape recorder and transcribing the tape word by word.

3. Effect of the interview on the informant:

This is a possible factor in all interview situations. The informant may mistrust the intention of the interview and dodge certain questions or give misleading answers**. For example:** in a survey on alcoholism you ask school children: 'Does your father sometimes get drunk?' Many will probably deny that he does, even if it is true. Such bias can be reduced by adequately introducing the purpose of the study to informants, by phrasing questions on sensitive issues in a positive way, by taking sufficient time for the interview, and by assuring informants that the data collected will be confidential (see **Module 10B**).

It is also important to be careful in the selection of interviewers. In a study soliciting the reasons for the low utilisation of local health services, for example, one should not ask health workers from the health centres concerned to interview the population. Their use as interviewers would certainly influence the results of the study.

4. Information bias:

Sometimes the information itself has weaknesses. Medical records may have many blanks or be unreadable. This tells something about the quality of the data and has to be recorded. For example, in a TB defaulter study the percentage of defaulters with an incomplete or missing address should be calculated.

Another common information bias is due to gaps in people's memory; this is called *memory* or *recall bias.* A mother may not remember all details of her child's last diarrhoea episode and of the treatment she gave two or three months afterwards. For such common diseases it is advisable to limit the period of recall, asking, for example, 'Has your child had diarrhoea over the past two weeks?'

Note:

All these potential biases will threaten the validity and reliability of your study. By being aware of them it is possible, to a certain extent, to prevent them. If the researcher does not fully succeed, it is important to report honestly in what ways the data may be biased.

Module 10B and **Module 12**, Annex 1 provide more information on how you can prevent biases in your research.

IV. ETHICAL CONSIDERATIONS

As we develop our data collection techniques, we need to consider whether our research procedures are likely to cause any physical or emotional harm. Harm may be caused, for example, by:

- violating informants' right to privacy by posing sensitive questions or by gaining access to records which may contain personal data;
- observing the behaviour of informants without their being aware (concealed observation should therefore always be crosschecked or discussed with other researchers with respect to ethical admissibility);
- allowing personal information to be made public which informants would want to be kept private, and
- failing to observe/respect certain cultural values, traditions or taboos valued by your informants.

Several methods for dealing with these issues may be recommended:

- obtaining **informed** consent before the study or the interview begins;
- not exploring sensitive issues before a good relationship has been established with the informant;
- ensuring the confidentiality of the data obtained; and
- learning enough about the culture of informants to ensure it is respected during the data collection process.

If sensitive questions are asked, for example, about family planning or sexual practices, or about opinions of patients on the health services provided, it may be advisable to omit names and addresses from the questionnaires.

REFERENCES

Abramson JH (1990, 4th ed.) *Survey Methods in Community Medicine.* London: Churchill-Livingstone. (In particular Chapter 6)

Moser CA, Kalton G (1989, 2nd ed.) *Survey Methods in Social Investigation*. Hants, UK: Gower Publishing Company. (In particular Chapters 10-13)

Patton MQ (1990, 2nd ed.) *Qualitative Evaluation and Research Methods.* Newbury Park, USA: Sage Publications. (In particular Chapters 17-29)

Pretty JN, Guyt I, Thompson J, Scones I (1995) *Participatory Learning & Action. A Trainer's Guide*. London: International Institute for environment and Development (IIED) (In particular Chapters 4 and 5 on semi-structured interviewing, diagramming, ranking and scoring.)

EXERCISE: Selection of study types and data collection techniques (in plenary)

Five health management problems for which studies must be developed are described below. For each problem you are asked to state:

- What type(s) of study you would propose.
- From whom (or from what) you would collect the data required for each study (your study populations).
- For each study population: which data collection techniques you would use.

1. You suspect that a large proportion of women in your region (pop. 1,000,000) is anaemic, in particular women at childbearing age. You would like to determine how big the problem is, and whether women perceive it as a problem. Furthermore you would like to know whether and how women themselves could contribute to improving their anaemic condition.

2. A district health team evaluated its malaria spraying programme by looking at available records and reports. It did not find significant flaws in the functioning of the services in different divisions and villages. Nevertheless, the incidence of malaria and mosquito counts show peaks in certain villages that the DHT cannot explain. It wants to find out if there is something wrong with the services.

3. A community survey in your region (pop. 2,000,000) indicates that 12% of the adults (15 – 60) in the capital are HIV positive, as compared to 7% in road side settlements and 2,5% in the rural area (where 80% of the population live). You want to step-by-step introduce an intensive STD programme in all 11 hospitals and 180 health centres, hoping to decrease the HIV incidence. You would like to evaluate whether the STD programme has an effect. Apart from the three main questions stated above: How would you organise the evaluation time-wise? Are any ethics involved? What could be two important biases in the study?

4. You are a midwife who has just been appointed as head of a maternity unit in a district hospital. You suspect that the number of low birth-weight babies is increasing, and you would like to know more about the physical and socio-economic conditions of the mothers to see if remedial actions should be taken. The clinic records at present are not complete enough to draw conclusions and you have neither the time nor the money to do a district-wide community survey.

5. There are long queues (waiting times), at the out-patient department of your district hospital. You are concerned about this and you would like to find out to what extent the problem may be related to the organisation and management of the department and whether certain bottlenecks can be identified. In a later stage of the research you would like to try to eliminate some of the bottlenecks and see whether there is improvement.

GROUP WORK: Selection of study type and data collection techniques

1. Decide what type(s) of study you will apply in your own research proposal.

   • Make your choice on the basis of your research objectives and the variables you would like to include in the study. (Hang objective and variables on a wall or flip chart board so the whole group can see them during the current group work session.) Review pages 1-2 of Module 9, to assist you in your choice of study type(s).

2. Determine what data-collection techniques you will use for each variable in your study.

   • Display the table that was prepared during the group work on selection of variables. For each variable determine the source of data and method(s) of data collection.

   • For some variables it may be necessary to collect additional data in order to be able to define the variable and determine the scale of measurement.

Example:	Factor:	Inadequate knowledge of patients about TB treatment
	Variable:	Knowledge of patients
	Scale of measurement:	Percentage or number of items of advice that are recounted by patients

   To determine the 'items' that should be used in the variable, it may be necessary to do a focus group discussion (FGD) with the staff treating TB patients and/or with a panel of experts on TB. Subsequently the results of the FGD could be used to develop questionnaires for interviews with patients. Hence the methods of data collection would be a FGD (to develop knowledge indicators) and face-to-face interviews with patients.

   • Display the results in the following table:

Variable*	Indicators* (If needed)	Definitions* (If applicable)	Data-collection technique	Source of data

   * These items were developed during the group work session for Module 8.

   * Preserve this table for presentation in plenary and for use in subsequent group work.

3. Summarise which data collection techniques you will use and which groups or records will form the sources of your data for each tool.

4. Decide whether there are any **ethical problems** with the type of study or data collection tools you propose.

Trainer's Notes

Module 10A: OVERVIEW OF DATA COLLECTION TECHNIQUES

Timing and teaching methods

1 hour	Introduction to data collection techniques and discussion
1 hour	Exercise: Selection of study types and data collection techniques
1 hour	Group work: Selection of study types and data collection techniques
1 hour	Plenary
4 hours	TOTAL TIME

Introduction and discussion

- Present an overview of the various data collection techniques. Give examples from the participants' fields of interest.
- If it would be useful for the participants' research projects, you can introduce additional research techniques using the material in **Module 10C** (Focus group discussions) or other sources, e.g., the volume on Participatory Action Research to be published by AFRO in the HSR series or Pretty et al. (1995) Participatory Learning and Action.
- Explain the difference between data collection techniques and data collection tools.
- Let the participants mention possible advantages and disadvantages of the various data collection techniques.
- Explain at what times qualitative research techniques are most useful and when quantitative techniques are more appropriate. Make sure that participants understand the advantages of combining quantitative and qualitative research techniques, preferably giving examples from one or more of the projects they are developing.
- Identify different possibilities for bias, using examples from the participants' own studies.
- Let the groups come up with examples of ethical issues that might play a role in their studies.

Exercise: Selection of study types and data collection techniques

- This exercise is designed to give participants some experience in choosing types of studies appropriate for typical situations before they have to select types of studies for their own proposals.
- Stress that objectives, if well formulated, should help to determine the appropriate study type(s).
- Ask participants to divide into sub-groups of 4-5 persons to do the exercise. Each sub-group should be assigned 2 topics. Allow 30 minutes for the groups to complete their work. Let them write the answers on an overhead transparency or flipchart.
- In the plenary (30 minutes) ask each group to answer the questions posed for one topic. Let other groups who discussed the same topic comment and add their own suggestions.

(An answer sheet for the exercise is presented in Annex 10A.1, on the next page.)

Group work: Selection of study type and data collection techniques

- This group work session is an important one, as it combines both the selection of study type and choice of data collection techniques. Work closely with your group, helping them to go step by step through the process described for the group work.

- Make sure they understand why it is important to refer back to their variables and the table they developed during **Module 8**, as they choose the data collection techniques they need.

Annex 10A.1: Answer sheet to exercise on selection of study types and data collection techniques

Proposed types of study	Proposed data collection techniques
Topic 1 (Anaemia)	
(a) *Exploratory study* on perceptions, symptoms and causes of anaemia and remedial action mothers may take + possible resources for action.	a) *Focus group discussions* with women of child bearing age, selected from different socio-economic backgrounds.
(b) *Cross-sectional community-based survey* to determine the size of the problem and high risk groups, or *clinic-based screening* of pregnant women at ANC (if high ANC coverage).	b) *Clinical observations*, *Hb testing* + *simple questionnaire to* obtain background data of women 15-45 years.
(c) *Cross-sectional survey*, community- or clinic-based (sub-sample of survey mentioned under (b)).	(c) More elaborate *interviews* among subsample of (b) on resources available to reduce risk factors (e.g., improved diet) and acceptability of different possible interventions.
Topic 2 (Malaria spraying)	
Exploratory study	*Participant observation (concealed)*
A number of observers receive a short training course in spraying procedures and mix among the spraying teams. They find out that the sprayers dump most of the insecticide in the morning, so that their load is lighter in the afternoon. The villages sprayed in the afternoon are underserved. (Foster GM (1987) World Health Organization behavioural science research: Problems and prospects. *Social Science and Medicine*, 24:709-717.)	
Topic 3 (STD intervention and HIV)*	
Experimental study in a number of experimental and control health centres + surrounding villages in the region, matched for socio-economic background. (This is possible because for logistic reasons the intervention cannot start everywhere at the same time).	- *HIV and STD tests* before start of intervention and 2 years later of all persons between 15-54 years in the selected areas, followed by comparison of cases and controls. - *Small questionnaire* on personal data, history of STDs and circumcision of all tested persons. - *Interviews* on sexual behaviour of sub-sample (1:8). - *Analysis of STD treatment cards* in HCs of the experimental villages on location of patients

Time-wise, you might wait two years before starting the STD treatment intervention in the control HCs (assuming that STD patients in the control villages would during the two-year interval use whatever treatment available: traditional; STD clinic in town; private doctors or quacks; sometimes HC).

* See Grosskurth H, Mosha F, Todd J, Mwijarubi E, Klokke A, Senkoro K, Mayaud P et al (1995) Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomised controlled trial. *The Lancet* **346**: 530-536.

This delay of two years would form an *ethical problem* as soon as the preliminary results would indicate that the extra input in STD treatment in the intervention HCs indeed reduces the risk of HIV infection. If that is the case, experiments are sometimes stopped in order to allow the control group to profit from the intervention as well.

One important *bias* could be that STD patients in control villages would have heard about the available treatment for STD in the experimental HCs and go there for treatment. To be able to control this potential bias, the research team analysed the treatment registers in the experimental HCs for the location of all STD patients who received treatment and deducted patients from the control villages (which appeared to be exceptions).

Another bias may be caused by intensified health education on STD symptoms and treatment over radio and TV, which could influence the treatment seeking behaviour of members of the experimental as well as of the control group.

Proposed type of study	**Proposed data collection techniques**
Topic 4 (Low birth-weight babies)	
Cohort study, examining all mothers who come for antenatal care over, say, 6 months, and following them up until after they deliver. or: *Case-control study* (mothers with low birth-weight babies and mothers with babies of normal weight) N.B.: If the problem is big (e.g., in 100 mothers you expect some 20 with an underweight baby) you can do a cohort study. Otherwise you may do a case-control study.	Thorough *history taking*; *measuring* mothers' body mass index (W/H^2)) and growth during pregnancy; lab tests on Hb, sugar, protein, blood smear for malaria; request to mothers who deliver at home to have babies weighed and examined one week after birth (if don't show up for examination, follow up). *Interviews* with all the mothers on socio-economic and cultural factors, or *Interviews* with mothers who gave birth to low birth-weight babies and a control group of mothers who gave birth to babies of normal weight, concerning socio-economic and cultural factors
Topic 5 (Queues/long waiting times at OPD in District Hospital)	
Exploratory/descriptive study	*Observation* of OPD procedures *Interviews* with staff on causes and solutions *Interviews* with patients
Or:	
Before–after study (quasi-experimental, as no control group)	Same as above, but in between the two studies you have taken steps to improve the situation.

Module 10B

DESIGN OF RESEARCH INSTRUMENTS; INTERVIEW GUIDES AND INTERVIEW SKILLS

Module 10B: DESIGN OF RESEARCH INSTRUMENTS; INTERVIEW GUIDES AND INTERVIEW SKILLS

OBJECTIVES

At the end of this session you should be able to:

1. **Distinguish** between various stages in questionnaire design.
2. **Demonstrate** appropriate techniques for wording questions and designing interview guides, questionnaires and checklists to ensure maximum quality of responses.
3. **Identify** appropriate data-collection techniques for your study.
4. **Prepare** your data-collection tools, taking care that you cover all important variables.
5. **Effectively carry out** an interview and train research assistants in interview skills.

I. Introduction

II. Types of questions

III. Steps in designing a questionnaire/interview guide

IV. Checklists

V. Interview skills

I. INTRODUCTION

The quality of research depends to a large extent on the quality of the data collection tools. Interviewing and administering questionnaires are probably the most commonly used research techniques. Therefore designing good 'questioning tools' forms an important and time-consuming phase in the development of most research proposals.

Once the decision has been made to use these tools, the following questions should be considered before designing them:

- What exactly do we want to know, according to the objectives and variables we identified earlier?
- Is questioning the right technique to obtain all answers, or do we need additional techniques, such as *observations* or *analyses of records*?
- Of whom will we ask questions and what techniques will we use? Do we understand the topic sufficiently to design a questionnaire, or do we need some loosely structured interviews with *key informants* or a *FGD* first to orientate ourselves?
- Are our informants mainly literate or illiterate? (If illiterate, the use of self-administered questionnaires is out of the question.)
- How large is the sample that will be interviewed? Studies with many respondents often use shorter, highly structured questionnaires while smaller studies allow more flexibility and may use interview guides or questionnaires with a number of open-ended questions.

II. TYPES OF QUESTIONS

In this module we will concentrate on interactive (usually face-to-face) interviews. Before examining the steps in designing a questionnaire, we need to review the types of questions used in interviews. Depending on how questions are asked and recorded we can distinguish two major possibilities:

- open-ended questions, (allowing for completely open as well as partially categorised answers), and
- closed questions.

1. Completely open-ended questions

OPEN-ENDED QUESTIONS permit free responses which should be recorded in the respondents' own words.

Such questions are useful for obtaining in-depth information on:

- facts with which the researcher is not very familiar,
- opinions, attitudes and suggestions of informants, or
- sensitive issues.

For example:

'What is your opinion on the services provided in the ANC?' (Explain why.)

'What do you think are the reasons some adolescents in this area start using drugs?'

'What would you do if you noticed that your daughter (school girl) had a relationship with a teacher?'

The answers to these questions are written down as closely as possible in the words of the respondents.

Advantages of completely open-ended questions

- Allow you to probe more deeply into issues of interest being raised.
- Issues not previously thought of when planning the study may be explored, thus providing valuable new insights on the problem.
- Information provided in the respondents' own words might be useful as examples or illustrations, which add interest to the final report.
- Often, re-reading an answer in a later phase of the analysis offers the possibility for different interpretations in relation to other data collected, which would have been impossible if the answer had been pre-categorised.

Risks of completely open-ended questions

- Skilled interviewers are needed to get the discussion started and focused on relevant issues and to record all information collected. A big risk is incomplete recording of all relevant issues covered in the discussion.
- Analysis is time-consuming and requires experience; otherwise important data may be lost.

Suggestions to improve use of completely open-ended questions

- Thoroughly train and supervise the interviewers or select experienced research assistants. (See section V of this module.)
- Prepare a list of further questions to keep at hand to use to 'probe' for answer(s) in a systematic way.
- Pre-test open-ended questions and, if possible, pre-categorise the most common responses, leaving enough space for other answers (see 2).

2. Partially categorised questions

In interviews questions are often **asked as open-ended questions**, but to facilitate recording and analysis, some answers can already be **pre-categorised**.

For example:

'How did you become a member of the Village Health Committee?'

	Details stated:
1. Volunteered	☐.....
2. Elected at a community meeting	☐.....
3. Nominated by community leaders	☐.....
4. Nominated by the health staff	☐.....
5. Other (specify):	.....

In this case the first four categories of answers are known, but there may be other possibilities. Therefore there is a category 'other' where other answers can be recorded. During the analysis, these responses can still be further categorised.

For open-ended questions, more than one answer is usually allowed. The interviewers will have to be trained to wait for additional answers

Advantages of pre-categorised answers

- Answers can be recorded quickly, and
- Analysis is easier.

Risks of pre-categorised answers

- If one pre-categorises too early, a lot of interesting and valuable information may never be recorded, or may end up in the category 'other'.
- Interviewers may try to force the information into the categories that are listed and, by merely ticking these, additional valuable information will be lost.
- Interviewers may stop after receiving the first answer, whereas more than one response could be applicable.
- Sometimes, if the respondent hesitates in answering, the interviewer may be tempted to present some possible answers, thereby causing bias.
- Frequently, questionnaires have very little space for recording full responses under 'other', forcing the interviewer to write down a response that summarises the respondent's answer, thereby losing valuable information.

Suggestions to minimise risks associated with pre-categorised answers

Clear guidelines have to be provided to interviewers on important issues. For example:

- If a question leads to an interesting discussion, it should be written down as completely as possible, *in addition to being coded. Space* should be *reserved to record these discussions*.
- Interviewers should be trained to solicit discussion when questions allow respondents to choose more than one option. The different options may be elaborated in subsequent questions.

- In case of non-response, the interviewer should repeat or rephrase the question without providing options for answers. The interview guidelines should provide suggestions for further probing which all interviewers should follow (see section V).

- Adequate space should be provided so that 'other' responses can be recorded as close as possible to the respondents' own words. Otherwise categorisation of these responses may be difficult afterwards.

Note:

There may be open-ended questions for which all possible categories of responses are known, e.g., some health practices (family planning methods: currently using, ever used, and never used). These may be fully categorised. However, these questions, when used in HSR, will always be followed by other questions asking for elaboration on reasons and conditions for use or non-use of the practice.

3. Closed questions

CLOSED QUESTIONS have a list of possible options or answers from which the respondents must choose.

Closed questions are most commonly used for background variables such as age, marital status or education, although in the case of age and education you may also take the exact values and categorise them during data analysis (see **Module 13**).

Closed questions may be used to get the respondents to express their opinions or attitudes by choosing rating points on a scale.

For example:

What is your opinion on the following statement:
'Women who have induced abortion should be severely punished.'

1. Strongly agree ☐
2. Agree ☐
3. Not sure/no opinion ☐
4. Disagree ☐
5. Strongly disagree ☐

Closed questions may also be used if one is only interested in certain aspects of an issue and does not want to waste time obtaining more information than one needs.

For example, a researcher who is only interested in the sources of protein in a family diet may ask:

'Did you eat any of the following foods yesterday?' (Circle yes if at least one item in each set of items is eaten.)

- Peas, beans, lentils Yes No
- Fish or meat Yes No
- Eggs Yes No
- Milk or cheese Yes No
- Insects Yes No

Using attitude scales in face-to-face interviews with literate respondents is most objectively carried out if the various options for each answer are provided on different cards. The respondents can be asked to put the cards in the order preferred by them while making their choice. If the researcher only reads the options, the respondents might not consider all options equally and the scale will not accurately measure the attitudes.

Advantages of closed questions:

- It saves time
- Comparing responses of different groups, or of the same group over time, becomes easier.

Risks of closed questions:

- In case of illiterate respondents, the interviewer may be tempted to read the list of possible answers in the given sequence, thereby influencing the choice of response and introducing bias.
- If there is no question to elaborate on the informant's reasons for choosing a certain rating, uniformity in rating may still be deceptive, as there may be considerable variation in reason for choosing the same ratings.

Suggestions to minimise risk associated with using closed questions:

- Develop picture codes can be used for illiterates as well as literates (e.g., five, four, three, two and one stars indicating a 5-point scale).
- First present the extremes and then the values in between so that the respondent is straight away aware of the range of answers.
- Ensure inclusion of follow up questions to elaborate on reasons for choosing a given rating.

Note:

Sometimes it is useful, especially in small-scale studies, to use pictures or drawings when asking certain questions in order to get the discussion going. In the case of illiterates, a questionnaire may even consist exclusively of pictures (See **Annex 1**).

III. STEPS IN DESIGNING A QUESTIONNAIRE/ INTERVIEW GUIDE*

Designing a good questionnaire always takes several drafts. In the first draft we should concentrate on the **content**. In the second, we should look critically at the **formulation and sequencing of the questions**. Then we should scrutinise the **format** of the questionnaire. Finally we should do a **test-run** to check whether the questionnaire gives us the information we require and whether interviewers as well as respondents feel at ease with it. Usually the questionnaire will need some further adaptation before we can use it for actual data collection.

* For the sake of simplicity we take semi-structured questionnaires used (containing completely open as well as partially pre-categorised and closed questions) in face-to-face interviews as an example. The same steps apply to designing other 'questioning tools'.

Step 1: Content

Take your objectives and variables as a starting point.

Decide what questions will be needed to measure or (in the case of qualitative studies) to define your variables and reach your objectives.

When developing the questionnaire, you should reconsider the variables you have chosen and, if necessary, add, drop or change some. You may even change some of your objectives at this stage.

Step 2: Formulating questions

Formulate one or more questions that will provide the information needed for each variable.

Take care that questions are specific and precise enough so that different respondents don't interpret them differently. For example, a question such as: 'Where do community members usually seek treatment when they are sick?' cannot be asked in such a general way because each respondent may have something different in mind when answering the question:

- One informant may think of measles with complications, and say he goes to the hospital, another of cough, and say he goes to the private pharmacy
- Even if both think of the same disease, they may have different degrees of seriousness in mind and thus answer differently
- In all cases, self-care may be overlooked

The question therefore, as a rule, has to be broken up into different parts and made so specific that all informants focus on the same thing. For example, one could:

- Concentrate on illnesses that have occurred in the family over the past 14 days and ask what has been done to treat these from the onset; or
- Concentrate on selected diseases, ask whether they have occurred in the family over the past x months (chronic or serious diseases have a longer recall period than minor ailments) and what has been done to treat each of them from the onset.

Check whether each question measures one thing at a time.

For example: the question, 'Do you think that the war situation leads to mental problems that require treatment by health staff?' brings three topics, which should be split up in

- mental problems resulting from the war,
- treatment required, and
- who should provide the treatment.

Avoid leading questions.

A question is leading if it suggests a certain answer. **For example:** the question, 'Do you think that people have to give bribes at hospital X to be seen by a doctor?' hardly leaves room for 'no' or for other options. A better question would be: 'Have you recently visited hospital X?' This would be followed by a series of other probing questions such as, 'By whom were you seen?' 'What were the complaints?' 'How much were you asked to pay?' 'Are the fees fixed?' 'Do they include an examination by the doctor if the condition of a patient is serious?' 'Do all patients have equal access to a doctor in case of need?' 'Was this what you expected?'

Sometimes, a question is leading because it presupposes a certain condition. **For example:** 'What action did you take the last time your child had diarrhoea?' presupposes the child has had diarrhoea. A better set of questions would be: 'Has your child ever had diarrhoea?' (If yes:) 'When was the last time?' 'Did you do anything to treat it?' (If yes:) 'What?'

Avoid words with double or vaguely defined meanings or that are emotionally laden. Concepts such as dirty (clinics), lazy (patients), or unhealthy (food), for example, should be omitted.

Ask sensitive questions in a socially acceptable way:

Questions relating to abortion, sexual practices of adolescents, or AIDS and mental illness in the family are usually sensitive. Such questions should be formulated in such a way that the question does not judge or embarrass the respondent. Making the subject more socially acceptable and then asking a question about the respondent's own preference or practice could lead to a more accurate response. **For example**: 'Many teenagers have had abortions for unwanted pregnancies. Do you know girls who had this problem? Have you ever had an abortion?'

Another way to deal with sensitive questions (as indicated by the respondent's hesitation in answering the question) is by asking the question *indirectly.* **For example** you could ask: 'If your friend was considering abortion for her daughter who became pregnant while in school, what would you advise her?'

A common weakness in questionnaires is the **inappropriate transformation of research questions into interview questions**. You cannot ask informants: 'Would a woman's educational status influence her health?' That is exactly what you have to find out by relating individual women's level of education to a number of health conditions.

Note:

Ask the questions to yourself or to a friend and check whether the answers you get are the type of responses you want.

Step 3: Sequencing the questions

Design your interview schedule or questionnaire to be 'informant friendly'.

- The sequence of questions must be logical for the informant and allow, as much as possible, for a **'natural' conversation**, even in more structured interviews.
- At the beginning of the interview a limited number of questions concerning 'background variables' (e.g., age, education, marital status) may be asked.
- As informants may be reluctant to provide 'personal' information and may become worried about confidentiality, or bored by a list of unrelated and, to them, senseless questions, you should restrict yourself to an essential minimum. You may postpone questions on religion until later when cultural questions are being asked. Socio-economic status/ occupation/ income questions can also better be postponed until later when you can link them to problems (e.g., low service utilisation).
- Start with an interesting but non-controversial question (preferably open) that is directly related to the subject of the study. This type of beginning should help to raise the informants' interest and lessen suspicions concerning the purpose of the interview.

- Pose more sensitive questions as late as possible in the interview (e.g., questions pertaining to income, political matters, sexual behaviour, or stigma experienced in case of stigmatising diseases).

- Use **simple, everyday language.**

- If interviews are carried out in English (or any other secondary language), local terminology should be used for crucial concepts that do not have the exact equivalent in the secondary language.

Step 4: Formatting the questionnaire

When you finalise your questionnaire, be sure that:

- A separate, introductory page is attached to each questionnaire, explaining the purpose of the study, requesting the informant's consent to be interviewed and assuring confidentiality of the data obtained.
- Each questionnaire has a heading and space to insert the number, date and location of the interview, and, if required, the name of the informant. You may add the name of the interviewer, to facilitate quality control.
- Layout is such that questions belonging together appear together visually. If the questionnaire is long, you may use subheadings for groups of questions.
- Sufficient space is provided for answers to open-ended questions, categories such as 'other' and for comments on pre-categorised questions.
- Boxes for pre-categorised answers are placed in a consistent manner (e.g., on the right half of the page). (See example in **Module 13, Annex 13.1**.)

If you use a computer, the right margin of the page should be reserved for boxes intended for computer codes. (Consult an experienced facilitator when designing your questionnaire.)

Your questionnaire should not only be INFORMANT - but also RESEARCHER FRIENDLY!

Step 5: Translation

If interviews will be conducted in one or more local languages, the questionnaire should be translated in order to standardise the way questions will be asked.

After having it translated you should have it retranslated into the original language by a different person. You can then compare the two versions for differences and make decisions concerning the final phrasing of difficult concepts.

Self-administered (written) questionnaires

All steps discussed above apply to written questionnaires as well as to guides/questionnaires used in interviews. For written questionnaires, however, clear guidelines will have to be added on how the answers to questions should be filled in.

Self-administered questionnaires are most commonly used in large-scale surveys using predominantly pre-categorised answers among literate study populations.

As a response rate of 50% or less to written questionnaires is not exceptional, these tools will rarely be used in small-scale studies. In exploratory studies which require intensive interaction with informants in order to gain better insight in an issue, self-administered questionnaires would, more over, be inadequate tools.

However, written questionnaires may sometimes be useful in small-scale studies on *sensitive topics*. Even then, they are usually combined with other tools:

> **For example**, during a FGD on sensitive issues such as sexual behaviour, a facilitator cannot always be sure that all participants dare to speak freely. A short, written questionnaire with open questions may then be used to explore remaining sensitive questions or worries of participants without revealing their identities.

Sometimes, a self-administered questionnaire seems the only possibility:

> **For example**, on one of the smaller islands in the Indian Ocean, the shortage of nurses was acute and public complaints about the functioning of health facilities abounded. A relatively high number of nurses had been leaving the profession over the past ten years and the vacancies could not all be filled as less and less young people were choosing nursing as a career. The research team that investigated the problem (all MOH staff) appreciated the sensitivity of the topic. Aware that everyone on the island knew each other, it was decided to use a flexible self-administered questionnaire which members of the research team personally delivered at each health facility, stressing the importance of filling it in. In that way they hoped to ensure anonymity as well as increase participation in the study. (Still, the response rate was only slightly over 50%.)

IV. CHECKLISTS

Checklists can be used to systematically observe human behaviour or the condition of specific equipment (e.g., fridge, expiring dates of medicines, or completeness of records). Observations can be relatively open or can be predetermined comparisons of reality against fixed standards. Sometimes the aim is systematic content analysis (e.g., newspaper articles, health information system). The objectives of the study determine the content of a checklist. A checklist includes all the items or points that must be considered during an observation in the field, or when extracting data from existing records.

An observation checklist for rating the cleanliness and use of Blair/Ventilated Improved Pit Latrines (VIP) in a study on hygiene practices is presented below, as an **example**:

Observations Ventilated Pit-latrines (Tick the appropriate boxes)	**Yes** (clear)	**Little** (hardly noticeable)	**No**
Evidence of use: worn footpath to VIP	☐	☐	☐
Toilet smells	☐	☐	☐
Flies in and around the toilet	☐	☐	☐
Fouling around the toilet	☐	☐	☐
Defecation in the area around the homestead	☐	☐	☐

V. INTERVIEW SKILLS

1. The interviewer-informant relationship and interview conditions

An interviewer needs to have the skills of a detective. (S)he should carefully, step-by-step, delve for the truth. Here the comparison stops, because the truth researchers are looking for, in general, has nothing to do with criminality. To the contrary, our **respondents are possible partners** who with their information help solve a shared problem, or at least help to better understand why people behave as they do.

To turn an informant into a partner, a researcher has to invest in the relationship. First of all, (s)he has to be clear to the informant about the purpose of the interview and the study. Enough information should be given to raise the interest of informants and to enable them to judge whether they would like to participate or not. Consent has to be obtained before the interview. (See Module 10A.) On the other hand, not too many details should be given about what will be asked and why. Otherwise, the researcher runs the risk informants become selective in what they tell, and conceal information in order to 'help' or please him/her (interviewer bias).

When sensitive topics are being explored anonymity should be ensured, for example, by not including a name in the interview notes. If a second interview would be useful, one may ask a name and address, but make sure this will never appear in the report.

Further, before the research starts, the possibility of **involving informants in the discussion of the results and recommendations** should be considered and discussed with informants. Feedback sessions with informants are usually rewarding for both informants and researchers. If it is impossible to organise such sessions for all participants, an interviewer can at least summarise the main results of the interview at the end and check if this is what the informant meant to say. Usually the informant will react, elaborate on the responses already given, or sometimes withdraw a statement if it seems too personal.

Partnership between interviewer and informant implies that the interviewer will **try to minimise the social distance** between him/herself and the informant. Interviewers should try to blend in the environment. In order to do so:

- Clothing of interviewers should be culturally acceptable and as simple as possible (no fancy dresses, high heels or tight jeans in rural areas).

- Sitting arrangements for interviewer(s) and informant(s) should preferably be at the same height (no straight chair when the informant is sitting on a mat on the floor) and beside each other, forming an angle of 90 degrees, rather than opposite each other.

- Gender relations have to be respected. When interviewer and informant are of opposite sex, more physical distance will usually be required than when they are of the same sex. In some societies, a male researcher may have to ask consent from a woman's husband or father before he can interview her, or may have to employ a female research assistant to interview women. One has to ask advice from key informants and try out what works, as 'opposite sex interviews' sometimes also produce interesting information that in 'same sex interviews' would not so easily be obtained. Differences in age and culture, if exploited with care, may have the same surprising effect of generating useful information on sensitive topics.

- A general rule is that, through his or her behaviour, an interviewer should **show interest** in what the informant says, be at ease (never in a hurry) and **make the informant feel at ease**. This implies that the interviewer should **never show any disapproval** of the information received during the interview. Otherwise the informant will close up. The interviewer should only try to listen and understand WHY people do what they do, even if the practice seems dangerous. Only after the interview or after the research is finished, can the interviewer try to address problems identified through the interviews.

For example: A research team headed by a Regional Medical Officer (RMO) carried out a study into reasons for low utilisation of delivery services in a district in Tanzania. When interviewing Traditional Birth Attendants (TBAs), the team discovered that TBAs delivered services in case of highly complicated deliveries which could be dangerous to the mothers. The RMO was surprised about the practices of the TBAs and asked many details about what they did with respect to various complications, but did not show his surprise. After finalising the study he organised a training course for TBAs in which he took their level of knowledge as a point of departure and explained why some practices were dangerous and could better be replaced by safer alternatives.

The **environment should** also **be supportive** of the interview situation. Anything that disturbs (noise, other people listening, a formal surrounding) should if possible be avoided.

For example: A hospital-ward, with nurses running in and out, may not be a good place to ask a patient's treatment history including traditional treatment, or his/her opinion about the treatment in the hospital. An adolescent home is not an appropriate environment for an interview with adolescents about their sexual behaviour, even if the parents are absent. A quiet place outside would be better.

If it appears impossible to prevent others from joining the interview, even after having made clear that one would like to speak undisturbed with the informant, there are different possibilities.

For example:

- If the information is personal but not very sensitive, one may just continue, hoping that others may get bored and disappear.
- Depending on the sensitivity of the data and who joins, one can turn the interview into a mini focus group discussion, asking indirect instead of personal questions (e.g., 'how would a husband react if he found out that his wife was using the pill to prevent pregnancy without his knowledge', instead of 'how would your husband react...').
- One may try to make a new appointment and hope the informant will be able to assure more privacy then.
- Or one may tell the visitors that they will be visited later.

Tape recording may be an enabling or a disturbing environmental factor in the interview. If the informant accepts the argument of the interviewer that (s)he will not be able to remember everything that is said nor write it down, it is an asset. Playing with it at the beginning of the interview and letting informants hear themselves may help to overcome hesitation among informants who have never experienced the use of a tape recorder. Informants should be asked for their consent before use, and be assured confidentiality. If they hesitate or refuse, the only solution is to take notes. This is in any case advisable, as the tape-recorder may not always function correctly. Usually, informants forget quickly that the recorder is on, but if they appear disturbed despite their consent that it be used, it should be stopped.

Note:

- The interviewer should always make notes on disturbances or 'enabling factors' that emerge during the interview. Likewise, it should be noted when an informant asks for assistance from others in answering certain questions.
- Remember that an interviewer can also be a 'disturbing factor' in the daily activities of informants. Identify the most suitable time for the interview and share in tasks that lend themselves to that purpose.
- The interviewer should be very conversant with the interview guide/questionnaire so that eye contact is maintained throughout the interview. If the discussion touches on sensitive issues, few or no notes should be taken, in order not to lose eye contact. The information collected should then be written down **immediately** after the interview.

2. The interviewer's tasks

Apart from the introduction to the interview, which includes building a partnership that needs to last throughout the interview and even beyond, the interviewer has a number of other tasks:

- Posing questions
- Evaluating answers and probing for elaboration or more precision in case of a superficial or invalid answer
- Noting down answers
- Leading the discussion, but at the same time encouraging the informant to give spontaneous information relevant to the topic by letting him/her talk

Unless the interview is carried out in pairs, the interviewer will have to carry out all these tasks at the same time.

(1) The **introduction to the interview** should be written as the interview guide is being developed, even if there is only one researcher, but certainly if a research team conducts the interviews. This will ensure consistency, which prevents bias due to interviewer variation.

We have already paid much attention to
(2) the art of **posing questions** in the preceding parts of this module. Not yet covered is the issue of what to do when an answer is unsatisfactory. Clearly, the quality of an interview will be determined by the skills of the interviewer in
(3) **evaluating answers** and diagnosing what went wrong and why, so that corrective action can be taken on the spot.

An answer can be unsatisfactory in many ways:

- There may be *no answer at all*, either because the informant did not understand the question or because the question touched on sensitive information that the informant hesitated to divulge.

  Possible remedy: If you suspect that the question was not understood, you may say: 'Perhaps, I was not so clear. What I meant to ask was …', followed by the question, phrased slightly more elaborately. If you suspect the informant is hesitant to answer, then repeat that if (s)he wishes (s)he can skip this question. If the informant has no objection to continuing, stress again that the information will remain in confidence. Then repeat the question.

- The *answer* may be *incomplete,* or *unclear.*

  Possible remedy: Repeat what was said and encourage the informant to continue with phrases such as: 'I do not fully understand what you mean. Could you elaborate a bit?' or: 'Could you mention more possibilities?'
  Or, if the informant has started to give an example but did not finish, you can probe with questions such as: 'When was that?', 'Why did you do that?', 'How often does that happen?', 'Why do you think so?', etc.

- The *answer* may be *irrelevant* to your question.

  Possible remedy: Say something such as 'That is not exactly what I meant to ask about' and repeat the question, slightly elaborated.

- You suspect the *informant* is *not telling the truth* (invalid answer), either because various parts of the answer contradict each other or because the informant knows the 'desirable' answer and gives it to please the interviewer. Both can happen at the same time.

For example: in an interview about sexual behaviour, a young man reports that he has used condoms last month. Later he says that he only uses condoms with casual partners. When asked to elaborate a bit on casual partners (who they are, where he finds them, how often, when the last time was) he states that he has not had a casual partner over the past three months, since he has a regular partner. Probably the first answer was a 'socially acceptable' statement to please you, but the contradicting data may also mean that the truth (actual sexual behaviour) is quite complicated. It may be 'socially unacceptable' to his new girlfriend that he still has other sexual contacts; he may be denying those contacts, and the first answer (that he used condoms with a casual partner last month) may be right.

Possible remedy: Summarise the three contradicting statements, and ask the informant: 'How should I interpret this: You said that you have used condoms last month. You also said that you use condoms only with casual partners, and that since three months you have not had casual partners as you have a regular partner now…' Then wait and see what comes. Avoid the slightest sign of disbelief or irritation. Make clear by an attentive and interested look that you would appreciate it highly if he told something more about himself and when he does and does not use condoms with his different girlfriends.

Note that **when something goes 'wrong' in an interview the interviewer should always take the blame**. (S)he is the one who 'did not ask the question clearly enough', or who 'did not understand the answer', and so is asking the question again. The interviewer is the one who receives a 'present' in the form of information and has to show by word and body language that (s)he appreciates this highly. Nodding, humming, or repeating the last few words may all be perceived as encouragement. A silence from the side of the interviewer may also work as encouragement for the informant to continue.

(4) **Noting down answers** should never go at the cost of the cost of the eye contact with the informant, even if the information collected is not sensitive. Scribbles in a small notebook are preferable over more extensive notes, provided they are elaborated straight after each interview. It is advisable to take brief notes even when a tape recorder is used, just in case the recording fails. Elaborating notes after an interview takes more time than the interview itself, and the full transcription of a taped interview may take between 5-10 times the duration of the interview, and often longer.

(5) **Keeping control over the interview** without imposing oneself is a skill each researcher has to learn. Even though an interviewer may encourage the provision of (often valuable) spontaneous information, some spontaneous information may distract from the main issues. The interviewer has to take care that there is not only a good beginning but also a satisfactory end to the interview. In between, all topics should have been adequately covered. Otherwise the information provided by different informants will not be comparable or, in case of key informants, not be comprehensive. Therefore an interviewer will also sometimes have to stop informants who go enthusiastically off track. This can be done politely: 'Thank you, this is interesting, but do you mind if I go back to the previous (or next) question? This (specify) is not yet fully clear to me.'

At the end, the interviewer should not only **summarise the interview**, which may lead to a valuable new discussion, but also respond to questions that came up during the interview, give advice (if necessary or asked for) and give an opportunity for further questions of the informant. Such **'after-interview' discussions and questions should *always* be recorded, like all spontaneous information**, because discussions can shed light on complicated, not yet fully clear issues from many preceding interviews. (Remember the researcher's function of detective!)

3. Training the research team/assistants

Those who for the first time conduct a face-to-face interview will need training in the application of all the advice provided above. Good quality interview data form the heart of an HSR study. Obtaining such data presupposes more than appropriate knowledge and skills; also specific attitudes (e.g., not imposing one self) and insight (e.g., appreciating signs of discomfort in an informant, and recognising evading or, to the contrary, highly illuminating answers) are required.

The learning process therefore preferably consists of reading and listening/observing, as well as doing. 'Doing' may involve a written exercise, a role-play, or actual field practice during a pre-test. The following exercise may be helpful:

EXERCISE: Interview training

1. **Written exercise:** Write introductions for one of the interviews your team will be developing for your study. Then meet in small groups and critique each other's introductions.

2. **Role-play:** Six participants will be asked to participate in a role-play. Three will act as interviewers - staff members from a health centre with a high percentage of unsupervised deliveries who wish to find out the reasons for the low coverage. The other three will play the roles of mothers who have delivered in the last six months. Each of the mothers will be given a particular role to play by the facilitators. The rest of the group will know what the roles are, but the interviewers will have to find out while they are interviewing and try to cope as best they can.

   While the three interviews are taking place, observe the interactions closely. Following the interviews a brief discussion will be held focusing on the attitudes of the interviewers and their skills. We will discuss what has been done well in each of the interviews, and suggestions for how the process could be improved so as to gain more accurate and complete information.

3. **Real interview:** As in 2, but now on a topic relevant to all participants. (For example, what made them join medical training.)

GROUP WORK (4 hours or more)

1. **Prepare your data-collection tools (instruments**), **taking care that you cover all important variables**. Refer back to the table that your group prepared during the group work session at the end of **Module 10A**, which specifies the methods of data collection you must use. (You might divide up the work, assigning different members of the group to design the various data collection instruments required.)

2. Take care that you have an optimal mix between open-ended and pre-categorised questions.

3. If methods are needed other than those presented in **Modules 10A** and **10B**, refer to **Module 10C** on FGD or other text-books on (participatory) methods.

4. Discuss the possibilities for bias, which may occur when using the data-collection tools. Try to avoid bias as much as possible.

EXERCISE: Review of data collection tools

1. Review in detail data collection tools of one other research team in relation to their objectives and variables and prepare suggestions for improving them. Be prepared to present your comments in plenary.

2. If there is time, review the data collection tools of the other research teams in the course as well.

REFERENCES

Briggs CL (1992) *Learning how to ask. A sociolinguistic appraisal of the interview in social science research.* (In particular chapter 3, analysis of communicative blunders) Cambridge: Cambridge University Press. (5th edition)

Lee RM (1993) *Doing Research on sensitive topics.* London: Sage Publications.

Pretty JN, Guyt I, Thompson J, Scones IL (1995) *Participatory Learning & Action. A Trainer's Guide*. London: International Institute for environment and Development (IIED) (In particular Chapters 4 and 5 on semi-structured interviewing, diagramming, ranking and scoring.)

Rubin HJ, Rubin IS (1995) *Qualitative Interviewing. The art of hearing data.* Thousand Oaks: Sage.

Seidman I (1998) *Interviewing as qualitative research. A guide for researchers in education and the social sciences.* New York: Teachers College Press. (In particular chapters 5-7.)

Spradley JP (1979) *The ethnographic interview.* New York: Holt, Rinehart and Winston.

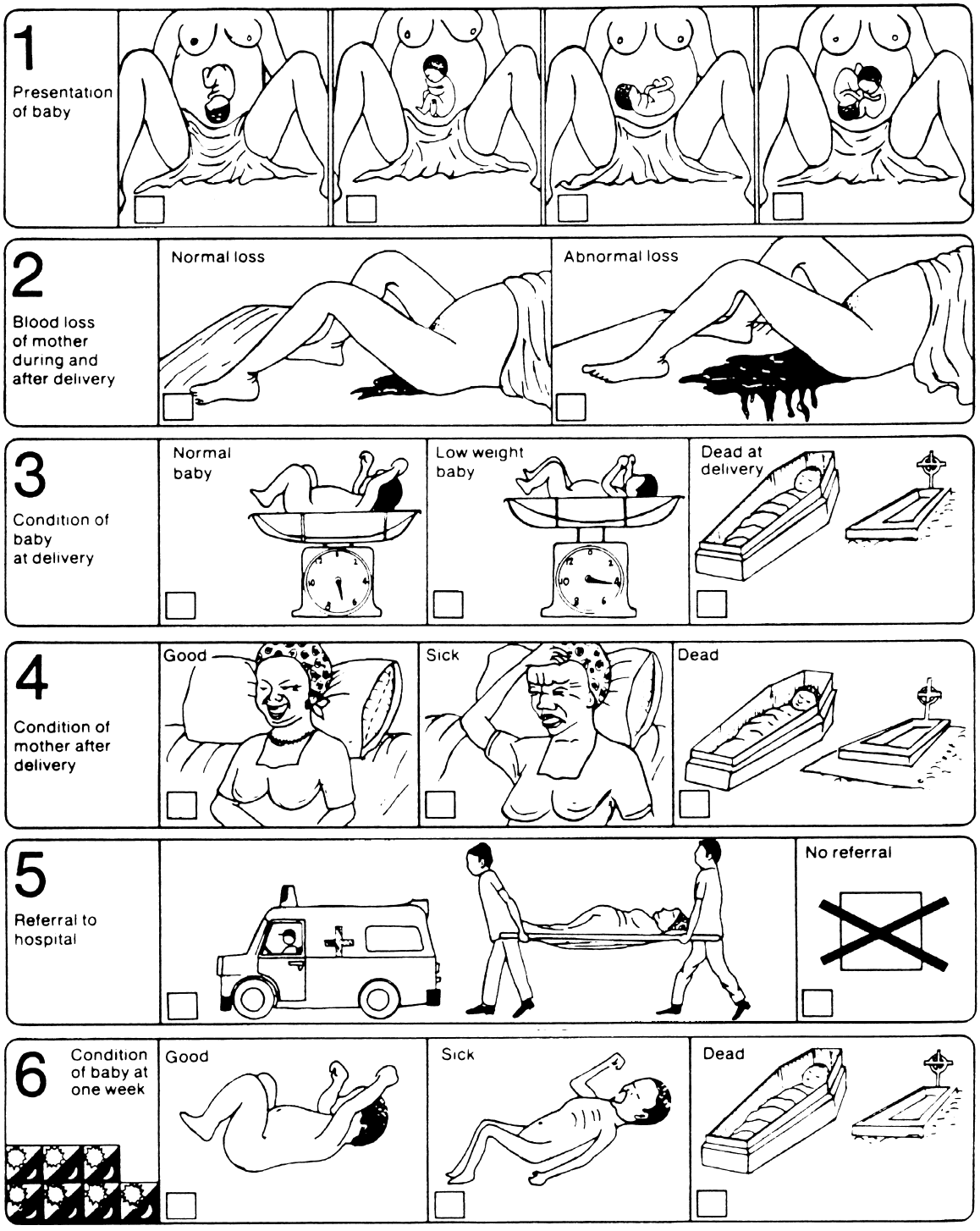
Annex 10B.1: MATERNAL RECORD*

MATERNAL RECORD

Name of mother _____ Name of TBA _____

Age of mother _____ Date of delivery _____

Parity _____ Number _____



* This questionnaire which was made for use by non-literate health workers was shared by Dr. Peter Lamptey.

Trainer's Notes

Module 10B: DESIGN OF RESEARCH INSTRUMENTS; INTERVIEW GUIDES AND INTERVIEW SKILLS

Timing and teaching methods

1 hour	Introduction to designing questionnaires and interview skills
45 minutes	Exercise: Interview training
4 hours+	Group work
2 hours	Exercise: Comment on the data collection tools of other groups
2 hours	Plenary on data collection tools
2 hours	Revision of data collection tools
11 hours+	TOTAL TIME

Introduction and discussion

- The introduction should be straightforward but interactive, providing participants with an opportunity to comment on obviously poor questions and come up with suggestions for improvement.

- The handling of half open-ended, partly pre-categorised questions, which appear in many questionnaires, may need special attention. Participants should be aware of the danger of bias if these questions are not asked uniformly with guidelines for probing.

- The formatting of questionnaires should be illustrated with examples.

Exercise: Interview training

1. **Written exercise:** Ask each participant to write an introduction for one of the interviews his or her team will be developing for its study. Then ask each team to meet in a small group and critique each other's introductions.

2. **Role-play:** Ask six participants to volunteer to participate in the role play - three as interviewers (health care staff working in an area with a high percentage of unsupervised deliveries) and three as mothers who have delivered in the past 6 months. Announce that the task of the three interviewers is to interview the mothers to determine the reasons for the low percentage of deliveries supervised by the health centres and possible ways to increase coverage. Ask the three interviewers to go out of the room and each separately jot down ideas on the questions he or she might ask the mothers in five minutes interviews. While the interviewers are doing this, discuss the roles of the three mothers with the participants who have volunteered to play them and the rest of the class. (One mother may have given birth at home and honestly will tell why; another may have given birth at home and is hesitant to disclose why; and the third may have given birth in the hospital, but is not an easy talker; the interviewer will have to draw the information out of her to get a complete picture.)

Once the 'stage is set', ask each of the interviewers to interview one of the mothers in turn, while the rest of the class watches. When the three interviews are completed, lead a discussion on the attitudes of the interviewers, their skills, what has been done well and suggestions for improving the interview. The 'mothers' can be asked what their feeling were about the interviews and what the interviewers could have done to obtain more accurate and complete information.

Another possibility is to take a 'real life' situation and let three pairs of participants interview each other on a topic of common interest. Discuss this first with the participants and ask for their suggestions for a topic.

Group work

- All facilitators should be aware that the quality of the data collection tools determines the quality of the data with which the participants return from the field. Skilful guidance of the groups is therefore essential.

- If participants are relatively inexperienced in research, the first version of an interview guide/ questionnaire is often too general and has too many closed questions. It is extremely important that the groups try their tools out in a 'real life' situation during the workshop.

- The time needed to develop the data collection tools may exceed four hours. Usually groups continue working in the evening. There should be two opportunities in the workshop program for the teams to revise the data collection tools: after the exercise on reviewing the data collection tools in class and after the exercise in the field, so that they can include what they learned from their own experience and the comments of facilitators and other course participants. In their own research situations the teams will then do full pre-tests (see **Module 14**).

Exercise: Review of data collection tools

- If possible, have all groups review and critique the data collection tools of all other groups, with special attention to the tools of one group. In plenary, the group that has the primary responsibility for reviewing a particular group's tools should comment first. Then other groups should be asked to give any additional suggestions, after which the group whose tools have been discussed can reply, if necessary.

- When groups have many different data collection tools (questionnaires, checklists, schedules for FGDs), it may not be possibly to discuss all tools of all groups in two hours. Then two groups could swap their tools and discuss them in mini plenaries of the two groups. **Facilitators, however, should read and comment on the data collection tools of all groups**.

Note:

When groups are preparing their data collection tools they should have ideas concerning which tools they would like to pre-test and where. At this point of time, **the course management team should therefore start organising the pre-test.**

Designing and Conducting Health Systems Research Projects
Part I: Proposal Development and Fieldwork

Module 10C

FOCUS GROUP DISCUSSIONS

Module 10C: FOCUS GROUP DISCUSSION

OBJECTIVES

At the end of the session you should be able to:

1. **Identify** the purpose, uses and limitations of the Focus Group Discussion (FGD) as a method of data collection in research.

2. **Conduct** a FGD, analyse the data and report on the results.

I. Characteristics and uses of focus group discussions

II. How to conduct a focus group discussion

III. Analysis of results

IV. Report writing

I. CHARACTERISTICS AND USES OF FOCUS GROUP DISCUSSIONS

A FOCUS GROUP DISCUSSION (FGD) is a group discussion of approximately 6 - 12 persons guided by a facilitator, during which group members talk freely and spontaneously about a certain topic.

A FGD is a qualitative method. Its purpose is to obtain in-depth information on concepts, perceptions and ideas of a group. A FGD aims to be more than a question-answer interaction. The idea is that group members discuss the topic among themselves, with guidance from the facilitator.

FGD techniques can, for example, be used to:

1. **Focus research** and develop relevant research hypotheses by exploring in greater depth the problem to be investigated and its possible causes.

   For example:

   A district health officer had noticed that there were an unusually large number of cases of malnutrition of children under 5 reported from one area in her district. Because she had little idea of why there might be more malnutrition in that area she decided to organise three focus group discussions (one with leaders, one with mothers from the area and one with health staff from the area). She hoped to identify potential causes of the problem through the FGDs and then develop a more intensive study, if necessary.

2. **Formulate appropriate questions** for more structured, larger scale surveys.

   For example:

   In planning a study of the incidence of childhood diarrhoea and feeding practices, a focus group discussion showed that in the community under study, children below the age of 1 year were not perceived as having 'bouts of diarrhoea' but merely 'having loose stools' that were associated with milestones such as sitting up, crawling, and teething. In the questionnaire that was developed after the FGD the concept 'diarrhoea' was therefore carefully described, using the community's notions and terms.

3. **Help understand** and solve **unexpected problems in interventions**.

   For example:

   In District X, the recent national (polio) immunisation days (NID) showed widely different coverage's per village (50-90%) and in a number of villages a marked decrease in coverage was observed compared to last year. Eight FGD were held with mothers, two in town, three in rural villages with a marked decrease in NID coverage and three in villages with a high coverage throughout. It appeared that overall, the concept NID had raised confusion. Most people believed that this mass campaign strengthened the children's immunity against *any* (childhood) disease, including malaria and Respiratory Tract Infections. In the villages with a low NID coverage there had been a high incidence of malaria in children immediately after the previous NID campaign and several children died. Mothers therefore believed that the NID campaign was useless.*

* This is an adapted version of an (as yet unpublished) study carried out in Bushenyi District, Uganda, by Nuwaha et al.

4. **Develop appropriate messages for health education programmes** and later evaluate the messages for clarity.

   For example:

   A rural health clinic wanted to develop a health education programme focused on weaning problems most often encountered by mothers in the surrounding villages and what to do about them. The focus group discussion could be used for **exploring relevant local concepts** as well as for **testing drafts** when developing the messages. The messages should be developed and tested in different socio-economic groups of mothers, as weaning practices may differ with income, means of subsistence and education of the mothers. Also ethnic differences may have to be taken into account.

5. **Explore controversial topics**.

   For example:

   Sexual behaviour is a controversial topic in the sense that males and females judge sexual relations and sexuality often from very different perspectives. Sexual education has to take this difference into account. Through FGDs, first with females, then with males, and then with a mixed group to confront both sexes with the different outcomes of the separate discussions (listed on flip charts) it becomes easier to bring these differences in the open. Especially for teenagers, who may have many stereotypes about the other sex or be reluctant to discuss the topic openly (particularly girls), such a 'multi-stage' approach is useful.

Strengths and limitations

Implementation of FGDs is an *iterative* process; each focus group discussion builds on the previous one, with a slightly elaborated or better-focused set of themes for discussion. Provided the groups have been well chosen, in terms of composition and number (see below), FGDs can be a powerful research tool which provides valuable spontaneous information in a short period of time and at relatively low cost.

FGD should *not* be used for quantitative purposes, such as the testing of hypotheses or the generalisation of findings for larger areas, which would require more elaborate surveys.

However, FGDs can profitably complement such surveys or other, qualitative techniques. Depending on the topic, it may be **risky** to use FGDs as a **single tool**. In group discussions, people tend to centre their opinions on the most common ones, on 'social norms'. In reality, opinions and behaviour may be more diverse. Therefore it is advisable to combine FGDs with at least some key informant and in-depth interviews. Explicitly soliciting other views during FGDs should be routine as well.

In case of very **sensitive topics**, such as sexual behaviour or coping with HIV/AIDS, FGDs may also have their limitations, as group members may hesitate to air their feelings and experiences freely. One possible remedy is the selection of participants who do not know each other (e.g., selection of children from different schools in FGDs about adolescent sexual behaviour), while assuring absolute confidentiality.

It may also help to alternate the FGD with other methods, for example, to precede it by a self-developed role play on sexual behaviour, or to administer a written questionnaire immediately after the FGD with open questions on sexual behaviour in which the participants can anonymously state all their questions and problems. This worked in Tanzania and Nepal.*

* The Tanzania-Netherlands Support Programme on AIDS, Mwanza Region, Tanzania (1990-2000+) and the Family Planning Association of Nepal Project on adolescent health in five districts of Nepal (1999-2003). The adolescent health section of WHO/HQ has developed a Narrative Research Method which is very well suited to help adolescents develop narratives and role plays about their interpretations of sexuality which can profitably precede the single sex and mixed FGDs: World Health Organization (1992) *A story of the sexual experience of young people in eleven African countries; The Narrative Research Method.* Geneva: WHO; World Health Organization (1993) *The narrative research method; Studying behaviour patterns of young people by young people. A guide to its use*, Geneva: WHO.

Another way to ensure confidentiality in a FGD on a sensitive topic is giving participants an option to introduce themselves under any name they would like to use (not necessarily their own). Further, before the discussion, it should be stressed that they may bring up experiences of friends and brothers/sisters as well as their own, and that it is not necessary to bring painful personal experiences in the open.*

II. HOW TO CONDUCT A FOCUS GROUP DISCUSSION

Determine the purpose

A FGD can be regarded as a mini-study. It therefore requires one or two clear objectives. (See **Module 6**.) These objectives will guide the research team in the formulation of discussion questions.

Situation analysis

Any FGD requires good knowledge of local conditions. Communities are seldom or never homogeneous. There are always differences between community members, for example in education, political power, gender, economic status and ethnic group. These differences will be reflected in their perceptions of the problems they suffer from and possible solutions. A researcher must be aware of these differences, otherwise (s)he may miss important groups of participants or obtain a hotchpotch of information. Similarly, (s)he must know which key persons or organisations could be good entry points for the selection of participants in the FGDs (e.g.: women's groups, parent associations, youth clubs, etc.). If a FGD forms part of a bigger study, or project, it may be easy to define target groups for the discussions. Otherwise, the first task of the researcher(s) will be to explore the area and identify possible target groups. Interviews with some key informants and a rudimentary situation analysis are then indispensable. The situation analysis should preferably be carried out in a participatory way, with representatives of the study population on which the FGD focuses.

For example:

In an intervention study on sexual health among out-of-school youth in an urban area, the researcher first planned some interviews with key informants. He selected the leaders of a political youth club and of a Christian youth club and some teachers, with whom he thoroughly discussed his research topic. Through them he came in contact with youth of different backgrounds. He let each of the three groups, separated into boys and girls, draw maps of the town and asked them to mark places which they thought riskful in terms of sexual behaviour (easy contacts, unprotected sex). The drawings formed a good basis for further FGDs but also helped him to identify wider networks of adolescents at risk who had to be included in the study.

Points to be considered when preparing the FGD

Recruitment of participants:

- Participants should be roughly of the same socio-economic group or have a similar background in relation to the issue under investigation. The age and sexual composition of the group should facilitate free discussion.

  Often you therefore need to obtain information on a topic from several different categories of informants who are likely to discuss it from different perspectives in separate FGDs, though in a later stage groups may be joined (see examples 3,4 and 5).

* Always ensure confidentiality of opinions: Ask co-operation from the group members as well, to keep what has been discussed confidential. If group members present very personal problems and need advice or help, this should be followed up after the FGD.

Participants should be invited at least a day or two in advance, and the general purpose and procedures of the FGD should be explained, in order to obtain their **consent to join**.

- **Selection of participants:**

If you are an outsider in the research area, you may have to rely on your key informants for the first selection of participants in FGDs. Your key informants to whom you have explained thoroughly the purpose and the process of the FGD might each suggest some individuals who could be invited to a focus group discussion.

Note that the key informants may select persons similar to themselves so that you do not get an adequate variety of views in your discussion group. So in your explanations be sure to emphasise that you want a group of people that can express a *range* of views, to be able to have a proper discussion. Participants in a first FGD may assist to find relevant participants for other groups.

Another way of getting participants is to conveniently select individuals in a *systematic* way, to try and ensure a range of views. You might, for example, ask every third or fourth person you find. This method might be more suitable in urban areas.

- **Physical arrangements:**

Communication and interaction during the FGD should be encouraged in every way possible. Arrange the chairs in a circle. Make sure that there will be no disturbances, sufficient quietness, adequate lighting, etc. Try to hold the FGD in a neutral setting which encourages participants to freely express their views. A health centre, **for example**, is not a good place to discuss traditional medical beliefs or preferences for other types of treatment.

- **Preparation of a discussion guide:**

There should be a **written** list of topics to be covered. It can be formulated as a series of open-ended questions. Guides for different groups gathered to discuss the same subject may vary slightly, depending on their knowledge or attitudes and how the subject should first be explored with them.

Conducting the session

One of the members of the research team should act as 'facilitator' or 'moderator' for the focus group discussion. One should serve as 'recorder'. The facilitator should preferably be as close as possible to the participants in their characteristics (same sex, roughly same age, etc.).

Functions of the facilitator

The facilitator should NOT act as an expert on the topic. His or her role is to stimulate and support discussion.

- **Introduce the session**

Introduce yourself as facilitator and introduce the recorder. Let participants introduce themselves with whatever names they wish to use. Put the participants at ease and explain the purpose of the FGD, the kind of information needed, and how the information will be used (for the planning of a health programme, an education programme, etc). Ask permission to use a tape-recorder, let people hear their own voices before the session starts. You might offer drinks and allow some informal discussion before the actual session starts.

- **Encourage discussion**

  Be enthusiastic, lively, and humorous and show your interest in the groups' ideas. Formulate questions and encourage as many participants as possible to express their views. Remember there are **no** 'right' or 'wrong' answers. **React neutrally** to both verbal and non-verbal responses.

- **Encourage involvement**

  Avoid a question-and-answer session. Some useful techniques include:

  - Asking for clarification:
    'Can you tell me more about...?'

  - Reorienting the discussion when it goes 'off the track':
    Saying: 'Wait, how does this relate to...?'
    Saying: 'Interesting point, but how about...?'
    Using one participant's remark to direct a question to another, for example, 'Mrs. X said ..., but how about you, Mrs. Y?'

  - When dealing with a dominant participant, avoiding eye contact or turning slightly away to discourage the person from speaking, or thanking the person and changing the subject.

  - When dealing with a reluctant participant, using the person's name, requesting his/her opinion, making more frequent eye contact to encourage his/her participation.

- Deal correctly wiht **sensitive issues**. If you notice that the discussion stops when dealing with a sensitive topic, you could ask participants (if literate) to anonymously write down their responses or opinions on the topic. Alternatively, you could summarise for the group some of the opinions from previous focus group discussions, focusing on one or two major contrasting opinions. Still another strategy is to form *sub-groups*, and to get a member of the sub-group to summarise and present the opinions of their sub-group members after which the whole group can still discuss these opinions.

- **Build rapport, empathise**

  Observe non-verbal communication. Ask yourself, 'What are they saying? What does it mean to them?' Be aware of your own tone of voice, facial expressions, body language, and those of the participants.

- **Avoid being placed in the role of expert**

  When asked for **your** ideas or views by a respondent, remember that you are not there to educate or inform. Direct the questions back to the group by saying: 'What do you think', 'What would you do?' Set aside time, if necessary, after the session to give participants the information they have asked for.

  Do not try to comment on everything that is being said. Don't feel you have to say something during every pause in the discussion. Wait a little and see what happens.

- **Control the rhythm of the meeting, but in an unobtrusive way**

  Listen carefully, and move the discussion from topic to topic. Subtly control the time allocated to various topics so as to maintain interest. If participants spontaneously jump from one topic to another, let the discussion continue for a while since useful additional information may surface; then summarise the points brought up and reorient the discussion.

- **Take time at the end of the meeting to summarise, check for agreement and thank the participants**

  Summarise the main issues brought up, check whether all agree and ask for additional comments. Thank the participants and let them know that their ideas have been a valuable contribution and will be used for planning the proposed research, intervention, or health education materials.

- Listen for **additional comments** and spontaneous discussions which occur after the meeting has been closed.

Functions of the recorder

The recorder should keep a record of the content of the discussion as well as emotional reactions and important aspects of group interaction. Assessment of the emotional tone of the meeting and the group process will enable you to judge the validity of the information collected during the FGD.

Items to be recorded include:

- Date, time, place
- Names and characteristics of participants
- General description of the group dynamics (level of participation, presence of a dominant participant, level of interest)
- Opinions of participants, recorded as much as possible in their own words, especially for key statements
- Emotional aspects (e.g., reluctance, strong feelings attached to certain opinions)
- Vocabulary used - particularly in FGDs that are intended to assist in developing questionnaires or health education materials
- Spontaneous relevant discussions during breaks or after the meeting has been closed

It is highly recommended that a tape-recorder be used to assist in capturing information. Even if a tape-recorder is used, notes should be taken as well, in case the machine malfunctions and so that information will be available immediately after the session for discussion.

If there is no reliable tape-recorder available, it is advisable to have **two recorders**.

A **supplementary role** for the recorder could be to assist the facilitator (if necessary) by drawing his or her attention to:

- missed comments from participants
- missed topics (the recorder should have a copy of the discussion guide during the FGD)

If necessary, the recorder could also help resolve conflict situations within the group that the facilitator finds difficult to handle on her own.

Number and duration of sessions

- **Number of sessions**

  The number of focus group sessions to be conducted depends upon project needs, resources, and whether new information is still coming from the sessions, (that is, whether contrasting views within and between various groups in the community are still emerging). If not, you may stop.

  One should plan to conduct at least two FGDs for each sub-group (for example, two for males and two for females). Otherwise you have no way of assessing whether the information you get from the first FGD is representative for that group.

- **Duration**

  A focus group session typically lasts up to an hour and a half. Generally the first session with a particular type of group is longer than the following ones because all of the information is new. Thereafter, if it becomes clear that all the groups have a similar opinion on particular topics, the facilitator may be able to move the discussion along more quickly to other topics which still elicit new points of view.

III. PROCESSING AND ANALYSIS OF RESULTS

- After each focus group session the facilitator and recorder should meet to review and **complete the notes** taken during the meeting. This is the right moment to **evaluate** how the focus group went and what changes might be made in the topics when facilitating the next focus group.

  Immediately afterwards a full report of the discussion should be prepared which reflects the discussion as completely as possible, using the participants' own words. List the key statements, ideas, and attitudes expressed for each topic of discussion.

- After the transcript of the discussion is prepared, **code**, following your topics, the participants' statements right away, using the left margin. Make finer sub-codes. **Write comments** (your first interpretation of the data) in the right margin. Formulate additional questions if certain issues are still unclear or controversial and include them in the next FGD. Further categorise the statements for each topic, if required. (See **Annex 10C.2.)**

- When you have all the data, **summarise** it in a **compilation sheet** organising the findings per topic for each. Number the FGD interviews and use key words to summarise group statements in the compilation sheet so that you can always go back to the full statement. If you have different categories of informants, e.g., male and female, you can summarise the information from the male and female groups on two **separate** compilation sheets. (See **Module 23** for an example.)

- You should then to do a **systematic comparison** between groups on all topics. Use your **objectives** and problem analysis diagram as a framework for analysis and comparison.

- The next step could be to put the major findings for different study populations on one sheet. You may want to use some of these sheets in your research report.

- Sometimes you may also wish to use diagrams when summarising the causes or components of the problem understudy. (See **Module 23** for more details.)

- Only now can you report the major findings of the FGDs in a narrative.

IV. REPORT WRITING

Start with a description of the purpose of the FGDs, the selection and composition of the groups of FGD participants and a commentary on the group process, so the reader can assess the validity of the reported findings.

Present your findings, following your list of topics and guided by the objective(s) of your FGD. Include quotations whenever possible as illustrations, particularly for key statements.

EXERCISE (3 hours total)

Conducting an FGD (75 minutes)

Participants working in groups of 6-12 conduct an FGD among themselves.

First let each group select a facilitator and a reporter

- Preparation of discussion guides (15 minutes)
- Discussion (60 minutes)

NB: It may be instructive to let the facilitator and reporter prepare the discussion guide for their group together with members of another group. Then the facilitator and reporter could come back to their own group with the guide when the FGD is to start. This resembles the real situation, where FGD members do not know which questions will be asked.

Analysis of data (30 minutes)

The reporter and facilitator analyse the notes and prepare the report.

Plenary (75 minutes)

The plenary sessions may include the following steps for each group:

1. The recorder presents the report of the FGD of his or her group.

2. Recorders may then ask for comments and reactions from members of the group.

3. A discussion can be held concerning the effects of the role-played by the facilitator, the group process, and the skills of the recorder on the validity of the report of the FGD.

4. If different groups discussed the same topic, the plenary can try to identify the different perspectives from which each group approached the topic.

If the group is big enough, one or more participants may act as observers and comment on the group process. Otherwise workshop facilitators can take that role. Sociograms are useful tools to record the flow of the discussion. (See **Annex 10C.3**)

REFERENCES:

Debus M, Porter Novelli (1986) *Handbook for Excellence in Focus Group Research,* Washington: Academy for Educational Development, Healthcom.

Helitzer-Alland, Makhambera Mano, Wangel AM (1994) Obtaining sensitive information: The need for more than Focus Groups. *Reproductive Health Matters* **3**:75-82

Khan ME, Anker M, Patel BC, Barge S, Sadhwani H, Kohle R (1991) The Use of Focus Groups in Social and Behavioural Research: Some Methodological Issues. *World Health Statistical Quarterly*, **44**:145-149.

Krueger RA (1988) *Focus Groups. A practical guide for applied research.* Newbury Park: Sage Publications. (In particular pp. 51-71 and 91-105.)

Morgan DL (1993) *Successful Focus Groups: Advancing the State of the Art*. California: Sage Publications Inc.

Powell RA and Simple HM (1996) Methodology Matters V: Focus Groups. *International Journal for Quality in Health Care* **8** (5):499-504.

Annex 10C.1: Example of a FGD on reasons why nurses leave the profession; discussion guide

1. When did the leaving of nurses become a problem to you? When did it start?
2. What, do you think, were (are?) the major reasons why they left the profession?
3. What problems do you experience due to this 'exodus' of nurses?
4. How did the Ministry of Health react?
5. Did you take any initiatives to alert the Ministry to your problems?
6. Is it only because so many nurses have left that you experience problems in your work, or are other factors playing a role as well?
7. Do you have any concrete suggestions to the Ministry to improve your working situation?

Annex 10C.2: Example of a FGD on reasons why nurses leave the profession; transcribed text from tape-recorder (Group of four nurses)

(codes)

(5), (4)

Nurse A (angry): We have been complaining about how things are in this Ministry but nobody listens to us. Every time we complain, we are told that things are being looked at and that soon things will work out. Nothing!

(own remarks)

(5), (4) Blaming represen-tatives

Nurse B: Yes! During our last strike on conditions of service, we were told that a committee had been set up to look at how our working conditions can be improved. Strange thing is that we had two representatives. I think we ourselves are also not very active in pushing our needs – how can we have representatives in the committee and nothing happens in two years?

NB. * Interview representa-tives !!

Nurse C: Representatives in the committee, I think, they are promised better things for themselves and forget about everybody else's needs. I tell you, we ourselves are killing the nursing profession in this country.

trust

Nurse A: I agree with you. We have asked these colleagues who are representing us to meet with us and tell us what has happened. We met them, but you don't hear what they are saying in the meetings. I really think that they are not expressing our views. No wonder they do not tell us what they are saying.

*

Balanced views, own resp. Action MOH? (4)

Nurse D (an elderly nurse who has not said anything yet): That is not fair to the representatives. I think they are trying their best. We all know how difficult it is to negotiate for these things. They are up against the whole Ministry. After all, most of the changes we want will take a long time. Of course we are suffering now but it is clear that the Ministry is trying to change things. We were told by the representatives that our salaries may be reviewed but that nothing can be done now. It will have to be at the beginning of the new fiscal year. I personally think we should not blame our representatives. If they are not performing, it is our fault for selecting the wrong people.

* Get overview of salary scales and changes over past 5 yrs

(5)

Facilitator: What then do you think needs to be done to ensure that nurses do not leave and who should do this?

Nurse A: I think we have done all we can. I told you, we have begged, asked, striked, and still ..., nothing.

Facilitator: What do you think are the major reasons why nurses are leaving their jobs?

(2), (4) Blaming MOH

Nurse A: I really think nurses are leaving because the Ministry is not interested in the welfare of nurses in this Ministry. Everything about nurses is wrong. We could write books about this! Even the President stated in his New Years speech that the nurses in our community are not performing as they should.

MOH blames nurses instead of looking for structural causes of problem

(Nurses start to discuss the New Years speech and go slightly off the track.)

Facilitator: What do you mean by welfare of nurses?

(2), (6), Low salaries

Nurse B: You know we get very low salaries here. I have still the same salary since I came to this Ministry five years ago. How can I pay for transport, the school fees for my children, food and rent – I don't even have any subsidised rent. You buy food and pay for transport and the money is finished.

* Get News-paper with New Year's speech!

Nurse A: I think the amount of salary we get shows the Ministry does not value our work. I have had increments- but what can you do with an annual increment of R20.00? It's nothing. It is a joke on our efforts.

Nurse B: Look at what X is getting since she left for Australia. At least three times as much as what we have, and things are expensive here, not much difference with Australia.

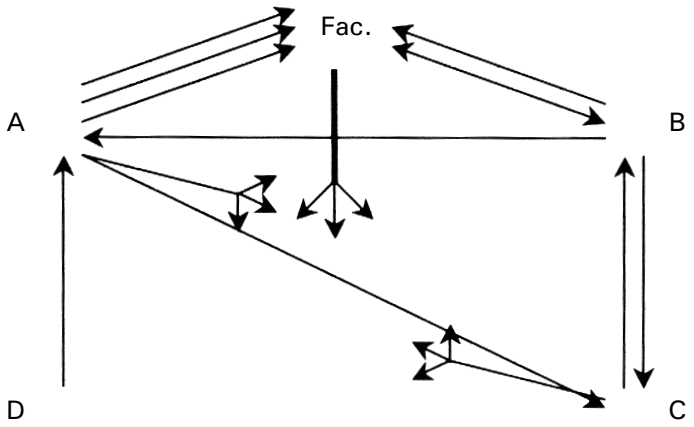
Nurse C: But we all know, the problem is not only salaries. Our working conditions are horrible. No consideration is given to our working conditions. Sometimes you are called back on your free days because a nurse is sick and there is no one to replace her.

Nurse B: Yes, even our welfare, nobody cares. I can give examples of many people who have big complaints about their welfare here.

Facilitator: I am still not so sure what you mean by welfare. Could you explain? (Asks nurse B to continue)

Nurse B: Actually what we are discussing now is the welfare. You have no life and there is no consideration of your life outside your performance. We are sometimes transferred to places where there are no schools, no shops – nothing. You are not even given transport to come home. I was transferred to a very rural clinic and stayed there for five years – five years!! – they had completely forgotten about me.

Annex 10C.3: Sociogram of FGD (5 minutes)



Remarks

Nurse A appears most talkative (and most angry). Nurse D is least talkative but important in the discussion as she tries to find a balance.

Trainer's Notes

Module 10C: FOCUS GROUP DISCUSSIONS

Timing and teaching methods

½ hour	Presentation on FGD
3 hours	Exercise: Focus group discussions
3½ hours	TOTAL TIME

Introduction and discussion

- Start with a 30 minute presentation on FGDs;
- Present a sample:

  - Guide for an FGD and
  - Report of an FGD

Exercise: Focus group discussion

- Prepare for the exercise by placing the participants in homogeneous groups of 8-12 persons, for example one group of males and one group of females. Try to select a topic on which males and females might react differently (for example: the most efficient way to propagate condom use as a means to prevent AIDS; target groups for propagating condom use; possible effects).

  Or: place the participants in three homogeneous groups of 8-12 persons (not necessarily according to sex) and give each of them a different, controversial discussion topic.

- Inform each group of their assigned topic. Instruct the group to appoint a FGD facilitator and recorder. Instruct the groups on preparation of discussion guides (15 minutes).

- Request each member to develop a written guide for the discussion. (NB. This part of the exercise should enable all participants to develop skill in writing a guide.)

  Or: Swap facilitators and reporters between the groups. Let the facilitator and reporter of group A develop the questions for group A with members of group B, and let the facilitator and reporter of group B develop the questions for their group together with members of group A.

- Allow 1 hour for the FGD. During the FGD, one of the workshop trainers or workshop facilitators should be assigned to observe each group.

- The workshop trainer/facilitator should observe and record the group process. It is useful to record the interaction (i.e., who talks/to whom) and the time frame as well as the process, that is:

  - The skills and limitations displayed by the facilitator;
  - The behaviour of group members; and
  - The influence of the group interaction on the development of the discussion.

- During the plenary, invite the participants to comment on the extent to which the recorder's report reflects their own opinions and feelings. This will help them appreciate the potential and limitations of a FGD and also the crucial role of the facilitator and recorder of a FGD.

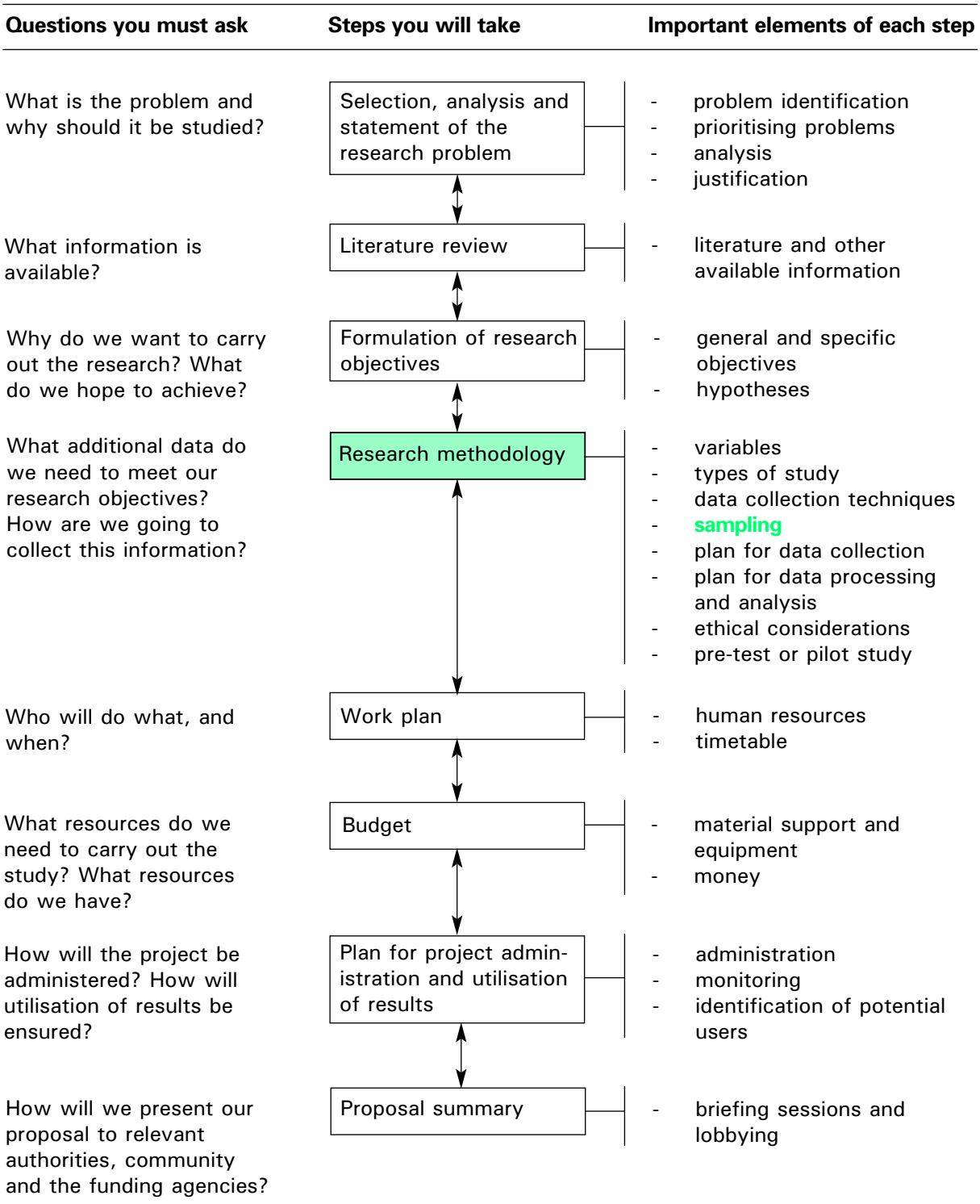
Designing and Conducting Health Systems Research Projects

Part I: Proposal Development and Fieldwork

Module 11

SAMPLING

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 11: SAMPLING

OBJECTIVES

At the end of the session you should be able to:

1. **Identify and define** the population(s) to be studied.
2. **Identify and describe** common methods of sampling.
3. **Discuss** problems of bias that should be avoided when selecting a sample.
4. **List** the issues to consider when deciding on sample size.
5. **Decide** on the sampling method(s) and sample size(s) most appropriate for the research design you are developing.

I. Introduction

II. Sampling procedures

1. Sampling methods for *qualitative* data

2. Sampling methods for *quantitative* data

3. Bias in sampling

4. Ethical considerations

III. Sample size

I. INTRODUCTION

WHAT is sampling?

SAMPLING is the process of selecting a number of study units from a defined study population.

Some studies involve only small numbers of people and thus all of them can be included. Often, however, research focuses on such a large population that, for practical reasons, it is only possible to include some of its members in the investigation. We then have to draw a **SAMPLE** from the total population.

In such cases we must consider the following questions:

- What is the group of people (STUDY POPULATION) we are interested in from which we want to draw a sample?
- How many people do we need in our sample?
- How will these people be selected?

The study population has to be clearly defined (for example, according to age, sex, and residence.) Otherwise we cannot do the sampling. Apart from persons, a study population may consist of villages, institutions, records, etc.

Each study population consists of STUDY UNITS. The way we define our study population and our study unit depends on the problem we want to investigate and on the objectives of the study.

For example:

Problem	Study population	Study unit
Malnutrition related to weaning in District X	All children 6-24 months of age in District X	One child between 6 and 24 months in District X
High drop-out rates in primary schools in District Y	All primary schools in District Y	One primary school in District Y
Inappropriate record keeping of hypertensive patients registered in hospital Z	All records of hypertensive patients in hospital Z	One record of a hypertensive patient registered in hospital Z

Representativeness

If researchers want to draw conclusions which are valid for the whole study population, which requires a **quantitative** study design, they should take care to draw a sample in such a way that it is **representative** of that population.

A REPRESENTATIVE SAMPLE has all the important characteristics of the population from which it is drawn.

For example:

> If you intend to interview 100 mothers in order to obtain a complete picture of the weaning practices in District X you would have to select these mothers from a representative sample of villages. It would be unwise to select them from only one or two villages as this might give you a distorted (biased) picture. It would also be unwise to only interview mothers who attend the under-fives clinic, as those who do not attend this clinic may wean their children differently.

When using **qualitative** research approaches, however, representativeness of the sample is NOT a primary concern. In exploratory studies which aim at getting a rough impression of how certain variables manifest themselves in a study population or at identifying and exploring thus far unknown variables, you may try to select study units which give you the richest possible information: you go for INFORMATION-RICH cases!

For example:

> Key informants should *never* be chosen *at random*, but *purposively* from among *those who have the best possible knowledge, experience or overview* with respect to topic of your study. Moreover they should be willing to share this information with you.

II. SAMPLING METHODS

As the rationale for the use of specific sampling methods in qualitative study designs is very different from the rationale underlying sampling methods in quantitative studies, we will discuss them separately.

Purposeful sampling strategies for qualitative studies

Qualitative research methods are typically used when focusing on a limited number of informants, whom we select *strategically* so that their in-depth information will give optimal insight into an issue about which little is known. This is called *purposeful sampling.* There are several possible strategies from which a researcher can choose. Often different strategies are combined, depending on the topic under study, the type of information wanted and the resources of the investigator(s).

This section owes much to Michael Quinn Patton who in his book *Qualitative Evaluation and Research Methods* (1990: 169-186) discusses a wide range of purposeful sampling techniques.

(1) Extreme case sampling

We have already discussed this type of sampling several times in **Modules 9** and **10**. Selection of e*xtreme cases*, such as good or very poor compliers to treatment, is a powerful and rapid strategy to identify contributing factors to poor compliance. In the same way, selection of well nourished children of the same age will help to identify contributing factors for malnutrition, and systematic comparison of a poorly and a well functioning district health team will give insight into factors that may contribute to the satisfactory functioning of DHTs.

Also *'thick' (elaborate) description of single, deviant cases* may be useful. In that way AIDS was discovered in California (USA) as a newly emerging disease.

(2) Maximum variation sampling

If a researcher wants to obtain as complete as possible insight in a certain issue in all its variations, maximum variation sampling will be used.

> **For example**, the stigma (see **Module 8** Part III) of leprosy, or TB, HIV, epilepsy, is considered a complicating factor in the control of these diseases. In order to obtain insight in how stigma manifests itself in different cultures in *males and females,* in *rural* and *urban* areas, in *well-to-do* and *poor* patients, or in *educated* and *illiterate* ones, a researcher has to take care that all these groups are included in the sample. To assess whether *social distance* influences stigma, one could also interview blood relatives (parents or children), spouses, friends, near neighbours of patients and more distant community members. For leprosy and TB patients, it is useful to interview *patients on treatment* as well as *patients declared cured*, to assess if any reversal of stigma is experienced when patients' conditions improve. If a researcher is interested in specific groups and interviews a fixed number per group, this type of sampling is also called ***quota sampling.***

Note:

Purposeful sampling should **not** be **haphazard**. Care should be taken that for different categories of informants, **selection rules** are developed to prevent the researcher from sampling according to personal preference.

In the study on stigma it would, for example, be possible to select the categories of patients one would like to focus on (male/ female, on treatment and released from treatment) in a systematic way from patient registers in a number of selected clinics. One could e.g., take all, or every second or third patient in the desired categories. The clinics may have been selected purposefully (town/rural, specific ethnic areas, poor/rich areas) but if there is choice the researcher may make lists for each category and select at random (see next section) the desired number of clinics. For identifying social distance variation, selection rules will be much more pragmatic: spouses, relatives and neighbours who are available will be interviewed in a careful, indirect way in order not to hurt the patient's interests. Still, however interesting data maximum variation sampling may generate, highlighting different factors and different perspectives, it does **not** provide **representative** data for the **total population**.

(3) Homogeneous sampling

Sometimes a researcher would like to have specific information about *one particular group only*, for example, a group that, for unclear reasons, is more at risk than others:

> In country S, death registers indicate that suicide among adolescents is on the increase at an alarming rate. Within that group twice as many boys as girls commit suicide. Researchers may therefore want to concentrate on the boys to identify what factors may be contributing to these suicides, conducting in-depth interviews with parents, other close relatives, teachers and friends of a number of boys who committed suicide.

In focus group discussions (FGDs), we usually select homogeneous groups because participants discuss more freely when they are amongst people of similar social status. (See **Module10C**.)

(4) Typical case sampling

It is sometimes illustrative to describe in-depth some cases which are *'typical'* for the group one is interested in. For example, one may describe a 'typical' family in a rural village in country A, or a 'typical' young school leaver who migrates from the rural areas to town in search of work, or 'typical' health problems of miners or malnourished children.

Such descriptions are merely illustrative; they cannot be generalised for the whole group. Typical examples can either be selected with co-operation of key informants who know the study population well, or from a survey that helps to identify the normal distribution and the modus of the characteristics we are interested in.

(5) Critical case sampling

Critical cases are those who 'can make the difference' with respect to an intervention you want to introduce or to evaluate.

> **For example**, you have developed a local weaning food that, you hope, is affordable to all mothers. Before propagating it at a larger scale through MCH clinics you first interview and observe some low-income mothers as 'test cases'. If they manage to produce and use it, this will indicate that it is affordable to the whole group.

(6) Snowball or chain sampling

This approach is particularly suitable for locating key informants or critical cases. You start with one or two information-rich key informants and ask them if they know persons who know a lot about your topic of interest. If a particular person is recommended to you by two or three different people you can be quite sure that he or she will be a valuable key informant.

The same approach can be used if an in-depth interview leads to discoveries, which seem rewarding to follow-up by a number of interviews with an additional group of informants.

> **For example**, in an exploratory study on coping behaviour among AIDS orphans it seemed that child-headed households managed by girls survived better than those managed by boys. The researcher then interviewed more adolescent boys and girls heading households, to see whether this gender difference in ability to cope was real, and how it could be explained.

Patton (1990: 179) labels this kind of additional sampling *during* the study ***opportunistic sampling.***

Flexible sampling procedures, steered by the data one collects (in relation to the objectives) forms a major opportunity for qualitative researchers to optimally exploit the field situation and explore 'in-depth' interesting issues which present themselves. It is exactly the opposite of the random sampling techniques discussed in the next section of this module, which are used in quantitative research to ensure representativeness of the sample for the total population. Still, if qualitative researchers can choose from a group of seemingly similar informants they will also sample at random (see example 2).

Note:

Purposeful sampling is NOT the same as convenience sampling.* CONVENIENCE SAMPLING is a method in which for convenience sake the study units that happen to be available at the time of data collection are selected in the sample. This *may* happen at the beginning of a study when researchers are merely orienting themselves, or, when there are many similar informants and the researchers do not (yet) have a preference for specific categories. When there seems no other choice (no one else available for an interview) researchers may also sample conveniently.

* In the earlier version of this module, following epidemiological tradition, all 'non-random' sampling methods were categorised under the headings of 'convenience' or 'quota' sampling. This injustice towards purposeful sampling techniques has been corrected in the present version. In HSR, purposeful and random sampling techniques are used equally.

2. Random sampling strategies to collect quantitative data

If the aim of a study is to *measure* variables distributed in a population (e.g., diseases) or to *test hypotheses* about which factors are contributing significantly to a certain problem, we have to be sure that we can generalise the findings obtained from a sample to the total study population. Then, purposeful sampling methods are inadequate, and **probability-** or **random sampling methods** have to be used.

PROBABILITY SAMPLING involves using random selection procedures to ensure that each unit of the sample is chosen on the basis of chance. All units of the study population should have an equal, or at least a known chance of being included in the sample.

Probability sampling requires that a listing of all study units exists or can be compiled. This listing is called the **sampling frame**.

The following probability sampling methods will be discussed:

- Simple random sampling
- Systematic sampling
- Stratified sampling
- Cluster sampling
- Multistage sampling

(1) Simple random sampling

This is the simplest form of probability sampling. To select a simple random sample you need to:

- Make or search for an existing numbered list of all the units in the population from which you want to draw a sample
- Decide on the size of the sample (this will be discussed in section III)
- Select the required number of sampling units, using a 'lottery' method or a table of random numbers (**Annex 11.1** explains how to use a table of random numbers.)

For example, a simple random sample of 50 students is to be selected from a school of 250 students. Using a list of all 250 students, each student is given a number (1 to 250), and these numbers are written on small pieces of paper. All the 250 papers are put in a box, after which the box is shaken vigorously, to ensure randomisation. Then, 50 papers are taken out of the box, and the numbers are recorded. The students belonging to these numbers will constitute the sample.

(2) Systematic sampling

In SYSTEMATIC SAMPLING individuals are chosen at regular intervals (for example every fifth) from the sampling frame. Ideally we randomly select a number to tell us where to start selecting individuals from the list.

For example, a systematic sample is to be selected from 1200 students of a school. The sample size selected is 100. The sampling fraction is:

$$\frac{100\ (=\ \text{sample size})}{1200\ (=\ \text{study population})} = \frac{1}{12}$$

The sampling interval is therefore 12.

The number of the first student to be included in the sample is chosen randomly, for example by blindly picking one out of twelve pieces of paper, numbered 1 to 12. If number 6 is picked, then every twelfth student will be included in the sample, starting with student number 6, until 100 students are selected: the numbers selected would be 6, 18, 30, 42, etc.

Systematic sampling is usually less time consuming and easier to perform than simple random sampling. However, there is a risk of bias, as the sampling interval may coincide with a systematic variation in the sampling frame. For instance, if we want to select a random sample of days on which to count clinic attendance, systematic sampling with a sampling interval of 7 days would be inappropriate, as all study days would fall on the same day of the week (e.g., Tuesdays only, which might be a market day).

(3) Stratified sampling

The simple random sampling method described above has as disadvantage that small groups in which the researcher is interested may hardly appear in the sample.

If it is important that the sample includes representative study units of small groups with specific characteristics (for example, residents from urban and rural areas, or different religious or ethnic groups), then the sampling frame must be divided into groups, or STRATA, according to these characteristics. Random or systematic samples of a pre-determined size will then have to be obtained from each group (stratum). This is called STRATIFIED SAMPLING.

Stratified sampling is only possible when we know what proportion of the study population belongs to each group we are interested in.

An advantage of stratified sampling is that we can take a relatively large sample from a small group in our study population. This allows us to get a sample that is big enough to enable us to draw valid conclusions about a relatively small group without having to collect an unnecessarily large (and hence expensive) sample of the other, larger groups. However, in doing so, we are using unequal sampling fractions and it is important to correct for this when generalising our findings to the whole study population.

For example, a survey is conducted on household water supply in a district comprising 20,000 households, of which 20% are urban and 80% rural. It is suspected that in urban areas the access to safe water sources is much more satisfactory. A decision is made to include 100 urban households (out of 4000, which gives a 1 in 40 sample) and 200 rural households (out of 16000, which gives a 1 in 80 sample). Because we know the sampling fraction for both strata, the access to safe water for all the district households can be calculated after the study (by multiplying the findings for the urban households by 40 and those for the rural households by 80, and then calculating statistics for the total sample).

(4) Cluster sampling

It may be difficult or impossible to take a simple random sample of the units of the study population at random, because a complete sampling frame does not exist. Logistical difficulties may also discourage random sampling techniques (e.g., interviewing people who are scattered over a large area may be too time-consuming). However, when a list of groupings of study units is available (e.g., villages or schools) or can be easily compiled, a number of these groupings can be randomly selected.

The selection of groups of study units (clusters) instead of the selection of study units individually is called CLUSTER SAMPLING.

Clusters are often geographic units (e.g., districts, villages) or organisational units (e.g., clinics, training groups).

For example, in a study of the knowledge, attitudes and practices (KAP) related to family planning in rural communities of a region, a list is made of all the villages. Using this list, a random sample of villages is chosen and all study units in the selected villages are interviewed.

(5) Multi-stage sampling

In very large and diverse populations sampling may be done in two or more stages. This is often the case in community-based studies, in which people are to be interviewed from different villages, and the villages have to be chosen from different areas. This type of sampling is frequently used in HSR.

For example, in a study of utilisation of pit latrines in a district 150 homesteads are to be visited for interviews with family members as well as for observations on types and cleanliness of latrines. The district is composed of 6 wards and each ward has between 6 and 9 villages.

The following four-stage sampling procedure could be performed*:

1. Select 3 wards out of the 6 by simple random sampling.

2. For each ward select 5 villages by simple random sampling (15 villages in total).

3. For each village select 10 households. Since simply choosing households in the centre of the village would produce a biased sample, the following sampling procedure is proposed:

   - Go to the centre of the village.

   - Choose a direction in a random way: spin a bottle on the ground and choose the direction the bottleneck indicates.

   - Walk in the chosen direction and select every (or, depending on the size of the village, every second or every third) household until you have the 10 you need. If you reach the boundary of the village and you still do not have 10 households, return to the centre of the village, walk in the opposite direction and continue to select your sample in the same way until you have 10. If there is nobody in a chosen household, take the next nearest one.

4. Decide beforehand whom to interview (for example the head of the household, if present, or the oldest adult who lives there and who is available).

* This is an adaptation of the method developed by the EPI division in WHO Geneva to measure EPI coverage in districts.

A MULTI-STAGE SAMPLING procedure is carried out in phases and it usually involves more than one sampling method.

The main **advantages** of cluster and multi-stage sampling are that:

- A sampling frame of individual units is not required for the whole population. Existing sampling frames of clusters are sufficient. Only within the clusters that are finally selected is there a need to list and sample the individual units (if not using the bottle spinning method).

- The sample is easier to select than a simple random sample of similar size, because the individual units in the sample are physically together in groups, instead of scattered all over the study population.

The main **disadvantage** of this type of sampling is that:

Compared to simple random sampling, there is a larger probability that the final sample will not be representative of the total study population. The likelihood of the sample not being representative depends mainly on the number of clusters that is selected in the first stage. The larger the number of clusters, the greater is the likelihood that the sample will be representative. Further, the sampling units at community level should be selected randomly (avoid convenience sampling!).

3. Bias in sampling

BIAS in sampling is a systematic error in sampling procedures, which leads to a distortion in the results of the study.

Module 10 discussed how the use of faulty data collection tools would lead to biased results. Bias can also be introduced as a consequence of **improper sampling procedures,** which result in the sample not being representative of the study population.

For example, a study was conducted to determine the health needs of a rural population in order to plan primary health care activities. However, a nomadic tribe, which represented one third of the total population, was left out of the study. As a result the study did not give an accurate picture of the health needs of the total population.

There are several possible sources of bias that may arise when sampling. The most well known source is **non-response**.

Non-response can occur in any interview situation, but it is mostly encountered in large-scale surveys with self-administered questionnaires. Respondents may refuse or forget to fill in the questionnaire. The problem lies in the fact that non-respondents in a sample may exhibit characteristics that differ systematically from the characteristics of respondents.

There are several ways to deal with this problem and reduce the possibility of bias:

- Data collection tools (including written introductions for the interviewers to use with potential respondents) should be pre-tested. If necessary, adjustments should be made to ensure better co-operation.

- If non-response is due to absence of the subjects, follow-up of non-respondents may be considered.

- If non-response is due to refusal to co-operate, an extra, separate study of non-respondents may be considered in order to identify to what extent they differ from respondents.

- Another strategy is to include additional people in the sample, so that non-respondents who were absent during data collection can be replaced. However, this can only be justified if their absence was very unlikely to be related to the topic being studied.

Note:

The bigger the non-response rate, the more necessary it becomes to take remedial action. It is important in any study to mention the non-response rate and to honestly discuss whether and how the non-response might have influenced the results.

Other sources of bias in sampling may be less obvious, but at least as serious:

- **Studying volunteers only.** The fact that volunteers are motivated to participate in the study may mean that they are also different from the study population on the factors being studied. Therefore it is better to avoid using non-random selection procedures that introduce such an element of choice.

- **Sampling of registered patients only.** Patients reporting to a clinic are likely to differ systematically from people seeking alternative treatments.

- **Missing cases of short duration.** In studies of the prevalence of disease, cases of short duration are more likely to be missed. This may mean missing fatal cases, cases with short illness episodes and mild cases.

- **Seasonal bias.** It may be that the problem under study, for example, malnutrition, exhibits different characteristics in different seasons of the year. For this reason, data should be collected on the prevalence and distribution of malnutrition in a community during all seasons rather than just at one point in time. When investigating health services' performance, to take another example, one has to consider the fact that towards the end of the financial year shortages may occur in certain budget items which may affect the quality of services delivered.

- **Tarmac bias.** Study areas are often selected because they are easily accessible by car. However, these areas are likely to be systematically different from more inaccessible areas.

4. Ethical considerations

If the recommendations from a study will be implemented in the entire study population, one has the ethical obligation to draw a sample from this population in a representative way. If during the research new evidence suggests that the sample was not representative, this should be mentioned in any publication concerning the study, and care must be taken not to draw conclusions or make recommendations that are not justified.

GROUP WORK, PART I (2 hours)

1. Develop in your working group:
   - a definition of your (different) study population(s);
   - a definition of your (different) study units (people, clinics, records, etc);
   - appropriate sampling procedures for your study, taking into account whether you use qualitative and/or quantitative research methods. State how you will try to avoid possible bias.
2. Prepare a summary on a flipchart for use in the exercise 'Commenting on each others sampling procedures' and in the plenary discussion (after group work on sample size).

III. SAMPLE SIZE

Having decided **how** to select our sample, we now have to determine our **sample size**.

1. Sample size in qualitative studies

There are no fixed rules for sample size in qualitative research. The size of the sample depends on WHAT you try to find out, and from what different informants or perspectives you try to find that out.

> **For example**, if you want to explore how you can involve mothers in your HC catchment area more effectively in early detection and treatment of pneumonia, you might decide to conduct some FGDs to assess mothers' knowledge, attitudes and practices with respect to pneumonia. You could start with two FGDs among lowly educated mothers and two among mothers with more education (who usually are of higher socio-economic status). If the different data sets reconfirm each other you may stop at this point and start a small scale intervention; otherwise you conduct one or two FGDs more till you reach the point of *redundancy**: no new data comes up any more.
>
> If your research objective is more complex e.g., attitudes of males and females towards family planning, and has policy implications for a larger area, your sample will be bigger. You might start with four FGDs, two among males and two among females, subdivided according to socio-economic status. Among male participants you could then select 5-10 users (or spouses of female users) for in-depth interviews on RH history of the couple and reasons for use and non-use of FP. For women, the same procedures can be followed. If possible, you may interview some couples, first separately, then together. Depending on ethnic differences and urban-town differences in attitudes and practices, the 'clusters' of FGDs and in-depth interviews can be expanded.

In exploratory studies, the sample size is therefore **estimated** beforehand as precisely as possible, but **not determined.**

Patton (1990: 183-186) stresses that richness of the data and analytical capability of the researcher determine the validity and meaningfulness of qualitative data more than sample size. Still, sampling **procedures** and **sample size** should always be carefully explained in order to avoid the allusion of haphazardness. Careful analysis of different complementing data sets can result in some plausible generalisations but without 'proving' them in a mathematical sense.

---

* Lincoln and Guba, in Patton (1990: 185)

2. Sample size in quantitative studies

For quantitative studies, calculations can be made which indicate the desirable sample size. The principles of such calculations will be discussed below.

It is a widespread belief among researchers that the bigger the sample, the better the study becomes. This is **not necessarily true**. In general it is much better to increase the **accuracy and richness** of data collection (for example by improving the training of interviewers or by better pre-testing of the data collection tools) than to increase sample size after a certain point. Also, it is better to make extra efforts to get a **representative** sample rather than to get a very large sample.

The following general rules may help to determine the **desirable sample size** of any given study:

- The desirable sample size depends on the *expected variation in the data* (of the most important variables): the more varied the data are, the larger the sample size we would need to attain the desired level of accuracy.

- The desirable sample size also depends on the *number of cells* we will have in the cross-tabulations (see **Module 13**) which we need to analyse the results. A rough guideline is to have at least 5 to 10 study units per cell.

For example, after conducting FGDs and in-depth interviews in the study on attitudes of men and women towards family planning (see Section III.1 of this module) you might decide to conduct a bigger survey. If your exploratory study revealed that age and education appear to be important factors determining FP use, you will be interested in comparing FP use in groups with different levels of education and of different ages. If you split each of these variables up in three categories, and you select four categories of informants (male users/ spouses of female users; female users; male non-users, female non-users) you would have 12 cells in each table. In order to obtain 5-10 answers per cell you would require 60-120 informants in each research area.

As other variables may have more categories, you may attempt to select 120 informants in rural areas and 120 in urban areas. However, as FP use may not be equally distributed in a population (e.g., 25% users, 75% non-users) your sample will have to be bigger than 120 in order to obtain the desired sample size. HOW big it should be can be calculated. Still, the desirable sample size can not always be achieved for lack of resources such as time, manpower and money. This constraint applies to quantitative as well as qualitative studies.

Therefore the eventual sample size is usually a compromise between what is DESIRABLE and what is FEASIBLE.

Sample size calculations

In quantitative studies, researchers will perform **sample size calculations** before embarking on the project to find the desirable sample size. The formulae for calculating a desired sample size are listed in **Annex 11.2.** They are divided into two categories, depending on whether the study:

- seeks to measure **one single variable** (e.g. a mean, a rate or a proportion) in one group with a certain precision, or

- tries to demonstrate a **significant difference** between two groups.

The formulae can only be used if you have a rough idea about the outcome of the study, which is not always the case. It is always advisable to call upon a statistician or an experienced researcher who can help you in choosing and using the appropriate formulae.

We will look at a few examples to highlight some important issues.

(1) Descriptive studies with one group

Example

In a descriptive study in a certain village we want to measure, with a certain precision, the proportion of children aged 12-23 months who are immunised against measles, using a simple random sample. The following steps should be taken:

1. Estimate how big the proportion might be (say 80%).

2. Choose the **margin of error** you will allow in the estimate of the proportion (say ± 10%). This means that, if the survey reveals that indeed 80% of the children have been vaccinated, this proportion will probably be between 70 and 90% in the whole study population from which the sample was drawn.

3. Choose the **level of confidence** at which you want to be able to state that the vaccination coverage in the whole population is indeed between 70 and 90%. You can never be 100% sure. Do you want to be 95% sure? or 99%? A commonly used confidence level in HSR studies is 95%.

   The formula for calculating the sample size for the single proportion expressed as a percentage is presented in **Annex 11.2 (1.3)**.

However, at present it is usually not necessary to calculate the desirable sample size by hand. There are computer programmes to assist us (for example Epi Info). In **Annex 11.3** two tables are presented, which will help you to calculate the desired sample size in the most common quantitative HSR studies. Both take a confidence level of 95% ($P<0.05$) as point of departure. In the **measles example**, you have to go down the first column till you find the 80% for the estimated vaccination coverage. If you allow a margin of error of 10 + or – the estimated 80%, you look in that column and find a desired sample size of 64. However, if you would allow a margin of error of only + or - 5%, the sample size would increase to 256, and if you would like to be 95% sure that the measles vaccination coverage is between 79 and 81%, you would need to check 6400 children. The smaller your margin of error (also called confidence interval) the bigger your sample size needs to be. If you would like to increase your confidence level from 95 to 99%, the sample would have to increase even more. But then researchers of course start to look at feasibility: do we really need that precision? Most HSR studies will be satisfied with the 95% confidence level.

Note also that in general you need more precision (or a smaller margin of error) if the estimated proportion is very small. This may be the case, for example, for the proportion of HIV* women or the maternal mortality rate in a population.

Table 11.1: Required sample size for studies of HIV prevalence in pregnant women

	Estimated proportion HIV^+ in women	**Margin of error (at 95% confidence level)**	**Required sample size**
District A	1%	± 0.5%	1584
District B	10%	± 5%	144

The table shows that in district A, where HIV is less prevalent, a smaller margin of error is desired and, therefore, the required sample is larger. (**Annex 11.4** explains how these sample sizes were calculated).

Table (a) in **Annex 11.3** will, however, help you again to identify the required sample sizes in a simple way. The first line of table (a) gives you the required sample sizes for an estimated prevalence of 1%. You would need a sample of 1584 women to be 95% sure that HIV prevalence in District A would be between 0.5 and 1.5% (so 1% plus or minus 0.5%). For district B, with 10% estimated HIV prevalence, and a margin of error plus or minus 5%, the required sample size would be 144.

(2) Comparing two groups for a significant difference

In comparative studies one usually wants to demonstrate that there is a **significant difference** between two groups. In this type of study the sample size depends primarily on the **estimated size of the difference** between the two groups that are compared. **The larger the difference, the smaller will be the sample** that is needed to show this difference. Second, the sample size depends on how large we want the **probability** to be that we indeed will find a significant difference.

The larger the sample size, the larger will be the probability of finding a significant difference. In the case of many variables, the one with the smallest estimated difference between the groups should be used as the basis for calculating the sample size, as it requires the largest sample. Still, researchers will discuss whether that variable is important enough to maintain if the sample would become too big to handle. They may also decide to measure only some variables in a large sample while measuring variables that require a smaller sample in a *sub-sample* to save costs and ensure the *feasibility* of the study.

Example:

In a study, a comparison will be made between the feeding patterns of well-nourished and malnourished children of 12 to 17 months. It is expected that of the well-nourished children 90% are breastfed whereas of the malnourished children approximately 50% are breastfed. The sample size in each group of children needs to be at least 15 to show a significant difference.

However, if 90% of the well-nourished children and 80% of the malnourished children were breast fed, the sample size would need to be at least 175 in each group to show a significant difference. (**Annex 11.4** explains how these sample sizes can be calculated).

Table (b) in **Annex 11.3** will again help you to identify the required sample size in a simple way. Compare 90% in the column with 50% in the row (or vice versa) and you will find a required sample size of 22. In this case a slightly different formula was used as the one presented in presented in **Annex 11.4**. Also the calculated sample size for the second example would be bigger (262) when using the table.

Note that it may be useful to conduct sample size calculations for each of the objectives of the study. These calculations may reveal, for instance, that some but not all objectives can be met. Or they may indicate that some variables need only to be measured on a sub-sample.

GROUP WORK, PART II (1 hour)

1. Determine the sample size requirements for the study population(s) defined in the previous group work session. Consider the issues discussed in the module when establishing the desirable sample size (s). Use the tables in **Annex 11.3** if sample size calculations have to be made.

2. Determine the feasible sample size after taking into account available time, manpower, transport and money.

   If there is a large discrepancy between the desirable and the feasible sample size you should look for a compromise and, if necessary, adjust the objectives of your study.

3. Put a summary of your group's work on flipchart for use in the exercise below and in the plenary discussion that will follow.

4. It will be easier for you to develop a realistic **Plan for Data Collection (Module 12)** as well as the **Budget** for your project (**Module 16**) if you already know as precisely as possible **where** and **from whom** data are going to be collected. Therefore it is important that you select the sample for your study **immediately** after the plenary session on sampling. If you choose a multi-stage sampling strategy you may find at this stage that it is only possible to partially draw the sample.

EXERCISE (½ hour)

1. Examine the definitions of study population and study units, the sampling procedures and the proposed sample size developed by another group.

2. Identify possible sources of bias in sampling and suggest improvements.

3. Put your comments on flipchart for presentation in plenary.

REFERENCES

All epidemiological and social science research handbooks mentioned in the references in **Module 9** are dealing with sampling procedures and sample size. Moreover you can consult:

Swinscow TDV, Revised by MJ Campbell (1998) *Statistics at Square One.* London: BMJ Publishing Group. (9th ed.)

Campbell MJ, Machin D (1993) *Medical Statistics: A Common Sense Approach.* Clichester: John Wiley. (2nd ed.)

Annex 11.1: How to use a random number table*

1. First, decide how large a number you need. Next, count if it is a one, two or larger digit number. For example, if your sampling frame consists of 10 units, you must choose from numbers 1-10, (inclusive). You must use **two** digits to ensure that 10 has an equal chance of being included.

   You also use two digits for a sampling frame consisting of 0-99 units.

   If, however, your sampling frame has 0-999 units, then you obviously need to choose from **three** digits. In this case, you take an extra digit from the table to make up the required three digits. For example, the number in columns 10,11, row 27: 43, would become 431; going down, the next numbers would be 107, 365 etc.

   You would do the same if you needed a **four**-digit number, for a sampling frame 0-9999 units. In our example of the number on columns 10, 11, 12, row 27 of the table: 431, this would now become 4316, the next down 1075, and so on.

2. Decide beforehand whether you are going to go across the page to the right, →, down the page ↓, across the page to the left ←, or up the page. ↑

3. Without looking at the table, and using a pencil, pen, stick, or even your finger, pin-point a number.

4. If this number is within the range you need, take it. If not, continue to the next number in the direction you chose before-hand, (across, up or down the page), until you find a number that is within the range you need.

   For example if you need a number between 0-50 and you began at column 21, 22, row 21 you get 74 which is obviously too big. So you could go down (having decided beforehand to go down) to 97, also too big, to 42, which is acceptable, and select it.

* The random number table on the following page has been taken from Hill AB (1977) *A Short Textbook of Medical Statistics*. London: Hodder and Stoughton, 1977:306-7.

Random sampling numbers

	1 2 3 4	5 6 7 8	9 10 11 12	13 14 15 16	17 18 19 20	21 22 23 24	25 26 27 28	29 30 31 32
1	8 0 9 4	2 5 2 5	8 2 4 7	1 3 4 7	7 4 3 3	3 6 2 0	1 8 9 7	2 1 3 4
2	3 5 6 3	2 1 9 8	8 2 1 1	9 0 4 5	2 6 1 8	2 7 5 1	2 6 2 7	1 0 9 5
3	1 3 3 0	6 3 3 1	3 7 5 3	9 6 9 3	8 7 3 8	6 8 1 5	1 5 3 8	8 5 4 3
4	3 5 6 5	0 0 1 6	2 2 4 3	6 4 3 2	4 7 9 6	6 0 9 5	5 2 8 3	1 6 2 0
5	7 8 5 0	5 9 2 5	5 5 8 8	7 3 1 1	2 1 9 2	4 5 4 5	3 5 3 0	5 5 8 9
6	4 4 9 0	5 4 1 7	9 7 2 7	6 1 5 3	5 9 0 1	4 8 7 8	9 9 8 0	9 8 7 7
7	6 5 4 5	9 1 0 4	9 3 1 8	8 8 1 9	7 5 3 7	2 7 8 5	9 3 7 3	2 4 4 5
8	3 6 2 8	5 9 9 5	1 2 1 5	9 7 5 3	9 2 2 3	5 6 5 8	2 9 4 4	2 8 9 9
9	4 6 6 5	4 8 2 0	7 5 5 4	0 6 1 2	9 6 8 3	4 2 5 1	9 1 3 8	1 7 0 9
10	6 4 9 8	7 5 1 9	0 4 7 4	7 8 1 8	6 8 3 2	9 6 8 3	9 8 7 2	4 0 9 0
11	6 7 2 2	9 8 6 9	9 3 6 1	7 8 7 5	4 8 8 3	1 3 1 5	9 6 7 9	8 8 3 4
12	9 7 4 8	5 9 3 2	5 1 1 5	2 7 2 1	0 0 3 3	9 3 0 3	9 7 1 3	4 0 1 2
13	5 6 4 1	1 4 1 7	1 4 1 9	7 4 3 4	8 1 6 5	7 3 6 8	1 2 1 8	5 0 3 9
14	7 4 4 4	9 2 0 0	8 8 4 0	5 8 8 2	4 3 9 8	3 9 0 4	9 1 9 9	9 3 3 6
15	8 2 7 9	3 0 1 9	4 6 7 2	3 7 4 3	3 9 7 9	4 6 8 9	9 0 2 1	6 9 9 0
16	0 1 6 1	7 6 1 7	1 0 2 4	2 3 8 7	2 8 9 1	6 6 7 7	1 5 8 5	2 4 8 2
17	7 3 8 8	9 7 5 9	7 5 5 5	6 6 2 4	9 9 7 7	2 0 0 8	5 5 9 6	9 7 4 0
18	7 8 3 0	4 7 1 4	3 6 9 5	2 9 1 9	1 8 0 4	4 0 4 4	1 0 3 4	2 5 9 7
19	9 8 8 7	4 2 1 6	6 5 2 6	4 5 3 5	8 4 3 0	5 2 7 0	9 6 0 5	0 7 6 8
20	1 2 6 1	2 5 1 6	8 5 6 9	2 3 1 0	3 9 3 9	8 7 0 3	9 8 4 1	0 3 5 3
21	3 9 4 7	4 9 3 7	7 6 3 4	2 5 4 3	6 2 3 9	7 4 5 5	2 0 5 5	7 7 9 5
22	4 5 5 0	8 1 0 3	1 2 5 0	2 3 0 4	1 1 3 8	9 7 8 8	9 1 4 4	4 5 2 6
23	1 3 4 4	9 6 9 7	2 3 8 3	6 9 7 6	6 2 5 1	4 2 0 1	2 0 3 8	6 5 5 2
24	8 9 7 6	5 8 2 3	8 4 8 7	0 4 5 0	3 1 0 6	9 1 6 6	2 7 1 7	7 6 0 1
25	7 7 1 0	9 9 4 3	6 9 7 8	8 2 7 3	9 7 1 4	9 7 0 0	1 5 6 6	2 8 8 9
26	6 9 5 9	6 0 0 8	8 4 4 2	2 2 8 2	1 5 2 4	2 5 1 7	5 8 1 8	0 0 8 1
27	7 9 4 1	2 3 1 2	2 4 3 1	6 7 0 2	9 9 8 4	3 4 6 9	3 0 8 5	4 7 6 2
28	2 2 8 4	0 8 9 6	9 1 0 7	5 5 4 2	7 3 1 9	3 7 8 2	1 0 6 8	9 5 7 4
29	9 5 9 4	7 4 1 6	9 3 6 5	6 0 4 5	1 1 8 3	5 9 1 6	9 5 9 9	1 1 4 3
30	4 6 1 3	8 5 4 9	6 3 6 9	3 2 0 8	5 1 0 9	9 6 8 0	1 1 6 8	6 1 3 3

Annex 11.2: Formulae for calculating sample size*

The formulae for calculating required sample size are divided in two categories:

1. For studies trying to measure one variable with a certain precision.

2. For studies seeking to demonstrate a significant difference between two groups.

1. Measuring one variable

In the formulae below the following abbreviations are used:

n - sample size
s - standard deviation
e - required size of standard error
(in the text of the module the term 'margin of error' is used for ±2 times the size of the standard error if a precision of 95% is required)
r - rate
p - percentage

1.1 Single mean

In a study the mean weight of newborn babies will be determined. The mean weight is expected to be 3000 grams. Weights are approximately normally distributed and 95% of the birth weights are probably between 2000 and 4000 grams; therefore the standard deviation would be 500 grams. The desired 95% confidence interval is 2950 to 3050 grams, so the standard error would be 25 grams. The required sample size would be:

$$n = \frac{s^2}{e^2} = \frac{500^2}{25^2} = \frac{250000}{625} = 400 \text{ new born babies}$$

1.2 Single rate

The maternal mortality rate in a country is expected to be 70 per 10,000 live births. A survey is planned to determine the maternal mortality rate with a 95% confidence interval of 60 to 80 per 10,000 live births. The standard error would therefore be 5/10,000. The required sample size would be:

$$n = \frac{r}{e^2} = \frac{70/10000}{(5/10000)^2} = 28.000 \text{ live births}$$

1.3 Single proportion

The proportion of nurses leaving the health services within three years of graduation is estimated to be 30%. A study that aims to find causes for this, also aims to determine the percentage leaving the service with a confidence interval of 25% to 35%. The standard error would therefore be 2.5%. The required sample size would be:

$$n = \frac{p\,(100 - p)}{e^2} = \frac{30 \times 70}{2.5^2} = 336 \text{ nurses}$$

* Modified from Kirkwood B (1988) *Essentials of Medical Statistics*. Oxford: Blackwell Scientific Publications, 1988.

1.4 Difference between two means (sample size in each group)

The difference of the mean birth weights in district A and B will be determined. In district A the mean is expected to be 3000 grams with a standard deviation of 500 grams (as in 1.1). In district B the mean is expected to be 3200 grams with a standard deviation of 500 grams. The difference in mean birth weight between districts A and B is therefore expected to be 200 grams. The desired 95% confidence interval of this difference is 100 to 300 grams, giving a standard error of the difference of 50 grams. The required sample size would be:

$$n = \frac{s_1^2 + s_2^2}{e^2} = \frac{500^2 + 500^2}{50^2} = 200 \text{ new-born in each district}$$

1.5 Difference between two rates (sample size in each group)

The difference in maternal mortality rates between urban and rural areas will be determined. In the rural areas the maternal mortality rate is expected to be 100 per 10,000 and in the urban areas 50 per 10,000 live births. The difference is therefore 50 per 10,000 live births. The desired 95% confidence interval is 30 to 70 per 10,000 live births giving a standard error of the difference of 10/10,000. The required sample size would be:

$$n = \frac{r_1 + r_2}{e^2} = \frac{100/10{,}000 + 50/10{,}000}{(10/10{,}000)^2} = 15{,}000 \text{ live births in each area}$$

1.6 Difference between two proportions (sample size in each group)

The difference in the proportion of nurses leaving the service is determined between two regions. In one region 30% of the nurses are estimated to leave the service within three years of graduation, in the other region 15%, giving a difference of 15%. The desired 95% confidence interval for this difference is 5% to 25%, giving a standard error of 5%. The sample size in each group would be:

$$n = \frac{p_1 (100 - p_1) + p_2 (100 - p_2)}{e^2}$$

$$= \frac{30 \times 70 + 15 \times 85}{5^2} = 135 \text{ nurses in each region}$$

2. Significant difference between two groups

In the formulae below the following abbreviations are used:

n, samples size
s, standard deviation
e, required size of standard error
r, rate
p, percentage
u, one-sided percentage point of the normal distribution, corresponding to 100% - the power. The power is the probability of finding a significant result. (e.g. if the power is 75%, u = 0.67).
v, percentage point of the normal distribution, corresponding to the (two-sided) significance level. (e.g. if the significant level is 5% (as usual), v = 1.96.)

2.1 Comparison of two means (sample size in each group)

The birth weights in district A and B will be compared. In district A the mean birth weight is expected to be 3000 grams with a standard deviation of 500 grams. In district B the mean is expected to be 3200 grams with a standard deviation of 500 grams (see 1.4). The required sample size to demonstrate (with a likelihood of 90%) a significant difference between the mean birth weights in district A and B would be:

$$n = \frac{(u + v)^2 (s_1^2 + s_2^2)}{(m_1 - m_2)^2}$$

$$= \frac{(1.28 + 1.96)^2 (500^2 + 500^2)}{(3200 - 3000)^2} = 131 \text{ new-born babies in each district}$$

2.2 Comparison of two rates (sample size in each group)

The maternal mortality rates in urban and rural areas will be compared. In the rural areas the maternal mortality rate is expected to be 100 per 10,000 and in the urban areas 50 per 10,000 live births (compare to 1.5). The required sample size to show (with a likelihood of 90%) a significant difference between the maternal mortality in the urban and rural areas would be:

$$n = \frac{(u + v)^2 (r_1 + r_2)}{(r_1 - r_2)^2}$$

$$= \frac{(1.28 + 1.96)^2 (100/10{,}000 + 50/10{,}000)}{(100/10{,}000 - 50/10{,}000)^2} = 6299 \text{ live births in each area}$$

2.3 Comparison of two proportions (sample size in each group)

The proportion of nurses leaving the health service is compared between two regions. In one region 30% of nurses is estimated to leave the service within three years of graduation, in the other region it is probably 15%.

The required sample size to show with a 90% likelihood that the percentage of nurses is different in these two regions would be:

$$n = \frac{(u + v)^2 \{p_1 (100 - p_1) + p_2 (100 - p_2)\}}{(p_1 - p_2)^2}$$

$$= \frac{(1.28 + 1.96)^2 (30 \times 70 + 15 \times 85)}{(30 - 15)^2} = 157 \text{ nurses in each group}$$

Annex 11.3: Tables for calculating sample size in the most common HSR studies

(a) Sample size for measuring proportions in one group

Percentage 95% confidence level (+ or – the percentage of allowed error in column heading)

	0.50%	1%	3%	5%	10%	20%
1%	1584					
5%	7600		1900	211		
10%	14400	3600	400	144		
15%	20400	5100	567	204		
20%	25600	6400	711	256	64	
25%	30000	7500	833	300	75	19
30%	33600	8400	933	336	84	21
35%	36400	9100	1011	364	91	23
40%	38400	9600	1067	384	96	24
45%	39600	9900	1100	396	99	25
50%	40000	10000	1111	400	100	25
55%	39600	9900	1100	396	99	25
60%	38400	9600	1067	384	96	24
65%	36400	9100	1011	364	91	23
70%	33600	8400	933	336	84	21
75%	30000	7500	833	300	75	19
80%	25600	6400	711	256	64	
85%	20400	5100	567	204		
90%	14400	3600	400	144		
95%	7600	1900	211			
99%	1584					

The percentages in the column heading are 2 x standard error (e in formula)

(b) Sample size for comparison of proportions in two groups

	Percentage 1									
	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Percentage 2										
5%	577	97	43	25	15	10				
10%		262	79	38	22	14				
15%	913	1207	157	62	32	19	12			
20%	262		388	105	48	26	16			
25%	129	1459	1669	199	73	37	21	12		
30%	79	388		472	121	52	28	16		
35%	53	181	1837	1963	223	79	37	20	11	
40%	38	105	472		514	126	52	26	14	
45%	29	68	213	2047	2089	227	77	35	17	
50%	22	48	121	514		514	121	48	22	10
55%	17	35	77	227	2089	2047	213	68	29	13
60%	14	26	52	126	514		472	105	38	16
65%	11	20	37	79	223	1963	1837	181	53	19
70%		16	28	52	121	472		388	79	24
75%		12	21	37	73	199	1669	1459	129	31
80%			16	26	48	105	388		262	42
85%			12	19	32	62	157	1207	913	59
90%				14	22	38	79	262		94
95%				10	15	25	43	97	577	199
100%					10	16	24	42	94	

$u = 1.28$ Power = 90%
$v = 1.96$ Significance: $P < 0.05$
$(u+v)^2 = 10.5$

Annex 11.4: Explanation of sample size calculations given in the text

1. Prevalence of HIV (p.13)

District A: the estimated HIV$^+$ **proportion** is 1% = 0.01. As the 95% confidence interval is the proportion ± 2 x the **standard error**, the standard error is 0.25% = 0.0025.

$$n = \frac{0.01 \times 0.99}{(0.0025)^2} = 1584$$

District B: the estimated HIV **proportion** is 10% = 0.1
Procedures followed are the same as above

2. Feeding patterns in malnourished and well-nourished children (p.14)

Formula used: no. 2.3 in Annex 11.2

$$n = \frac{(u + v)^2 \{p_1 (100 - p_1) + p_2 (100 - p_2)\}}{(p_1 - p_2)^2}$$

$$n_1 = \frac{(0.67 + 1.96)^2 (10 \times 90 + 50 \times 50)}{(90 - 50)^2} = 15$$

$$n_2 = \frac{(0.67 + 1.96)^2 (10 \times 90 + 20 \times 80)}{(90 - 80)^2} = 173$$

If the power is 75%, u = 0.67 and $(u + v)^2 = 6.9$;
if the power is 90%, u = 1.28 and $(u + v)^2 = 10.5$.
(The power is the probability of finding a significant result).

If the power is increased from 75% to 90%, the sample size is increased 10.5/6.9 (i.e. 1.5 times.)

Trainer's Notes

Module 11: SAMPLING

Timing and teaching methods

The topic on sampling has two major components (sampling procedures and sample size), which preferably should be presented in two separate sessions. These sessions will require 6½ hours in total.

Materials

- Calculators
- Paper

Introduction to Sampling Procedures (Part II of Module 11)

Timing and teaching methods

1 hour Introduction and discussion
2 hours Group work

Introduction and discussion

- When presenting part I of this module make sure that everyone understands what sampling is and why it is done. Carefully explain the merits of purposeful sampling in small-scale qualitative studies and of random sampling in quantitative studies.

- In the presentation of sampling methods (Part II), use examples from the groups' own protocols as much as possible. You may do an exercise to show the differences between the different sampling methods with the participants themselves as a group.

  For example, you may sample 6 or 8 persons from your audience using simple random sampling and systematic sampling (from an alphabetical list of participants and facilitators). Ask the participants to name the sampling method applied and discuss the advantages and disadvantages of each method. (As names tend to cluster according to origin, it is likely that the systematic sampling will turn out to be less representative than the simple random sampling.)

- Allow time during and after the presentation for questions and discussion.

Group work, Part I

- Have the working groups choose the appropriate sampling methods for their own projects. The methods should be worked out in as much detail as possible.

Introduction to Sample Size (Part III of Module 11)

Timing and teaching methods

½ hour	Introduction and discussion
1 hour	Group work
½ hour	Exercise
1-½ hours	Group reports in plenary

Introduction and discussion

- Stress that one does not always have to do calculations to determine the desired sample size. Actually, in many (exploratory) HSR studies one you would not do any calculations, though for the sampling procedures a plan must be worked out which should be adhered to (e.g. selection of extremes from a list of patients, or snowball sampling through different key informants according to criteria set together).

- The formulae for calculating a desired sample size are therefore put in an annex. You are not expected to go into technical details of sample size calculations during your presentation unless the participants are familiar with statistics (all relevant concepts will only be explained during the data analysis workshop). Rather use the tables in **Annex 11.3** and make sure participants understand from those tables how sample size and precision (confidence level) go together.

Group work, Part II

- Let each group determine the sample size for the proposal it is working on.

- Participants should be advised to consult experts when they think they will need to calculate sample size but do not know how to go about it. Make sure, for this reason, that there is a statistician present who can be consulted during group work and plenary presentations.

- If a group plans to measure statistical entities such as infant or maternal mortality rates in its study, it should definitely consult a professional with statistical training.

Exercise

- At the end of this group work session each group should examine another group's chosen sampling procedures and sample size. Ask the groups to look for possible sources of bias and make suggestions for reducing it.

Plenary

- Have each group present their sampling methods and sample size, immediately followed by the comments of the group that examined the sampling methods for bias. A discussion can follow each presentation or be held after all the group presentations.

- Emphasise that, after having incorporated useful suggestions from the plenary discussion, the groups should actually select their samples, as far as possible (e.g. sampling of districts, villages, clinics). This will be useful for the next group work sessions in preparation of the fieldwork (especially **Modules 12** and **16**).

Module 18

FINALISING AND REVIEWING THE RESEARCH PROPOSAL

Flowchart: Steps in the development of a health systems research proposal

Questions you must ask	Steps you will take	Important elements of each step
What is the problem and why should it be studied?	Selection, analysis and statement of the research problem	- problem identification - prioritising problems - analysis - justification
What information is available?	Literature review	- literature and other available information
Why do we want to carry out the research? What do we hope to achieve?	Formulation of research objectives	- general and specific objectives - hypotheses
What additional data do we need to meet our research objectives? How are we going to collect this information?	Research methodology	- variables - types of study - data collection techniques - sampling - plan for data collection - plan for data processing and analysis - ethical considerations - pre-test or pilot study
Who will do what, and when?	Work plan	- human resources - timetable
What resources do we need to carry out the study? What resources do we have?	Budget	- material support and equipment - money
How will the project be administered? How will utilisation of results be ensured?	Plan for project administration and utilisation of results	- administration - monitoring - identification of potential users
How will we present our proposal to relevant authorities, community and the funding agencies?	Proposal summary	- briefing sessions and lobbying

NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 18: FINALISING AND REVIEWING THE RESEARCH PROPOSAL

OBJECTIVES

At the end of the session you should be able to:

1. **Finalise the research proposal** for the presentation to the relevant authorities
2. **Write a brief summary** of the completed research proposal
3. **Prepare** a letter of intent to send to potential funding agencies

I. Finalising the research proposal

II. Writing a summary of the research proposal

III. Presenting the research proposal to the relevant authorities and potential funding agencies

I. FINALISING THE RESEARCH PROPOSAL

When you have finished the methodological section of your research proposal and have pre-tested the methodology or at least reviewed it thoroughly (**Module 14)**, you can start preparing the final draft of various parts of your research proposal.

If relevant to the problem, you should first review again whether major cross-cutting issues have been addressed in a consistent way. These may include:

- Gender;
- Equity (e.g. in access to services and quality of care provided);
- Participation of the target group in formulating the problem, major findings (through giving feedback) and recommendations.

You should have considered these issues when analysing and prioritising your research problem (see **Module 4**); when developing your research objectives (see **Module 6**), when developing your data collection techniques (see **Module 10**), when selecting your study populations (see **Module 11**), and when planning for data analysis and utilisation of results, which should be reflected in your work plan and budget (**Modules 13, 15, 16, 17**).

You can start working on part I of **Module 18** any time after the group work of **Module 14** has been completed.
The outline of your research proposal, as presented in **Module 1**, is as follows:

SUMMARY

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Background information (context of the problem)
1.2 Statement of the problem
1.3 Literature review (partly or fully integrated into 1.1 and 1.2)

2. OBJECTIVES

3. METHODOLOGY

3.1 Study type, data collection techniques used.
Variables or research themes can either be annexed or (if succinct) included in the text after study type
3.2 Sample and sampling procedures
3.3 Plan for data collection
3.4 Plan for data processing and analysis
3.5 Ethical considerations
3.6 Pre-test

4. WORK PLAN (including description of project staff)

5. BUDGET (including explanatory note on major budget posts)

6. PLAN FOR ADMINISTRATION, MONITORING, AND UTILIZATION OF RESULTS

REFERENCES

ANNEXES

Annex 1. List of abbreviations, if applicable (can also go to the beginning of the proposal).
Annex 2. Interview guides/ questionnaires (and/or other data collection tools).

How should you proceed?

1. The first section of your proposal contains **background information, the statement of the problem** and **literature review**. This section should convince the reader of the relevance of the study (magnitude, severity of the problem). It should provide enough background data for an outsider to understand the different aspects of the problem, or the different factors influencing the problem and the context in which it occurs. Your review of available literature and reports should further illustrate why the problem is important, not only in your own working area but probably also beyond.

   Remember that, if your problem has a strong global dimension, (e.g., HIV/AIDS, family planning) you might consider a separate section for reviewing international literature as an introduction to your statement of the problem (see **Module 5**). However, in many cases (especially if you are focussing on smaller, local research problems) it will be possible to include all literature you consulted in the 'Background information' and 'Statement of the problem' sections.

> **Note:**
>
> Remember that review of literature is an ongoing activity that will continue during fieldwork, data analysis and even report writing.

   You can justify your study by pointing to the gaps in available information, which you hope to fill in with the data from your planned research. Finally, you can increase the interest of your readers by summarising what results you hope will emerge from your study and how you plan to use them to help solve or alleviate the problem on which your study concentrates.

You now have to thoroughly review the various pieces of text that you have produced during earlier sessions of the workshop, and rewrite them to form a coherent study design.

> **Note:**
>
> When developing your research methodology you may have somewhat revised the focus on your research problem; you may have become more specific as you proceeded, adding certain factors, or omitting others. Any adjustments should be included throughout the text of your proposal because all parts of the study should be consistent and logically connected to each other.
>
> **When revising your proposal you can best work backward**, from the data collection tools to the objectives and then to the statement of the problem, to check for consistency.

2. Critically review your **objectives**. Check whether they still cover what you have planned to study, and whether they are specific enough.

3. The next section presents the **methodology**. You have already prepared small sections focusing on various aspects of your methodology. You should now check the text for clarity of wording and logical coherence. (An outsider must be able to understand what you mean)

4. Discussion of various **ethical issues** affecting your study may be scattered in different parts of your draft. Identify the most important issues and discuss them in a separate section. (Include, for example, issues relative to the selection of your topic, your methodology, and the collection of your data, where you will have to pay attention to by asking consent and avoiding any harm to your informants.)

5. The last sections of the research proposal, which will focus on **project management,** the **work plan**, the **budget** and the **plan for administration, monitoring and utilisation of results,** are quite clear-cut. When writing them, consult the directions presented in the GROUP WORK sections of the respective modules.

6. **References** should follow in a separate section at the end of the research proposal. Take care that you list the references in a consistent way. (Author(s), year, title, place, publisher, etc., see **Module 5**.)

7. **Annexes**. You should add a list of abbreviations, if there are many, here or after the summary. Your data collection tools should be annexed, each with a number, so that you can easily refer in the text to the various instruments.

8. Finally, prepare a title page, summary, table of contents and, if appropriate, acknowledgements.

GROUP WORK

1. Prepare a final draft of your proposal following the guidelines presented above. It is advisable to work in groups of one or two persons, each with the responsibility for one or more sections.

   Take care that you number the sections, for example as in the outline presented in this module.

2. Two persons should be responsible for final editing. They should review and revise the text so that it flows smoothly from one section to the next.

3. All members of the group, including the facilitator, should have read all sections of the proposal before the final manuscript is handed in for typing.

4. The team leader should be responsible for co-ordinating the production of the final draft of the proposal.

   It is useful to prepare a list of all sections that have to be written (see Table of Contents) and make a note as group members go through each step in the production process.

For example:

	Gone for first processing	First draft ready	Discussed and revised	Gone for finalising	Finished
1. Introduction 1.1 Background information					
1.2 Statement of the problem					
2. Objectives					
Etc.					

II. WRITING A SUMMARY OF THE RESEARCH PROPOSAL

When you have completed writing your research proposal, there is usually a need for the protocol to be reviewed by senior authorities and policy makers or funding agencies. For the purpose of obtaining approval from policy makers or very busy administrators it is advisable to add a summary (of no more than two pages) to the proposal.

A summary usually includes:

One page containing essential information such as:

- **Title** of the research proposal
- **Duration** (proposed dates of onset and completion of the project)
- **Total budget** (in local currency and US$)
  - Contribution of Ministry of Health
  - Contribution of donor
  - Additional resources to be mobilised
- **Research team** (names and functions)
  - Team leader
  - Research team members
- **Name of the project administrator**

A brief narrative summary of one page that could contain the following elements:

- One paragraph on the **statement of the problem**
- **General objective**
- **Study populations, sample sizes** and **data collection techniques** used
- Indications concerning **what major results may be expected** from the study and their possible contribution for solving the problem being researched.

You should **put the summary at the beginning of the proposal**, although it is the last thing you prepare.

After the summary, a **table of contents** should follow. Adding numbers to the pages of your report and including them in your table of contents is one of the last activities involved in preparing your proposal.

Then a **title page** should be prepared, containing the title of your study, the names of the researchers with their titles, the name of the institution that has organised the course (Ministry of Health, or Health Research Unit of the Ministry of Health, for example) and date of issue. Add also this is a **research proposal** in order to distinguish it clearly from the research report that will later appear, probably under the same name and with the same authors. Finally you may add a page of **acknowledgements**, thanking all who enabled you to develop and implement this study.

III. PRESENTING THE RESEARCH PROPOSAL TO THE RELEVANT AUTHORITIES

Before a research project can be implemented, the health systems research proposal usually has to be:

- approved by the relevant health authorities,
- approved by the appropriate research committee or council, and
- given the funding.

In certain circumstances some of the above steps may be combined.

The procedure for approval may require that the research proposal be submitted with an accompanying letter or prescribed form for the relevant authority. In addition, the researchers may be requested to make a brief verbal presentation or 'defend' the proposal in person.

At the end of the proposal development workshop, participants will make 7-10 minutes presentations of their research proposals to a panel so they will gain an appreciation for the concerns of the various approving agencies and acquire the skills to respond briefly and succinctly to questions relating to particular aspects of the proposal.

Presentation to a panel

The panel should consist of experienced researchers who will comment on the research aspects of the proposal and, on the other hand, health managers who are familiar with the problem that is being investigated and will therefore be competent to comment on the focus, scope and usefulness of the proposed study. The panel members should be given a copy of the summary of the research proposal before the presentation begins.

Each participant group should prepare a presentation covering briefly the salient points in each section of its proposal. Participants should be encouraged to use the overhead projector and to practice and time their presentation prior to the actual presentation to the panel. (See Trainer's Notes on 'Course Management during the Workshop' for further details on presentation skills that should be stressed.)

The main points that should be emphasised in the presentation include:

1. Title of the study
2. A brief description of the problem, why the study is needed, what information is needed and how such information will be used
3. Objectives of the study
4. A brief statement on the type of study design, sample(s) and methods of data collection
5. A summary of how and when the study will be implemented (where, by whom, when, etc.)
6. A summary of how data will be analysed to provide the required information
7. A summary of the main resources required (e.g., manpower, budget, transport)
8. A brief summary of ethical considerations, and plan for project administration, monitoring and utilisation of results.

Although the presentation itself should be brief, participants should be prepared to respond to detailed questions on any of the aspects of the proposal that have been presented. Keep a list of research themes and/ or variables at hand for questions on content of the study.

Submission of the proposal

Accompanying letters should contain the title, the name of the team leader and project administrator, and the period over which the study will be carried out. If a letter is going to the national research council or a similar group you may briefly refer to your study's methodology and expected results and mention where further details can be found in your proposal.

Prepare a letter of intent that clearly summarises your research proposal and the estimated resources required, to send to potential funding agencies. In letters to potential donors you should not only state the total amount required but also the account to which the money, if granted, should be credited. Furthermore, it may be advisable to develop a follow-up proposal with a budget line, if implementation of your research recommendations will require additional resources.

Trainer's Notes

Module 18: FINALISING AND REVIEWING THE RESEARCH PROPOSAL

Timing and training methods

½ hour Introduction and discussion
8 hours + Group work

Introduction and discussion

- Part I of this module should be presented as soon as one or more of the groups is ready to start compiling their final document. Parts II and III may be presented in the same session or somewhat later in a separate session. The presentation of the proposal to the panel of relevant authorities, which will come near or at the end of the course, should be arranged early so that key authorities will not have scheduling conflicts.

- Stress, with the participants, the importance of preparing the final draft of their research proposal in such a way that it reads well to outsiders not fully familiar with their topic. It should be comprehensive, to the point, and coherent.

- A brief summary is required for decision-makers who have little time to study the whole research proposal. It is best to write this summary when the proposal is more or less finalised. The team should pay extra attention to writing the summary, as it is the 'eye catcher' for their proposal.

To ensure that the research proposals are prepared, word processed, and duplicated in time for the panel presentation, facilitators should monitor each group closely, making sure that the sections of the proposal that have been completed are promptly submitted to the secretariat or group members for word processing. Encourage groups to use the list of proposal sections presented in the group work as a checklist to keep production of the final draft co-ordinated.

Group work

The major task of the facilitator is to assist the team leader in distributing writing tasks among group members, in editing, in organising the word processing and making corrections in the different drafts. All group members should be involved in writing, either working in pairs or individually. They should read all sections written by others. The group as a whole should, in particular, discuss the summary.

If time allows, it is highly recommended that a draft of the proposal (or of the most important sections of the proposal) be given to a facilitator of one of the other groups for comments before the text is finalised.

Presentation of the research proposal to a panel

Trainers should use the presentations to the panel as an opportunity to establish interaction between participants, health managers and experienced researchers. This interaction will be beneficial to all parties. For example:

- The **participants** will acquire a better understanding of the concerns of managers and of research councils. Furthermore they will acquire confidence in presenting and defending research proposals.

- The **managers** will be exposed to a systematic approach to problem solving and acquire a better appreciation of research information.

- **Experienced researchers** will be exposed to the practical concerns of health systems research and acquire a better understanding of the approaches and potential of this type of research.

Briefing of panel members

The course facilitator should brief panel members so that they are familiar with the purpose of the presentation. It is useful to request that they behave as though they were members of a research committee that is responsible for approving the projects.

Time allocated

Allow 10 minutes for each presentation and 10-15 minutes for questions and discussion.

Role of facilitators

During the presentation, course facilitators should refrain, as far as possible, from intervening unless it is obvious that an important point is being misunderstood or overlooked. However, panel members or participants may require assistance on specific issues, and the facilitators should serve as resource persons for this purpose.

Module 12:

PLAN FOR DATA COLLECTION

Flowchart: Steps in the development of a health systems research proposal

Questions you must ask	Steps you will take	Important elements of each step
What is the problem and why should it be studied?	Selection, analysis and statement of the research problem	- problem identification - prioritising problems - analysis - justification
What information is available?	Literature review	- literature and other available information
Why do we want to carry out the research? What do we hope to achieve?	Formulation of research objectives	- general and specific objectives - hypotheses
What additional data do we need to meet our research objectives? How are we going to collect this information?	Research methodology	- variables - types of study - data collection techniques - sampling - **plan for data collection** - plan for data processing and analysis - ethical considerations - pre-test or pilot study
Who will do what, and when?	Work plan	- human resources - timetable
What resources do we need to carry out the study? What resources do we have?	Budget	- material support and equipment - money
How will the project be administered? How will utilisation of results be ensured?	Plan for project administration and utilisation of results	- administration - monitoring - identification of potential users
How will we present our proposal to relevant authorities, community and the funding agencies?	Proposal summary	- briefing sessions and lobbying

NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 12: PLAN FOR DATA COLLECTION

OBJECTIVES

At the end of this session, you should be able to:

1. **Identify** and discuss the most important points to be considered when starting to plan for data collection.
2. **Determine** what resources are available and needed to carry out your study.
3. **Describe** typical problems that may arise during data collection and how they may be solved.
4. **Prepare** a plan for data collection for the research proposal you are developing.

I. Introduction

II. Stages in the data collection process

I. INTRODUCTION

Where are we now in the development of our research proposal?

Look again at the flowchart in **Module 7** that introduces research design. We have just finished four crucial theoretical sessions, in which we have defined:

- what information we want to collect to answer the research questions implied by our objectives (**Module 8:** Variables)
- what **approach** we will follow to collect this information (**Module 9:** Study types)
- what **techniques** and **tools** we will use to collect it (**Module l0:** Data collection techniques)
- **where** we want to collect the data, **how** we will select our sample and **how many** subjects we will include in our study (**Module 11:** Sampling)

Now we enter a new phase in the development of our research methodology: planning our field-work. We have to plan concretely **how** we will **collect** the data we need (**Modules 12** and **15**), how we will **analyse** it (**Module 13**), and how we can **test** the most crucial parts of our methodology (**Module 14**). Finally, we will have to develop a work plan to budget the **resources** necessary to carry out the study (**Module 16**).

A PLAN FOR DATA COLLECTION can be made in two steps:

1. Listing the tasks that have to be carried out and who should be involved, making a rough estimate of the time needed for the different parts of the study, and identifying the most appropriate period in which to carry out the research (the present module).
2. Actually scheduling the different activities that have to be carried out each week in a work plan (**Module 15**).

Before the proposal development workshop is finished, a pre-test of the data collection and data analysis procedures should be made. The advantages of conducting a pre-test **before** we finalise our proposal is that we can draft the work plan and budget based on more realistic estimates, as well as revise the data collection tools before we submit the proposal for approval. However, further pre-testing and planning of the research have to be done in the field (**Module 14**).

Why should you develop a plan for data collection?

A plan for data collection should be developed so that:

- you will have a clear overview of what tasks have to be carried out, who should perform them, and the duration of these tasks;
- you can organise both human and material resources for data collection in the most efficient way; and
- you can minimise errors and delays which may result from lack of planning (for example, the population not being available or data forms being misplaced).

It is likely that while developing a plan for data collection you will identify problems (such as limited manpower), which will require modification of the proposal. Such modifications might include adjustment of the sample size or extension of the period for data collection.

II. STAGES IN THE DATA COLLECTION PROCESS

What are the main stages in the data collection process?

Three main stages can be distinguished:

Stage 1: Permission to proceed
Stage 2: Data collection
Stage 3: Data handling

Stage 1: PERMISSION TO PROCEED

Consent must be obtained from the relevant authorities, individuals and the community in which the project is to be carried out. This may involve organising meetings at national or provincial level, at district and at village level. For clinical studies this may also involve obtaining written informed consent.

Most likely the principal investigator will be responsible for obtaining permission to proceed at the various levels. The Health Research Unit in the Ministry of Health or the institution organising the course may assist in obtaining permission from the national level.

Note:

In many countries research proposals have to be screened for scientific and ethical integrity by national research councils. However, proposals developed during workshops may be exempted from this procedure if the research is considered as a training exercise and the research council is assured that the course facilitators and Health Research Unit have screened the methodology during the workshop.

Stage 2: DATA COLLECTION

When collecting our data, we have to consider:

- Logistics: who will collect what, when and with what resources
- Quality control

1. Logistics of data collection

WHO will collect WHAT data?

When allocating tasks for data collection, it is recommended that you first list them. Then you may identify who could best implement each of the tasks. If it is clear beforehand that your research team will not be able to carry out the entire study by itself, you might plan to look for research assistants to assist in relatively simple but time-consuming tasks.

For example, in a study into the effects of improvements in delivery care on utilisation of these services the following task division could be proposed:

Task	To be carried out by
Record study (hospital + HCs)	Research team, with research assistant
Focus group discussions with health staff before and after individual staff interviews	Research team
Individual health staff interviews	Research team
Shadowing MCH nurses	Principal investigator
Interviews with mothers (community-based) before and after delivery	Research team, with research assistants

HOW LONG will it take to collect the data for each component of the study?

Step 1: Consider:

- The time required to reach the study area(s);
- The time required to locate the study units (persons, groups, records); If you have to search for specific informants (e.g., users or defaulters of a specific service) it might take more time to locate informants than to interview them.)
- The number of visits required per study unit. For some studies it may be necessary to visit informants a number of times, for example if the information needed is sensitive and can only be collected after informants are comfortable with the investigator or if observations have to be made more than once (for example, follow-up of pregnant mothers or malnourished children). Time needed for follow-up of non-respondents should also be considered.

Step 2: Calculate the number of interviews that can be carried out per person per day (e.g., 4)

Step 3: Calculate the number of days needed to carry out the interviews. For example:

- you need to do 200 interviews,
- your research team of 5 people can do 5 x 4 = 20 interviews per day,
- you will need 200:20 = 10 days for the interviews.

Step 4: Calculate the time needed for the other parts of the study, (for example, 10 days)

Step 5: Determine how much time you can devote to the study. Since the research team usually consists of very busy people, it is unlikely that team members can spend more than 30 working days on the entire study:

- 5 days for preparation (including pre-testing and finalising questionnaires),
- 20 days actual field work,
- 5 days data processing + preliminary analysis.

If the team has 20 days for fieldwork, as in the example above, it could do the study without extra assistance. However, if the research team has only five days available for the interviews, they would need an additional five research assistants to help complete this part of the study.

Note:

The recruitment of research assistants for data collection may, on the one hand, relieve the research team, but, on the other hand, the **training and supervision of research assistants require time (**see **Annex 12.1)**. The team has to carefully weigh advantages and disadvantages. If none of the team members has previous research experience, they might prefer designing a study, which they can carry out themselves, without or with only minimal assistance.

If research assistants are required, consider to what extent local health workers can be used. They have the advantage of knowing the local situation. They should never be involved, however, in conducting interviews to evaluate the performance of their own health facility. Local staff from related services (teachers, community development) or students might help out. Sometimes village health workers or community members can collect certain parts of the data.

Note:

It is always advisable to slightly overestimate the period needed for data collection to allow for unforeseen delays.

In WHAT SEQUENCE should data be collected?

In general, it is advisable to start with analysis of data that is already available. This is essential if the sample of respondents is to be selected from the records. Another rule of thumb is that qualitative research techniques (such as key informant interviews, focus group discussions) that are devised to focus the content of questionnaires for interviewing larger groups of informants should be carried out *before* finalisation of these questionnaires. FGDs designed to provide feedback on issues raised in larger surveys, should, logically, be conducted *after* preliminary analysis of the questionnaires.

To use time and transport efficiently, data to be drawn from different sources but in one locality should be collected at the same time. (For example, interviews with health staff in a Health Centre, observations of equipment available in the Health Centre and interviews with mothers living around the Health Centre should be scheduled together.)

WHEN should the data be collected?

The type of data to be collected and the demands of the project will determine the actual time needed for the data to be collected. Consideration should be given to:

- availability of research team members and research assistants,
- the appropriate season(s) to conduct the field work (if the problem is season-related or if data collection would be difficult during certain periods),
- accessibility and availability of the sampled population, and
- public holidays and vacation periods.

Note:

The field visit to obtain consent from local authorities for the research may also be used to obtain necessary details about the best period for data collection and availability of local resources (research assistants, transport), if required.

2. Ensuring quality

It is extremely important that the data we collect are of good quality, that is, reliable and valid. Otherwise we will come up with false or misleading conclusions.

In the previous modules **possible sources of data distortion** (bias) have been discussed which we should try to prevent. They include:

- *Deviations from the sampling procedures* set out in the proposal.
- *Variability or bias in observations or measurements made* because:
  - Our study subject changes his/her behaviour as a consequence of the research. For example, a subject may act more positively while being observed; blood pressure and pulse may increase when the subject is apprehensive.
  - We use unstandardised measuring instruments.
    For example, we may use unstandardised weighing scales or imprecise or no guidelines for interviewing.
  - Researchers themselves vary in what they observe or measure (**observer variability**). For example, researchers may be selective in their observations (observer bias); measure, question or note down answers with varying accuracy or follow different approaches (one being more open, friendly, probing than the other).
- *Variations in criteria for measurement or for categorising answers* because we changed them during the study.

There are a number of possible measures that can be taken to prevent and to partly correct such distortions. Remember: prevention is FAR better than cure!! Cure is usually surgery: you may have to cut out the bad parts of your data or, at best, devise crutches.

Measures to help ensure good quality of data:

- **Prepare a field work manual for the research team as a whole,** including:
  - Guidelines on **sampling procedures** and what to do if respondents are not available or refuse to co-operate.(See **Module 11,** section II-3),
  - A clear **explanation** of the purpose and procedures of the study which should be used to introduce each interview, and
  - **Instruction sheets** on how to ask certain questions and how to record the answers.

- **Select your research assistants, if required, with care.** Choose assistants that are:

  - from the same educational level;
  - knowledgeable concerning the topic and local conditions;
  - not the object of study themselves; and
  - not biased concerning the topic (for example, health staff are usually not the best possible interviewers for a study on alternative health practices).

- **Train research assistants carefully in all topics covered in the field work manual as well as in interview techniques** (see **Annex 12.1** and **Module 10B** part V) and make sure that all members of the research team master interview techniques such as:

  - asking questions in a neutral manner;
  - not showing by words or expression what answers one expects;
  - not showing agreement, disagreement or surprise; and
  - recording the answers precisely as they are provided, without sifting or interpreting them.

- **Pre-test research instruments and research procedures** with the whole research team, including research assistants (see **Module 14**).

- **Take care that research assistants are not placed under too much stress** (requiring too many interviews a day; paying per interview instead of per day).

- **Arrange for on-going supervision** of research assistants. If, in case of a larger survey, special supervisors have to be appointed, guidelines should be developed for supervisory tasks.

- **Devise methods to assure the quality** of data collected by all members of the research team. For example, quality can be assured by:

  - requiring interviewers to check whether the questionnaire is filled in completely before finishing each interview;
  - asking the supervisor to check at the end of each day during the data collection period whether the questionnaires are filled in completely and whether the recorded information makes sense; and
  - having the researchers review the data during the data analysis stage to check whether data are complete and consistent.

Stage 3: DATA HANDLING

Once the data have been collected and checked for completeness and accuracy, a clear procedure should be developed for handling and storing them:

- At some stage questionnaires and other research tools will have to be numbered. Decide if this should be done at the time of the interview or at the time the questionnaires are stored. Usually each tool used will get its own numbers starting from 1. However, if some data sets are linked, e.g., you interview leprosy patients as well as their relatives and neighbours to analyse their interaction and possible stigma from different perspectives, you better link the numbers. For example, if you have interviewed a patient, P9, his neighbour will be N9 and his wife or son/daughter will be R9.

- Identify the person responsible for storing data and the place where it will be stored.

- Decide how data should be stored. Record forms should be kept in the sequence in which they have been numbered.

GROUP WORK (1¾ hours)

Make a plan for data collection, considering the points below:

1. **Permission to proceed** (10 minutes)

   - Which organisations or individuals should be approached to obtain permission to proceed with the research project?
   - Who will ask for permission? When? What procedures will be followed?

2. **Data Collection** (1-¼ hours)

   - List the different components of your study and the number of interviews, observations or measurements required.
   - Calculate for each component how many interviews or observations one person can do per day.
   - Decide if you need extra assistance, considering the fact that you, as a research team, will probably not be able to spend more than approximately 20 working days per person in the field and 5 days per person for preparation of the fieldwork.
   - If you need research assistants: For which components of the research? How many research assistants? Who would be the right persons to assist you and for how many days will you need them?
   - How will you train them? (place, timing, content, duration, trainers)
   - How will you ensure their supervision?
   - How will the quality of the data be checked and by whom?

3. **Data handling** (5 minutes)

   - How will the questionnaires/checklists be numbered?
   - How will the data be stored and who has the final responsibility for storing the data?

4. **Ethical considerations** (15 minutes)

Take care that your data collection process is ethical in all respects:

   - How have you planned to obtain informed consent from your informants? Are there any categories of informants that need special consideration (e.g., children, sick persons, mentally disabled individuals)?
   - Are certain parts of the research focused on sensitive issues? How will you handle problems that may arise?
   - Do certain parts of your research require extra attention to assure confidentially? How will you handle the issue?

5. **Summarise the outcome** of your group work on a flip chart. Record the details of your discussions so that you can use them for the development of your work plan (**Module 15**).

Annex 12.1: Training research assistants (as well as other research team members with little research experience)

1. Interviewers' tasks

During the fieldwork interviewers (or research assistants) may work independently or together with one of the researchers. If they go out independently they may have to carry out the following tasks:

- Do the **sampling in the field** (for example sampling of households within a village and/or sampling of individuals to be interviewed within households).

- Give a clear **introduction** to the interviewee concerning the purpose and procedures of the interview and ask for permission to interview. (An introductory sheet should be prepared and attached to interview guide or questionnaire, so that each research team member approaches the informants in a similar way.)

- **Perform the interviews.** Usually it is best to give research assistants more or less structured questionnaires to administer with clear instructions for open questions. It is not wise to assign the more difficult tasks of performing highly flexible interviews or focus group discussions to research assistants unless they are mature, experienced persons.

It is imperative that the researchers train their research assistants so they can carry out their tasks accurately and correctly, according to the procedures developed by the researchers. Research assistants should not be left to develop their own procedures; otherwise bias will be almost certain to result.

The training of the research team may take 2 to 3 days. The first day may be devoted to theory, followed by 1 or 2 days of practical training, depending on the local circumstances and the nature of the study.

2. Theoretical training

Research assistants must be thoroughly familiar with the objectives of the research project and the methodology. Therefore, it is recommended that they be provided with copies of the research protocol and that the most relevant sections be discussed thoroughly, including:

- statement of the problem,
- objectives,
- data collection tools to be used (an overview),
- sampling procedures (if sampling has to be done in the field),
- plan for data collection, and
- plan for data analysis.

It is important at this stage that the research assistants get an ample opportunity to ask questions.

Then a more in-depth discussion should follow concerning the data collection tools (interview FGD guides, questionnaires and possibly checklists) that are to be used by the research-team members. For each and every question all should know WHY the information is required.

Research assistants should be taught basic interview techniques, such as

- asking questions in a neutral manner;
- not showing by words or expression what answers one expects;
- not showing agreement, disagreement or surprise; and
- recording answers to open questions precisely as they are provided, without sifting or interpreting them.

Clear instructions should be given as well, concerning to what extent an interviewer is allowed to alter the phrasing of questions, if it seems necessary, and whether he/she should probe for answers. For questions which have pre-categorised answers it should be made clear whether the possible answers should be mentioned by the interviewer during the interview or not. (Usually they are not to be mentioned.) There should be no misunderstandings concerning how to record answers and observations.

Finally, explanations should be given concerning how an interviewer should introduce him or herself to the interviewee, what to say concerning the purpose of the study, how to ask for consent, and how to close the interview.

3. Practical training

Practical training in interviewing is essential. This may be provided in two stages. (See **Module 10B part V,** interview exercise.)

Firstly, **role-plays** can be performed, during which one trainee assumes the role of the interviewer and another plays the interviewee. Other trainees and the trainers (researchers) should observe carefully what happens and give constructive feedback right after the role-play. Then roles can be changed, until each trainee has had a chance to practice each type of interview at least once.

Secondly, a **field test** should be conducted, which will serve two purposes: the training of the entire research team including research assistants, and a (further) test of the data collection tools. A test of the tools is essential if a previous field-test resulted in important changes and/or if the questionnaire(s) were translated into a local language after the first field test. If the research assistants will be involved in the proper phrasing of questions this will definitely strengthen their interest and commitment.

The field test is best carried out in groups of 2 to 3 persons (see **Module 14**), with each team including at least one trainer and one trainee. Note that after the field test constructive critique of each interview should be made, starting from the moment the interviewee was approached up to the farewell.

4. Supervision of research assistants

Even if research assistants are used, responsibility for the research remains with the research team as a whole. To guarantee the quality of the data collected it is important to supervise the research assistants' performance, especially at the beginning of the data collection period. If they are going out into the field independently, plans could be made to accompany them on selected visits and/or to question a small sample of the interviewees as a 'control group' concerning key aspects of the interview.

As a further quality control check it is important that an interviewer's name (or interviewer's code) appears standard on each questionnaire/checklist so that it is possible to ask for clarification if certain information is not clear.

Trainer's Notes

Module 12: PLAN FOR DATA COLLECTION

Timing and teaching methods

¾ hour	Introduction and discussion
1¾ hours	Group work
1 hour	Plenary
3½ hours	TOTAL TIME

Introduction and discussion

- Explain that we are entering a new phase in the development of the methodology, in which we go from theory to practice: the concrete planning of the data collection process.

- Of the three phases that can be distinguished in the planning for data collection (i.e., permission to proceed, actual data collection and data handling), planning for actual data collection will require most attention, both in the presentation and in the group work.

- The **logistics** of data collection is one major aspect of the planning: WHO will collect WHAT and WHEN. It is important that relatively inexperienced researchers give thought to this aspect of data collection **before** the pre-test, so that they can consciously check whether their plans are realistic.

- As the participants begin planning for data collection they should clearly distinguish and list the **different components** of their study. Next, they should consider the **time** it will take to carry out the different components, so that they can decide whether they can do the study alone or whether they need assistance.

- Take the most labour-intensive part of one of the studies being developed in the workshop and follow the steps outlined in the module to determine the time needed.

- If you do not want to favour one of the groups, you can use the example of the utilisation of delivery care provided in the module or any other example that resembles the topics participants are working on.

- Discuss the advantages and disadvantages of using research assistants and of using local health service personnel as research assistants.

- Then proceed to the second major aspect of the planning for data collection, that of ensuring quality.

- Stress the importance of preparing a fieldwork manual (irrespective of whether research assistants are recruited or not) and of using appropriate interview techniques. Refer to Annex 12.1 for more detailed guidelines on how to train interviewers.

Group work

- Make sure that participants have understood the principle of calculating the time required for each component of the study and that they calculate carefully the time needed for the most labour-intensive parts of the study, so that they can decide:

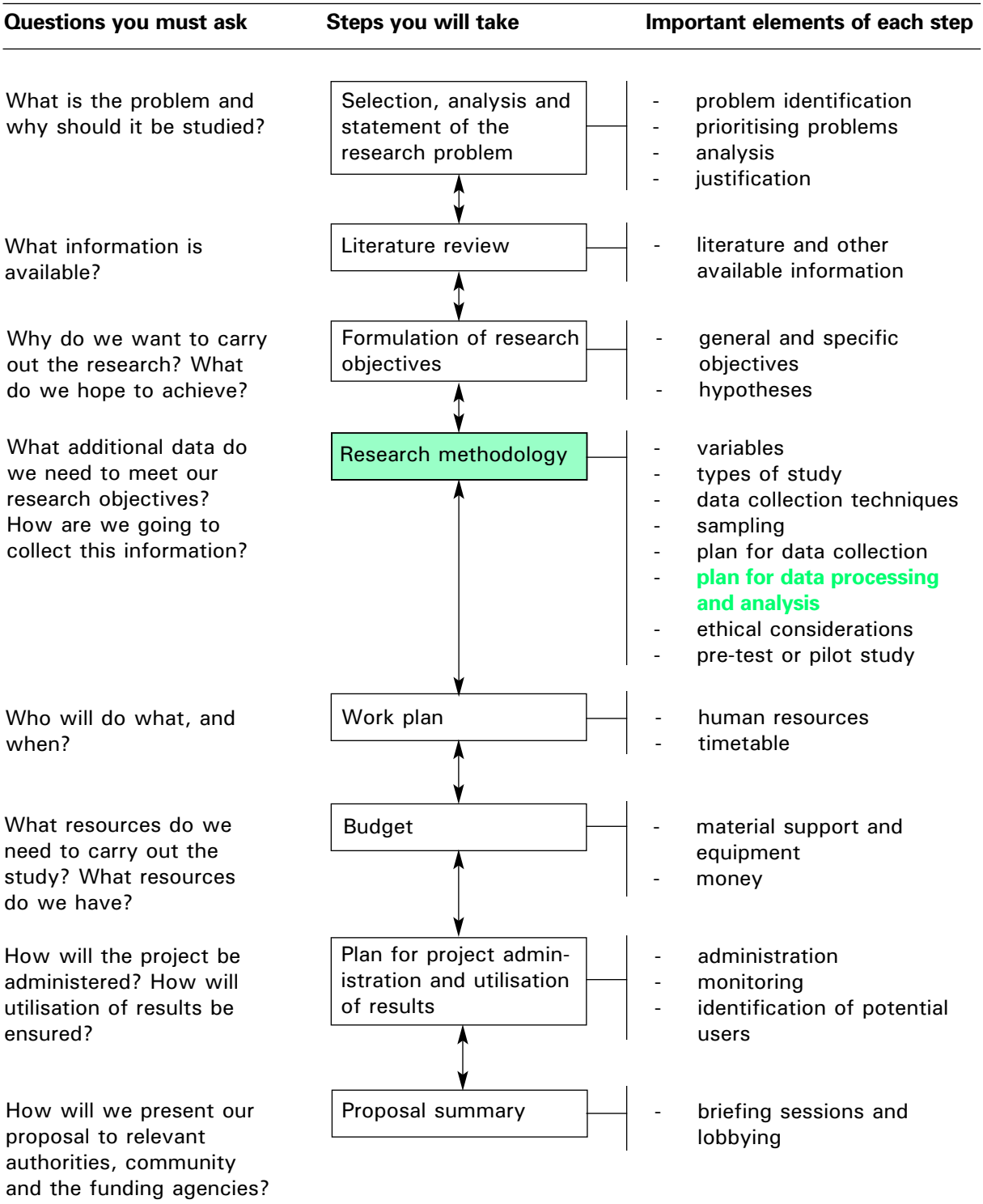
  - whether they need assistance, and
  - who would be appropriate research assistants, if required.

- Let them summarise ethical issues involved in data collection, which have been discussed during previous presentations. Participants need to include a section on ethical considerations in their research proposal (see **Module 18** or **Module 1** for content of the research proposal).

Module 13

PLAN FOR DATA PROCESSING AND ANALYSIS

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

MODULE 13: PLAN FOR DATA PROCESSING AND ANALYSIS

OBJECTIVES

At the end of this session, you should be able to:

1. **Identify** important issues related to sorting, quality control, and processing of data.
2. **Describe** how data can be best be analysed and interpreted based on the objectives and variables of the study.
3. **Prepare** a plan for the processing and analysis of data (including data master sheets and dummy tables) for the research proposal you are developing.

I. Introduction

II. Sorting data

III. Performing quality checks

IV. Data processing - qualitative data

V. Data analysis - quantitative data

VI. Processing and analysis of qualitative data

I. INTRODUCTION

Data processing and analysis should start in the field, with checking for completeness of the data and performing quality control checks, while sorting the data by instrument used and by group of informants. Data of small samples may even be processed and analysed as soon as it is collected.

Why is it necessary to prepare a plan for processing and analysis of data?

Such a plan helps the researcher assure that at the end of the study:

- all the information (s)he needs has indeed been collected, and in a standardised way;
- (s)he has not collected unnecessary data which will never be analysed.

The plan for data processing and analysis must be made after careful consideration of the objectives of the study as well as of the tools developed to meet the objectives.

The procedures for the analysis of data collected through qualitative and quantitative techniques are quite different.

- For **quantitative data** the starting point in analysis is usually a description of the data *for each variable* for all the study units included in the sample. Processing of data may take place during data collection or when all data has been collected; description and analysis are usually carried out *after* the fieldwork has been completed.

- For **qualitative data** it is more a matter of describing, summarising and interpreting the data obtained *for each study unit* (or for each group of study units). Here the researcher starts analysing *while* collecting the data so that questions that remain unanswered (or new questions which come up) can be addressed before data collection is over.

Preparation of a plan for data processing and analysis will provide you with better insight into the feasibility of the analysis to be performed as well as the resources that are required. It also provides an *important review of the appropriateness of the data collection tools* for collecting the data you need. That is why you have to plan for data analysis *before* the pre-test (Module 14). When you process and analyse the data you collect during the pre-test you will spot gaps and overlaps which require changes in the data collection tools before it is too late!

What should the plan include?

When making a plan for data processing and analysis the following issues should be considered:

- Sorting data,
- Performing quality-control checks,
- Data processing, and
- Data analysis.

II. SORTING DATA

An appropriate system for sorting the data is important for facilitating subsequent processing and analysis.

If you have different study populations (for example village health workers, village health committees and the general population), you obviously would number the questionnaires *separately*.

In a comparative study it is best to sort the data right after collection into the two or three groups that you will be comparing during data analysis.

> **For example**, in a study concerning the reasons for low acceptance of family planning services, users and non-users would be basic categories; in a study of the reasons why nurses object to being posted in rural areas, rural and urban nurses would be basic categories; in a case-control study obviously the cases are to be compared with the controls.

It is useful to number the questionnaires belonging to each of these categories *separately* right after they are sorted.

> **For example**, the questionnaires administered to users of family planning services could be numbered U1, U2, U3, etc., and those for the non-users N1, N2, N3, etc.

In a cross-sectional survey it may also be useful to sort the data into two or more groups, depending on possible sub-groups you would like to compare.

III. PERFORMING QUALITY CONTROL CHECKS

Usually the data have already been checked in the field to ensure that all the information has been properly collected and recorded. Before and during data processing, however, the information should be checked again for **completeness and internal consistency**.

If a questionnaire has not been filled in completely you will have MISSING DATA for some of your variables. If there are many missing data in a particular questionnaire, you may decide to exclude the whole questionnaire from further analysis.

- If an inconsistency is clearly due to a mistake made by the researcher/research assistant (for example if a person in an earlier question is recorded as being a non-smoker, whereas all other questions reveal that he is smoking), it may still be possible to check with the person who conducted the interview and to correct the answer.

- If the inconsistency is less clearly a mistake in recording, it may be possible (in a small scale study) to return to the respondent and ask for clarification.

- If it is not possible to correct information that is clearly inconsistent, you may consider excluding this particular part of the data from further processing and analysis as it will affect the validity of the study. If a certain question produces ambiguous or vague answers throughout, the whole question should be excluded from further analysis. (Normally, however, you would discover such a problem during the pre-test and change the wording of the question.)

Note:

A decision to exclude data of doubtful quality is ethically correct and it testifies to the scientific integrity of the researcher. You should keep track of any questions you had to exclude because of incompleteness or inconsistency in the answers, and discuss it in your final report.

For computer data analysis, quality control checks of data must also include a verification of how the data has been transformed into codes and subsequently entered into the computer. The same applies if data are entered into master sheets (see next page).

IV. DATA PROCESSING – quantitative data

Decide whether to process and analyse the data from questionnaires:

- **manually**, using data master sheets or manual compilation of the questionnaires, or
- **by computer**, for example, using a micro-computer and existing software or self-written programmes for data analysis.

Data processing in both cases involves:

- categorising the data,
- coding, and
- summarising the data in data master sheets, manual compilation without master sheets, or data entry and verification by computer.

1. Categorising

Decisions have to be made concerning how to categorise responses.

For **categorical variables** that are investigated through closed questions or observation (for example, observation of the presence or absence of latrines in homesteads), the categories have been decided upon beforehand.

In interviews the answers to open-ended questions (for example**,** 'Why do you visit the health centre?') can be pre-categorised to a certain extent, depending on the knowledge of possible answers that may be given. However, there should always be a category called 'Others, specify ...', which can only be categorised afterwards.

These responses should be listed and placed in categories that are a logical continuation of the categories you already have. Answers that are difficult or impossible to categorise may be put in a separate residual category called 'others', but this category should not contain more than 5% of the answers obtained.

For **numerical variables,** the data are often better collected without any pre-categorisation. If you do not exactly know the range and the dispersion of the different values of these variables when you collect your sample (e.g., home-clinic distance for out-patients, or income), decisions concerning how to categorise and code the data at the time you develop your tools may be premature. If you notice during data analysis that your categories had been wrongly chosen you cannot reclassify the data anymore.

> **For example**, in a study into utilisation of health services, the research team wanted to establish whether income was related to utilisation. They had pre-coded income into three categories. When they analysed the data they discovered that over 80% fell in the lowest income category. In hindsight they would have preferred a five point scale in order to distinguish different grades of poverty, but as the raw data had not been recorded it was impossible to reclassify the data, and the variable was almost useless.

2. Coding

If the data will be entered in a computer for subsequent processing and analysis, it is essential to develop a CODING SYSTEM.

For computer analysis, each category of a variable can be coded with a letter, group of letters or word, or be given a number. For example, the answer 'yes' may be coded as 'Y' or 1; 'no' as 'N' or 2 and 'no response' or 'unknown' as 'U' or 9.

The codes should be entered on the questionnaires (or checklists) themselves. When finalising your questionnaire, for each question you should insert a box for the code in the right margin of the page. These boxes should not be used by the interviewer. They are only filled in afterwards during data processing. Take care that you have as many boxes as the number of digits in each code.

If analysis is done by hand using data master sheets, it is useful to code your data as well (see section 3 below)

Coding conventions

Common responses should have the same code in each question, as this minimises mistakes by coders.

For example:

Yes (or positive response)	code - Y or 1
No (or negative response)	code - N or 2
Don't know	code - D or 8
No response/unknown	code - U or 9

Codes for open-ended questions (in questionnaires) can be done only after examining a sample of (say 20) questionnaires. You may group similar types of responses into single categories, so as to limit their number to at most 6 or 7. If there are too many categories it is difficult to analyse the data. (For details, see section V part 2 of this module and **Module 23**.)

Finally it should be borne in mind that the personnel responsible for computer analysis should be consulted very early in the study, i.e., as soon as the questionnaire and dummy tables are finalised. In fact the research team needs to work closely with the computer analyst or statistician **throughout** the design and the implementation of the study.

2. Summarising the data in data master sheets, manual compilation, or compilation by computer

(1) Data master sheets

If data are processed by hand, it is often most efficient to summarise the raw research data in a so-called **DATA MASTER SHEET**, to facilitate data analysis. On a data master sheet all the answers of individual respondents are entered by hand.

To illustrate the use of master sheets, we will give an example of a rapid appraisal carried out by students of a nursing school about the smoking habits of the inhabitants of their town. Smoking was perceived as a big problem, and the study had been designed in order to develop an anti-smoking campaign. The (24) students divided the map of their town in 24 roughly equal parts. Each student put herself in the centre of the part of town designed to her, and interviewed six persons of 15 years and above, three males and three females, so 144 in total. (Another group of students conducted FGDs observations and individual interviews in schools to obtain insight into the onset of smoking behaviour.) The questionnaire had only 17 questions (see Annex 13.1), of which 9 were asked of everyone, 4 exclusively to smokers and 4 exclusively to non-smokers. It was therefore decided to process the data by hand, divided in two groups: smokers and non-smokers, which were again subdivided in males and females. For each of the four groups, master sheets were prepared, on which all the answers of individual respondents could be recorded.

Master sheets can be made in different ways. For short simple questionnaires you may put all possible answers for each question in headers at the top of the sheet and then list or tick the answers of the informants one by one in the appropriate columns.

For example, the straightforward answers of the smoking questionnaire for male smokers could be processed as follows (see Table 13.1):

Table 13.1: Master sheet for smokers (males)

No.	Q1 Sex	Q2 Age		Q6 No of cig.		Q7 Age onset		Q9 Tried to				Q14 Cough > 2wks		Q14 Cough/ chest ever	
								reduce		stop					
		Yrs	Cat	No	Cat	Yrs	Cat	Yes	No	Yes	No	Yes	No	Yes	No
1	M	18	(1)	10	(2)	12	(2)	1x			√		√		√
2	M	35	(3)	30	(4)	20	(4)		NR	1x			√	√	
3	M	54	(4)	15	(2)	14	(2)	10x		3x		√		√	
Etc.															
Total	31	Av	35	Av	20	Av	18	26	4 + 1NR	19	12	5	26	11	20

Categories

Age
15 - 24 = 1
25 - 34 = 2
35 - 44 = 3
45 - 54 = 4
55 + = 5

No. of cigarettes/day
< 10 = 1
10 - 19 = 2
20 - 29 = 3
30 - 39 = 4
40 + = 5

Age onset smoking
< 10 = 1
10 - 14 = 2
15 - 19 = 3
20 - 24 = 4
25 + = 5

Note that for age and number of cigarettes smoked both the raw data and the categories have been entered. This makes it easier to control for coding mistakes and allows for calculating averages. There are 31 male smokers; if there are less than 31 answers, there must be some non-responders (NR), as happened in Q9, or a mistake was made. If you work with *two persons,* one reading and one writing, the risk of mistakes will be reduced, as you can discuss the answers and control for mistakes while filling in the data.

Even this limited data suggest that male smokers start usually when they are teenagers, that the informants on average smoke one package of cigarettes a day and that attempts to stop appear to increase with age.

Some answers, however, require more elaborate coding and have more categories. For example, Q4 on education and Q5 on occupation could be summarised as follows (see **Table 13.1**, continued):

Table 13.1 (continued): Master sheet for smokers (males)

No	Education (Q3)										Occupation (Q4)													
	None	Highest level reached						Still in school		NR	Self							Head Household						
		years	Type																					
			PS	SS	DT	UY	O	Yes	No		0	1	2	3	4	5	6	1	2	3	4	5	6	U	NR
1		4	√						√			√								√					
2		9		√					√					√					NA					
3	√								NA						√				NA					

Categories

Education

PS = Primary school
SS = Secondary school
OT = Occupational training
UY = University level
O = Other

U = No response/unknown
NA = Not applicable

Occupation/income

0 = No source of income
1 = Irregular income from petty trade, handicraft, farming on borrowed land (not enough to live on)
2 = Unskilled labourer
3 = Farmer (with land)
4 = Trader/beer brewer/coffee-shop or taxi owner etc.
5 = Teacher, nurse, civil servant
6 = Doctor, lawyer, government administrator

It will be clear that including all possible categories for one question in the headers of the master sheet may take a lot of space. If a questionnaire has only 13 questions, that does not matter; with 64 questions it would mean that you would need several sheets to include all answers. As the idea of master sheets is to have an easy overview of all data, you may then look for another solution, and enter the different codes for one question in one column instead of having different columns of which you tick one. For education, there would in this way only be four columns: none, highest level reached in years, highest type of school reached, and still in school; for occupation only two: self and head household.

No	Education (Q3)				Occupation (Q4)	
	Y/N	Highest level		Still in school	Self	Head HH
		Years	Type			
1	Y	4	PS	N	1	3
2	Y	9	SS	N	4	NA
3	N	NA	NA	NA	5	NA
4		U	PS	Y	0	2

This data reveals that Smoker 1 is, like No. 4, still dependent on his father, though No. 1 earns some money. This means the money for cigarettes is strongly limited. Numbers 2 and 3 are financially independent, and smoke more heavily, though 3 has tried to reduce or stop many times.

Finally, there are open questions such as Q8, 'Why do you smoke?' or Q10, 'Why not?', or Q17, soliciting suggestions on how the students could best approach smokers in a campaign to discourage them from smoking. Coding and analysis of such qualitative data will be dealt with in section V of this module.

Note:

In any small-scale study processed by hand in which groups will be compared, a **different master sheet** should be made for each of those groups, e.g., good and poor compliers to treatment. As gender is an important cross-cutting theme, it is usually also advisable to subdivide males and females within each of the groups that are being compared.

See **Annex 13.2** for example of full master sheet.

(2) Compilation by hand (without using master sheets)

When the sample is small (say less than 30) and the collected data is limited, it might be more efficient to do the compilation manually.

Certain procedures will help ensure accuracy and speed.

i. If only one person is doing the compilation use **manual sorting**. If a team of 2 persons work together use either manual sorting or **tally counting**.

   Manual sorting can be used only if data on each subject is on a different sheet of paper/ entered in a separate questionnaire.

ii. To do **manual sorting** the basic procedure is to:

   - Take one question at a time, for example, 'use of health facility',
   - Sort the questionnaires into different piles representing the various responses to the question, e.g., hospital/ health centre/ traditional practitioners) and
   - Count the number in each pile.

   When you need to sort out subjects who have a certain combination of variables (e.g. females who used each type of health facility) sort the questionnaires into piles according to the first question (gender), then subdivide the piles according to the response to the other question (use of health facility).

iii. To do **tally counting** the basic procedure is:

   - One member of the compiling team reads out the information while the other records it in the form of a tally (e.g., III representing 3 subjects, ~~IIII~~ or representing 5 subjects who present a particular answer).

   - Tally count for no more than two variables at one time (e.g., sex plus type of facility used).

   If it is necessary to obtain information on 3 variables (e.g., sex by time of attending a health facility by diagnosis), do a manual sorting for the first question, then tally count for the other two variables.

   - After tally counting, add the tallies and record the number of subjects in each group.

iv. After doing either manual or tally counting, **check** the total number of subjects/responses in each question to make sure that there has been no omission or double count.

Note:

One can tally in two ways, 𝍸 or ⎺ ⌐ ◸ ◹ ◺. The latter way is preferable as it reduces the possibility of error.

It should be noted that hand tallying is often used in combination with master sheet analysis when the relationship between two or three variables needs to be established, or details analysed. (For example, the questionnaire forms of non-smokers whose close relatives or co-workers are smoking (Q11 and 12) and who feel disturbed by the smoke (Q13) may be selected to analyse more in detail by tallying the (health) problems these non-smokers are experiencing.

Note:

Researchers often assume that hand compilation is merely 'common sense' and do not train their staff in the correct procedure. Subsequently many hours of work are wasted in trying to detect the source of errors due to double counts, wrong categorisation, and omissions.

(3) Computer compilation

Before you decide to use a computer, you have to be sure that it will save time or that the quality of the analysis will benefit from it. Note that feeding data into a computer costs time and money. The computer should not be used if your sample is small and the data is mainly generated by open questions (qualitative data), unless there is a resource person who is competent in using a program for qualitative data analysis such as Qualitan or SPSS. The larger the sample, the more beneficial in general the use of a computer will be.

Computer compilation consists of the following steps:

i. Choosing an appropriate computer program
ii. Data entry
iii. Verification or validation of the data
iv. Programming (if necessary)
v. Computer outputs/prints

i. Choosing an appropriate computer program

A number of computer programs are available on the market that can be used to process and analyse research data. The most widely used programs are:

- Epi Info (version 6), a very consumer friendly program for data entry and analysis, which also has a word processing function for creating questionnaires (developed by the Centre for Disease Control, Atlanta, USA and World Health Organization, Geneva),
- LOTUS 1-2-3, a spreadsheet program (from the Lotus Development Corporation),
- dBase (version III plus or IV), a data-management program (from Ashton-Tate), and
- SPSS, which is a quite advanced Statistical Package for Social Sciences (SPSS Inc.).

If you intend to use a computer, you may ask advice from an experienced person concerning which program is the most appropriate for your type of data. Note that Epi Info may be freely used and copied. All the other programs have copyrights.

ii. Data entry

To enter data into the computer you have to develop a data entry format, depending on the program you are using. However, it is possible to enter data using dBase (which is relatively good for data entry) and do the analysis in LOTUS 1-2-3 or SPSS.

After deciding on a data entry format, the information on the data collection instrument will have to be coded (e.g., Male: M or 1, Female: F or 2). During data entry, the information relating to each subject in the study is keyed into the computer in the form of the relevant code (e.g., if the first subject (identified as 001) is a male (code 1) aged 25, the data could be keyed in as 001125).

Note that data entry can be done through the private sector, which may be fast and not too expensive. Health office staff who are not accustomed to this work tend to be slow and make many errors in entry.

iii. Verification

During data entry, mistakes will definitely creep in. The computer can print out the data exactly as it has been entered, so the printout can be checked visually for obvious errors, (e.g., exceptionally long or short lines, blanks that should not be there, alphabetic codes where numbers are expected, obviously wrong codes).

Example:

- Codes 3-8 in the column for sex where only 1(F) and 2 (M) are possible
- Codes above 250 when you had only 250 subjects

If possible, computer verification should be built in. This involves giving the appropriate commands to identify errors.

Example:

The computer can be instructed to identify and print out all subjects where the 'sex' column has a code different from 1 (F) or 2 (M).

iv. Programming

If you use computer personnel to analyse your data, it is important to **communicate effectively** with them. Do not leave the analysis to the computer specialist! You as a researcher should tell the computer personnel:

- the names of all the variables in the questionnaire;
- the location of these variables in relation to the data for one subject (i.e., the data format);
- how many subjects are to be analysed and which groups are to be compared;
- whether any variables are to be re-coded or calculated; and
- for which variables you need straight tabulations and which variables you would like to cross-tabulate.

A certain amount of basic knowledge of computer programming is needed to give the appropriate commands.

v. Computer outputs

The computer can do all kinds of analysis and the results can be printed. It is important to decide whether each of the tables, graphs, and statistical tests that can be produced makes sense and should be used in your report. That is why we PLAN the data analysis BEFOREHAND! (See section V.)

V. DATA ANALYSIS – quantitative data

Analysis of quantitative data involves the production and interpretation of frequencies, tables, graphs, etc., that describe the data.

1. Frequency counts

From the data master sheets, simple tables can be made with **frequency counts** for each variable. A frequency count is an enumeration of how often a certain measurement or a certain answer to a specific question occurs.

For example,

Smokers	51
Non-smokers	93
Total	144

If numbers are large enough it is better to calculate the frequency distribution in percentages **(relative frequencies)**: 51/144 x 100 = 35% are smokers and 93/144 x 100 = 65% non-smokers. This makes it easier to compare groups than when only absolute numbers are given. In other words, percentages standardise the data.

It is usually necessary to summarise the data from numerical variables by dividing them into categories. This process may include the following steps:

(1) Inspect all the figures: What is their range? (The range is the difference between the largest and the smallest measurement.)

(2) Divide the range into three to five categories. You can either aim at having a reasonable number in each category (e.g. 0-2 km, 3-4 km, 5-9 km, 10+ km for home-clinic distance) or you can define the categories in such a way that they are each equal in size (e.g., 20-29 years, 30-39 years, 40-49 years, etc.). Sometimes one looks actively for a 'critical' value when making different categories. For example, in a study relating family income to prevalence of diarrhoea over a certain period, there appeared to be no statistical relation when income was arbitrarily subdivided into four categories. When the *average* income was calculated, however, this appeared to be a critical value. The children in families with an income above average had had significantly less diarrhoea than the children in families with an income below average.

(3) Construct a table indicating how data are grouped and count the number of observations in each group.

2. Cross-tabulations

Further analysis of the data usually requires the combination of information on two or more variables in order to describe the problem or to arrive at possible explanations for it.

For this purpose it is necessary to design CROSS-TABULATIONS.

Depending on the objectives and the type of study, two major kinds of cross-tabulations may be required:

- Descriptive cross-tabulations that aim at describing the problem under study.
- Analytic cross-tabulations in which groups are compared in order to determine differences, or which focus on exploring relationships between variables.

A **descriptive cross-tabulation** would, for example, relate smoking behaviour to sex or occupational background:

Table 13.2: Smoking by sex

Sex	Smoking		Not smoking		Total	
Males	31	(43%)	41	(57%)	72	(100%)
Females	20	(28%)	52	(72%)	72	(100%)
Total	51	(35%)	93	(65%)	144	(100%)

The males appear to be smoking more (43%) than females (28%).

An **analytic cross-tabulation** serves to investigate if there is a relationship between smoking (independent variable) and persistent cough, or chest complaints (dependent variables/problems).

Table 13.3: Smoking in relation to persistent cough over the past 2 weeks

Smoking behaviour	Cough		No cough		Total	
Smoking	10	(77%)	41	(32%)	51	(35%)
Not smoking	3	(23%)	90	(68%)	93	(65%)
Total	13	(100%)	131	(100%)	144	(100%)

Of the informants with a cough, the majority (77%) is smoking, whereas among those without a cough, only one third (33%) are smokers. The expected relationship between smoking and chest problems seems therefore confirmed.

When the plan for data analysis is being developed the data, of course, is not yet available. However, in order to visualise how the data can be organised and summarised it is useful at this stage to construct so-called DUMMY cross-tabulations.

A DUMMY TABLE contains all elements of a real table, except that the cells are still empty.

In a research proposal dummy tables should be prepared to describe the study population in order to show the crucial relationships between variables.

For the study on smoking behaviour, for example, you would have to prepare a number of descriptive dummy tables (describing characteristics of smokers and non-smokers, their behaviour and attitudes towards smoking), such as Table 13.2 but without numbers or percentages. Further you would make two analytic dummy tables, one on the relationship between smoking and persistent cough (see Table 13.3) and one on the relationship between smoking and long-term chest problems.

Some practical hints when constructing tables:

- If a dependent and an independent variable are cross-tabulated, the headings of the dependent variable are usually placed horizontally (see Table 13.3: 'cough' and 'no cough'), and the headings of the independent variable vertically: ('smoking' and 'not smoking' in the same table).
- All tables should have a clear title and clear headings for all rows and columns.
- All tables should have a separate row and a separate column for totals to enable you to check if your totals are the same for all variables and to make further analysis easier.

- All tables related to a certain objective should be numbered and kept together so the work can be easily organised and the writing of the final report will be simplified.

To further analyse and interpret the data, certain calculations or **statistical procedures** must usually be completed. Especially in large cross-sectional surveys and in comparative studies, statistical procedures are necessary if the data is to be adequately interpreted. Statistical tests should, for example, indicate whether the gender differences in smoking behaviour are true differences or due to chance. When conducting such studies it is advisable to consult a person with statistical knowledge from the start in order that:

- correct sampling methods are used and an appropriate sample size is selected;
- decisions on coding are made that will facilitate data processing and analysis; and
- a clear understanding is reached concerning plans for data processing, analysis and interpretation, including agreement concerning which variables need to be cross-tabulated.

Some elementary statistical procedures will be taught in the second workshop after field work is completed. An elementary knowledge of statistics will help you better understand the whole process of data analysis and interpretation.

VI. PROCESSING AND ANALYSIS OF QUALITATIVE DATA

Qualitative data may be collected through open-ended questions in self-administered questionnaires, in individual interviews or focus group discussions or through observations during fieldwork. For a detailed description of the analysis of qualitative data see **Module 10C** and in particular **Module 23**, which specify the methods most often used. Here we will concentrate on the analysis of responses obtained from open-ended questions in interviews or self-administered questionnaires.

Commonly solicited data in open-ended questions include:

- opinions of respondents on a certain issue;
- reasons for a certain behaviour; and
- descriptions of certain procedures, practices or perceptions with which the researcher is not familiar.

The data can be analysed in seven steps:

Step 1: Take a sample of (say 20) questionnaires and list all answers for a particular question. Take care to include the source of each answer you list (in the case of questionnaires you can use the questionnaire number), so that you can place each answer in its original context, if required.

Step 2: To establish your categories, you first read carefully through the whole list of answers. Then you start giving codes (A, B, C, for example or **key words**) for the answers that you think belong together in one category, and write these codes in the left margin. Use a pencil so that it is easy to change the categories if you change your mind.

Step 3: List the answers again, grouping those with the same code together.

Step 4: Then interpret each category of answers and try to give it a label that covers the content of all answers. In the case of data on **opinions**, for example, there may be only a limited number of possibilities, which may range from (very) positive, neutral, to (very) negative.

Data on **reasons** may require different categories depending on the topic and the purpose of your question. In the exercise below you will be asked to categorise the reasons why people smoke by grouping them in such a way that it is easy to find entry points for health education aimed at reducing smoking.

After some shuffling you usually end up with 5 to 7 categories.

Step 5: Now try a next batch of 20 questionnaires and check if the labels work. Adjust the categories and labels, if necessary.

Step 6: Make a final list of labels for each category and give each label a code (keyword, letter or number).

Step 7: Code all your data, including what you have already coded, and enter these codes in your master sheet or in the computer.

Note again that you may include a category 'others', but that it should be as small as possible, preferably used for less than 5% of the total answers.

If you categorise your responses to open-ended questions in this way you can:

- Analyse the content of each answer given in particular categories, for example, in order to plan what actions should be taken (e.g., for health education). Gaining *insight* in a problem, or in possible interventions for a problem, is the most important function of qualitative data.

- Report the number and percentage of respondents that fall into each category; so that you gain insight in the relative weight of different opinions or reasons.

Questions that ask for descriptions of procedures, practices, or beliefs usually do not provide quantifiable answers (though you may quantify certain aspects of them). The answers rather form part of a jigsaw puzzle that you have to put together in order to obtain insight in your problem/topic under study.

IN CONCLUSION, a plan for the processing and analysis of data may include:

- a decision on whether all or some parts of the data should be *processed by hand or computer*;
- *dummy tables* for the description of the problem, the comparison of groups (if applicable) and/or the establishment of relationships between variables, guided by the objectives of the study;
- a decision on the *sequence* in which tables or data from different study populations should be analysed;
- a decision on how *qualitative data* should be analysed;
- an estimate of the *total time needed* for analysis and how long particular parts of the analysis will take;
- a decision concerning whether *additional staff* will be required for the analysis; and
- an estimate of the *total cost* of the analysis.

GROUP WORK

Prepare your plan for data processing and analysis, considering the following points:

1. Sorting and quality control of data (10 minutes):

- How will the sorting be done? When?
- What quality checks should be made? Who will do them? When?

2. Processing of data (50 minutes):

- How will you do it (by hand or by computer)? If by computer: do you have enough experience and is the necessary equipment available?
- Prepare data master sheets for your proposal (preferably on flip charts).
- How many open-ended questions do you have that require categorising or coding? Who will do the categorising or coding? How much time will be required for data processing (taking into account the sample size)?

3. Analysing and interpreting the data (1 hour):

- Using the specific objectives and the list of variables, prepare dummy tables in which you relate variables to each other to analyse possible (causal) relationships. Select the dummy tables that you plan to fill in before we have our workshop on data analysis and reporting.
- Make estimates of the time and materials required for the analysis (in our case, only for the period up to the second workshop during which we will continue the analysis of data).

4. Prepare to present in plenary your master sheet, three dummy tables, a list of other important variables that you would like to cross-tabulate, and rough estimates of manpower, time, and materials required for the analysis of data (15 minutes).

EXERCISE: Analysis of responses to open-ended questions

Please analyse and interpret the following answers from the study of student nurses on smoking which were given to Question 7: **'Why do you smoke?'**

1. I have tried to give up so many times but I have been unable to.
2. I like the feel of the cigarette in my hand.
3. Because it gives me pleasure.
4. I do not see why I should give up smoking!!
5. Because I like to blow the smoke through my mouth and nose.
6. Because I feel confident and in charge when I am smoking.
7. It helps me to think better.
8. I like the image that comes with smoking.
9. I feel that people respect me more as a smoker.
10. All my friends are smokers.
11. It helps to make people more friendly and comfortable, especially when offering a cigarette.
12. Why not?!!
13. Smoking makes me feel like a man.
14. I like to blow smoke rings.
15. I like the taste.
16. It is too difficult to give up.
17. It helps me to relax.
18. It helps me to reduce the pressure and tension at work.
19. My wife likes a man who smokes.

Analyse and interpret these answers as follows:

- Develop a coding system by categorising the answers. First read all answers carefully. Then make rough categories of answers that seem to belong together. Try to limit yourself to 5-7 groups. Label each group of answers with a key word that seems to characterise the answers.
- List all answers again, but now in 5-7 groups under the labels you have selected.
- Discuss the groups, for example in terms of types of messages you would use to convince these smokers to stop smoking. This helps to identify whether the answers indeed belong together. You may split up some groups of answers and combine others. Find an appropriate 'label' for each category and count how many answers you have for each category.
- Come up with suggestions for interventions.

In reality you would have tried the coding system out on another sample of answers. Thereafter you would most likely still have slightly adjusted the coding system, after which you would have coded all answers on the questionnaires and entered the codes in the master sheet or computer.

Trainer's Notes

Module 13: PLAN FOR DATA PROCESSING AND ANALYSIS

Timing and teaching methods

1 hour	Introduction and discussion
½ hour	Exercise
2½ hours	Group work
1 hour	Plenary
5 hours	TOTAL TIME

Introduction and discussion

- Give a brief introduction to the topic, starting with an overview of your presentation.
- Explain and clarify terms such as sorting, quality control, processing, categorising, coding and tallying.
- Emphasise the importance of an adequate numbering system for the different data collection tools: they can be either numbered before going into the field or afterwards or both.
- Discuss the pros and cons of using a computer if one or more groups intend to use it. If computers are not being used in the course, the section on computer compilation can be omitted and participants can be asked to read it themselves if interested.
- Discuss the importance of having different data master sheets for different categories of informants.
- Pay special attention to how to deal with missing data: missing data should be recorded on the data master sheet so the groups can arrive at correct total figures. FOR EACH QUESTION/ITEM the total number of answers and the total number of missing values should add up to the total number of interviewees. If the totals are not correct, groups go astray when processing their data. This should be avoided by all means.
- Prepare an overhead sheet for dummy tables 13.2 and 13.3, and another sheet with the data filled in to put on top of the previous sheet. Let the participants interpret the data.
- In your conclusion, summarise the different components that should be included in a plan for data processing and analysis.

Exercise: Analysis of responses to open-ended questions

This can be done in the small working groups, before starting on the group work assignment. Assist the participants in analysing and interpreting the 19 answers given to the open-ended question. One way you might categorise and interpret the answers is given on the next page. Explain that the codes can be inserted on the master sheets.

Annex 13.1: Questionnaire for adults (15+) on smoking behaviour

Introduction

May I talk with you for a few minutes? We are students from the nursing school, and we are conducting a small study on smoking. We ask about 10 questions. Would you mind participating? (State you would like to interview non-smokers as well as smokers.)

If the informant agrees:

Everything we discuss will be treated confidentially (no one else will know who said what). But you should feel free to remain silent if you hesitate to answer a particular question. And if you have any questions for us, please feel free to ask them.

Questionnaire

Questionnaire number ☐

Date _____

Location _____

Interviewer _____

Put √ in the applicable box, unless otherwise indicated

1. Sex: male ☐
   female ☐

2. Age ☐ (number of years)

3. Occupation (If not self-supporting: add occupation of head of household) _____

4. Highest degree of education reached

(a) Total number of classes completed

(b) Highest type of school attended:

Primary school ☐

Secondary school ☐

Occupational training ☐ (specify) _____

Other ☐ (specify) _____

Is informant still in school?

Yes ☐

No ☐

5. Do you smoke?

Yes ☐ (go to Q6)

No ☐

If no: did you ever smoke?

Yes ☐

No ☐

(go to Q10)

If smoking

6. How many cigarettes do you smoke a day? ☐ (number of cigarettes)

7. How old were you when you started smoking? ☐ (number of years)

8. Why do you smoke?

_____

_____

_____

9. Have you ever tried to stop or reduce smoking since you started to smoke? (Probe for the number of times (s)he tried.)

Tried to reduce ☐ (state number of times)

Tried to stop ☐ (state number of times)

Comments of informant

_____

_____

_____

If not smoking

10. Why not? (Probe whether this was a conscious decision or not.)

_____

_____

_____

11. Do any of your close relatives (husband, wife, children; parents, brothers, sisters) smoke?

Yes ☐ (a) Define relationship of smoker(s) to informant:

(b) How close do they stay?
(Use one box per smoking relative. Make more boxes to tick if required.)
- Same house ☐ ☐ ☐ ☐
- Different house ☐ ☐ ☐ ☐

No ☐

12. If informant is working (see Q3):
Does anyone in your close working environment (e.g. same room) smoke?

Yes ☐ (give details) _____

No ☐

13. Do you ever feel disturbed by people in your close surrounding (at home, at work) who are smoking?

Yes ☐ (give details) _____

No ☐

For all

14. (a) Are you suffering from persistent cough (at least 2 weeks) at present?

Yes ☐

No ☐

(b) Have you ever had persistent chest complaints/cough?

Yes ☐

No ☐

(If yes, please explain what, when)) _____

_____

_____

15. Do you have children? Yes ☐ No ☐

If yes: How old are they? Age youngest ☐ Age oldest ☐

Is smoking an issue you will discuss (have discussed) with them when they reach(ed) adolescence?

☐ Strong yes (probe for details: positive or negative sanctions used)

☐ (Will) try, but

☐ No (explain)

16. Would you appreciate it if the nursing school started an anti-smoking campaign?

☐ Yes (explain)

☐ Indifferent (explain)

☐ No (explain)

17. How would you advise us to proceed?

(a) Which groups should we target(+why)

(b) In which way could we best approach them?

(c) With what messages?

Annex 13.2: Master sheet for female non-smokers

No. quest.	Q1 sex	Q2 age		Q3 Education			Q4 Occupation		Q5 Smoking		Q10 Why not?	Q11 Relatives smoking			Q12 working env. smoking
					High. level				Now	Ever					
		Yrs	Cat	Y/N	Yrs	Type	Self	Head HH	Y/N	Y/N		Y/N	Nr	Same hs	Y/N
1	F	18	1	N	NA	NA	0	3	N	N	2, 3, 4	Y	3 (Fa, Br, Br)	3	NA
4	F	46	4	Y	4	PS	1	3	N	N	3, 5, 6. Priest strict	Y	3 (Hu, Br, So)	2	NA
5	F	27	2	Y	8	PS	0	3 + 4	N	Y	1, 4	N	NA	NA	NA

Categories
15-24 = 1
25-34 = 2
35-44 = 3
45-54 = 4
55 + = 5

Categories
PS = Primary school
SS = Sec. school
OT = Occ. training
UY = University
O = Other

Categories
0 = No
1 = Irregular Insufficient
2 = Unskilled
3 = Farmer (with land)
4 = Small trader/crafts
5 = Teacher, nurse etc.
6 = Doctor, lawyer etc.

Categories
1 Bad taste
2 Parents forbid
3 Religion forbids
4 Waste of money
5 Never thought of it
6 Other (write on sheet)

Annex 13.2: Master sheet for female non-smokers (continued)

No quest.	Q13 Feeling disturbed		Q14 Cough		Q15 Children				Q16 Apprec. smok camp.	Q17 Target groups	Approach
			Now	Ever		>10	Discour.	Reason			
	Y/N	Reason	Y/N	Y/N	Y/N	Y/N	+ + / ± / -		Y/N		
1	N	You cannot change it	N	?	N	N	NA		Y	Youth + adult men	Theatre Posters
4	Y	1,2,3	N	Y	Y	Y	+ +	1,2	Y	Youth	Church, School
5	NA		N	N	Y	N	+ +	1,2,4	Y	Youth	School

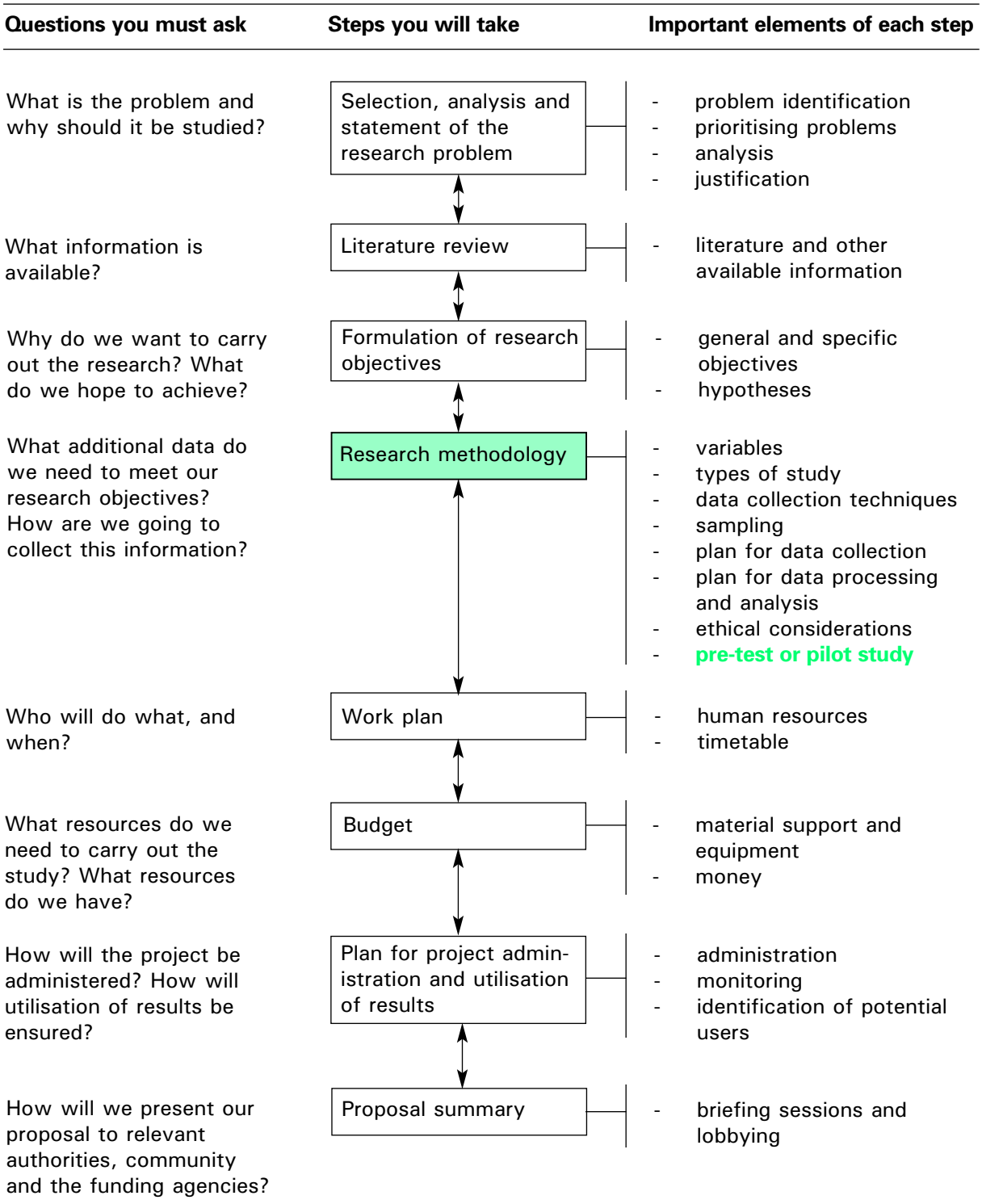
Categories
1. Bad example for others (children)
2. Worried about their health
3. Forbidden by church
4. Worried about own health (makes you cough)

Categories
1. Bad for health
2. Waste of money
3. Religion forbids
4. Bad habit
5. Cannot prevent

Module 14

PRE-TESTING THE METHODOLOGY

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 14: PRE-TESTING THE METHODOLOGY

OBJECTIVES

At the end of this session you should be able to:

1. **Describe** the components of a pre-test or pilot study that will allow you to test and, if necessary, revise your proposed research methodology before starting the actual data collection.
2. **Plan** and **carry out** pre-tests of research components for the proposal being developed.

What is a pre-test or pilot study of the methodology?

A PRE-TEST usually refers to a small-scale trial of particular research components.

A PILOT STUDY is the process of carrying out a preliminary study, going through the entire research procedure with a small sample.

WHY do we carry out a pre-test or pilot study?

A pre-test or pilot study serves as a trial run that allows us to identify potential problems in the proposed study. Although this means extra effort at the beginning of a research project, the pre-test and/or pilot study enables us, if necessary, to revise the methods and logistics of data collection before starting the actual fieldwork. As a result*, a good deal of time, effort and money can be saved* in the long run. Pre-testing is simpler and less time-consuming and costly than conducting an entire pilot study. Therefore we will concentrate on pre-testing as an essential step in the development of research projects.

WHAT aspects of your research methodology can be evaluated during pre-testing?

1. **Reactions of the respondents to the research procedures** can be observed in the pre-test to determine:

   - availability of the study population and how respondents' daily work schedules can best be respected;
   - acceptability of the methods used to establish contact with the study population;
   - acceptability of the questions asked; and
   - willingness of the respondents to answer the questions and collaborate with the study.

2. The **data-collection tools** can be pre-tested to determine:

   - Whether the tools you use allow you to collect the information you need and whether those tools are reliable. You may find that some of the data collected is not relevant to the problem or is not in a form suitable for analysis. This is the time to decide not to collect this data or to consider using alternative techniques that will produce data in a more usable form.

   - How much time is needed to administer the interview guide/questionnaire, to conduct observations or group interviews, and/or to make measurements.

   - Whether there is any need to revise the format or presentation of interview guides/ questionnaires, including whether:

     - The sequence of questions is logical.
     - The wording of the questions is clear.
     - Translations are accurate.
     - Space for answers is sufficient.
     - There is a need to pre-categorise some answers or to change closed questions into open-ended questions.
     - There is a need to adjust the coding system.
     - There is a need for additional instructions for interviewers (e.g., guidelines for 'probing' certain open questions).

3. **Sampling procedures** can be checked to determine:

   - Whether the instructions concerning how to select the sample are followed in the same way by all staff involved.
   - How much time is needed to locate individuals to be included in the study.

4. **Staffing and activities of the research team** can be checked, while all are participating in the pre-test, to determine:

   - How successful the training of the research team has been.
   - What the work output of each member of the staff is.
   - How well the research team works together.
   - Whether logistical support is adequate.
   - The reliability of the results when instruments or tests are administered by different members of the research team.
   - Whether staff supervision is adequate.

Note:

The pre-test can be seen as a period of extra training for the research team in which sensitivity to the needs and wishes of the study population can be developed.

5. **Procedures for data processing and analysis** can be evaluated during the pre-test. Items that can be assessed include:

   - Appropriateness of data master sheets and dummy tables and the ease of use.
   - Effectiveness of the system for quality control of data collection.
   - Appropriateness of statistical procedures (if used).
   - Clarity and ease with which the collected data can be interpreted.

6. The proposed **work plan and budget for research activities** can be assessed during the pre-test. Issues that can be evaluated include:

   - Appropriateness of the amount of time allowed for the different activities of planning, implementation, supervision, co-ordination and administration.
   - Accuracy of the scheduling of the various activities.

When do we carry out a pre-test?

You might consider:

- Pre-testing at least your data collection tools, either during the workshop, or, if that is impossible, immediately thereafter, in the actual field situation.

- Pre-testing the data collection and data-analysis process 1-2 weeks before starting the fieldwork, with the whole research team (including research assistants, if required) so that you have time to make revisions.

Which components should be assessed during the pre-test?

1. **Pre-test during the workshop**

Depending on how closely the pre-test situation resembles the area in which the actual field work will be carried out, it may be possible to pre-test:

- The reactions of respondents to the research procedures and to questions related to sensitive issues.
- The appropriateness of study type(s) and research tools selected for the purpose of the study (e.g., validity: Do they collect the information you need? and reliability: Do they collect the data in a precise way?).
- The appropriateness of format and wording of questionnaires and interview schedules and the accuracy of the translations.
- The time needed to carry out interviews, observations or measurements.
- The feasibility of the designed sampling procedures.
- The feasibility of the designed procedures for data processing and analysis.

Even if you cannot assess all these components fully, the field experience will provide information that will be quite valuable to you in reviewing the methodological aspects of your proposal and in planning your work plan and budget.

2. **Pre-test in the actual research area**

All the issues mentioned above will have to be thoroughly reviewed during a pre-test in the actual field situation. Other issues, such as the functioning of the research team, including newly recruited and trained research assistants, and the feasibility of the work plan, can only be tested in the research area. *An important output of the pre-test should be a fully developed work plan*.

If choices have to be made as to what to include in the pre-test, the following considertions may be helpful:

- **What difficulties do you expect in the implementation of your proposal?** Think of possible sources of bias in data collection techniques and sampling and ethical issues you considered during the preparation of your plan for data collection (**Module 12**). Can some of these potential problems be overcome by adapting the research design?

- **Inexperience with a certain data-collection technique** is also a reason to include it in the pre-test.

- **Which parts of your study will be most costly and time-consuming?** Questionnaires used in large surveys, for example, should always be tested. If many changes are made, the instruments should be pre-tested again. If an interview guide or questionnaire has been translated into a local language, the translated version should be pre-tested.

Note:

It is highly recommended that you analyse the data collected during the pre-test right away. Then finalise and adjust the master sheets, if necessary. Make totals for each variable included in the master sheets. Fill in some dummy tables and prepare all the dummy tables you need, considering your research objectives.

Do all this even if you plan to analyse the data by computer. You will detect shortcomings in your research tools that you can still correct!

Who should be involved in the pre-test or pilot study?

- The research team, headed by the principal investigator.
- Any additional research assistants or data collectors that have been recruited.

How long should the pre-test or pilot study last?

The time required for a pre-test or pilot study will be determined by a number of factors:

- The size and duration of the research project. (The longer the study will take, the more time you might reserve for the test run.)
- The complexity of the methodology used in the research project.

Keep in mind that this is the last chance you will have to make adjustments which will help to ensure the quality of your fieldwork. If you have a 20 day field work period you might reserve at least 3-5 days for pre-testing your data collection tools, analysing the results of the pre-test, finalising your tools and elaborating the work plan.

GROUP WORK I – to prepare the pre-test during the workshop (1½ hours+)

Only half a day will be available for conducting a pre-test of your methodology during the course:

1. Determine what parts of the methodology you would like to test. Include all data collection tools, if possible.

2. Decide with your facilitator and course manager where in the local area you could best carry out the pre-test.

3. Decide which members of your team will conduct various aspects of the pre-test. You are advised to work in pairs, so that you can discuss observations during the pre-test.

4. Prepare a short list of questions you wish to answer during the pre-test. (See **Annex 14.1** for suggestions.)

GROUP WORK II – after completion of the pre-test (4 hours)

1. Answer the questions you developed for the pre-test.

2. Determine whether your pre-test experience indicates that you need any:

   - changes in your research proposal
   - changes in your data collection tools

   Assign various group members to make these changes.

3. Determine what aspects of the study you would like to pre-test (again) in your research area and why, with whom, when and where.

4. After completing items 1-3, summarise the major points on a flip chart and in one or two paragraphs for your research proposal.

GROUP WORK III – (instead of Group Work I and II, if no pre-test is possible during the workshop, 3 hours)

1. Decide what aspects of the study you would like to pre-test in your research area and why, with whom, when and where. Summarise this information in one or two paragraphs in your research proposal.

2. Instead of doing a pre-test during the workshop, you might take the time to carefully review your research methodology and your data collection tools, using the checklist provided in **Annex 14.2**.

Annex 14.1: Summary of points to assess during a pre-test or pilot study

1. Reactions of respondents to your research procedures	Acceptable	Not acceptable	Suggestions
Availability of sample needed for full study			
Work schedules of population that may affect their availability			
Desire of population to participate			
Acceptability of questions			
Clarity of the language used			

2. The data collection tools	Acceptable	Not acceptable	Suggestions
Whether the tools provide the information you need and are reliable			
The time needed for administering each of the data collection tools			
Presentation of questions and format of questionnaire			
Accuracy of translation			
Pre-categorising of questions			
Coding system and coding guidelines			
Handling and administering the tools			

3. Sampling procedures	Acceptable	Not acceptable	Suggestions
Whether the instructions to obtain the sample are used uniformly by all staff			
The time needed to locate the individuals to be included in the study			

4. Preparation and effectiveness of research team	Acceptable	Not acceptable	Suggestions
Adequacy of staff training			
Output of each team member			
Team dynamics			
Reliability of tools when administered by different team members			
Accuracy of interpretation			
Appropriateness of plan for supervision			

5. Procedures for data processing and analysis	Acceptable	Not acceptable	Suggestions
Use of data master sheets			
Effectiveness of data quality control			
Appropriateness of statistical procedures			
Ease of data interpretation			

6. Schedule for research activities	Acceptable	Not acceptable	Suggestions
Amount of time allowed for: • field trips for data collection			
• supervision			
• administration			
• analysis of data			
Sequence of activities			

Annex 14.2: Summary of possible fallacies in the design and implementation of studies

As we have now gone through all steps of the study design, including the planning of data processing and analysis, it may be useful to summarise the critical points at which a researcher can go wrong:

- in the SELECTION of RESPONDENTS or study elements, and
- in the COLLECTION of data.

These potential errors should be reviewed while pre-testing your research methodology.

1. Errors in selection of respondents or study elements.

In the selection of respondents we may distinguish several major possibilities for error.

Too limited (or inappropriate) definition of the study population or use of incorrect sampling procedures, for example by:

- Studying registered patients only;
- Obtaining responses from male opinion leaders only (if one needs the opinion of the whole community);
- Choosing a sample because it is close to a road or in some other way easier to access (tarmac bias); or
- Conducting the study during only one season of the year (when results may be biased by not including other seasons or because access is difficult).

Errors in the assignment of research subjects to study groups in analytic and experimental studies:

- Defective matching in case-control studies;
- The inclusion of volunteers for study groups in cohort studies;
- Non-randomisation in experimental studies; or
- If randomisation is impossible: failure to develop a quasi-experimental design that corrects as much as possible for 'rival explanations'.

Selective dropouts or non-response

Dropouts or subjects who do not respond to selected questions may represent a special category of respondents. If attrition is high or the rate of non-response excessive, results may be biased.

In cohort studies, follow-up of individuals can be difficult overtime. Bias in follow-up emerges if there is a differential drop out between those exposed to the risk and those not exposed.

2. Errors in data collection

We may obtain:

Invalid data, by applying indicators and measuring techniques or instruments which do not adequately measure what we want to measure.

Unreliable data due to:

- Variation in the characteristics of the research subject measured, as a consequence of the research;
- The use of unstandardised measuring instruments; or
- Differences between observers/interviewers in asking questions or observing.

Reliability of data collected is always required, but it is of crucial importance if we want to measure changes over time. If we find changes we must be sure that these are not caused by errors in our research methods that could have been prevented.

All the above-mentioned shortcomings may threaten the **validity** of your findings and conclusions.

The shortcomings can be prevented to some degree by being alert to them when designing and implementing the study; otherwise they have to be mentioned in the study design.

Trainer's Notes

Module 14: PRE-TESTING THE METHODOLOGY

Timing and teaching methods

½ hour	Introduction and discussion
1-½ hours	Group work I (to prepare the pre-test)
4 hours	Actual pre-test
4 hours	Group work II (to discuss pre-test results and revise the data collection tools)
1 hour	Plenary
11 hours	TOTAL TIME

Note: If no pre-test can be carried out during the workshop, consider using the Group Work III alternative (3 hours) instead of Groups Works I and II.

It is important to note that well before this session (preferably after **Module 10**) the course manager should ask the groups what data collection tools they want to pre-test and on whom. Each research team, with the assistance of its facilitator, should decide how many interviews or observations will be conducted and should begin making arrangements for obtaining necessary copies of the instruments and other supplies needed. The course manager should make arrangements for all the groups (i.e., look for suitable sites to do the pre-test, inform (health) authorities and/or local leaders of the plans, ask for their consent, and arrange transport).

Introduction and discussion

- Discuss the concept and process of pre-testing or conducting a pilot study of the methodology, covering the questions posed in the module.
- Refer to **Annex 14.1**, 'Summary of points to assess during a pre-test or pilot study.' Using the annex, review briefly the important aspects of pre-testing covered in the presentation.
- Discuss the pre-test that will be undertaken during the workshop and make sure that each research team knows where it is going to do the pre-test and with whom it is going to meet. Explain how much time each team has for the preparation of the pre-test (Group Work I), the actual pre-test, its evaluation (group work II) and for reporting and discussion in plenary. Stress the importance of working in pairs during the pre-test, so that experiences can be shared.
- Emphasise that it is important to make notes on all observations during the pre-test so that they can be discussed afterwards.

Group work I

Ask the participants to meet in their working groups to design the pre-test for their project.

Pre-test field exercise

If necessary, an instruction sheet should be prepared for the field exercise including information such as:

- How the field exercise will be organised, i.e.:

  - Where each working group will go;
  - What pairs of participants will work together;
  - What formalities need to be observed with community leaders, directors of health facilities, and respondents;
  - What explanation should be given concerning the purpose of the pre-test and whether any feedback will be given to those participating as respondents;
  - How many interviews or observations should be conducted;
  - Time available for the exercise, etc.

- Points that should be assessed during the pre-test. (These could include some or all of the points listed in **Annex 14.1**.)

- When and where the working groups should reassemble after the field exercise for the group work II session.

Group work II

Arrange for each group to meet after its members have come back from the field to discuss and analyse their experiences and to revise their data-collection tools and possibly other aspects of the research methodology. Ask each group to prepare a short report of its main findings and conclusions.

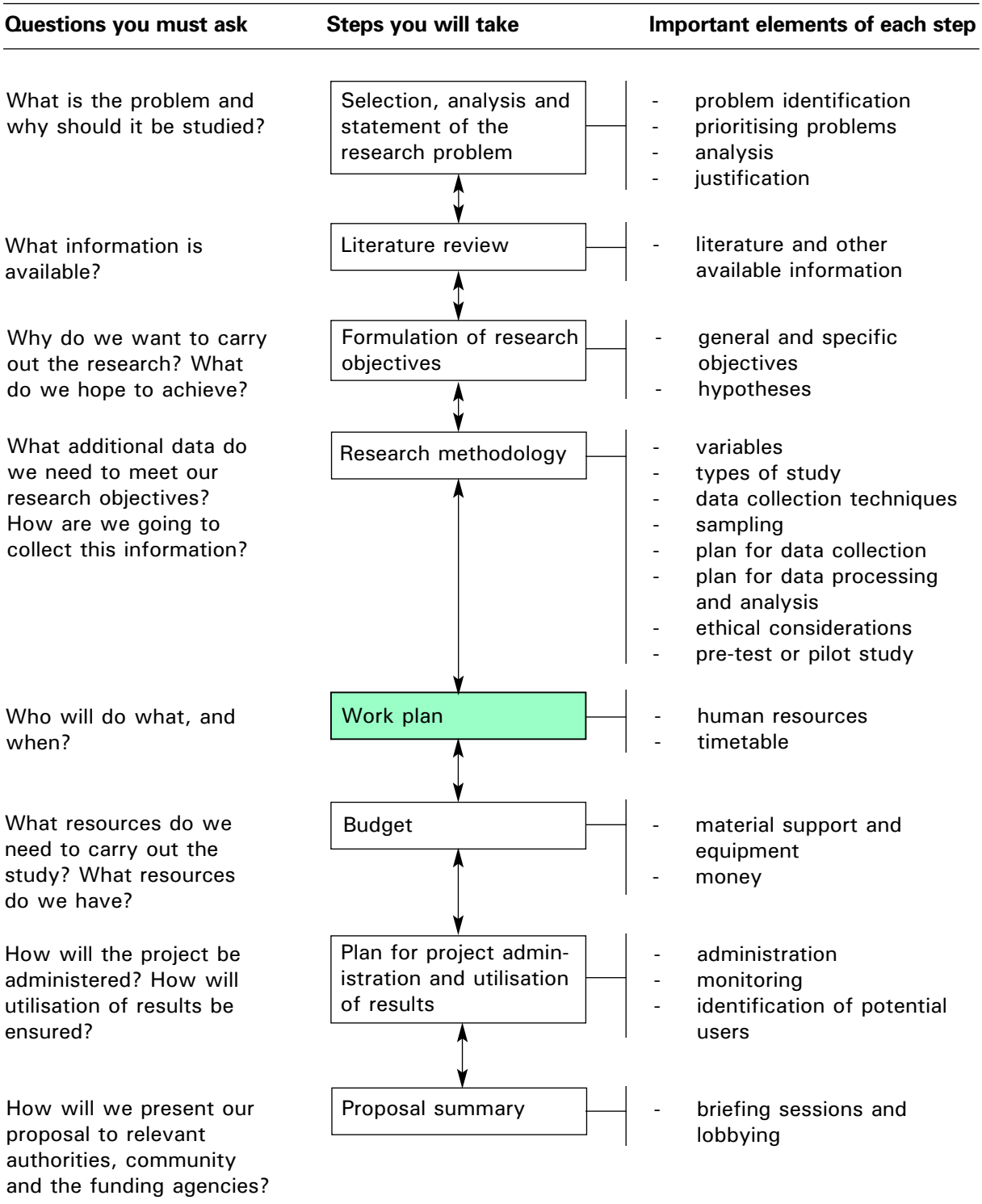
Plenary

One member of each group should report in plenary on the main findings and conclusions of the pre-test.

Module 15

WORK PLAN

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 15: WORK PLAN

OBJECTIVES

At the end of this session you should be able to:

1. **Describe** the characteristics and purposes of various project planning and scheduling techniques such as 'work scheduling' and 'GANTT charting';

2. **Determine** the various tasks and the staff you need for your project and justify any additional staff (research assistants, supervisors) apart from the research team that developed the proposal, where you will recruit them, for how long a period you need them, and how you will train and supervise them.

3. **Prepare** a work schedule, GANTT chart and staffing plan for the project proposal you are developing.

I. Introduction

II. Work scheduling and planning techniques

I. INTRODUCTION

What is a work plan?

A WORK PLAN is a schedule, chart or graph that summarises the different components of a research project and how they will be implemented in a coherent way within a specific time-span.

It may include:

- The tasks to be performed;
- When and where the tasks will be performed; and
- Who will perform the tasks and the time each person will spend on them.

II. VARIOUS WORK SCHEDULING AND PLANNING TECHNIQUES

1. The work schedule

A WORK SCHEDULE is a table that summarises the tasks to be performed in a research project, the duration of each activity and who is responsible for the different tasks.

The version of a work schedule given on the following page includes:

- The tasks to be performed;
- The dates each task should begin and be completed;
- Research team, research assistants and support staff (drivers, typists) assigned to the tasks
- Person-days required by research team members, research assistants and support staff (The number of **person-days** equals the number of working days **per person**)

Note:

The period for field research for the course project should not exceed 6 months. Week 1 is the first week after completion of the present workshop.

This work schedule was developed for a study of factors associated with maternal mortality in Mongo district, Zambia. The research team consisted of five persons (mainly district health team members). The study consisted of two main parts: (1) analysis of health facility-based maternal deaths; and (2) analysis of community-based maternal deaths.

EXAMPLE OF A WORK SCHEDULE:

Factors associated with maternal mortality, Mongo district, Zambia

TASKS TO BE PERFORMED	DATES	PERSONNEL ASSIGNED TO TASK	NO. OF PERSON-DAYS REQUIRED
1. Finalise research proposal and submit to MOH for clearance	2-7 March 1999	Team Team leader (TL)	6x2 = 12 days
2. Translate questionnaires, typing, multiplying	23-27 March 1999	Team (partly) secretary	3x2 = 6 days 1x2 = 2 days
3. Recruit research assistants	6-25 April 1999	TL.	1x1 = 1 day
4. Obtain clearance and orient of DHOs, health institutions and village health workers	6-25 April 1999	TL. secretary	1x5 = 5 days 1x2 = 2 days
5. Train research assistants	5-7 May 1999	Team + assistants	5x3 = 15 days 5x3 = 15 days
6. Pre-test study	11-30 May 1999	Team + assistants + facilitator + drivers	5x2 = 10 days 5x2 = 10 days 1x2 = 2 days 2x2 = 4 days
7. Collect data	1 July 1999 - 31 December 1999	Team + assistants + 2 drivers	5x20 = 100 days 5x20 = 100 days 2x50 = 100 days
8. Process data + make preliminary interpretation	End of each month of data collection	Team + facilitator	5x6 = 30 days 1x2 = 2 days
9. Analyse data and write report	3-15 Feb. 2000	Team + facilitator + secretary	5x15 = 75 days 1x15 = 15 days 1x5 = 5 days
10. Disseminate and discuss research findings and preliminary recommendations with community members, health staff	End of Feb. 2000	Team + secretary + drivers	5x1 = 5 days 1x1 = 1 day 2x1 = 2 days
11. Disseminate and discuss of research findings and preliminary recommendations with policymakers/ managers/ others	End of Feb. 2000	Team	5x1 = 5 days
12. Draft preliminary Plan of Action	End of Feb. 2000	Team	5x2 = 10 days
13. Hold meetings with policy makers/managers/others to discuss plan of action for implementing recommendations	March 2000	Team	5x1 = 5 days
14. Follow up on implementation of Plan of Action	From March onwards	TL and rest of team	

You will notice that each team member roughly spent 30 working days on the research, except the Senior Clinical Officer who was the Team Leader. He spent time to recruit research assistants and to visit district authorities to obtain their support for the study. Though he integrated these tasks with his normal clinical duties, he spent about 6 more working days than the other team members did. Five research assistants were recruited to assist with the interviewing. The number of working days required (20) were multiplied by five (for the research team as well as for the research assistants) to arrive at the number of person-days. Another 21 days (6 + 15) were spent on data analysis and report writing, whereas three days were reserved for feedback of findings and preliminary recommendations to all parties concerned, and two for drafting a plan of action to implement them.

How to develop a work schedule

- Review and revise, if necessary, the list of tasks you prepared for your Plan for Data Collection (**Module 12**). Add to the list other tasks you must complete not related to data collection (such as clearance of proposal; data analysis and report writing; and feedback to authorities and target group). Number all tasks.

- Now review the staffing for the different tasks, taking into account your experience during the pre-test. Consider:

  - Who will carry out which tasks;
  - The amount of time needed per research unit (interview/observation/record) including travel time; and
  - The number of staff needed to complete each task in the planned period of time.

  Make revisions, if required. Complete the staffing for the tasks you have just added.

- Consider whether the use of short-term consultants is necessary for certain tasks. Always consider using local consultants. If consultants are used, involve them in the planning stage of the project so you can incorporate any useful suggestions they may have concerning the design of the methodology.

In reviewing your tentative staffing plan you should ask:

- Are the types of personnel and levels of expertise you require likely to be available for the project? **For example**, is there a sufficient range of disciplines available including, where appropriate, personnel from outside the health field?

- If special staff has to be recruited or reassigned from other ministries or agencies, what regulations or procedures will have to be followed?

- Is the staffing plan realistic, taking into account the project budget that is likely to be available?

- To what extent can community members, traditional healers, students or other non-professionals be involved in the study?

- What training would the research assistants/data collectors require? How long would the training last? Who would do the training? How do you intend to supervise the assistants/data collectors? Review what you have tentatively planned in **Module 12** and revise it, as necessary.

Then fix the dates (in weeks) indicating the period in which each task will have to be carried out and calculate the number of working days per person required to complete each task.

2. The GANTT Chart

A GANTT chart is a planning tool that depicts graphically the order in which various tasks must be completed and the duration of each activity.

The GANTT chart shown on the following page indicates:

- the tasks to be performed;
- who is responsible for each task; and
- the time each task is expected to take.

The length of each task is shown by a bar that extends over the number of days, weeks or months the task is expected to take.

How can a work plan be used?

A work plan can serve as:

- A tool for planning the details of the project activities and drafting a budget.
- A visual outline or illustration of the sequence of project operations. It can facilitate presentations and negotiations concerning the project with government authorities and other funding agencies.
- A management tool for the Team Leader and members of the research team, showing what tasks and activities are planned, their timing, and when various staff members will be involved in various tasks.
- A tool for monitoring and evaluation, when the current status of the project is compared to what had been foreseen in the work plan.

When should the work plan be prepared and when should it be revised?

- The first draft of the work plan should be prepared when the project proposal is being developed, so the schedule can be discussed easily with the relevant authorities.
- A more detailed work plan should be prepared after the pre-test in the study area.
- There should be no hesitation in revising work plans as necessary, based on a reassessment of what can be realistically accomplished in the coming months.

What factors should be kept in mind when preparing a work plan?

- It should be simple, realistic, and easily understood by those directly involved.
- It should cover the preparatory and the implementation phases of the project, as well as data analysis, reporting, dissemination and utilisation of results.
- The activities covered should include training, technical or research tasks; administrative, secretarial and other support tasks.
- The realities of local customs (local holidays, festivals) and working hours should be considered, when preparing the work plan.
- Also seasonal changes and their effect on travel, work habits, and on the topic you are studying (such as incidence of disease or nutritional status), should be kept in mind as the schedule is planned.

Example of a GANTT Chart: Factors associated with maternal mortality, Mongo district, Zambia

TASKS TO BE PERFORMED	1999 J F M A M J J A S O N D	2000 J F M A
1. Finalise research proposal and submit for clearance		
2. Translate questionnaires, type, multiplie		
3. Identify research assistants		
4. (After obtaining clearance:) Inform DHOs, health institutions and VHWs		
5. Train research assistants		
6. Pre-test study and finalise procedures/tools		
7. Collect data		
8. Process data + make preliminary interpretation		
9. Analyse data + write report		
10. Feedback, discuss and disseminate research findings and recommendations with policy makers, staff and community members		
11. Draft preliminary plan of action		
12. Meet with policy makers/managers/others to discuss action plan for implementing the recommendations		
13. Follow-up implementation		→

Computer-assisted formulation of work plan:

There are a number of computer software packages on the market that one can use to prepare and monitor the implementation of a work plan. Microsoft Project Manager, Excel, and Lotus are among the commonly used software packages.

Budget preparation

The work plan is the starting point for developing your budget. Specify, for each activity in the work plan, what resources are required. Determine for each resource needed the **unit cost** and the **total cost**.(See **Module 16**.)

GROUP WORK (3 hours)

Prepare a work plan for inclusion in your proposal, following the steps below:

1. Start with the development of a work schedule:

- List all tasks to be carried out, completing and revising the list of tasks you prepared for your plan for data collection.
- Consider who will carry out each task, the number of working days required per person to complete each task, the number of staff you will need to finish each task in a given period of time, and the period in which you plan to actually carry out each task.
- Look at a calendar and note any public holidays or other important activities scheduled for the period (about 6 months) in which you plan to conduct the fieldwork.
- Include your facilitator in stages of the fieldwork where you feel you would require assistance (e.g., during training of research assistants or during the initial period of data collection in the field). If needed, schedule the use of a local consultant.
- Do not forget to include support staff required (typists, drivers, for example).

2. Consider whether the number of days each member of the research team plans to invest in the fieldwork is adequate for the task and will be acceptable. (It should most likely not exceed 30 working days.)

3. Prepare a GANTT Chart to include in your proposal.

4. Include two or three paragraphs on the staff required for your research and their tasks in your work plan, including:

- Composition of research team itself and the tasks of various members;
- Reasons for recruiting research assistants/data collectors/supervisors, where you will recruit them, what their tasks will be, for how long you will need them, and how you will train and supervise them;
- The role of facilitators during the field work and when they will be needed; and
- Whether any other consultants will be needed and, if so, what skills they should have and what their tasks would be.

5. Copy your work schedule and GANTT chart on flipcharts or overhead sheets, for use in the exercise below and in the plenary discussion.

EXERCISE (Optional): Project work plan

Review the work plan that another group developed for their research proposal and provide constructive criticism.

Trainer's Notes

Module 15: WORK PLAN

Timing and teaching methods

½ hour	Introduction and discussion
3 hours	Group work
¼ hour	Exercise (optional)
1 hour	Plenary
4½ hours	TOTAL TIME

Introduction and discussion

- Introduce and discuss the aims and uses of a work plan, encouraging participants who have experience in making work plans to contribute actively.

- Pay special attention to the concept of person-days. It is important that everyone understands the concept as the groups need to calculate the number of person days for various tasks when making their work plans as well as preparing the budgets for their research proposals.

- The importance of having a detailed and realistic work plan that is at the same time flexible should be stressed.

Group work

Ask the participants to prepare work plans for their research proposals taking into account the plans for data collection that they already prepared (**Module 12**). Ask them to start by listing the tasks to be performed in the correct sequence. Then they should estimate the time involved for each task and assign the tasks to various staff members and consultants (if needed). Encourage each group to think seriously about what staffing pattern would be most cost-effective and efficient for its particular research project.

Exercise: Project work plan (optional)

Ask each group to look at the work plan developed by one of the other groups and provide constructive criticism.

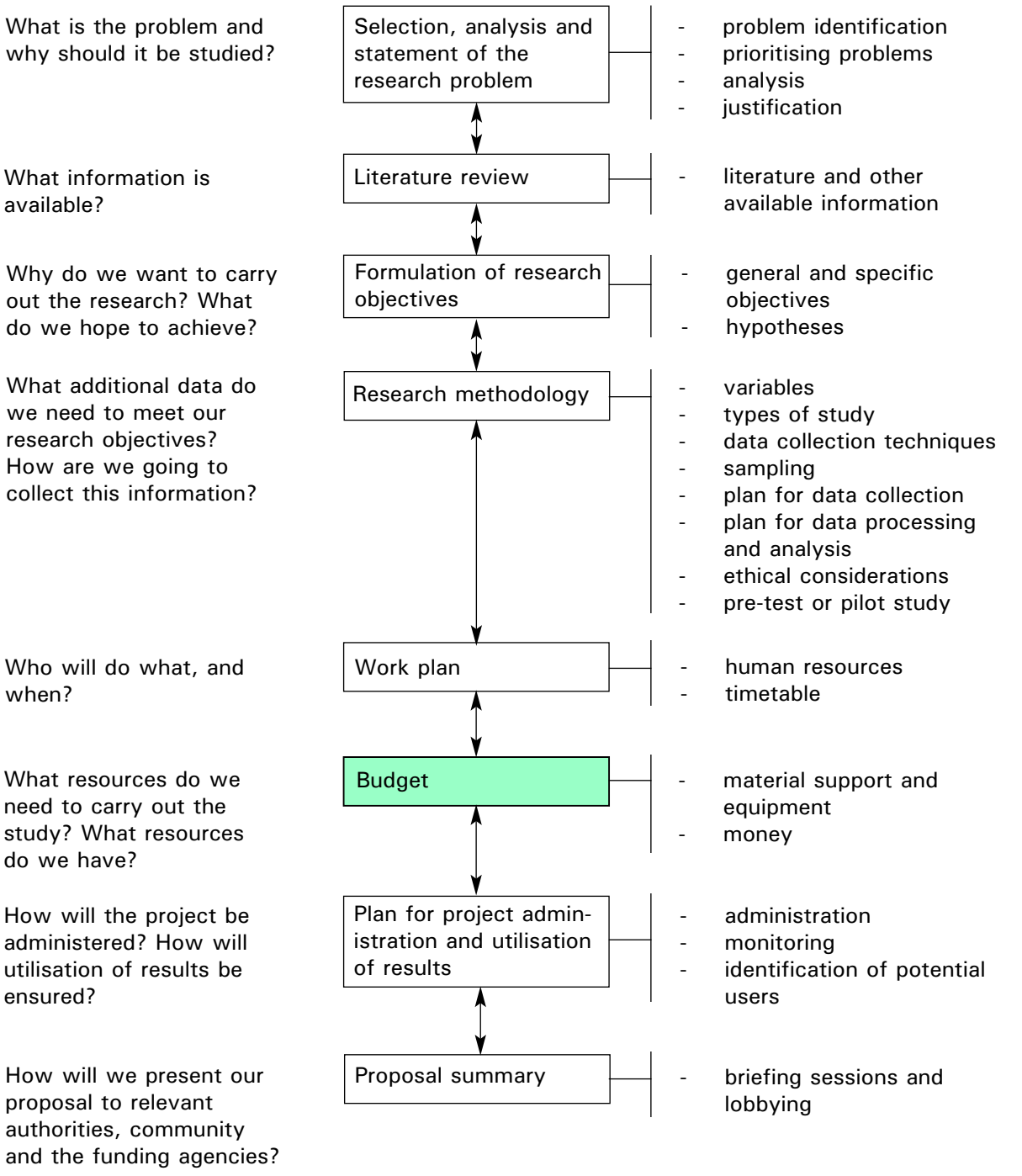
Plenary

Ask each group to present its work schedule, Gantt chart, and staffing plan followed by a short discussion.

Module 16

BUDGET

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 16: BUDGET

OBJECTIVES:

At the end of the session you should be able to:

1. **Identify** major categories for a budget.

2. **Make** reasonable estimates of the expenses in various budget categories.

3. **List** various ways a budget can be reduced, if necessary, without substantially damaging a project.

4. **Prepare** a realistic and appropriate budget for the project proposal being developed during the course.

Why do we need to design a budget?

- A detailed budget will help you to identify which resources are already locally available and which additional resources may be required.

- The process of budget design will encourage you to consider aspects of the work plan you have not thought about before and will serve as a useful reminder of activities planned, as your research gets underway.

When should budget preparation begin?

A complete budget is not prepared until the final stage of project planning. However, cost is usually a major limiting factor and therefore must always be kept in mind during planning so that your proposals will not have an unrealistically high budget. (See **Module 4**: Analysis and statement of the problem.) Remember that both ministries and donor agencies usually set limits for research project budgets.

The use of locally available resources increases the feasibility of the project from a financial point of view.

How should a budget be prepared?

It is necessary to use the work plan (**Module 15**) as a starting point. Specify, for each activity in the work plan, what resources are required. Determine for each resource needed. the **unit cost** and the **total cost**.

Example:

In the work plan of a study to determine factors affecting maternal mortality in the Mongu District, Western Province, Zambia, it is specified that 5 research team members will each visit 20 households, one per working day, as 100 households of deceased mothers will need to be visited. With two drivers this amounts to 50 working days per driver. Each research team member will be accompanied by one of the research assistants.

The budget for the fieldwork component of the work plan will include funds for personnel, transport and supplies.

Note:

that UNIT COST (e.g., per diem or cost of petrol per km), the MULTIPLYING FACTOR (number of days), and TOTAL COST are required for all budget categories.

EXAMPLE OF A BUDGET: Costs involved in fieldwork for a maternal mortality study

CATEGORY	UNIT COST (Kwachas)	MULTIPLYING FACTORS	COSTS (Kwachas)
I. ALLOWANCES			
a. Researchers (5)			
• Training research assistants	40 per day	5x3 = 15 days	600
• Field work during pilot study	60 per day	5x2 = 10 days	600
• Field work during actual study	60 per day	5x20 = 100 days	6 000
b. Research assistants (5)			
• Training	25 per day	5x3 = 15 days	375
• Field work during pilot study	40 per day	5x2 = 10 days	400
• Field work during actual study	40 per day	5x20 = 100 days	4 000
c. Secretary (1)			
• Typing questionnaire	25 per day	1x5 = 5 days	125
• Finalising report*	25 per day	1x5 = 5 days	125
• Typing and sending circulars	25 per day	1x2 = 2 days	50
• Typing and sending invitations	25 per day	1x1 = 1 day	25
d. Drivers (2)			
• Field work during pilot study	20 per day	2x5 = 10 days	200
• Field work during actual study	20 per day	2x50 = 100 days	2 000
• Feedback	20 per day	3x1 = 3 days	60
e. Facilitator (1)			
• Evaluation pilot study	40 per day	1x2 = 2 days	80
SUBTOTAL (I)			MK 14 620
II. **TRANSPORT COSTS**			
a. Field work during pilot study (10 cases, 150 km per case)	0.50 per km	1 500 km	750
b. Field work during actual study (100 cases, 150 km per case)	0.50 per km	15 000 km	7 500
c. Travel costs for facilitator	0.50 per km	400 km	200
d. Travel costs for feedback in field	0.50 per km	200 km	100
e. Travel costs for those who attend meeting in March 2000			400
SUBTOTAL (II)			MK 8 950
III. SUPPLIES			
a. Stationary			2 000
b. Lunch for those who attend meeting in March 2000			600
SUBTOTAL (III)			MK 2 600
TOTAL (I+II+III)			MK 26 170
5% CONTINGENCY			1 310
GRAND TOTAL			MK 27 480

The Ministry of Health will make available the necessary manpower to conduct the study, as well as vehicles. The total budget required from external sources amount to Kwacha (Kw) 27,480. At an exchange rate of 2,75 to US$ 1, this is equivalent to US$ 9,990.
If more than one budget source will be used (e.g., the Ministry of Health and a donor), it would be

* NB: If team members do not write the report themselves on computers, one secretary will be required for 10 days during the actual writing of the report.

useful to indicate in the budget which source will pay for each cost. Usually a separate column is used for each funding source. The Ministry of Health may, for example, provide salary costs and operational costs for vehicles, whereas the donor agency is asked to provide per diem (according to local regulations), petrol/public transport expenditures and stationary.

Advice on budget format

The type of budget format to be used may vary depending upon whether the budget will be supported by your own organisation or submitted to the Ministry of Health or a donor organisation for funding. Most donor organisations have their own special project forms, which include a budget format. If you intend to seek donor support it is advisable to write to the potential funding organisation as early as possible during the period of project development.

Advice on budget preparation

- Keep in mind the tendency to underestimate the time needed to complete project tasks in 'the real world'. Include a 5% contingency fund if you fear that you might have budgeted for the activities rather conservatively. (If inclusion of a contingency fund is not allowed, an alternative is to slightly over-budget in major categories.)

- Do not box yourself in too tightly with very detailed categories and amounts, especially if regulations do not allow adjustments afterwards. Ask the supervising agency to agree that, if necessary there may be some transfer between 'line items' in the budget.

- If your government or department has agreed to contribute a certain amount for the project, try to arrange that the contribution be administered separately, so that the administrators remain aware of the commitment. This may also ensure easier access to the funds.

- If the budget is for a period longer than a year, build in allowances for inflation before the project begins and in subsequent years by increasing costs by a set percentage. (If inflation is high in the local economy, you may have to build in allowances for even shorter projects.)

Budget justification

It is not sufficient to present a budget without explanation.

The budget justification follows the budget as an explanatory note justifying briefly, in the context of the proposal, why the various items in the budget are required. Make sure you give clear explanations concerning why items that may seem questionable or that are particularly costly are needed and discuss how complicated expenses have been calculated. If a strong budget justification has been prepared, it is less likely that essential items will be cut during proposal review.

How can budgets be reduced?

- Explore whether other health-related institutions are willing to temporarily assign or second personnel to the project.

- When possible, use local rather than outside personnel. If consultants are needed at the beginning, train local personnel as soon as possible to take over their work.

- Explore the use of students or community volunteers, where appropriate.

- Plan for strict control of project expenditures, such as those for vehicle use, supplies, etc.

Obtaining funding for projects

To conduct research, it is usually necessary to obtain funding for the research project. Such funding may be available from local, national or international agencies. In addition to preparing a good research proposal, the following strategies are useful for researchers who need to obtain their own funding:

1. Familiarise yourself with the policies and priorities of funding agencies. Such policies and priorities may be:

   - Explicit, i.e., available from policy documents issued by the agency.
   - Implicit, i.e., known to officials in the agency and to other local researchers who have previously been funded by that agency.

   Obtain the names of such persons and make direct contact with them.

   The funding policies of many agencies may emphasise:

   - a priority given to research aimed at strengthening a particular program (e.g. MCH, PHC)
   - institution building (i.e., building the capacity of an institution to do research)
   - research credibility

   Annex 16.1 gives a list of some prominent research funding agencies.

2. Identify the procedures, deadlines and formats that are relevant to each agency.

3. Obtain written approval and support from relevant local and national health authorities and submit this together with your proposal.

4. If you are a beginning researcher, associate yourself with an established researcher. Host agencies scrutinise the 'credibility' of the researcher to whom funds are allocated. Such credibility is based on previous projects that have been successfully completed.

5. Build up your own list of successfully completed projects (i.e., your own reports, publications, etc.).

GROUP WORK (2½ hours)

1. Prepare a budget for your project. Keep in mind the importance of having a realistic budget, for which resources can actually be found.

2. Examine the work plan in your project proposal and consider See the expenses involved in completing each component. Local rules should be followed for calculating per diems, travel cost and overtime (if required).

3. Indicate for each item, the UNIT COST as well as the NUMBER OF UNITS. Justify large budget items, travel and allowances in one or two paragraphs attached to the budget.

4. Consider the 'cost-effectiveness' of various budget levels. Will the final results be worth the expense?

5. Consider the budget level that possible funding authorities would consider appropriate:

   - Examine their guidelines.
   - If appropriate, talk with donor representatives about their policies.

6. If additional funding is requested from an outside donor, make clear what contribution the Ministry of Health and your own institution is making.

Annex 16.1: INTERNATIONAL SOURCES OF FUNDING FOR RESEARCH

1. International Multi-Lateral Agencies

WHO and Associated Special Programmes:

IMCI (Integrated Management of Childhood Illnesses)
MCH, Reproductive Health, Adolescent Health, etc.
RBM/AIM (Roll Back Malaria/African Initiative on Malaria)
TDR (Tropical Disease Research)
UNAIDS
WHO Headquarters
WHO Regional Offices

African Development Bank
Asian Development Bank
IARC (International Agency for Research on Cancer)
UNICEF (United Nations Children's Fund)
UNFPA
World Bank etc.

2. Bilateral Agencies

ADAB (The Australian Development Assistance Board)
BOSTID (Board on Science and Technology for International Development)
CIDA (Canadian International Development Agency)
DFID (Department for International Development), United Kingdom
DGIS (Directorate for International Co-operation), The Netherlands
GTZ (Deutsche Gesellschaft für Technische Zusammenarbeit, Germany)
IDRC (International Development Research Centre), Canada
JICA (Japanese International Co-operation Agency)
SAREC (Swedish Agency for Research Co-operation with Developing Countries)
SIDA (Swedish International Development Agency)
USAID (United States Agency for International Development) etc.

3. Private Foundations

Carnegie Corporation of New York
Ford Foundation (Child Health) (USA)
International Health Policy Program (USA)
Kellogg Foundation (Health Services; primary interest in Latin America)
Rockefeller Foundation (USA)
Welcome Trust (UK) etc.

4. National Sources

This will vary from country to country.

NB: You can find addresses through Internet or through the Embassies of the respective countries.

Trainer's Notes

Module 16: BUDGET

Timing and teaching methods:

¾ hour	Introduction and discussion
2½ hours	Group work
1 hour	Plenary
4¼ hours	TOTAL TIME

Introduction and discussion

- Introduce and discuss important issues related to project budgeting.
- Invite input from those participants who have experience in making budgets.
- Explain the concepts of UNIT COST (e.g., 50 kwachas/km) and MULTIPLYING FACTOR (e.g., 1500 km as the total distance), and make sure that everyone understands them.
- Refer to the budget example in the text.
- Emphasise the importance of a budget justification.
- Discuss useful strategies for reducing a budget that is too high.

Before asking participants to work on the budgets for their own projects, it is important to announce STANDARD unit costs both for transport (distance or fuel) and for allowances. These should conform to Ministry of Health regulations. If there are limits to budgets, these should be agreed upon beforehand as well.

Group work

Ask the participants to meet in their groups and prepare the budgets for their own projects, based on the work plan (**Module 15**). Ask them to specify what the contribution of their own institution/ministry will be to the project and what the external donor is requested to contribute.

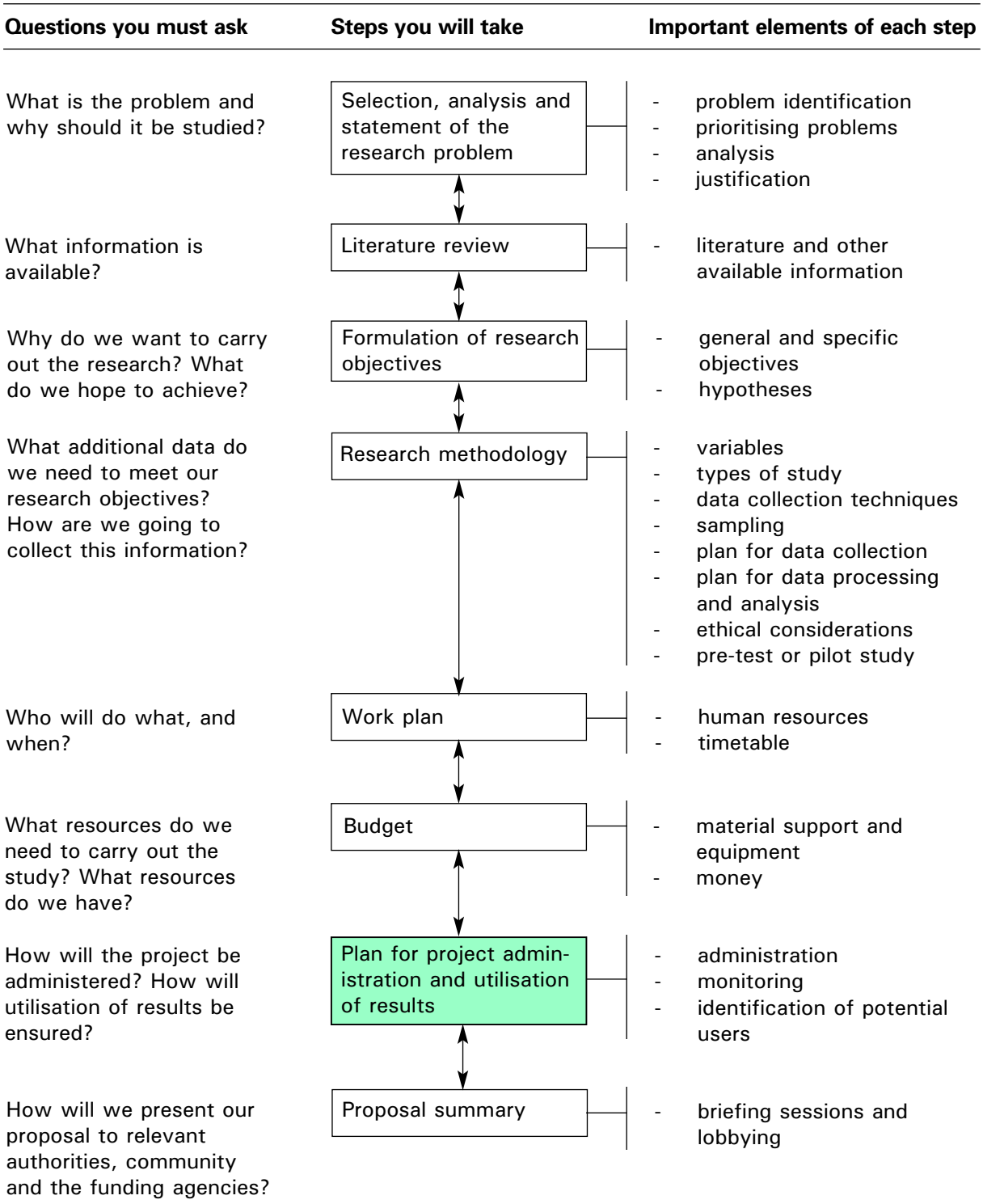
Plenary

Ask each group to present its budget proposal in plenary. Allow sufficient time for discussion after each presentation.

Module 17

PLAN FOR PROJECT ADMINISTRATION MONITORING, AND UTILISATION OF RESULTS

Flowchart: Steps in the development of a health systems research proposal



NB: Development of a research process is a cyclical process. The double-headed arrows indicate that the process is never linear.

Module 17: PLAN FOR PROJECT ADMINISTRATION, MONITORING AND UTILISATION OF RESULTS

OBJECTIVES

At the end of this session you should be able to:

1. **List** the responsibilities of the team leader and project administrator related to the administration and monitoring of an on-going project.

2. **Prepare** a brief plan for administration and monitoring of the research project being developed.

3. **Prepare** a plan for actively disseminating and fostering the utilisation of results for the project proposal being developed.

I. Administrating research projects

II. Project monitoring

III. Planning for the dissemination, communication and utilisation of research results

I. ADMINISTRATING RESEARCH PROJECTS

What is project administration?

PROJECT ADMINISTRATION is the term for all the activities involved in managing the human, material, financial and logistical resources of a project.

Why is good administration important in a research project?

- It allows for orderly and accurate purchase and procurement of equipment, payment of bills, and preparation of financial reports.

- It allows researchers to foresee the need for and to make timely requests for resources in order to avoid unwanted breaks in the implementation of the project.

- It allows researchers to devote most of their time to the technical and scientific aspects of the project.

What administrative issues should be considered as the project proposal is being finalised?

As a team developing a research project, you should now consider the following issues:

- One of the research team members should be selected by you as **team leader** (TL). A team leader is the 'first among equals'; (s)he is ultimately responsible for implementing the proposal as planned and for solving possible problems that may arise. The TL should be selected based on his or her leadership skills, commitment to the project, availability, and the ability to facilitate the smooth implementation of the project. The TL should co-ordinate the team's official contacts with the Ministry of Health and with other relevant (funding, research or service) institutions. If the TL is not available or is for some reason unable to provide the leadership needed, the research team, in consultation with the HSR unit or facilitators, may select another team leader.

- An organisational unit or official has to be identified outside the team, who has the power to receive and handle funds: a **project administrator** (PA). The research team has to consider what service unit is best able to:

  - work in collaboration with the team leader and funding authorities to ensure an adequate flow of funds, including petty cash for minor expenses; and
  - avoid creating unnecessary bureaucratic or administrative difficulties that may hinder the implementation of the study.

- Procedures for ensuring the smooth procurement and flow of funds should have been worked out immediately after the workshop between the research ream and the MOH, the facilitators and possible external donors, so that the research teams can start working as soon as possible after official approval to implement the study has been obtained. Since the documents that must be prepared for this purpose require the signatures of the TL and PA, they need to be finalised during the workshop.

What would the tasks of the team leader, related to project administration, include?

- Supplying the project administrator or the administrative team with a copy of the research proposal and making sure they understand the work of the researchers and when funds are needed.

- Delegating selected administrative tasks to other members of the research team.

- Alerting administrative officials in a timely fashion concerning staff, materials, equipment and funds needed during various stages of the project.

- Supervising the flow of funds, project accounting and preparation and submission of financial reports.

- Discussing with the relevant authorities in the Ministry of Health (Health Research Unit for example) any difficulties encountered in the project and attempting to identify appropriate solutions.

What administrative operations need to be supervised by the team leader at the end of the project?

- Working with project administration to plan for 'end of project' activities, such as arranging for termination or transfer of staff, making an inventory of supplies and equipment and dispensing them, if required, and arranging for any final payments and financial accounting.

- Overseeing the preparation and distribution of the final administrative/financial report.

- Making sure that all financial obligations are met.

II. PROJECT MONITORING

What is project monitoring?

MONITORING is the on-going process by which information is gathered concerning the implementation and evolution of the research project. Monitoring involves activities designed to keep track of resources available and used and the quantity and quality of the operations carried out during each phase of the project so that its objectives will be met.

Monitoring should continue throughout the project and be organised so that it is helpful in alerting staff to problems that develop and changes needed. It is a valuable management and learning tool for everyone concerned.

During monitoring sessions you will review:

- The resources needed for the project, including staff, equipment, supplies, logistical support and funds, to assess if they are available when needed and being appropriately used;

- The activities of each team member and their relations to the project as a whole, to assess if the work plan is being carried out as planned and what delays or difficulties, if any, have emerged that need to be addressed;

- The flow and quality of the data that are being collected; and to what extent they meet the objectives or answer the research questions; and

- The research team's communication and co-ordination with the study population, other collaborating groups, and funding authorities.

Note:

Monitoring will usually take place at team meetings during field activities. If there is a gap in the fieldwork, it may be necessary to convene a special meeting.

It is advisable to keep close track of changes in the work plan and problems encountered and solved (or not solved) so that you can inform your facilitator and superiors, and include this information in your preliminary report. (See **Module 20**.)

III. PLANNING FOR THE DISSEMINATION, COMMUNICATION AND UTILIZATION OF RESEARCH RESULTS

Before you finish drafting your research proposal you should start planning how the results of your study could be used.

Why should researchers be concerned about dissemination, communication and utilisation of research results?

The fundamental reason for undertaking health systems research is to obtain results that can be used to improve health and health care.

Who will be interested in the results?

Depending on the topic you selected, the results may be useful to the community, staff and managers of health and health-related services and to researchers and donor agencies in your own country, as well as others.

However, above all, you as a research team and your program should benefit from the results, as you have developed the proposal to help solve one of your own priority problems.

What strategies can you follow to ensure that the results of your study will be used?

1. Involve relevant authorities, staff and community members in the selection of your topic and in the definition of your problem.

If possible, these groups should be consulted before the proposed development workshop begins. If the final decision for a certain topic is made during the workshop, however, not all parties concerned may have been consulted. If not, they should be consulted immediately after the workshop.

2. List two or three major recommendations you expect to obtain from your study and identify who should be involved in their implementation.

Here we must distinguish between two categories of people who should be involved:

- Those who authorise you to implement the recommendations, and
- Partners in the implementation process.

Most likely you will be authorised to implement certain recommendations yourself but for others you will need the approval of your superiors and/or of decision-makers from other sectors. Some authorities may merely need to give their approval, but you may need the active collaboration of others during the application of the results. Furthermore you will need to identify from which colleagues, subordinate staff and target groups in the community co-operation will be required for the formulation and implementation of the study's recommendations.

3. Identify which communication channels already exist which can be used to discuss and disseminate results.

Channels for discussing and disseminating results may include, for example:

- Provincial or District Development Team meetings;
- Provincial or District Health Team meetings;
- Supervisory visits to health facilities involved; staff meetings;
- Mobile clinics or other health activities carried out in villages included in the study; monthly meetings of village health workers when they collect drugs; meetings of village health committees.

Keep relevant parties informed of progress during implementation of the study and plan to obtain their input when study findings and recommendations are being drafted.

4. Determine what written materials should be prepared to keep relevant parties informed. They may include:

- A one to two page summary of your project proposal that includes details on expected results, to distribute when you introduce the project to policy makers and staff concerned.

- An introductory statement to use with interview guides and questionnaires, explaining to informants the purpose and procedures of the study, as well as expected results. This introduction could also be used when you introduce the project to policy makers in the village.

- A progress report of four to five pages, including preliminary findings and recommendations, which you will prepare for presentation of the data analysis and report writing workshop. This report can also be used to inform authorities that will be crucial to utilisation of project results.

- The draft report of findings and recommendations, prepared during the data analysis workshop. The summary of this report can be used for discussion with policy makers and staff. However, to obtain feedback from decision makers and target groups in the community, you will need a different summary, concentrating in simple words on the findings and preliminary recommendations that directly concern them.

Make sure that summaries of your findings and preliminary recommendations are adapted to the level of understanding and interests of different audiences. This will increase their motivation to provide thorough feedback and to participate in the implementation of the final recommendations collectively agreed up on.

5. Determine whether additional actions should be taken or mechanisms developed to discuss the study results with all parties concerned and obtain their input, approval and co-operation for the implementation of the recommendations. These may include, for example:

- Special visits to top policy maker(s) by the team leader or the whole research team to report on progress during the fieldwork and/or to discuss preliminary results and recommendations.

- The invitation of the most crucial persons for implementation of your recommendations to the last day of the data analysis workshop, when you will present your preliminary findings and recommendations in plenary for discussion.

- Special discussions with policy makers, staff and representatives of the target groups concerned to finalise the findings and recommendations of the study and develop a plan for action.

For complex studies of relatively long duration it may be advisable to have a Project Advisory Committee, representing the major parties involved. Since the projects developed during workshops will in general not last longer than six months, you may be able to keep key individuals or representatives informed through ad hoc or even routine meetings.

Note:

Never forget to report the findings to the subjects/community /organisation studied before the report is finalised. This should be done to fulfil an obligation to those studied, to obtain information on possible errors in your draft report, and to discuss your proposed recommendations and obtain useful feedback. (See **Module 33**)

GROUP WORK (1½ hours)

1. Develop a plan for administrating and monitoring your project. Consider the following questions as you develop your plan:

   Administration

   - Who will be the team leader for your project?
   - Which organisational unit or which official would be best able to administer the project? (Remember that the team leader cannot also be the project administrator.)
   - Which authorities are likely to fund the project?
   - How can a smooth flow of funds be assured?
   - Who will do the project accounting and file and submit receipts?

   Monitoring

   - What aspects of the project will be monitored and who will be responsible?
   - How will the monitoring activities be organised and when will they take place?

2. List two or three major recommendations expected from your study in a table and identify who should be involved in their implementation:

Expected recommendations	Is the research team authorised to implement these recommendations?	Is further authorisation reguired? If so, from whom?

3. Determine what channels or mechanisms you will use (or develop) to keep parties from whom you require authorisation and/or co-operation for implementing recommendations informed concerning the project (1) before starting the field work; (2) after completing the field work; (3) after preparing the draft report of findings and recommendations.

4. Identify the one or two authorities who are most crucial for implementation of your recommendations so that they can be invited for the presentation and discussion of your findings and recommendations at the end of the data analysis workshop.

5. Present the results of your group work on a flipchart and prepare several paragraphs on project administration, monitoring and the utilisation of results for inclusion in your research proposal. Do not forget to include the dissemination and utilisation of the results in your work plan and, if necessary, in the budget.

Trainer's Notes

Module 17: PLAN FOR PROJECT ADMINISTRATION, MONITORING AND UTILISATION OF RESULTS

Timing and teaching methods:

¾ hour	Introduction and discussion
1½ hour	Group work
1 hour	Plenary
3¼ hours	TOTAL TIME

Introduction and discussion

- Give a brief overview of the topics that will be covered in the presentation.
- Stress that the team leader is not necessarily the most senior group member, but should be the best organiser.
- Give a brief introductory presentation on project administration and monitoring and its importance. Highlight administrative activities that should be undertaken by the team leader before, during and at the end of the research project.
- It is important to discuss how a smooth flow of project funds can be guaranteed. These procedures should have been agreed upon with the relevant authorities in the Ministry of Health and with possible external donors **immediately after** the workshop.
- Introduce and discuss the importance of drafting a plan for the dissemination, communication and utilisation of research results and what such a plan should contain.

Group work

Ask the participants to meet in their working groups and develop a plan for project administration, monitoring and utilisation of results. A brief summary should be made for presentation in plenary and for inclusion in the research proposal.

Plenary

Have each group present its plan for project administration, monitoring and utilisation of results, followed by discussion.

Module 19

FIELDWORK ACTIVITIES

The fieldwork phase

Questions you must ask	Steps you will take	Important elements of each step
Will managers, health staff and target groups provide support?	Administrative and modivational preparation	Give briefing; obtain permission; form research advisory groups
Are data collection tools and instruction manuals ready? Have research assistants been trained?	Preparation for data collection	Logistic preparations Pre-testing and revisions Training research team Arrangements for supervision and quality control
Is data collection of good quality and on schedule?	data collection	Checking for completeness and consistency; Sorting of data
For qualitative data: Are more (or different) data needed?	Preliminary analysis of data	Categorisation and coding of qualitative data Preparation and filling in of master sheets for manual processing of data Preliminary interpretation of data Coding instructions for entry of data in computers

MODULE 19: FIELDWORK ACTIVITIES

PROCEDURAL GUIDE

for course participants, so that during the field period they will have:

1. **Briefed** managers, health service staff and community members regarding the project.
2. **Obtained** the necessary permission to collect data.
3. **Identified** and obtained the resources (manpower, materials etc.) needed to collect data.
4. **Reviewed** the availability of subjects and information and organised logistics for data collection.
5. **Trained** interviewers/research assistants/supervisors.
6. **Refined, pre-tested and revised** the research instruments and procedures for data collection.
7. **Collected** the required data.
8. **Processed** the data.

The activities to be carried out between the workshop on proposal development and the workshop on data analysis and report writing consist of field operations and data processing. The facilitator/ resource person for the project should visit the research team at least once but preferably twice, including:

- during the training of research assistants (if required) and pre-test, and
- at the onset of data processing.

(S)he should also be available for consultation by telephone.

The activities that should be completed during the inter-workshop period are discussed below:

1. Briefing of managers, health service personnel and community members concerned

The **purpose** of the briefing is to obtain support for the project. Such support is necessary in order to obtain resources as well as to obtain permission to collect data. Attention to the following points will increase your chances of obtaining permission to conduct the study and being allocated adequate resources:

- **Selecting the relevant audience(s) for the briefings**

It may be necessary to obtain resources and permission for the study at several levels and from different organisations. Briefings should be conducted with:

- your direct superiors;
- managers or key persons in the institutions, organisations, and communities that are being studied; and
- other key persons or organisations that will be involved in the research or utilise the results.

Note that 'key persons' should include the official leaders (e.g., village headmen) as well as unofficial opinion leaders (leader of women's cultivation group). In a hospital, the nursing sisters in charge of wards are important opinion leaders although they may not be designated as hospital managers.

Because of the differing interests of these audiences it may be advantageous to brief them separately, using a different emphasis in each presentation.

• **Winning support**

Present the project as an institutional project with yourself (team leader) as the advisor or leader of the project. Do NOT present it as a project you have to do as a training or research *exercise*. The briefing should enable the audience to recognise the benefits the project will bring to their own unit or community. This will encourage them to 'adopt' the project and provide support.

Develop strategies to overcome resistance and generate support. For example, identify a high level officer who is likely to be supportive and invite him or her to the briefing. Evidence of his/her interest will influence others who are lukewarm.

2. Identifying and obtaining project resources

Identify and obtain the resources (manpower, materials etc.) needed to collect data. Refer back to your project document so as to make sure that all the items needed for the study are included.

• **Requesting assistance**

Identify specific types of assistance that will be needed and present these requests diplomatically during the briefings. For example, 'Do you think health centre staff could help fill in two questionnaires per day over a period of 6 weeks?' is more likely to receive a positive response than 'I need manpower' or 'I need nurses for this study'.

3. Reviewing availability of subjects and information*

It is important to make a personal visit to every site where data will be collected in order to understand the physical and manpower limitations, constraints and special circumstances that could influence data collection. During the visit:

- **Discuss** with the staff/community members who are on site any routine procedures and patterns of behaviour (e.g., working hours, holidays) which may affect availability of subjects.

- **Observe the physical conditions and procedures** that are being followed to determine how they will affect your proposed data collection procedures. Remember that data collection will be reliable only if it does not overburden busy staff members or disrupt routine procedures. If researchers familiarise themselves with the actual situation at the site, it is often possible to design data collection procedures that do not interfere with on-going activities.

- Try to **use local personnel** for data collection, as they are more aware of local customs and problems. They may be less expensive, require less training, and will be less disruptive. Don't use them, however, in interviews about the quality of their own work with informants who may know them (see **Module 10A** III (bias)).

- If sources of data include registers, cards, etc., **inspect a sample of the data sources** so that you are able to modify the data collection tools (compilation sheets) in order for the research team to obtain the data with a minimum waste of time.

- **Identify suitable members of staff** who can be research assistants and additional supervisors, if the research team itself will not be able to do all the supervision. If data collection has to be done after office hours, remember to devise a system of supervision for those hours as well.

* Sections 3 - 9 are partly based on the following: The Population Council (1970) *A manual for Surveys of Fertility and Family Planning: Knowledge, Attitude and Practice*. New York, and: Institute Kesihatan Umum (1986) *National Health and Morbidity Survey, Supervisors' Manual.* Kuala Lumpur. (Monograph document)

4. Organising logistics for data collection

After finishing an inventory of available resources, the logistics for data collection should be organised. This will involve planning in detail how, where and when data collection will be carried out, elaborating the work plan prepared earlier.

5. Preparing fieldwork manuals

Manuals or instruction sheets should be prepared for:

• **Interviews**

The manual for interviews should have instructions concerning the

- Purpose of the study
- Role of the interviewers
- Way interviewers should introduce themselves to informants
- Interview techniques
- Interview guide/questionnaire:
  · general format
  · clarification of terms and what the research units are (e.g., household, family, respondent)
  · instructions regarding how to ask complicated questions (e.g., whether to mention pre-categorised answers or not and whether to probe for more than one answer or not)
  · instructions concerning how to fill in answers (e.g., the need to write answers to open-ended questions using the words of the informants)
- Use of the map (if any)
- Sampling procedures (and what to do if informant is absent, etc.)

• **Other data collection techniques**

Guidelines should be prepared for the implementation of any focus group discussions and interactive or projective research techniques that will be used, so that all members of the research team, including research assistants, will follow the same approach.

There should be guidelines concerning any measurements that will be made, including instructions on:

- what to measure and how, and
- how to properly calibrate measuring instruments

• **Supervision**

In addition to all instructions given above, the manual should include a separate section on supervision, with directions, for example, on:

- maintaining a record of attendance of research team members
- safe-keeping of data and records
- recording the number of interviews/FGDs/observations, etc., completed each day
- ensuring the quality control of field work
- dealing with non-responses and incomplete interviews, and
- reporting progress at specific intervals, to superiors and/or funding agencies.

6. Training of research team members, including assistants as well as supervisors

The research team including, in particular, research assistants who join just before the pre-test, must be given explicit training. They should not only be able to collect data properly but also understand other procedures such as the selection of sampling units, map reading and data handling. They may also be involved in the pre-test and in the adjustment of instruction sheets and data collection tools after the pre-test.

The training programme usually consists of:

- discussion on the objectives and methodology of the study,
- reading of manuals or instruction sheets prepared for the study,
- interview training (see **Module 10B section V** and **Annex 12.1**),
- field experience (this should include participation in the pre-test described below), and
- discussion on data-collection tools and instruction sheets and how they need to be adjusted (based on field-testing).

The research assistants should be trained together with the whole team, including possible additional supervisors.

7. Conducting the pre-test in the research location, with preliminary data analysis and revision of data collection tools

- The pre-test should assess the validity of the data-collection instruments and procedures, as well as the sampling procedures.

- Reread **Module 14** before planning your pre-test.

- Arrange for your **facilitator** to **visit during the training of interviewers and pre-test.**

- The study may involve the use of a variety of methods of data collection such as

  - collection of data from recorded sources,
  - face-to-face interviews using interview guides/questionnaires,
  - focus group discussions, and
  - measurements or observations.

- **Plan to pre-test all your methods.**

- **Analyse the data you collect during the pre-test**. Finalise and fill in master sheets, including quantitative as well as qualitative data (using key words). Fill in some of the cross-tables. This process will help you make a realistic assessment of the entire data collection and analysis process and will invariably lead to revisions of some of the tools.

- **The pre-test should identify scientific as well as logistical problems and constraints**. Discuss these with your facilitator.

- **Revise the data-collection and data analysis tools and procedures** after the pre-test. Arrange for typing and copying or duplicating of the tools. Check all forms for accuracy before duplicating (see **Module 10B**). Make sure that sufficient materials and manpower are available for this process. If a computer will be used for analysis, prepare a coding manual.

8. Collecting data

Having obtained permission for the study, and having

- obtained the necessary resources,
- trained the team members,
- organised the logistics, and
- pre-tested and modified the data collection tools and procedures,

the data collection can now be carried out.

9. Processing data

After collecting and sorting the data, all data should be checked for errors. The content may be converted into codes or keywords for processing by computer or using master sheets.

The steps during this process include:

1. editing/cleaning,
2. categorising and coding,
3. summarising data on data master sheets, or
4. if a computer is used, writing instructions to the computer analyst concerning data input and analysis.

Note:

Reread **Module 13** for more information on steps in data processing and analysis.

1. Editing

During editing look for:

- Completeness of responses. (Note that a blank space may mean 'no response' or 'don't know' unless you've made a category for each of these responses.)
- Logical inconsistencies, correcting them whenever possible.
- The possibility of combining responses, if that is more suitable for analysis.(See **Module 13** for guidelines on making scores.)

Editing should be done by the research team or under its direction. If several persons are involved in editing, as in the case of large surveys, an **editing manual** should be compiled beforehand.

2. Categorising and coding

A coding manual or coding instructions on the data collection tools, if required, should have been completed when the questionnaire was finalised after the pre-test. Look at **Module 13** and **Module 23** for coding instructions and for instructions on how to process data from open-ended questions, and **at Module 10C** for processing qualitative data from FGDs with key words.

3. Summarising data on master sheets

After the data have been edited and coded they may be summarised on data master sheets.

- Review the master sheets you developed during the proposal development workshop. Have your questionnaires changed since you made your master sheets? Can you categorise the answers to certain questions that you were unable to categorise before?
- Remember that you can use letters to represent the different categories of your variables (e.g., M for male, F for female), and key words.
- Then fill in your data on the master sheets. Do not forget to include information on missing data and non-responses.
- Prepare frequency counts for the variables tabulated in your master sheets and check if they match with the number of respondents in your sample.

4. Computer analysis

If the study is large, or if there are other reasons for the use of a computer, instructions should be written for the computer analyst.

After editing, coding and summarising the data, a preliminary analysis can be made by hand or using the computer. (See **Module 20.**)

Final note:

Despite all the advise presented in this module, some EMERGENCIES may arise during the field work. WHAT SHOULD YOU DO?
1. Use common sense
2. Consult your principal investigator and co-researchers
3. Consult your research proposal
4. Consult the modules
5. Write/phone/fax/email your facilitator
6. Other (specify)

Trainer's Notes

Module 19: FIELDWORK ACTIVITIES

Timing and training methods:

10-15 minutes Presentation and discussion

Presentation and discussion (10 minutes)

This module has been prepared to provide course participants with a succinct guide that covers all the tasks they must complete during the fieldwork period. The module need not to be presented in detail, but participants should be made aware of its content so that they will remember to consult it during appropriate stages in their fieldwork.

Role of facilitators during the fieldwork

Research teams should receive at least one, and preferably two visits from a facilitator during the fieldwork period. If only one visit is possible, this visit should focus on:

1. Checking the progress of the project;
2. If necessary, assisting in obtaining managerial support;
3. Observing the real-life situation in which the project will be implemented, identifying problems and anticipating pitfalls;
4. Evaluating the proposed methodology for data collection with the group (by discussing as well as by pre-testing in the field) and advising any necessary modifications in the research design (sampling, data collection tools);
5. Assisting in the training of research assistants (if required); and
6. Finalising and trying out the procedures for data processing and analysis during the pre-test, and, if a second visit is possible, assisting at the onset of data processing and analysis.

Suggested checklist

1. Determine whether managers and health staff concerned have been adequately briefed.
2. Determine whether permission for data collection has been obtained. (As resource person you can provide support by making courtesy calls.)
3. Review each of the proposed methods of data collection, reassessing:
   - The sampling frame, sampling procedures and sample size
   - The data collection tools that have been developed. Make sure that the tools collect the necessary data for each variable or theme, while not collecting any unnecessary information.

4. Make sure that the research team has visited data collection sites and identified working procedures and conditions, constraints, possible additional resources.

5. Determine whether the pre-test of data collection has been well planned. (Assist in any further planning, if necessary.)

6. Assist with the pre-test and training of research assistants and advise on:

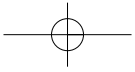
   - revision of data collection tools;
   - preparation and adjustment of the field work manual;
   - supervision of data collection;
   - editing and coding of the data collected; and
   - processing of data (data master sheets).

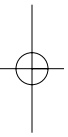
   NB: The sequence of these activities is arbitrary. Revision of instruments may occur twice, before and after pre-testing. Fieldwork manuals may be prepared before the pre-test but adjusted thereafter.

7. Make arrangements for telephone or written consultation during subsequent stages of the fieldwork, encouraging the participants to contact you when necessary.

Module 20

FIELDWORK REPORT





MODULE 20: FIELDWORK REPORT

OBJECTIVES

After reading this module you should be able to:

1. **Summarise** your field experiences and observations, including technical or logistical difficulties encountered in carrying out your research project.

2. **Assess** the extent to which you are able to answer the specific objectives with the data you have collected.

3. **Summarise** your main findings and preliminary conclusions for each objective.

4. **Identify** areas in which you need to do further analyses and specify in what sets of data you will find the data.

5. **Produce** a preliminary report which covers all the issues mentioned above.

I. Introduction

II. Content of the fieldwork report

I. INTRODUCTION:

Elaborating a fieldwork report is an entry point to HSR training on data analysis and research report writing. The research team, research assistants, if there were any, and facilitators should participate in this activity (see Module 15). Why should you prepare a fieldwork report that summarises your fieldwork experiences, observations, and preliminary conclusions? This will help you to:

- get a clear overview of the data collected (both qualitative and quantitative), your field observations and impressions, and consider how different sets of data work together to answer the research questions implied in your objectives;
- assess how well your research project was designed and thus the extent to which you can provide valid information to help solve the problem you investigated;
- develop the general approach you will use in reporting your findings and drawing conclusions;
- allow the facilitators and the other groups to provide you with feedback which will help you to identify what further analyses to make and how to organise the final report; and
- assess what you have gained from the data analysis workshop by comparing your preliminary and your final report.

II. CONTENT OF THE FIELDWORK REPORT

Your fieldwork report should include:

- A review of your field work experience, and
- A summary of preliminary findings.

1. Fieldwork experience

Review your fieldwork experience and evaluate how well you were prepared technically (in terms of the methodology developed in your research proposal), and organisationally (work plan, budget and administrative procedures). Summarise your experience and your evaluation of it in at most two pages. Address questions such as those posed below:

- **General**

- How did you function as a group? Were all group members active?
- Did you lose any members? Did you recruit any new members?
- What procedures did you follow to obtain permission for the research?
- Did you manage to obtain the research assistants, equipment, transport, and financial support needed?
- Were the resources you budgeted sufficient?

- **Technical preparations**

- What did you do to train your research assistants? Where and how did you do your pre-test or pilot study? How long did it take? Were any major revisions of the data collection tools and/or the research procedures necessary?

- **Fieldwork**

  - Did you do your sampling the way you had originally planned? Did you obtain the information and co-operation you wanted? How many interviews did you conduct? How does your planned sample size compare with the actual sample collected? (NB: If you have different categories of informants, specify for each group.) How many records were analysed?
  - Were your data collection tools adequate? Did they provide you with the information you wanted?
  - Were you able to follow your work plan? Did you correctly estimate the manpower and time needed to collect the data?

- **Technical support**

  - Did you receive support from your facilitator/resource person? In what phases of the fieldwork? Was the support timely? Was it sufficient or would more support have been helpful?

2. Preliminary research findings

When presenting research findings:

- First of all, **get an overview of all the data you collected and processed**.

  - Review any record forms or checklists you've completed. Has all the data you wanted to obtain been collected?
  - Review your master sheets or any computer printouts available. Are they complete?
  - Do the total number of responses for each question agree with the number of informants?
  - Have answers to open-ended questions been listed and categorised?
  - Have results from Focus Group Discussions if conducted), interviews with key informants, and/or field observations been transcribed in full, discussed and coded?

- **Reread the plan for data analysis** in your research proposal. Review the preliminary analysis of data you conducted during the pre-test (see **Module 14**) in the field. You may have done some useful ground work for data analysis which can be used now as your prepare your preliminary findings.

- **Reread your statement of the problem and the objectives**.

  Take the SPECIFIC OBJECTIVES as a starting point. Brainstorm as a group on the data you have collected and to what extent they appear to answer the research questions implied in your objectives.

  Consider not only the quantitative data from record forms and relevant sections of your questionnaires but also qualitative data and relevant observations you made or impressions you gained during the fieldwork.

  Discuss whether (and how) the data from various sources complement or contradict each other.

- **Record the details of these discussions**. This will help you structure the report you are going to write, keeping focused on major issues and yet not forgetting other relevant information.

- **Analyse** the VARIABLES that further describe the nature, size and distribution of your problem and make a brief summary.

- If you did an analytical study on a practical problem, prepare (at least) two **tables** for each objective, showing how crucial independent variables relate to dependent variables. (Review the dummy tables you prepared when developing your research proposal and determine which of them you can use).

  If you have mainly untabulated qualitative data, just **summarise** how crucial parts of the data you collected will answer the questions implied in your specific objectives.

Note:

If you have gone further with preliminary data processing and analysis, **state what you have done and what remains to be done**.

GROUP WORK

1. **Select a team member** to be responsible for taking notes on your discussions and **divide responsibilities** for writing specific parts of the preliminary report.
2. **Complete the review of your fieldwork experiences** as proposed in section 1 of this module, stressing the problems you experienced and whether and how they were overcome. This review should be very brief, at most two pages. Include a summary description of your sample populations (persons and/or records).
3. **Discuss your specific objectives one by one**, brainstorming on whether you have data to answer the research questions implied in your objectives. Remember to consider not only data obtained from record reviews or interviews but also from informal interviews with key informants, FGDs and your own observations during fieldwork.
4. **State tentative conclusions you can draw from your research at this stage**, using all data available.
5. Give a brief **overview of how far you have proceeded with data analysis** and what remains to be done:

   • **Processing of material:**

   - Have master sheets been filled (or if a computer is used, has data entering been completed)?
   - Has all qualitative data been listed and categorised/coded?

   • **Preliminary analysis:**

   - Have all straight frequency counts been done?
   - To what extent have cross-tabulations been made?
   - Has the interpretation of qualitative data been completed?

The preliminary report should not exceed four to five pages. Try to have it computerised or clearly written. It can be distributed at the beginning of the data analysis workshop. The main points from the report can be put on overhead sheets on the first workshop day for presentation in plenary.

Trainer's Notes

Module 20: FIELDWORK REPORT

Timing and teaching methods

15 minutes Introduction and discussion

It is necessary to present this module during the workshop. Facilitators should stress that participants should consult the module when they have filled in their master sheets or have received the first printouts from the computer. It will guide them in beginning data analysis and preparing the preliminary report that the team leader will present on the first day of the data analysis workshop.

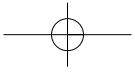
If a facilitator can pay two visits to the field, the second visit should preferably take place at the onset of data processing. He or she can then work with the group on this task, using **Modules 13 and 20** as reference material.

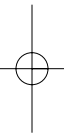
Note:

The discussion of **Modules 19** and **20** can easily be combined. The best opportunity to discuss them may be just before the course evaluation, when groups have handed in the last draft of their research proposals for finalising by a course secretary.

ANNEX to the modules: GUIDELINES FOR ORGANISING SHORT HSR COURSES; Part I: PROPOSAL DEVELOPMENT AND FIELDWORK*

* See also: The Population Council (1970) *A manual for Surveys of Fertility and Family Planning: Knowledge, Attitude and Practice.* New York and
Institute Kesihatan Umum (1986) *National Health and Morbidity Survey, Supervisor Manual* Kuala Lumpur, Malaysia. (Monograph document)





I. PLANNING FOR THE WORKSHOP

1. Selection of course facilitators

The course co-ordinator will be the chief organiser of the course. A course co-ordinator will usually be supported by at least four facilitators, depending on the number of research teams. Together, they will be responsible for planning the course content, preparing instructional objectives and guiding the learning process throughout the course. They will give lectures, facilitate group sessions and guide research projects. The course co-ordinator is usually attached to the HSR Unit of the Ministry of Health.

Course facilitators may be selected according to the following criteria:

- Experience in health systems research.
- Experience with participatory teaching.
- Availability for the full duration of the workshop and for providing supervision and support through field visits during the 5-6 months when research projects will be implemented.
- Experience in previous HSR workshops as a participant or facilitator, and/or in a Training of Trainers course for HSR facilitators (the majority should have both!).
- Ideally, the team should comprise a variety of disciplines, such as medical sociology, health management/public health, and epidemiology. At least one facilitator should have knowledge of experience with statistics, and one with collecting and analysing qualitative data.
- An equitable mix of male and female facilitators is recommended.

2. Course administrator

Although the course co-ordinator is responsible for the overall functioning of the course, it is highly recommended that he delegate administrative tasks to a course administrator. The course administrator will, for example, make administrative arrangements, supervise support staff, (typists, drivers), ensure that participants and facilitators receive the necessary support to travel to and from the course site, and make sure that necessary payments are made and various other support tasks during and after the workshop are carried out promptly. The course administrator should attend all meetings of facilitators so that logistic support for the participants can be arranged for appropriate times.

3. Requesting consent for conducting the workshop

To obtain approval and funding to conduct the workshop, a proposal may have to be submitted to the relevant authorities for approval and funding approximately 12–18 months prior to the workshop. The proposal should include:

- The title, a brief background statement, and summary of the rationale for the workshop;
- Objectives of the workshop;
- Number and types of participants;
- Tentative date, duration and venue;
- Budget requirements (see guidelines on budgeting in Annex 1); and
- Any assistance required in the form of consultancies from within and outside the country.

As workshops in general take place at the initiative of the HSR Unit of the MOH, a training institution or NGOs active in health, approval is usually a matter of rubber-stamping, if needed at all. However, sometimes (additional) funding has to be identified for the workshops or for implementation of the research proposals developed, which requires time and energy.

4. Request for consultants (if needed)

Requests for consultants should be based on specific terms of reference and made through the Ministry of Health or other initiating organisation to the relevant donor agency. The workshop proposal should be included with the request.

5. Pre-workshop preparatory activities

A course management team consisting of the course manager, a core group of facilitators, and a course administrator should be set up at least 5 months prior to the workshop, to:

- Determine who the participants will be (depending on the priorities of the organisation that takes the initiative);
- Identify content areas, methodology and the course schedule (see **Annex 1** of these Guidelines for sample schedules);
- Identify and arrange for the venue;
- Identify and make plans for procuring the required materials (e.g., stationery and other supplies and transport); and
- Identify resource persons who may be required to give special technical or logistical input (e.g., permission from administrative authorities for transport, if it is intended that workshop participants should do field work for pre-testing of data collection tools.)

(1) Selection of participants and research topics (about 4 months before the workshop)

The selection of participants is usually delegated by the HSR Unit or organising NGOs to the provincial levels, hence the need to give clear guidelines.

Number of participants: 20–25 (5–6 per team)

Criteria for selection:

The criteria for selection of the participants should be clearly defined, taking into consideration the types of participants who are available, their educational background, the feasibility of their incorporating research into their functions and their interest in research. The following factors have been found to be useful in selecting participants for the basic course described in this volume:

- Select small groups of participants from the same geographic or institutional locations, so that each group can develop and implement a research proposal as a team and support each other in the development of subsequent research projects.

- If HSR is in an early stage of development, give priority to the following types of participants:

  - Staff of relevant training institutions (e.g., public health or nursing schools) so as to rapidly create a pool of persons who can both do research and help others;
  - Staff who have had previous exposure to basic epidemiological or sociological research;
  - Participants with leadership qualities;
  - Participants from districts, institutions, or regions where the director or manager is strongly committed to HSR and is likely to provide leadership and support;
  - It is advantageous to have participants from different disciplines and with leading functions in the major health programs (i.e., maternal and child health, sanitation, nursing, rural development). Some junior social scientists may be included.

- In subsequent workshops a good mix of district- or provincial-level participants may be invited.

Communications with national and regional or provincial health managers or institutes that are invited to provide participants should be made 3-5 months before the workshop.

First, informal contact should be made with these authorities to inform them of the possibility of authorising staff to participate in an HSR training course, to enlist their interest and support in selecting problems for the research projects. The criteria for selection of participants and topics should be discussed, and the managers should be requested to explore potential topics with staff members whom they plan to select as participants. These should in turn consult other parties concerned (field staff, community members).

- Then official letters should be sent to the same authorities stating:

- The objectives of the training course;

- The structure and schedule (e.g., two workshops with a fieldwork period of 5-6 months (part time) for implementation of an HSR project in between);

- The venue of the workshop;

- The selection criteria for participants;

- The preparations the participants need to complete before attending the workshop;

- The deadline for the confirming participation in the workshop; and

- A request that the supervisors identify specific topics for HSR projects which the participants can consider/select during their workshop. A copy of **Module 1** and the relevant sections of **Module 3** could be sent to the authorities (and participants) to guide them through the process of selecting appropriate topics. It would be highly recommended to arrange for a facilitator to assist in the selection of research topics (see **Module 1**).

(2) Communication with selected participants

An information circular should be sent to all the participants selected, providing them with preliminary information on the workshop (similar to the one sent to their superiors). It should emphasise that they will be expected to do a research project themselves. The relevant sections of **Module 3** can also be sent, providing there will be guidance for the selection of topics. (See a sample of an information circular in Annex 3).

On receiving confirmation of their participation, the prospective participants can be sent background-reading materials on HSR.

(3) Discussions on training methodology and training procedures

It is extremely important that the course management team as a whole takes time to discuss the course content and methodology. All facilitators should be very familiar with the training materials. Consensus will have to be reached concerning who will introduce the different modules and the role of facilitators during group work and plenaries. The capabilities of each facilitator will have to be discussed in relation to the training requirements.

(4) Ordering or duplicating sufficient copies of the training modules

This can be done through WHO Geneva, AFRO or other regional offices, or IDRC Ottawa.

(5) Selection of additional local resource persons

Additional local expertise in disciplines such as epidemiology, statistics, qualitative data collection and analysis, may be required if these specialisations are not sufficiently represented in the course management team. A local librarian or a researcher who is presently involved in an interesting HSR project may also provide ad hoc assistance.

Outside resource persons should generally not be asked to present modules, unless they are very familiar with the course and its training methodology. However, it is useful to invite them to one or more of the course sessions to familiarise them with the course; to introduce them to the participants as valuable resource persons (both during the course and afterwards); to make their expertise available during group work; and, finally, to enlist their support for the implementation of the proposals being developed.

(6) Invitation of authorities to open or close the course

Usually a high official from the Ministry of Health, or, if appropriate, a representative from another agency supporting the course, should be invited to open the course. This is a useful strategy for making high-level officials aware of HSR and motivating them to support it.

Usually, the official opening of courses takes place on the first morning. It might be worthwhile, however, to officially open the course the evening of the day previous to the first day of the workshop. This will save time. Alternatively, the opening may take place in the later afternoon of the first day or, better, the morning of the second day, when participants can present the final selection of their research topics.

(7) Invitation of donors

If you are considering inviting donors to explain what types of research projects they presently support and to provide details on research priorities and funding procedures, it is advisable to invite them for a panel session all together one evening. Time could be scheduled after the panel for teams to talk with the donors individually.

(8) Invitation of panel members for the presentations of proposals and research results

The plenary sessions at the end of the first and of the second workshop are extremely important. At the end of Workshop I each group presents its research proposal while at the end of Workshop II each group presents its research report, including a first draft of major findings, conclusions and recommendations. Each of these gives the participants a chance to gain experience in presenting research proposals or research papers. It also provides the opportunity to invite senior managers, researchers, academicians, etc., as panel members so that they gain a better understanding of health systems research and can have an input in the formulation of conclusions and recommendations. This will likely enhance their support for implementing them. Selection of appropriate panel members is important. Health managers of different levels and interested researches should be invited. If possible (distant-wise), it is highly recommended to include the direct superior(s) of the research team will utilise the research findings in the panel.

(9) Selection of support staff

Support staff for the workshop should include two secretaries and one driver/messenger. For the last 3 days of the workshop, four full time secretaries would be desirable, unless participants themselves do the word processing of part of the protocol. Secretaries may have to work overtime to finish typing the research tools before the pre-test and to finalise research proposals. For the data analysis workshop, one secretary may be sufficient during the first week, but when the groups start writing their reports four secretaries should preferably by available, unless participants are (partly) self-reliant. Ensure that there is a computer available for each research team, and two diskettes per group.

(10) Site preparation

Space required:

- Plenary space for 30–35 persons plus two small meeting rooms for group work.
- Office facilities for 2–4 secretaries, computers and space for a photocopy- or duplicating machine.

Materials required:

- Access to a vehicle for the whole workshop period. During the pre-test, extra transport may be needed.
- See **Annex 2** for details concerning the materials needed.

II. COURSE MANAGEMENT DURING THE WORKSHOP

1. Course co-ordinator

The course co-ordinator will have overall responsibility for the workshop. Some of the essential functions include:

- Conducting opening and closing sessions;
- Making general announcements (concerning reading materials for the next day, work on weekends, deadlines for submission of materials for typing, etc.);
- Presenting the session on orientation to the course and doing a review of progress at the start of each day to enable participants to keep track of the workshop process;
- Introducing resource persons;
- Resolving specific problems that may arise; and
- Maintaining a chart of the progress of each group on submitting its drafts of the proposal/report for typing. (This chart should be on display throughout the course.)

2. Chairpersonship of plenary sessions

It may be useful to rotate the chairpersonship, depending on the subject being discussed. For instance, the facilitator presenting a specific module and guiding the following discussion could be the chairperson of that session.

3. Allocating facilitators to working groups

When the participants have selected their research topics, a final decision will have to be made as to which facilitators, considering their interest and expertise, would best be in charge of particular groups. Facilitators will, in principle, stay with the same groups throughout the course, in order to ensure continuity and the quality of the end product.

In addition, each facilitator may have overall responsibility for certain technical aspects of the research process in which he or she is specialised, and assist other groups as well. Also local resource persons may assist on an ad hoc basis (e.g., on sampling).

4. Facilitators' meetings

It is desirable to have a daily meeting of facilitators to monitor course progress and give an opportunity to the facilitators to discuss possible problems This meeting is best held in the evening and will usually last between half an hour and an hour. The course co-ordinator should be responsible for convening this meeting. It is probably helpful to have a secretary record at least the action points for each meeting.

5. Approval of projects

The national agencies that will need to endorse the research proposals (the National Research Council, for example) will also have to be mobilised before as well as after the workshop, to speed up procedures.

6. Workshop report

The official report of the workshop should be as brief as possible. After a one page introduction (when, where, why, organisers, sponsors, management team (functions), type of participants in the course), a summary report of 2-4 pages could follow, describing the training process, starting with the topics chosen and ending with evaluation results. A list of participants and course facilitators and their addresses (by research group) could be annexed to the report, as well as (summaries of) opening speeches.

The report should contain the final drafts of the groups' research proposals. It is highly recommended that the course facilitators as a group screen the proposals immediately after the workshop, because some items may have been dropped out or added to the proposal that would need some clarification. The final polishing up of the proposals can be completed in the month following the workshop. Usually procedures to obtain consent for implementing the proposals take up the first month, so there is some time to do any final editing needed.

7. Meeting the participants' needs for technical support during research implementation

The facilitators that have assisted in the development of the proposals will also assist the groups in the implementation of the proposals. However, sometimes additional support may be required, (e.g., the assistance of an experienced sociologist or statistician for data collection and data processing).

The participants will need to state this in their proposals and include it in the projects' budgets.

All groups will need assistance when they start sorting and processing their research data.

III. TRAINING METHODOLOGY

Sessions in this training course on health systems research contain the following components:

- Introduction and discussion;
- Group work;
- Exercises; and
- Plenary.

1. Introduction and discussion

The introduction and discussion period is used to briefly explain new concepts and their application. Inviting responses and suggestions from participants and listing these on a flipchart or using them as a starting point for discussion is an essential element of the training method. This increases the interest of participants and may bring up valuable points of view that would be missed in classical (pure lecture-type) classroom teaching. Encourage all to participate in the discussion.

Depending on the level of the participants, the facilitator can delete or add details in the introduction, preferably using the research proposals that are being developed as examples, in interaction with the respective teams.

> The text of the sessions as given in the training modules is not meant to be followed word by word. Each introduction and discussion period should not last longer than at most one hour to 75 minutes.

2. Group work

The purpose of the group work is to develop four to five research proposals (one per working group) that should be **ready for implementation** by the end of Part I of the training course. Thus the facilitators need to always keep in mind that the proposals being developed need to be feasible and of good quality.

To increase the efficiency of the group work, a chairperson and rapporteur should be appointed for each group. The chairperson is not only responsible for leading the discussion, but also for dividing the work among group members. It is recommended that, after discussion within the work group, the group should split up into groups of two or three persons to work on separate parts of the task to be completed. The work of each sub-group can then be discussed and amended before presentation in plenary.

Each facilitator should be responsible for one group throughout the course, in order to ensure continuity. Facilitators should only change groups if they have major problems in assisting their own groups. Other facilitators and resource persons, of course, can be consulted at any time on technical issues. The amount of time the facilitator spends with his/her group will depend on the needs and demands of that group. In the beginning of the course the needs may be greater than towards the end. In principle **facilitation is a full-time activity**. Even if a facilitator is not permanently participating in the group work, (s)he should be available at all times for consultation.

The facilitator's role in discussion is primarily to stimulate the group to find its own solutions. However, if the group is clearly going in the wrong direction the facilitator should provide more direct guidance. In the beginning a facilitator may have to keep the group from wasting time on less relevant issues, or prevent relevant issues brought up by group members from being dropped because not everyone sees their importance.

3. Exercises

There are two types of exercises. In some exercises groups practice the use of new concepts by working with case studies prepared in advance. For these, it is probably a good idea to organise groups of a different composition than for the group work, so that all participants get to know each other well. In the second type of exercise, each group will examine a component of the proposal that is being developed by another group and provide constructive criticism. Groups should be encouraged to put the summaries of their comments on flipcharts or transparencies for presentation in plenary and for reference by the group developing the proposal. Not all modules contain exercises. Exercises can be omitted or added, depending on the needs of the participants and the time available.

4. Plenary

Presentations of the results of group work or exercises in plenary require special skills. Before the first plenary (in which the research topics considered for the development of research proposals are presented) the importance of presenting the group reports clearly and audibly, and of using readable visual aids, should be discussed with the participants. The working groups can use either flipcharts or transparencies for the presentations. Flipcharts have the advantage that they can be easily referred to or elaborated on later in the working groups. However, if the plenary exceeds 25 persons, it may become difficult for all to read the flipcharts. The use of transparencies and an overhead projector, in that case, may be indicated.

It should be stressed that there are limitations to what one can put on a transparency or a flipchart. Prepare two examples, one of a readable and one of an unreadable transparency, and let the participants give suggestions concerning how much information a transparency should contain (12-14 lines is the limit).

Stress also that one should never turn one's back to the audience when presenting. (A pointer can be used to indicate various points on the transparency, rather than on the screen.)

In general, the presentation of one working group should not exceed 15 minutes, discussion included. Sometimes even less time is required. If necessary, the facilitator chairing the session should let presenters know when they have just a minute or so left.

IV. SUPERVISION OF RESEARCH PROJECTS: INTER-WORKSHOP ACTIVITIES

The activities during the inter-workshop period consist of preparing and implementing the studies and (preliminary) processing of data. The facilitator or resource person for the project should visit at least once, preferably during the pre-test, and if possible also during the preparations for data processing.

1. Activities that should be carried out by participants

During this period, the participants should:

1. Brief managers, health service staff and concerned community members regarding the project;
2. Obtain the necessary permission to collect data;
3. Identify and obtain the resources (manpower, materials, etc.) needed to collect data;
4. Review the availability of subjects or respondents, information, ethical aspects of the study and adjust the methodology if necessary;
5. Train research assistants, if required. Refine, pre-test and revise the research instruments and procedures for data collection and data analysis with the entire research team, preferably together with the facilitator;
6. Collect data;
7. Prepare for processing of data and do some of the processing by hand and by computer (if feasible and useful); and
8. Prepare a preliminary report.

2. Guidelines for participants

Modules 19 and **20** should be presented to participants at the end of Workshop I for their use during the inter-workshop period. The content of the modules should be reviewed before they leave for the field, so participants will know what they contain and why they are important.

Module 19 will serve as a checklist as well as a guideline for field activities.

3. Visit by the facilitator/resource person

The purpose of the visit by the facilitator near the beginning of the fieldwork is to:

1. Check the progress of the project;
2. Observe the real life situation in which the project will be implemented and identify problems and anticipate pitfalls;
3. Review the proposed methodology of data-collection;
4. Advise on modifications/adaptation of the research design, sampling, data collection procedures, etc.;
5. Assist in training team and possible research assistants in research- and interview techniques; and
6. If necessary, assist in obtaining managerial support.

The trainer's notes at the end of **Module 19** provide a useful checklist for facilitators to use during their site visits.

Annex 1: Example of a course schedule (used in southern Africa)

WORKSHOP I: PROPOSAL DEVELOPMENT AND FIELDWORK

Date/Time	Session	Responsible person(s)
Date (Sunday)		
Evening	Welcome address	Course co-ordinator
(1 hour)	Mutual Introduction of participants and facilitators	
Date (Monday)		
08.30 – 09.15	Opening ceremony	
09.15 – 10.00	Administrative remarks	Course co-ordinator
	Module 1: Course orientation	Facilitator
10.00 – 10.30	Tea	
10.30 – 11.30	**Module 2: Introduction to health systems research**	Facilitator
11.30 – 12.45	**Module 3: Identifying and prioritising problems for research** (including exercise)	Facilitator
12.45 – 14.00	Lunch	
14.00 – 16.00	Project group work	
16.00 – 16.30	Tea	
16.30 – 17.30	Group reporting in plenary of selection process and topic selected	
Date (Tuesday)		
08.00 – 09.00	**Module 4: Analysis and statement of the problem**	Facilitator
09.00 – 12.30	Project group work (including tea)	
12.30 – 13.30	Lunch	
13.30 – 14.30	Group reporting in plenary	
14.30 – 15.00	**Module 5: Review of available literature and information**	Facilitator
15.00 – 15.30	Tea	
15.30 – 17.30	Project group work	
evening	Structures and mechanisms for HSR in (host country) (optional)	

Date (Wednesday)

08.00 – 08.30	**Module 6: Formulation of research objectives**	Facilitator
08.30 – 10.30	Project group work	
10.30 – 11.00	Tea	
11.00 – 11.30	Exercise: assessing the statement of the problem and the objectives formulated by another group	
11.30 – 12.45	Group reporting in plenary	
12.45 – 14.00	Lunch	
14.00 – 15.00	**Module 7: Introduction to HSR methodology**	Facilitator
	Module 8: Variables (including exercise)	Facilitator
15.00 – 15.30	Tea	
15.30 – 18.00+	Project group work (revision of objectives, SoP, development of variables)	

Date (Thursday)

08.00 – 09.30	Project group work (continued)	
09.30 – 10.30	Group reporting in plenary	
10.30 – 11.00	Tea	
11.00 – 12.15	**Module 9: Study types**	Facilitator
12.15 – 13.30	Lunch	
13.30 – 14.30	**Module 10A: Overview of data collection techniques**	Facilitator
14.30 – 15.30	Exercise (Study types + data collection techniques)	
15.30 – 16.00	Tea	
16.00 – 17.00	Project group work (Mod. 9 + 10)	
17.00 – 18.00	Group reporting in plenary	

Date (Friday)

08.00 – 09.00	**Module 10B: Design of interview tools; interview techniques**	Facilitator
09.00 – 10.00	Exercise interviewing	
10.00 – 13.00	Project group work (including tea)	
13.00 – 14.00	Lunch	
14.00 – 14.30	**Module 10C: Focus group discussions**	Facilitator
14.30 – 16.00	Exercise on FGD	
16.00 – 18.00+	Tea and group work on tools (continued)	

Date (Saturday)

08.00 – 09.30	**Module 11: Sampling**	Facilitator
09.30 – 12.15	Project group work (including tea)	
12.15 – 12.45	Exercise: commenting on sampling procedures and sample size of another group	
12.45	Lunch	

Date (Monday)
(NB: In case of a 12 day workshop instead of 15 days this part of the program may have to be skipped)

08.00 – 09.30	Plenary presentations and discussion of sampling procedures and sample size	
09.30 – 12.00	Exercise (Module 1OB): Commenting on the data collection tools of other groups (including tea)	
12.00 – 13.00	Plenary discussion of data-collection tools (2 groups)	
13.00 – 14.00	Lunch	
14.00 – 15.00	Plenary discussion (continued, 2 remaining groups))	
15.00 – 15.30	Tea	
15.00 – 17.30	Project group work: Revision of data collection tools	

Date (Tuesday)

08.00 – 08.45	**Module 12: Plan for data collection**	Facilitator
08.45 – 11.30	Project group work (including tea)	
11.30 – 12.30	Group reporting in plenary	
12.30 – 14.00	Lunch	
14.00 – 14.30	**Module 14: Pre-testing the metholology**	Facilitator
14.30 – 16.00	Project group work to prepare pre-test	
16.00 – 16.30	Tea	
16.30 – 17.30	Project group work to finalise all data collection tools	

Date (Wednesday)

08.00 – 10.15	**Module 13: Plan for data processing and analysis** (including exercise on processing qualitative data)	Facilitator
10.15 – 12.30	Project group work (including tea)	
12.30 – 14.00	Lunch	
14.00 – 15.30	Group reporting in plenary	
15.30 – 16.00	Tea	
16.00 – 17.30	Project group work to revise first part of the research proposal (Background, Statement of the Problem, with literature review)	

Date (Thursday)

08.00 – 13.00	**Pre-test**
13.00 – 14.00	Lunch (if necessary packed lunch)
14.00 – 17.30	Project group work: evaluation of pre-test and revision of data collection tools

Date (Friday)

08.00 – 09.00	Group reporting in plenary of results of pre-test	
09.00 – 10.30	**Module 15: Work plan and Module 16: Budget**	Facilitator
10.30 – 13.00	Project group work (including tea)	
13.00 – 14.00	Lunch	
14.00 – 14.15	Group work (continued)	
14.15 – 15.15	Group reporting in plenary	
15.15 – 15.45	Tea	
15.45 – 16.15	**Module 18: Finalising and reviewing the research proposal**	Facilitator
16.15 – 18.00	Group work	

Date (Saturday)

08.00 – 08.45	**Module 17: Plan for project administration, monitoring, and utilisation of results**	Facilitator
08.45 – 10.30	Project group work	
10.30 – 11.00	Tea	
11.00 – 12.00	Group reporting in plenary	
12.00 – 13.00	Lunch	

Date (Monday)

08.00 – 08.15	**Module 19: Fieldwork activities**	Facilitator
08.15 – 08.30	**Module 20: Preparing a preliminary report**	Facilitator
08.30 – 13.00	Project group work: Finalising research proposals (including tea)	
13.00 – 14.00	Lunch	
14.00 – 18.00+	Finalising research proposals Preparation of public presentation of proposal	

Date (Tuesday)

08.00 – 10.00	Evaluation of the workshop Finalising presentation preparations	
10.00 – 13.00	Tea with visitors Presentation of proposals with discussion Closing of the workshop	

Annex 2: Guidelines for budgeting an HSR training course

The following items will probably have to be budgeted. Indicate for each item who will cover the cost (Ministry of Health or donor, for example). Salaries of local participants and transport are usually provided by the Ministry of Health, whereas accommodations and meals are usually covered by the donor.

(1) Accommodation and meals

Board and lodging or lodging and an allowance for meals for:

24 participants
5 facilitators
2 secretaries (perhaps 4 at the end of the workshop)

Make sure the workshop site has available:

- a large conference room
- two small meeting rooms
- a room for the secretaries/photocopying machine

Also include:

- **Coffee/Tea** for 13 days, twice a day, for 30-35 people

Consider the inclusion of a:

- **Reception** for 50 people, after the official opening

(2) Salaries and allowances

Facilitators	salary	at US$ ...../day x 14-16 days
	allowance	at US$ ...../day x 14-16 days
Participants	salary at US$ ...../day x 14-16 days (MOH)	
	allowance, if applicable at US$ ...../day x 14-16 days (according to MOH rules)	

(3) Transport

- For facilitators and participants to come to the workshop and return home;
- For participants to get to the workshop site each day if they are lodging somewhere else than the workshop site (not recommended, but it may sometimes be necessary);
- For pre-testing of the methodology, including field visits for four working groups of participants

(4) Supplies

If photocopying is used during the workshop for duplication of the research protocol, questionnaires/tools, final report; and other documents:

- 34 reams of photocopy paper
- 2 reams of typing paper
- toner for photocopying

If all duplicating is done using stencils:

- 500 stencils for use during workshop
- 200 stencils for the final report
- 34 reams of duplicating paper (500 sheets each)
- 1 ream of typing paper
- ink for stencilling

40 note pads, 40 pens, 40 pencils, 40 rubbers, 40 file holders

35 name tags

paper clips, staplers, staples, paper hole punchers, scissors, chalk

200 overhead sheets, markers

5 flip charts, markers

35 copies of the modules, to be ordered through KIT, Amsterdam, AFRO or IDRC Ottawa, at least three months before the beginning of the workshop

Annex 3: Example of information circular for course participants (used in Malaysia)

Background

Health systems research (HSR) has been identified as an important tool to provide managers with information they can use in decision-making processes aimed at improving health care. In this context 'managers' could be those responsible for planning or implementing health programmes at district, state or national level or those responsible for managing hospitals, clinical units within hospitals or clinical outpatient services in hospitals, clinics etc.

Objective

The aim of the course is to enable you to develop and implement health systems research projects assisting managers, including yourselves, to improve the effectiveness and efficiency of health services.

Expected outcome and future functions

Once you have successfully completed the course, you will be expected to incorporate the conduct of HSR into your regular duties. As staff members who have had training in research, you will design and supervise projects and train your staff to collect data, analyse it, etc. You will also serve as 'resource persons' for your programmes, hospitals, etc., and provide assistance with the analysis of problems, design of studies, preparation of study reports.

Research skills can be acquired only through real life practice. Therefore this training programme is designed to give both theory as well as practical experience in conducting research. The practical experience will be in the form of a project that you will carry out in your place of work as a supervised training exercise. The course will be conducted in two sequential 'parts' with an 'inter-workshop period' during which you will actually carry out your study.

The **structure of workshop** will be:

1. **Pre-workshop assignment:**

   - Background reading
   - (Tentative) selection of a suitable problem for the research training project

2. **Part 1: Protocol development workshop (12-14 days)**

   - Design of the research proposal
   - Design of research instruments

3. **Inter-workshop period (5-6 months)**

   - Collection of data. (This will be done at the place of work of each participant and will be done in conjunction with her/his other duties.)

4. Part 2: Data analysis/report writing workshop (14 days)

- Analysis of data
- Preparation of report
- Presentation and discussion of findings
  (The respective state/programme directors will attend these presentations and participate in the discussions.)

5. Time and research will have to be reserved for implementation the recommendations

Background reading

This workshop will be very intensive and you will need to do a considerable amount of background reading both before and during the workshop. Pre-workshop reading consists of: (List a number of relevant short papers on the concept and purpose HSR).

Selection of projects for the training exercise

Development and implementation of a research project will be the most important part of this training programme. The first step in doing research is to select a problem that is an appropriate topic for research. You will have to do this **before** you come to the workshop.

One of the basic principles of HSR is that research should focus on **priority problems**. Although the project that you will do during the workshop will be designated as a training exercise, the only difference from any other research project will be that the scope might have to be limited to enable you to complete the project before the Data Analysis/Report Writing Workshop. Therefore, the problems that are selected should conform to all the criteria that would be used in selecting projects for research, and the process of selection should be the same as in actual practice.

We suggest that you should meet with your state or regional/provincial and district medical officer, in order to identify one or two priority problems that **need additional information that can be provided through research**. (N.B. If sufficient information is already available either through routine data or from other studies, it is *not* suitable for a research project, even if the problem is a priority.)

Criteria for selecting a problem for research include:

- Is the problem a priority?
- Is the problem specific and can it be clearly stated?
- Is necessary data missing to help solve the problem?
- Can the research be carried out with the available resources? (4-8,000 US$) and in the available time (5-6 months) with a team of 5-6 persons?
- Will the research findings contribute important information that can be used to solve the problem?
- Is it likely that the recommendations of the study will be applied?
- How urgently are the results needed for making a decision?

If two or three problems emerge as suitable, a final decision concerning which to choose as the focus of the research can be made during the Protocol Development Workshop.

Preparation for the project

Before coming for the workshop, you should be able to answer the **following questions regarding** the **problem** that you will focus on in your research:

- What type of information will assist managers in making decisions regarding the problem? For example:

  - The causes of the problem?
  - The factors contributing to the problem?
  - The relative importance of various factors?
  - The comparative effectiveness of various solutions?

- Can existing information be analysed to provide part or all of the data needed? Does new data need to be gathered as well?

- How will managers use the information when they receive it? (i.e., What actions will the manager be able to take based on the results?)

- Can the research provide the type of information the managers needs?

Available data on the problem

- Collect and bring to the workshop all available data, copies of district or national annual reports, or circulars and guidelines, etc., on the problem that you will be investigating. This will help you prepare your research protocol.

- Also visit the health centres, hospitals and/or other sites where you may eventually collect data for this project and familiarise yourself with their systems of keeping registers, cards, appointment books, etc., so that you will know how to select the sample for your research project.

Manpower for data collection

Identify members of your staff such as nurses, medical assistants, health inspectors and others who can assist you in collecting data for your research project during the implementation period. You will need this information to help you determine how much data you can collect within the given time period.

Annex 4: WORKSHOP: DESIGNING AND CONDUCTING HEALTH SYSTEMS RESEARCH PROJECTS: Reply format for participant

1. Name of participant: _____

2. Position: _____

3. Mailing address: _____

_____

4. I am able to attend Workshop Part I ☐ and Part II ☐ and available for the research period in between ☐ (tick if yes)

5. The research project that will be designated as a training exercise during the workshop is: (give one if all agree, two if you are not yet sure)

_____

_____

_____

_____

6. Other participants who will work on the same project during the workshop include:

_____

_____

_____

_____

_____
Signature of Participant

I support the above project.

_____
Signature of Officer in-charge

_____
Position

Annex 5: Evaluation form

Evaluation questions: HSR Proposal Development Workshop in .....

1. Have the objectives of the workshop been achieved?

1.1 Have you gained the expertise to develop a research proposal? (explain)

1.2 Have you developed a proposal that you think can be carried out by your group within the coming 5 months? (explain)

1.3 Do you feel (1) motivated and (2) confident enough to start other small research projects in the future in your own working situation? (explain) ☐Yes ☐No

2. Do you have any comments on the course content?

Would certain parts need extension?

Should certain parts be slimmed down?

How clear were the presentations?

3. What is your opinion of the training methods used in the workshop (compared, for example, to the 'lecture' type of teaching)?

4. Did you find the division of time between lectures, group work and plenary satisfactory or would you propose more or less time for any of these three components of the course?

5. Were you satisfied with the type of assistance provided by the facilitators?
☐ Yes ☐ Partly ☐ No
Would you have any suggestions for improving the work of facilitators for similar courses in the future?

6. How did you function as a group?
Do you feel that every group member had an equal chance to gain from and contribute to the course? ☐ Yes ☐ Partly ☐ No
(Explain)

Would you have any suggestions for improving the opportunities for group members to gain from or contribute to similar courses in the future?

7. Acknowledging that you are all busy people, but that the course was quite compact: **Would you have liked the duration of the course longer, shorter or was it the right length as it was?**

8. What is your opinion on the organisation/accommodation/working conditions of the course?

Do you have any suggestions for the next workshop in this respect?

9. ANY OTHER COMMENTS

ABOUT THE AUTHORS

Corlien M. Varkevisser, MA, PhD, MPH, is a medical sociologist-anthropologist by profession who specialised in public health. As a staff member of the Royal Tropical Institute, Amsterdam, and former head of the Primary Health Care (PHC) Unit, she has gained extensive experience in health systems research and PHC management in sub-Saharan Africa. She was one of the co-initiators of the Joint HSR Project (WHO/Netherlands Ministry for Development Cooperation/ Royal Tropical Institute) for Southern Africa and was based at the WHO Sub-regional Office in Harare as manager of the Joint HSR Project from its onset in April 1987 till 1992. Thereafter she became manager of the MPH course at the RTI and professor in HSR at the Faculty of Political and Social Sciences, University of Amsterdam.

Indra Pathmanathan, MMBS, MPH, is a physician specialised in public health who, as Head of the HSR program of the Ministry of Health in Malaysia since its inception, has been responsible for developing and implementing several strategies for HSR that have been replicated in other countries. These included training programs in HSR and Quality Assurance for decision-makers in ministries, for physicians, and for staff in district health teams, hospitals, and universities. She was a member of the Advisory Group on HSR, WHO-Geneva and served on the editorial board of BRIDGE. Over the past ten years she has been a consultant to the World Bank in the field of health.

Ann Brownlee, MA, PhD, is a medical sociologist who specialised in HSR, planning and evaluation, and cross-cultural aspects of health care. She served as Research and Evaluation Coordinator for the Project for Strengthening Health Delivery Systems in West and Central Africa for a number of years, where she worked closely with WHO's Regional Office for Africa and with colleagues from Africa and elsewhere to develop an HSR training program and to publish the *HSR Training Course* that was a forerunner of this volume. She currently works as a consultant in international health for groups such as WHO, IDRC, and Wellstart and teaches at the University of California at San Diego.