A Generic Guide to Research Practice

Following discussion at Lilongwe workshop of GHIN African Teams

Global HIV/AIDS Initiatives Network **December 2006**

Table of Contents

A Generic Guide to Research Practice

Following discussion at Lilongwe workshop of GHIN African Teams

INTRODUCTION

The overall aim of the Lilongwe workshop was for countries to finalise protocols and plans for the district studies. A significant amount of time was taken up with discussion on the following:

- District selection
- Matching detailed research questions to research methods
- Discussion and training on qualitative and quantitative methods

There was rich discussion on the appropriateness of the different methods and how the various methods can be put into practice. This document should be seen as a 'pocket guide to good research practice' and deals with the specific issues which were discussed in Lilongwe. It is not intended to be an exhaustive document outlining the main principles and benefits of qualitative and quantitative methods.

1. DISTRICT SELECTION

There was much useful discussion at the Lilongwe workshop around district selection, including numbers and types of districts to be selected. The consensus was that each country study should have clear criteria to justify its approach, based on an analysis of the country context. Selection criteria should include and involve a balance of the following:

1.1 Complexity of (funders) HIV/AIDS control at the district level

Where the country HIV/AIDS context is *complex* involving multiple sources of parallel funding (from Global Fund, PEPFAR, MAP, Government and other sources of funding) to multiple funders of HIV/AIDS services at the district level, it is better to select a smaller number of districts for the study. Zambia is a typical country where one would choose to restrict the number of districts because GHI funds come through multiple Global Fund PRs (Principle Recipients), and also from PEPFAR and from World Bank MAP.

Where the country HIV/AIDS context is *simpler*, because HIV/AIDS funding is being channelled to districts through a single financing channel – through the SWAp or through the National AIDS Council – and there are fewer providers of HIV/AIDS services at the district level (often it is mainly Government and faith-based facilities) – then it may be feasible to conduct surveys in a larger number of districts (the number of districts will be country dependent). Malawi is a country where a larger number of districts could be sampled.

1.2 District HIV prevalence rates

Where data are available that categorise districts as 'high', 'medium' and 'low' HIV prevalence, one may make a convincing case for selecting perhaps two high prevalence and two medium prevalence districts. Low prevalence districts are likely to have less need for and less funding of HIV/AIDS services. District (or at least regional) HIV prevalence rates should probably be a selection criterion in all studies.

1.3 Districts where there is evidence of multiple GHIs / stakeholders operating

A purposive approach may be the most appropriate strategy based on information from a national context mapping or a HIV/AIDS resource tracking exercise. Such exercises may show that the purposive selection of certain districts for the study, based on sources, patterns and levels of funding, is likely to provide important insights into how support from the different GHIs to district HIV/AIDS scale-up works when large levels of funds come down to those districts through different channels: Government, Global Fund, PEPFAR, MAP and other sources. Tanzania has excellent up-to-date (July 2006) GIS maps that show the levels and recipients of funds for HIV/AIDS activities at the district level. Uganda is developing an approach to district selection based on 'numbers of key stakeholders' involved in HIV/AIDS control at the district level.

1.4 Feasibility

Research budgets, which have to cover cost and time to travel to districts and cost of maintaining teams of field workers in the field, is inevitably an important criterion; and this will also depend on country context. For these reasons (as well as complexity of the district context), Zambia is likely to select a smaller number of districts and Malawi a larger number.

2. DISTRICT MAPPING

Most of the GHIN African countries already have country maps that display all government health facilities including: district hospitals, rural and other (including faith based) hospitals, health centres and lower level facilities (sometimes called health posts). Mapping of all district-based providers of HIV/AIDS services within the selected districts is essential, even if a country study does not aim to conduct district surveys of all such providers. This requires geographical district-level mapping of facilities and types of HIV/AIDS services.

The GHIN network will need to develop *proformas* for the country studies to capture information on the following, including geographical location of the facilities and services below.

- Mapping of all *government facilities* providing HIV/AIDS services within the district, categorising facilities by their level in the health system. In Uganda this would be district hospital, Level 4, Level 3 and probably Level 2 health centres. In other countries, the facility levels may consist of district and other hospitals, health centres, and health posts (terminology will differ across countries).
- Mapping of all non-government *not-for-profit facilities*: *faith-based* (mission, Church, etc.) and also *NGO facilities*, mapping them by their level in the health system: district referral or lower level facilities

- Mapping of *organisations* providing out-reach services – these may include fixed facilities providing outreach and also organisations (usually NGOs) that are providing outreach HIV/AIDS services.
- Mapping of important Private for-Profit (PfP) providers of HIV/AIDS services is an option that some country teams may want to include, so as to assess if these providers make a significant contribution to HIV/AIDS service coverage at the district level. For example, Malawi National AIDS Control Programme supports about 30 private-for-profit facilities which deliver ART according to national protocols.
- Mapping the *types of HIV/AIDS services* provided by each of the above:
    - Family planning and HIV/AIDS prevention interventions: condoms, IEC (information, education communication), etc.
    - HIV counselling
    - VCT including HIV testing
    - PMTCT (Prevention of Mother to Child Transmission)
    - ART (Anti Retroviral Therapy), including ARV provision
    - Laboratory support (usually HIV testing but could include CD4 counts)
    - Non HIV/AIDS services

This mapping exercise will be an important preparation for selecting sites for quantitative and qualitative data collection and for identifying the key district coordination issues and questions and who to interview for modules A - C. The district health office or health management team may be a good starting point for obtaining or developing maps. However, the mapping of HIV/AIDS services delivered will require visits to facilities. Those initial visits will be an opportunity to undertake a number of important activities:

- Seeing how facility records are maintained and how tools may need to be adapted
- Identifying who to interview and compiling sampling frames
- And of course getting 'buy-in' from facility mangers.

3. SAMPLING FOR QUANTITAIVE RESEARCH

3.1 Sampling Strategies

The most appropriate approach to selecting a sample will depend on the population being surveyed:

3.1.1 Census

Where you are conducting a structured survey of front-line health workers (HWs) at the facility level, the numbers are small and you should aim to survey all eligible HWs. First you will need to compile a census of all facility staff. In practice the numbers of doctors and front-line nurses at the district hospitals, are likely to be small.

3.1.2 Random sample

Where you are sampling from a large population of eligible subjects where the views of each subject is of equal importance, e.g. in a community-based survey of health

seeking behaviour and of barriers to accessing services, a random sample that gives each member of the reference population (the community served by that hospital) an equal opportunity to be selected is the best approach. This is the most common type of probability sample. However, in practice, population-based community surveys are resource intensive exercises. You need to compile a sampling frame of eligible subjects (or a sampling frame of households) from which to randomly sample. This often requires a detailed mapping of households and may require an initial survey of all households to compile a list of eligible subjects.

3.1.3 Stratified random samples

This is a variation of the random sample where you compile sampling frames of eligible subjects, e.g. men and women, or adolescents (<20 yrs), 20-35 yr olds, and older adults (>35 yrs), and draw random samples from within each of these strata. You can also choose to weight the samples to give certain target groups (e.g. adolescents) a known greater chance of being included in the survey.

3.1.4 Systematic samples

This is an appropriate strategy to use when conducting an exit survey of patients in a large health facility such as a district hospital, e.g. to assess barriers and costs in accessing care. Because patients come to the OPD at different times of the day it may not be feasible to compile a sampling frame of all the patients so as to choose a random sample. A more feasible strategy, if there are say 200 patients attending the OPD on a given day, would be to sample every 4^{th} patient, giving a sample of 50. An exit survey conducted over 3 days would give an adequate sample of 150 patients. An additional advantage of a systematic sample is that it would include patients who arrived at the hospital early in the day (maybe they live closer) and those who arrived later.

3.1.5 Purposive samples

Purposive sampling is most appropriately used for qualitative methods. See section 8.1.2.

3.2 Sample Size Calculations

3.2.1 *Prevalence studies:*

Sample size determines the statistical power of a survey. Statistical significance is typically measured at the 95% level of confidence, which means that a given value or measurement can be considered – 95 times out of 100 – to be an accurate reflection of the value in the larger population from which the random sample was taken.

Most of the GHIN studies where larger samples are being surveyed will require the calculation of prevalence rates, i.e. the proportion of the sampled population with a given characteristic, e.g. (1) the proportion of patients who report a high level of satisfaction with HIV/AIDS services received, or (2) the proportion of health workers who report low levels of satisfaction with the incentive systems in place to retain them in their jobs. The sample size determines the precision of the prevalence measure,

which is normally calculated as a 95% confidence interval that measures the level of certainty and the margin of error around a given prevalence rate.

The size of the population from which the sample is drawn has an effect on the margin of error (the confidence interval). In measuring prevalence rates for certain behaviours or opinions and generalising these findings to all health workers of a particular category (let's say nurses in district hospitals and health centres) in your country, you are generalising to a large but finite number – maybe 2,000 or 5,000 nurses. If surveying doctors in district hospitals and generalising findings to district hospital doctors nationwide, the reference population is likely to be smaller – 200? 500? – depending on the size of the country. If you are measuring patients' behaviour, your reference population of patients attending district hospitals is probably larger – measured in hundreds of thousands. The following table is a guide to calculating sample sizes.

Sample sizes for different margins of error for populations of different size [*]

Margin of		**Size of**	**Populatio n**			
Error	**Large**	**5000**	**2500**	**1000**	**500**	**200**
± 20%	24	24	24	23	23	22
± 15%	43	42	42	41	39	35
± 10%	96	94	93	88	81	65
± 7.5%	171	165	160	146	127	92
± 5%	384	357	333	278	217	132
± 3%	1067	880	748	516	341	169

Patients attending a district hospital are typically drawn from a catchment area with a population of 100,000 – 200,000, i.e. a 'large' population. A sample size of 96 patients surveyed will give a margin of error of ± 10%, which is probably the maximum level of error acceptable. Given the possible problems of non-response and coding errors, etc., a patient exit survey (and a community survey) should therefore aim for a sample size of well over 100. Too small a sample size may mean that a large and important difference or association does not reach statistical significance.

A structured survey of health workers involves drawing a sample from a much smaller sampling frame. There are far fewer providers than patients at the district level, and it may only be possible to generalise the results of a survey of health workers to those employed in that facility. If you are conducting a HR (human resource) study which would benefit from collecting data in a structured way from a larger sample of health workers, you should probably aim to survey all eligible health workers at the district level, if this is the only way to reach a sample size of around 80 (± 10% margin of error). However, a structured survey that collects comparable 'hard' data from a sample of 40-60 health workers is still of value; and one should not assume that qualitative methods and tools are the only ways of getting useful data about health workers. Comparable measurements through applying the same tool (asking

[*] from Conroy R. Sample size calculation, a quick guide. Royal College of Surgeons in Ireland. October 2004.

the same question) to a sample of health workers can provide important results, even if they are not statistically significant.

4. ROUTINE DATA:

Can include:

- Facility Surveys
- Record Reviews / Proformas
- Inventories

4.1 Uses: to collect data on the following:

- Disease burden / types of diseases presenting at health facilitie
- Services delivered and levels of coverage (percentage of population who need each service and receive it) for key focal HIV/AIDS and non-focal diseases: a) *numbers of patients* receiving services and b) *numbers of services* delivered covering the following, specifically:
  - VCT
  - ART: numbers of patients started on ARVs; numbers on first line and second line regimens[1]
  - numbers and % of patients meeting ART adherence targets
  - PMTCT
  - Condoms distributed
  - whether ARVs are consistently in-stock in the pharmacy
- Commodities / drugs / HIV testing kits, covering: essential drugs for focal HIV/AIDS services and for non-focal priorities such as MCH and IMCI (e.g. antibiotics for Integrated Management of Childhood Illnesses).
- Facilities – buildings (VCT room) and equipment
- HR issues – staff numbers, salaries, incentives, hours of work, etc
- Others?

4.2 Relevance to GHIN country studies

The consensus at the Lilongwe workshop was that Module A – measuring HIV/AIDS control scale-up at the district level – would be a core component across all country studies. Data collected on Health Systems capacity (Module B), e.g. studies on Human Resources (HR) and capacity of districts to coordinate HIV/AIDS control, would be interpreted in the light of (and would in turn help to interpret) successes and/or failures in HIV/AIDS scale up, and the role of GHIs in supporting scale up. Similarly, studies on the equity and access effects of GHIs (Module C) necessitated first doing actual measurements on the scale up of different core components of a district HIV/AIDS service.

[1] it will be important to collect data on the actual ARV regimens (types, doses and combinations of drugs) being used in each facility; plus any evidence of: a) changes in regimens over time, which may be due to drug stock-outs; and b) differences in regimens supported or funded by different funding sources, government and GHI.

4.3 Approaches to collecting data / data format

4.3.1 National Monitoring and Evaluation (M&E) format and forms

Advantages:

- More likely to obtain historical disease and service coverage data at the facility and district levels
- Such data may already be aggregated by month or by quarter (3 months), as was the case in Malawi
- Much easier to relate district findings to national level data and communicate district findings to national level research users (NAC and NACP).

4.3.2 Global data format and global HIV/AIDS coverage indicators:

Advantages:

- Comparable data will facilitate comparisons across GHIN country studies
- Use of global indicators will facilitate communication of findings to global policy makers and research users.

In practice, it is likely that national and global formats for collecting data on disease burden and service coverage will be similar and often the same, in that there is a trend towards global harmonisation and alignment of M&E.

- Table 7 of the Monitoring and Evaluation Toolkit (pp 27-30), January 2006, contains Selected Programmatic Indicators for HIV/AIDS, which UNAIDS, WHO and Global Fund (among others) have signed up to.
- Table 15 (p.54) contains Selected Indicators for Health Systems Strengthening, developed by the Global Fund, which cover service delivery and human resources among others.

4.4 Collecting data at different levels of the system

Visits made to two facilities (a mission hospital and a Government rural hospital) near Lilongwe revealed useful data on service delivery that are routinely collected in Malawi, including all of the types of data identified above under 'Uses'. Facility-level data available included: a) patient registers; and b) copies of monthly and quarterly returns made to higher levels of the health system. While this quality of facility level data may not be routinely collected in this level of detail across all other countries, country teams will want to search for such data in the first instance.

a) Facility-level patient registers
b) Facility-level copies of monthly and quarterly service returns
c) summary district returns (aggregate of facilities)

d) summary national (aggregation of districts)

An added bonus would be if such data are aggregated at: c) the district level – District Health Management Team (DHMT) or similar district office – which would be the next port of call for study teams. It may also be possible in some countries to obtain d) disaggregated district (or even facility-level) data that are filed at the national level, at the National AIDS Control Programme (NACP) or National AIDS Council (NAC).

Where aggregated data are not available at the district level on a) numbers of patients receiving key HIV/AIDS services over a period of time (usually 1 month); and on b) numbers of services delivered, field workers will have to compile such data, usually by reviewing patient and department registers (for VCT, ART, PMTCT, plus pharmacy records) and summarising data on proformas designed by the country team for this purpose. Proformas should be designed to achieve the double advantage outlined above under *Approaches to Collecting Data / Data Format* above. Consultation with the GHIN coordination team would facilitate cross-country comparability.

Even where monthly data and service delivery summaries are available at the facility level, as is the case in Malawi, it would be a useful exercise for field workers to review patient and department registers to validate that such summaries are accurate.

4.5 Evidence of HIV/AIDS Scale-up – need for historical data

Much of the added-value of the GHIN country studies, far beyond what even the best district and national M&E systems could provide, will lie in:

- demonstrating *what scale-up* in HIV/AIDS control services is taking place
- explaining *how scale-up* is happening (and not happening)
- showing *the effects of GHI-supported scale up* on district health systems and on equity and access to HIV/AIDS control for important target groups.

Facility visits in Malawi showed that records are available that provide essential information on what scale-up is happening and on trends over time. Two approaches may be considered:

i. Where comprehensive patient registers and copies of monthly and quarterly summaries of services delivered are available, as in Malawi, recording ***monthly data over a period of one year*** or more is possible. This will provide the most useful information for proceeding to 2): semi-structured interviews of facility managers to explore trends and any marked changes (increases and reductions) in the delivery of different services over the course of a year, which a monthly trend analysis might demonstrate. For example:
    - a fall-off in the number of patients receiving VCT across 1-2 months might be explained (through an interview) by a lack of HIV testing kits during this period; or by the absence of some key hospital staff.
    - A reduction in some months in the number of patients being enrolled into an ART programme, or a fall-off in ART adherence targets, might be due to an interruption in ARV distribution and supplies, or due to staff absences. A trend analysis of patient registers / ART records would indicate the need

to explore these findings through interviews with facility managers, including the hospital pharmacist, to explore and establish what were the reasons. It might also indicate the need to explore and find out about other barriers to accessing HIV/AIDS services, e.g. a fall-off in patients attending the facility, which might require quantitative surveys or qualitative interviews of patients and community representatives.

ii. An alternative approach is to conduct ***repeat cross sectional surveys*** of services delivered over a shorter period of time, e.g. compile service delivery data for one month and repeat the exercise for the same calendar month, one year later (this is the approach used in one of the SWEF facility surveys). In practice, this may mean compiling the number of patients who received the different HIV/AIDS services (VCT, PMTCT, ART, etc.) in January 2006 and again in January 2007 (maybe also 2005 if relevant). Such data would provide evidence of scale up or not of focal services, but would be less likely to pick up significant upward and downward trends over the course of the year.

Field workers will require proformas that are designed for the collection of facility level data for key focal HIV/AIDS services, for key drugs (to detect any 'stock-outs') in the pharmacy, and for HIV tests (Module B3)

Where study teams are collecting information on other key services, such as TB and non-focal priority services (EPI, MCH, IMCI, etc.), can also be included in such proformas. Their design can be modelled on national M&E systems (see above on *Approaches to collecting data*). It was agreed on the last day of the Lilongwe workshop that country teams would look for the forms used by the MoH and NAC in their own countries to collect and report on such data at the facility, district and national level; and then send (preferably scan on to computer and email) copies of these forms to the GHIN coordinators. The GHIN coordinators can then advise on design of study data collection formats (proformas) across the country studies that would maximise cross-country comparability, as well as ensure consistency with globally collected indicators.

It may also be possible to design proformas for collecting data on health worker / staff arrivals and departures; and on their availability / absence from facilities over recent time (1 month? 3 months? 6 months? 1 year?). These proformas can also be used to capture information on Human Resources: staff, training, salaries, incentives etc. The GHIN coordinators are currently looking at this.

The importance of sequencing the facility record data collection and analysis in advance of interviews of facility managers was also discussed at Lilongwe. The recommendation was that:

- ***facility record analyses*** should be conducted to measure HIV/AIDS service delivery trends and identify any significant changes (especially reductions) in key services over the previous year;

- The results of these analyses should then be used to inform the design of semi-structured interview guides for ***follow-up interviews*** of the different managers at the facility level.[2]

The above proformas, facility surveys, inventories can be developed as one data gathering method. They do not exclude, but complement the use of other methods such as semi-structured interviews with facility and department heads. We recommend that proformas to demonstrate scale up and scale up gaps be done to inform semi-structured interview guides (see section 8.2 on qualitative methods).

5. STRUCTURED QUESTIONNAIRE SURVEYS

Structured questionnaires for:

- Facility staff / healthcare providers
- Patients
- Community Samples
- Others?

5.1 Key issues:

Validity = the key issue in *Questionnaire (tool)* design

Reliability *(repeatability)* = the key issue in selection, training and supervision of Interviewers

5.2 Validity in tool design

- Is the tool (questionnaire) measuring the right things?
- Are the questions clear? Not ambiguous (not open to different understandings)? Are they single questions (not two questions asked within the one question)?
- Are the questions asked in the right language; the type of English used by the respondents; or the local vernacular language which the respondents would normally use for discussing such issues.
- Will the respondents understand the questions? Will they be able to answer the questions? Are the questions appropriate to their educational level?

5.3 Drafting tools – Approaches and Stages in questionnaire design

(a) Writing new questions

When conducting original research and doing the research in your particular country context, it is unlikely that existing questionnaires will have included just the right questions to achieve all of your research objectives. In drafting new questions, it is good practice to first do some qualitative work (e.g. focus group discussions [FGDs]

[2] From the two facility visits we found that the answers to many of the questions we asked facility managers were found in the patient registers; hence there would be much greater value and better use of facility managers' time if study teams can target questions and interviews of managers towards exploring the reasons behind service trends.

and/or in-depth interviews) in the type of community where you will be doing the research, so as to learn how best to word the questions.

Preliminary qualitative work is more important where you are researching unfamiliar issues; and also if you are conducting surveys in communities with which you are not familiar. This is especially the case if the interviews are to be administered in a local language – the interviewers and the respondents need to have a common understanding about the meaning of key words and phrases. Questions must also be culturally sensitive, whether adapting from existing questionnaires or designing new questions.

In drafting new questions for a questionnaire to be administered to health workers or patients (e.g. exit interviews), in practice it may be sufficient to work at an early stage with a small group of experienced local health workers, especially nurses who have had many years working with patients. Reviewing draft questions with local nurses can go a long way to helping to find the right language that patients / service users / facility clients will understand.

(b) Adapting existing questionnaires

Take and adapt questions from existing questionnaires that have been used for similar purposes. These should be questionnaires that have been piloted and tested in the field.

Examples:

- SWEF facility surveys
- MEASURE SPA Surveys
- KAB / KAP (Knowledge, Attitudes, Behaviour / Practice) questionnaires that have been administered to health care providers to test their knowledge and care practice.
- Exit surveys of patients to explore health seeking behaviour and barriers / costs to accessing services.
- Bridging Gaps Surveys

6. RELIABILITY AND INTERVIEWER ISSUES FOR QUALITATIVE AND QUANTITAIVE METHODS

6.1 Selecting interviewers

Country team investigators may be able to employ field workers that they have trained and used on previous research studies. In such cases you will know their abilities (strengths and weaknesses) and be able to tailor the training based on this knowledge of their abilities

Where you need to select and train fieldworkers with little or no previous experience, their educational level and knowledge of the context in which they will be interviewing are important selection criteria. School leavers who have completed secondary level education are usually quick learners and – when trained to be respectful of older respondents – the age gap need not be a problem. The gender of the interviewer may be an important issue, especially when researching sensitive and stigmatising issues around sex and HIV/AIDS.

6.2 Training interviewers

One full week of training is a reasonable expectation for interviewers to gain the necessary expertise and to learn the standard operating procedures (SOPs) for the study; although it may be possible to shorten this to perhaps 3 days full training, where you are working with experienced field workers. Community surveys may require additional training time for field workers to learn how to (a) map households, (b) compile sampling frames and (c) draw probability samples.

Training needs to be structured and intensive, enabling interviewers to become familiar with:

- *The issues being researched*: ensuring that the interviewers have a thorough understanding of the overall aim and objectives of the research is the most important component of the training. This includes ensuring they have an understanding of the meaning of the different questions, what the questions aim to find out, and the types of answers that will 'make sense'.
- *Procedures for how to ask the questions*: reliability for structured questionnaires is about interviewers being disciplined in asking the questions in the same way under the same conditions with each respondent.
- *Respondents' level of understanding:* the interviewer needs to be able to consider each response given by a respondent and – if the answer is surprising or doesn't make sense – then check to see if the respondent fully understood the question. In such cases, so as to maintain reliability, the interviewer should repeat the question (best not to change the wording of the question but to ask it again slowly) and check if it was understood.
- *Interviewer role play*: one-on-one interviewer role play is a good training method. Arrange trainee interviewers into pairs, have one play the role of the interviewer and the other the respondent; and have the rest of the team of trainee interviewers observe the pair asking and answering questions. Then the whole team can review and comment on the role play.
- *Problems with the* questions: often the problems that emerge through group training and role play are because of problems with the questions, and not poor performance of the trainee interviewers. The training is an opportunity to fine-tune and make corrections to poorly worded questions.

6.3 Piloting the survey – interviewers and questionnaires

This essential phase can be overlooked when you have brought a team of interviewers into the field (maybe to a remote rural area) and you are under time and resource pressure. Consider first doing a pilot with typical respondents (doctors, nurses, patients) in a location convenient to your training centre, e.g. the nearest hospital or health centre.

Intensive supervision of the pilot is more important than piloting on a large number of respondents. If you plan to have 5 interviewers, a pilot of 10 respondents (2 per interviewer) closely observed is better than a pilot of 15 respondents with less intensive supervision. If you opt for two respondents per interviewer, one approach is for you (as the study leader) and for your supervisors to sit and observe and interview

with a real respondent; and then allow the interviewer to do a second interview without on-the-spot supervision, but with a follow-up de-briefing with the interviewer.

As part of the pilot, ask the interviewer to make notes during the interview about any questions that the respondent found difficult to answer. Afterwards, go through the completed questionnaire in detail with each interviewer, question-by-question, to try to work out if the problem was a) with the question, or b) with how the interviewer asked the question. At the pilot stage, you will usually find that there were problems with some of the questions that need further fine-tuning

It is often an effective strategy to follow up the pilot survey with at least a half day's further training, where you and the field worker supervisors, together with your team of interviewers, work through the questionnaire, question-by-question. This is so as to identify: a) questions that need to be improved; and b) if there is a need for clearer instructions to interviewers on how to ask questions and how to record responses.

The pilot is the opportunity to identify if there are particular *weaknesses with interviewers* that need to be addressed. You have to be prepared to drop interviewers from your team, if it appears that they are unable or unwilling to reach the standard that you demand. This is the time to do this before proceeding to the full survey. If you leave it until later during the full survey, you may find that a large percentage of your completed questionnaires (20% if using 5 field workers) are compromised. You may be forced to drop parts or all of these completed questionnaires from your analysis.

If, as the main investigator, you are not in a position to take time off and go into the field with your interview team and conduct the pilot in the locality where the survey is taking place, then consider doing a two-stage pilot: the first stage you oversee and the second stage of the pilot which your field worker supervisors oversee in the field.

The pilot survey will usually require you to:

- make further changes and fine-tuning to the questionnaire and questions;
- clarify instructions to your interviewers
- conduct further *individualised* training of your interviewers, addressing the different weaknesses of different interviewers.

6.4 Supervision and Quality Control

The most important thing in quality control is to ensure that there is *trust* between supervisors and interviewers. Interviewers must know and feel confident that they will not be blamed for making mistakes (e.g. if they overlook asking a question or omit to record a response on the questionnaire); but they must also know that they will be dismissed for any dishonesty, e.g. if they try to cover up their mistakes or omissions.

All completed questionnaires should be checked over by a trusted supervisor in the field before the end of the day. The supervisor should do this in the presence of the interviewer and check through every question and answer looking for:

- Questions *where the answer was not recorded*. The two most likely reasons for this are: a) the interviewer forgot to ask the question; and b)

he/she asked the question but forgot to record the answer. The latter is quite common in an interviewer administered structured survey. If there are clear ground rules about honesty and fairness, and trust has been established between supervisor and interviewers, then you can address the omission the same evening. Where the interviewer has a clear recollection of the respondent's answer, you may decide to record the response. Otherwise you have to leave it blank.

- *Consistency*: a careful reading of the completed questionnaire by the supervisor may reveal inconsistent or surprising responses (you may also have intentionally drafted consistency checks into the questionnaire, i.e. having two different questions at different points in the questionnaire, which are looking for the same type of information). Inconsistency, i.e. contradictory responses, may be a sign of some serious problems. It may mean that a) the respondent has not understood the questions; or b) the interviewer is asking the questions wrongly; or c) he/she is recording or coding the answers wrongly.

The importance of high standard in-the-field supervision of interviewers cannot be over-emphasised and this needs to be resourced and rewarded. One might want to consider performance-based incentives for supervisors and interviewers.

6.5 Data entry and cleaning

- Some software packages allow one to build in data entry checks that will pick up data entry errors (EPI Info has such checks).
- Independent double data entry (two different entry clerks and merging files to check for inconsistencies) is usually a good investment that will virtually eliminate data entry errors. Again EPI Info allows this facility.
- Doing data entry while in the field may help to identify and correct coding errors by interviewers.
- Frequency checks on each data entry field (each variable) will also show up data entry errors such as unusual outliers. If time has elapsed since the survey, the best you can usually do is check back to the completed hard copies of the questionnaires.

7. ANALYSIS OF STRUCTURED DATA

The type of analysis required for many of the surveys undertaken by GHIN African country teams will be basic analysis. Each of the following types of analysis can be done using EPI Info, although country teams may find other software packages, e.g. SPSS, better for some forms of analysis such as multiple logistic regression:

- *Frequencies* of variables
- *Prevalence rates* (proportions or percentages), e.g. of patients accessing key HIV/AIDS services such as VCT, PMTCT, ART, condoms, etc. The numerators (numbers of patients who received services) can be obtained from facility records (see *Method A*). The challenge will be to identify and find data on the size of the population in need of such services, which provide the denominator. One approach is to get figures for the district catchment population for each of the target groups, e.g. to

estimate the coverage of PMTCT services, you need to know for a period of time (preferably a calendar year, but could be a calendar month) the:

Numerator: number of HIV positive pregnant women who received PMTCT.

Denominator: this can be calculated from having estimates of:

- the number of women giving birth in the district each year *plus*
- the number of women in the child-bearing age (15-45 yrs? 20-45 yrs?) who are HIV positive.

- *Chi square* statistics, i.e. 2 by 2 tables to show the odds ratios and P (probability) values of the associations between variables, e.g. statistical association of levels of patient satisfaction (high versus low) with age of women giving birth (e.g. under 30 yrs versus over 30 yrs); or association of levels of satisfaction with VCT services with sex of the respondent (male versus female clients).

- *T Tests* may be used to compare age of respondent (in years) with level of satisfaction around quality of care (high versus low levels).

- *Rank sum tests* (e.g. Kruskal Wallis) might be used to compare levels of satisfaction in a Likert scale (see earlier), ranging from high to medium to low levels of satisfaction, testing the association with sex of the patient (male versus female).

8. QUALITATIVE RESEARCH IN PRACTICE

Qualitative interviewing (semi-structured and unstructured interviews with individuals and focus group discussions) is useful for in-depth exploration of issues, exploring issues unanticipated by the researcher and explaining patterns revealed in quantitative/ structured research methods. There are a number of uses of such methods:

i. As a primary data collection tool or approach, for example to explore barriers to accessing services (***Modules B1, B2 & C***). Barriers may be due to a combination of: a) individual and household factors, b) community factors; and c) health facility factors. There may be complex ranges of factors within each of these three sets of factors – a), b) and c) – that need to be explored in-depth through focus group discussions or in-depth interviews of key informants.

ii. To explore issues that have emerged through other methods, for example through: a) structured surveys; and b) record reviews. These methods are particularly useful for getting information from district and health facility managers, which can help explain *how* services are being delivered and the reasons for *why* these services are working well or *why not*. Examples:

- Service delivery records (***Module A 1***) may show that certain target groups or vulnerable groups (e.g. women, the young, less educated or poorer people, or those from rural areas) are not accessing services (*see Facility Surveys* earlier). These findings can be presented to key informants: facility managers, NGO managers,

PLWHA and patient advocates (***Module C 1 and C 2***), and to focus groups in the community, to explore the possible reasons for these findings (***Module C 3***).

- An historical analysis of service delivery records may show significant fall-offs in services delivered over a number of months. A potentially powerful tool for conducting in-depth interviews of facility managers – medical officers and clinical officers, nursing officers, human resource managers, hospital pharmacist – would be to collect monthly data on key services and activities (numbers of patients receiving VCT, ART, PMTCT, whether ARVs are consistently in-stocks in the pharmacy). One could present these historical data to a facility manager in the form of line graphs that show upward and/or downward trends in key services and activities over time. These initial analyses should inform the development of semi-structured interview guides, tailored to particular managers (*see below*) and inform the focus of these interviews (***Module B 2 and C 2***).
- A review of staff records, or information provided by the Human Resource (HR) manager on proportion of the staff complement present in the facility, and upwards / downward trends over time if such records are available, would provide the basis for an in-depth interview of the HR Manager to explore, for example, reasons for staff absences, etc. (***Module B 1***).

8.1 Sampling for qualitative research

Qualitative research is most often based on non-probability sampling, that is, individuals are not selected randomly as would be the case in most quantitative studies. There are a number approaches to sampling that are typical to qualitative research. These include:

8.1.2 Purposive sampling.

This involves selecting participants on the basis of their characteristics, roles or experiences. The purpose is not to generalise across a whole population but to shed light on a range of issues relevant to research questions. Often this is used where the research is interested in the perspectives of individuals who have specific roles, for example, Medical Officers in charge, human resource managers, District-DHMT-Directors and others. A variation of purposive sampling is *maximum variation sampling* that also involves recruiting research participants on the basis of their characteristics, roles or experiences. The aim is to interview *as diverse a range of individuals as possible*, and to focus on exploring the differences between individuals' accounts in the analysis. This approach may be appropriate when interviewing service users and communities.

8.1.3 Snowball sampling.

Snowball sampling involves asking interviewees to nominate other people they know who may be willing to participate in the research. This may be appropriate in settings in which it may be difficult to recruit research participants directly (for example 'hidden' populations not using services or involved in activities discriminated against or criminalised). This may also be appropriate approach to identifying and interviewing key policymakers or officials who may not be known at the start of the research project.

8.1.4 Sample size. In qualitative research the sample size is usually determined by the specific focus of the research. For example, in the case of purposive sampling the sample size may reflect the number of key individuals who have particular characteristics, roles or experiences that are of interest to the research. It may not always be possible to determine the precise number of interviewees before fieldwork commences since the researcher may not know which individuals they want to approach until starting fieldwork. *Saturation* is often seen as the point where an appropriate sample size has been reached meaning no additional perspectives or issues are emerging from additional interviews.

9. SEMI – STRUCTURED INTERVIEWS

A number of measures can be taken in maintaining quality at the data generation stage of qualitative research:

9.1 Recording the interview

Ideally qualitative interviews should be tape recorded so that a complete record exists of what was discussed. Tape recording should be subject to participants' consent. Tape recording is not without potential problems: in practice interviewees may be more forthcoming about controversial or sensitive issues if an interview is not tape recorded or after a tape recorder is turned off. It is usually recommended that detailed notes are taken (whether the interview is recorded or not) or a second researcher is present to make notes.

9.2 Field notes

Field notes can be very useful when analysing the data: it is important that interviewees write field notes straight after an interview or at the end of the day. This will strengthen the record that exists of the interview. Writing up relevant observations about the interview are also important. For example, did an interviewee's body language suggest they were being guarded or open? It is also useful to have a record of any problems encountered when conducting the interview that can be shared with other members of the research team. Interviewees should be encouraged to write up emerging issues, especially those that could usefully be explored in future interviews. Writing up observations such as those relating to how services are delivered in practice or systems of recording routine data provide additional insights outside formal interview settings that can also be very valuable.

9.3 Follow up interviews

A key advantage of qualitative approaches to interviewing is that it allows issues not originally anticipated by the researcher to emerge. Repeat interviews with selected individuals can be a useful way of exploring these issues, adding further depth to the research, as well as potentially breaking down barriers between the interviewer and interviewee.

9.4 Research assistant training and supervision

It is critical that research assistants/fieldworkers have good knowledge of the research topic being explored: this is especially important in qualitative research. This should enable them to probe key issues in depth by asking informed questions that are not

included on the interview schedule. Exploring issues in depth and gaining new insights are the hallmark of high-quality qualitative research. Ideally fieldworkers should have prior experience of conducting semi-structured and unstructured interviews. Training can clearly be used to extend their skills, clarify the nature of the research and identify the issues they should be aiming to explore in particular detail. Field supervision is important in drawing out key unanticipated issues for further exploration by the field team as a whole. Training and ongoing supervision are also an important part of maintaining the ethical integrity of fieldwork. Inexperienced fieldworkers may not know how to deal with sensitive issues that an interviewee finds distressing. It is important that the fieldwork team know when it is appropriate to probe further and when it is appropriate to move to a new line of enquiry.

10. CONDUCTING FOCUS GROUP DISCUSSONS

Focus group discussions (FGDs) are unstructured forms of group interview using a topic guide (a broad set of guiding themes for exploration). The method is often seen as having the following advantages:

- allowing participants to raise issues that they feel are important or significant
- allowing participants to agree with, probe or even disagree with/challenge each others views, often leading to the qualification, development or modification of views (potentially, getting nearer to actual experiences and perceptions);
- exploring how participants collectively make sense of phenomena. FGDs are rarely useful for generating data about experiences of a highly personal nature such as issues relating to living with HIV/AIDS.

Indeed, in settings where HIV+ status is discriminated against, FGDs may be an inappropriate method due to the risk that participants' status may be disclosed against their wishes. FGDs ideally comprise homogenous groups of 6-10 persons. When exploring issues relating to the use of HIV/AIDS services it is likely gender specific groups would yield more revealing sets of views. For some topic areas, e.g. utilisation of sexual health services, you may want to keep a) older and married women and b) teenage and young single women in separate groups. How you compose the groups should be determined by the research questions and an understanding of how these will impact on the target group. Guidelines on sample size correspond with that in 8.1.4. One error is to carry out too many FGDs and not analyse them properly. It is also important to choose the setting carefully: places perceived as neutral and private would be most appropriate (neither a clinical setting nor too close to participants' homes). It is essential that interviewers running FGDs have good facilitation skills and good knowledge of the issues being discussed, enabling them to probe emerging issues.

11. MAINTAINING QUALITY AT THE DATA ANALYSIS STAGE

There are also a number of means by which quality may be enhanced at the data analysis stage of research:

11.1 When to analyse data

Unlike analysis of quantitative data where analysis is typically carried out after data have been collected, analysis in qualitative research usually involves taking a *sequential* or *interim analysis* approach. This means that data are analysed soon after interviews take place and on an ongoing basis as further interviews are conducted. This approach allows the researcher to make tentative conclusions based on early analysis, as well as to review the performance of research instruments and if appropriate refine them. Based on this approach issues emerging can then be fed into subsequent rounds of interviewing for further exploration. Ideally data analysis should be carried out by the person conducting the interview who will be most familiar with the data.

11.2 Respondent validation

This involves revisiting interview respondents either when some or all interviews have been analysed or when research outputs are drafted to check that researchers' interpretations/conclusions correspond with what interviewees intended to convey. Giving respondents the opportunity to confirm, clarify or amend the final research output is likely to add confidence that findings are valid. One should first consider if a) this is feasible (community samples) or b) the respondent may want to bring pressure to bear on study (policy maker).

11.3 Triangulation

The *triangulation of methods* involves combining two or more methods to address a research question in order to cross-check results for consistency. The approach is useful for combinations of quantitative, qualitative or mixed (quantitative and qualitative) methods research. If triangulation produces consistent results, this may add confidence about the validity of findings and reduces the limits of individual methods. Similarly, *data triangulation* involves comparing different datasets (generated using the same method) for consistency. If triangulation produces inconsistent results it would be appropriate to either:

- re-analyse the data in order to understand the nature of the inconsistency or,
- generate additional research data to explain the inconsistencies (eg conduct further interviews). An additional approach to triangulation is known as *researcher triangulation*. This comprises researchers analysing the same datasets independently and then comparing results for consistency.

11.4 Working with transcripts

Whilst generating interview transcripts is a useful basis for detailed analysis it was recognised that this is a resource-intensive approach to analysis. It may therefore be appropriate to only transcribe some of the richest/most detailed interviews.

12 TWO APPROACHES TO ANALYSING QUALITATIVE DATA

12.1 Content analysis.

This is an approach to analysing qualitative data *quantitatively*. The approach is most commonly used if interviews are transcribed but could potentially be used to analyse tape recordings. It involves measuring the frequency of words or phases used by groups of interviewees and comparing this to the responses of other interviewees. Words or phases might be HIV/AIDS, gender, stigma, antiretrovirals etc. It could be used, for example, to compare female interviewees' accounts with those of male interviewees: do female interviewees talk more commonly about antiretrovirals than male interviewees? Content analysis is therefore useful for comparing the relative importance of those issues to different groups of interviewees. It is also good way of analysing policy documents. Using this method it would be possible, for example, to look at whether GFATM documents emphasise gender or stigma more than PEPFAR documents.

12.2 Framework approach.

The framework approach is a systematic way of analysing qualitative data that is relevant where many of the principal themes have been set in advance of fieldwork being conducted. It also enables emerging themes to be systematically analysed. The approach is best used for the analysis of transcribed interviews, although could also be used to analyse tape recorded interviews. The approach consists of five stages of analysis as follows:

i. **Familiarisation:** immersion in raw data (reading transcripts, listening to tapes); listing some key ideas and recurrent themes emerging from data. Numbering each transcript.
ii. **Developing a thematic framework:** identifying key themes (an 'index') - from original research questions and those emerging
iii. **Indexing:** applying the index systematically to all data (transcripts)
iv. **Charting**: reorganising the data from multiple interviews according to the indexes (themes). Simplifying and synthesising.
v. **Mapping and interpretation:** mapping/summarising the range of issues. Making connections between issues and providing explanations for the findings.