

Chapter 7: Collecting the Evidence

Henry Lucas
Institute of Development Studies

1. Data collection and resource allocation

In chapter 1 it was argued that where possible implementation researchers should be embedded as members of the implementation team, equally committed to ensuring a successful outcome and fully engaged in decision-making processes. One important area of decision-making concerns the allocation of scarce resources – which could alternatively be used to improve the scope or quality of the intervention – to data collection activities. Following the proposed definitions of chapter 1, these activities can be seen as directed towards three identifiable objectives:

1. Intervention management and accountability;
2. Operations Research intended to improve the current implementation;
3. Implementation Research to learn lessons from the current implementation that can be used in scale-up or re-location to a new context.

However, these objectives can be seen as highly interrelated, each involving the need to track implementation progress against the original intervention design, identify potential weaknesses in that design and test the initial assumptions on which it was based.

Every intervention should have budget lines intended to address objective 1, covering the costs of collecting the data required for intervention management and to demonstrate to those providing funds that resources are being allocated appropriately and outputs produced as intended. For convenience we can regard all such planned expenditures as falling under the general heading of intervention ‘monitoring and evaluation’. Often this will include a separate item to meet the cost of operations research studies to be undertaken in pursuit of objective two, for example testing alternative approaches to service delivery or behaviour modification in order to determine which would best serve the needs of the intervention. As discussed earlier, funding of the implementation research activities, including researcher time and data collection costs, will typically not be included in the intervention budget and would often be provided by another agency. However, much of the data required by the implementation researcher can be derived from that collected under the intervention budget, though the analysis of that data may well differ. This provides an opportunity to establish a mutually beneficial arrangement, with implementation researchers providing support, and possibly additional resources, to the intervention monitoring and evaluation system in return for full engagement in the design of that system.

From an implementation research perspective, the design of monitoring and evaluation systems requires not only an understanding as to how the data generated will allow rigorous analysis of the implementation process and the interaction of that process with key contextual factors, but an awareness of the types of evidence that will be acceptable to different stakeholders and audiences ([Murray 2007](#)). Apart from the other members of the implementation team, these might include national/ local policy-makers and officials; health workers; NGOs; donor and other international agencies; beneficiary communities and the general population. For example, donor agencies may demand ‘objective’ quantitative outcome indicators, while communities may be more impressed by qualitative evidence that reflects their own perceptions and concerns. A further consideration is that traditional monitoring and evaluation systems can tend to follow a routine reporting and analysis plan which is insufficiently responsive to rapidly developing potential opportunities and threats to the implementation process, especially in a CAS context. Both operations research and

implementation research activities may often benefit from ad-hoc, 'real-time' exercises, possibly undertaken in collaboration with service providers or intended beneficiaries, which can be effective in providing rapid feedback on access barriers and process bottlenecks as they arise. Such activities can combine information gathering with exercises that explore ways in which these barriers or bottlenecks may be overcome, both in the existing or in future implementations.

Choice of research methods

There is a tendency for researchers to think in terms of undertaking studies that can be easily categorized, using labels such as quantitative, qualitative, participatory, action-research, desk research, etc. A better approach is to start from a careful review of the various research questions that need to be addressed and then to assess which research methods might be able to deliver the required information on each of these questions. Resources should then be allocated in such a way as to best meet the overall research objectives, which will usually involve prioritising some questions over others, within whatever 'budget constraints' apply – which may relate not only to financial limitations but also to the limited availability of time, skilled/experienced personnel, access to data sources, etc. Such an approach may often involve unpalatable compromises relating to the scope, depth or precision of the intended research findings, but setting unattainable goals or attempting to 'make do' with inadequate resources will almost certainly degrade the *quality* of those findings.

One complicating factor in adopting such an approach, is that most researchers have a strong preference for primary data collection. They identify what they see as their requirements for specific data items and then assume that those requirements can only be met by the careful design and application of data collection instruments that are intended to deliver those items. However, before deciding to invest in any substantial data collection exercise, which will almost inevitably be costly and will typically prove substantially more costly than anticipated, it is almost always worthwhile to undertake a systematic inventory of relevant, accessible, secondary sources. These will almost certainly not provide precisely the data you *want* but may well provide data that can meet at least some of your underlying *needs*.

For example, reports and/or data from previous income or expenditure surveys may provide a reasonably adequate guide to current distributional questions if there is no reason to suspect that these may have changed radically since those surveys were undertaken. Even poorly maintained hospital financial records may provide better data for the estimation of inpatient treatment costs for a given condition than can be obtained from a survey that relies on the memories of former patients. Careful study of official reports, even if you are doubtful as to their reliability, will often enable you to be much more efficient in undertaking key informant interviews with senior policy-makers, allowing you to focus on questions which test the veracity of the information and opinions in those documents.

The general proposition here is that *all* potential sources of relevant data should be explored and their availability, accessibility, cost and potential value assessed before deciding on your research strategy. These would include:

- Documents: official reports, academic journals, media articles, internet blogs, etc.
- Routine data systems (RDS): financial data, personnel data, clinic records, etc.
- Existing survey data: national surveys, Demographic and Health Surveys ([DHS](#)) World Health Surveys ([WHS](#)), Multiple Indicator Cluster Surveys ([MICS](#)), etc.
- Implementation RDS: from the implementation monitoring system.
- New sample surveys: of patients, facilities, providers, community members, etc.
- Qualitative studies
- Rapid appraisal and/or Participatory exercises.

2. Secondary Data

Document Review

Document review should involve the *systematic* compilation and analysis of relevant printed and electronic material. In terms of health systems interventions, probably the most important sources will be legislative documents, policy statements and which set out the basic frameworks within which health systems function. There will also be a wide range of regulations, guidelines, manuals, protocols, etc., issued by ministries, other official agencies or by facilities themselves, which define the detailed operational procedures that should be followed in the management, administration and delivery of health services. These documents can be important even if the researcher is fully aware that they are widely disregarded, in that they can indicate what individuals perceive to be appropriate behaviour in terms of health service provision or at least what they perceive as being acceptable to the general population.

Organisations and individuals will often try to behave as set out in such documents even in the most difficult and chaotic circumstances, following procedures they know to be irrelevant simply because they have no well-defined alternative mode of operation. Working in Nigeria in the mid-1990s, when public health services were almost non-existent in many rural areas, the author had to work around a legal prohibition on the use of alternative forms for the collection of data on public facilities. This was often cited by providers even though it was clear to all those concerned that the official health information system had ceased to function. Similarly, state government officials would expend considerable efforts on the careful preparation of annual budgets, even though they knew that these would have limited effect in terms of controlling actual expenditures. Analysis of the gulf between what is contained in such official documents and the reality on the ground is often key to understanding the context within which interventions are undertaken.

A systematic document review should aim to at least consider, if not analyse, all those materials which may be relevant, from whatever sources. This will be time consuming and should not be seen as an activity which takes place only at the start of the implementation process but one which can be conducted at a steady pace over the research period. Increasing use of media outlets and in particular of the internet has dramatically increased the volume of information that is relatively easily accessible to the researcher. For example, reports from the international Demographic and Health Surveys, World Health Surveys and Multiple Indicator Cluster Surveys indicated above can be inspected or downloaded from their websites. In many countries census and survey reports are often made available in reasonably timely fashion through the internet sites of national statistical agencies or ministries. Reports from earlier periods, possibly useful in considering trends over time, may also have been archived on the International Household Survey Network ([HISN](#)) website. Expectations should be limited. Survey reports tend to provide relatively simple statistics, often at a high degree of aggregation, using variables that will almost certainly not have been defined as you might wish. Nonetheless, they can often provide a limited number of apparently relatively reliable indicators which may be very useful in terms of confirming or challenging information received from other sources.

Interesting insights into the concerns and intentions of relevant organisations can often be gained by examining their press releases, which again are now often made available via the internet. Given that they are almost always intended to present the organisation in a favourable light, these need to be subjected to careful analysis and interpretation, but can be extremely useful in determining the most effective strategy for exploring their underlying aims and objectives, for example in the design of key informant interviews. Media articles – in newspapers/magazines or on television or radio – provide another relevant source, which in this case will need to be assessed in terms of an informed judgement as to whether the author

can be seen as independent or biased in one direction or another. Such biases do not render the information useless – as long as its implications can fully incorporated into your analysis. Articles based on the opinions of those critical of your intervention can be of particular interest in terms of understanding the arguments that an implementation may have to address and in revising your stakeholder analysis.

A related and underutilised source are the internet 'blogs' that may be written by individuals within a healthcare agency or community based organisation, either on their own websites or on social media sites such as Facebook. The author gained valuable insights into the problems faced by an agency concerned with providing health advice from the activity on one such site, where it is easy for contributors to forget that their discussions are open to public view. Other sources of interest include the many advertisements for health providers and products, which may play a major role in influencing the attitudes of the local population as to the availability of treatments for a range of conditions. These may be found in the local media or on the internet but are also widely displayed on posters, either positioned by roads or in shop windows. Such 'documents' can now easily be captured and analysed in the same way as other materials using digital cameras. Such cameras can also be used to incorporate a range of relevant maps, charts and photographs which the researcher may encounter.

It should be emphasised that document review is a research activity and as such should be fully described in the research report. Details must be provided as to how different sources were explored and relevant materials identified, accessed and analysed. Stage one in such an analysis, as indicated above, is to understand the origins of the document and the reasons for its production. This should allow you to make an intelligent assessment as to how it may be interpreted. Was it an uncontentious attempt to codify existing practices to make sure that all providers followed a common approach or a highly contested regulation which imposed unwelcome constraints on income generating activities? Was it intended to demonstrate how careful a government agency had been in managing a social insurance fund or to attack the profligacy of political opponents? Stage two would involve seeking ways to verify any of the claims or estimates contained in the document. Can they be compared with those from any other source? Is the methodology adopted described in the document and if so does it seem appropriate? Is it possible to discuss the findings with those who had published them? Stage three involves the more difficult tasks of extracting relevant excerpts from each document summarising these without losing essential content and then combining these summaries under various themes and sub-themes. This is most frequently undertaken on a relatively pragmatic basis, relying on the experience and skills of the researchers. However, there are more rigorous methodological approaches which are usual described under the heading of 'content analysis' ([Hsieh and Shannon 2005](#)) and in recent times these have often been undertaken using a range of specialised computer software packages ([ESRC undated](#)).

Routine Data Systems

Routine (administrative) data systems (RDS) have the great potential advantage that they can deliver disaggregated, time-series data by geographical area (region, state, district, sub-district, etc.) ([Lagarde 2012](#)). Facility records, for example attendance registers, patient records, disease registers, prescriptions, insurance payments, financial accounts, etc. can be an important source of quantitative data, if they can be accessed by the researcher. They may be immediately of value or of use after further processing. For example they may require reorganisation, aggregation, disaggregation or other manipulation. In this case it is necessary to ensure that the nature of the data in terms of such aspects as definitions and collection procedures is thoroughly understood, as the possibilities for misinterpretation are considerable. It may in some cases be cost-effective to invest resources in measures which support improvement of the RDS. For example, in some countries many primary facilities still lack simple electronic calculators and may have to spend considerable time adding up many

columns of figures, often making mistakes. In one study the author found that the simple provision of higher quality attendance registers and prescription pads (the only documents available at this level) and a supply of pens, pencils and erasers, dramatically improved the recording of patient and treatment details over the course of the research. In recent years the provision of mobile phones does appear to have considerable potential ([Neupane et al. 2014](#), [Uganda 2012](#)).

Mainly because they can provide data which can be disaggregated to the particular location in which an intervention is undertaken, it can be more useful to analyse RDS than existing survey data. However, in most countries it suffer from well-known limitations, often in spite of many attempts at improvement. There are three major issues:

- coverage – focusing only on those who use services can be extremely misleading – we know that it is generally the poor and vulnerable who are most often excluded;
- general poor quality (accuracy, timeliness) – often reflecting indifference on the part of staff who have come to believe that senior health service managers, officials and politicians rarely make use of, or even consult, the data they provide;
- incentives to misreport (e.g. where providers receive performance related payments) – and an absence of effective audit systems that might detect misreporting.

The poor quality of the data may be improved to a limited extent by measures such as those indicated above, but typically relate to a widespread culture of indifference to reliable reporting which is not easily amenable to change, given the resources available to any specific intervention. Some financial data (e.g. payroll data, payments by health insurance agencies) may be more reliable because it is subject to audit procedures. One potentially useful activity is to explore the possibilities for combining routine data with other sources, such as surveys, to generate ‘best estimates’ ([Rowe 2009](#)). This implies the need for expectations to be limited and second-best options to be explored. For example, while such basic indicators as service utilisation, access and cost are not ideal, they may provide a reasonable basis for context analysis and to verify data from other sources.

RDS data quality is a particular concern when disaggregation within the intervention area is required. As a general rule, administrative data quality depends on the quality of administrators, and both tend to be correlated with the overall level of development. The poorest areas and facilities typically have the least reliable data. This is of particular concern in terms of indicators derived from information systems which are subject to the pressures associated with the provision of marketable goods and services. For example, rural health workers in poorer areas (given that their government salaries are sometimes barely sufficient to purchase basic food and clothing) have become very adept at providing information that satisfies higher levels of administration while not limiting their alternative income generating activities. It should be noted that variations in the quality of data, particularly administrative data, between areas and facilities may also influence aggregate estimates, as these are often based on partial coverage. Facilities in less developed areas not only tend to provide less reliable data, they often fail to provide data on time. As overall estimates are often derived by “grossing up” the information available when estimates are required – i.e. information from better resourced facilities, biases which tend to overestimate service utilisation, staffing levels, drug availability, etc. may be introduced.

Finally, note that many of the most important health indicators require the combination of service data from the RDS with overall or age-specific population estimates. These will reflect the ‘denominator problem’ of indicator construction – the fact that these estimates are typically crude estimates and/or outdated. The influence of changing population sizes and distributions, often due to internal migration, on access and utilisation measures can be substantial and will often need to be considered in the interpretation of trends over time. Again, poor regions may be particularly affected by both push and pull migration factors. The use of population

estimates also raises issues of data availability. Population estimates in years removed from that in which the census is taken will be derived from demographic models, often based on parameters estimated from DHS data. These models may be reasonably reliable at the national level but are not intended for sub-national estimation and typically do not allow for the effects of possibly large-scale internal migration.

Existing Survey Data

Anyone who has undertaken a reasonably large-scale sample survey will appreciate that it can be a daunting task. It primarily requires a range of managerial and administrative skills that are often lacking even in some of the most talented and experienced social science researchers. In particular, surveys usually involves the hiring, training and management of a substantial number of enumerators, supervisors and data entry staff who may have little interest in the survey objectives and need constant encouragement and oversight to ensure the quality of the data produced. They also typically involve a considerable investment in terms of both time and money. If it seems at all possible that relevant research questions can be addressed by secondary analysis of an existing survey data set which is known to be of reasonable quality, it would be a mistake not to at least seriously consider this option ([Boslaugh 2007](#)).

One key question to be addressed is how the quality of the existing data set is likely to compare to that from any new survey. Where surveys which have been conducted on a regular basis for a number of years by permanently employed staff members, for example from a national statistics office, their accumulated experience may well imply that the quality of the final product is likely to be considerably in advance of that from a newly designed survey conducted by a team of recent recruits employed on short term contracts. The sampling expertise and resources (e.g. computerised sampling frames) available within the agency that designed the existing survey may also have been far in advance of that available within the implementation research team. This may imply that there will be greater uncertainty as to the validity of estimating population parameters using sample statistics derived from the new survey.

In addition, the existing survey may have included questions on topics, for example incomes or expenditures, which would be of considerable value in any analysis but which could not realistically be included in a new survey given budget constraints. If it had been undertaken on a national or sub-regional basis, it could also provide an opportunity for direct comparison of data from the implementation sites with that from other areas, an important consideration when exploring the opportunities and challenges involved in scaling-up or relocating the intervention. Similarly, if the same questions have been asked in successive rounds of the survey over previous years, it may allow analysis of trends over time which provide insights that would not be available from a cross-sectional survey.

The above qualities are of course irrelevant if the survey data cannot be used to explore the questions that the research needs to address. An initial problem may be that it is difficult simply to gain access. For example, national statistical agencies will usually argue that survey data is collected on the basis that it will only be used for a specific purpose and that the respondents have been assured that it will not be shared with other organisations. Versions of the data from which any variables that can be used to identify respondents have been removed may be made available but often only with a considerable delay that reduces its value. Agencies may also require researchers to make a formal request for the data which involves a detailed explanation of the types of analysis to be performed and the intended uses of any findings, possibly requiring any resulting reports to be submitted to them before dissemination. In some cases they may also demand a substantial payment for use of the data. Note that the international agencies indicated above usually do make data freely available to researchers with minimal formality and it is also worthwhile to explore the International Household Survey

Network ([IHSN](#)) website, which does hold selected survey data sets, though many of these will be some years out of date.

The researcher will not only need to gain access to the data itself but also to the ‘meta-data’ which provides a detailed description as to how it may be analysed and interpreted. As a minimum this must include the questionnaires and coding manuals, but it will often be very useful also to have copies of the enumerator and supervisor manuals. For example, if respondents were asked if they visited a public or private clinic the answers may well differ substantially depending on the guidance (if any) provided by enumerators as to how to distinguish between these two types of facility. Having considered the precise nature of any variables of potential interest in the data, the researcher will then have to make a considered decision as to whether they can be used to at least provide insights into the original research questions. It is in the nature of secondary data analysis that the variables available are rarely those which the researcher would have chosen to analyse. The original question may not have been worded as you would have wished. The instructions to the enumerators may have resulted in an excessive number of missing responses. The coding system adopted may have lost information that would have been extremely useful. Even if these problems can be overcome, you may find that the sample size is too small to allow disaggregation to the extent necessary to provide relevant estimates for the intervention population. Nonetheless, all these potential limitations should not prevent you from exploring this option. The costs are often minimal and the potential benefits considerable.

3. Primary Data Collection

Qualitative or Quantitative?

A somewhat simplistic view of the appropriate uses of alternative approaches to primary data collection is shown in the following table:

Table 1: Matching methodologies to objectives

Objective	Methodology
Quantitative estimates representative of population parameters Knowledge of sampling errors	Formal surveys
Quantitative data with some understanding of processes Repeatable for trend assessment	Quantitative Rapid Appraisal
In-depth knowledge of behaviours, perceptions, attitudes, etc. Interpretation of existing quantitative data	Qualitative Methods
Very limited contextual knowledge Limited resources	Qualitative Rapid Appraisal

At one extreme, we might be concerned to produce estimates of specific population parameters, for example utilisation rates for health facilities or the frequency of given symptoms in young children over a previous period, together with the associated ‘errors of estimation’, which allow us to specify how confident we can be that those estimates fall within a given range. If we wish these estimates and confidence limits to be widely accepted as valid, we would be well advised to use formal surveys which follow the accepted principles of

statistical inference. If we are less concerned about the precision of such estimates and believe that we can derive them to an acceptable degree of accuracy using alternative and less resource intensive methods, for example by extrapolation from facility records, questioning key informants or focus groups, or techniques such as participatory ranking or mapping ([Chambers 2007](#), [Rifkin 1996](#)). In the above, such approaches are described as 'quantitative rapid appraisal'. Using standard procedures can allow comparison between areas and over time, but the extent to which such estimates are accepted will in this case depend on our ability to persuade others of their reliability.

If we need to understand not only *how* individuals and organisations behave but *why*, we may decide that some form of detailed qualitative study is required. This may involve long-term engagement with the study population, using a range of observational and interview techniques to formulate and then test alternative explanatory theories. Finally, if we know very little about the context within which we are working, a common situation at the start of any research activity, we might adopt an approach which can for convenience be labelled 'qualitative rapid appraisal', mainly using key informant interviews to enable us to at least frame relevant research questions.

However, it is often worthwhile to think 'outside-the box' when considering which methodologies and methods might be the most appropriate (or cost-effective) to meet data requirements in a specific context ([Holland 2007](#)). [Kanbur \(2003\)](#) suggests that we usually categorise qualitative and quantitative methods as having the following characteristics, locating them at the opposite ends of five 'dimensions':

Qualitative	➡	Quantitative
Non-numerical information	➡	Numerical information
Specific and narrow target groups	➡	Large general target population
Active engagement with respondents	➡	Passive involvement of target population
Inductive methods of inference	➡	Deductive/statistical inference
Description/generalisation/theory construction	➡	Hypothesis testing/econometric modelling

But there are no 'rules' that force you to accept this dichotomy. Given that researchers are always constrained by limited budgets, they should try to assess the costs and benefits of locating at different points along each of these dimensions in a specific research context, considering only how they will justify their decisions if challenged. For example:

- Traditional household surveys can be used to gather non-numerical information using 'open-ended' questions ([Rog et al. 2011](#)).
- Participatory methods can be used to generate numerical data – e.g. ranking of providers, estimated travel times to different facilities, etc. ([Chambers, 2007](#)).
- Qualitative studies can use probability sampling and large sample sizes to gain credibility ([Barahona and Levy 2002](#)).
- Qualitative studies may rely primarily on observational data, involving limited interaction with members of the targeted population ([Walshe et al. 2012](#)).
- Qualitative studies of social networks can use statistical methods and mathematical modelling techniques to generate network maps ([Bishai et al. 2014](#)).

Potential advantages and disadvantages of qualitative studies

One great attraction of qualitative approaches to many researchers is the extent to which they feel in control of the process. Sample sizes are typically relatively limited, allowing a small number of skilled, experienced researchers to take the time required to fully engage with those who are providing information. There can be considerable flexibility, with those researchers being trusted to make decisions as the research proceeds, for example selecting additional or alternative respondents, adapting questions or participatory as their knowledge of a situation increases and possibly opening up unplanned lines of enquiry if unexpected responses or observations suggest that these may be of importance.

Given sufficient expertise, researchers can undertake detailed investigations not only as to the knowledge of respondents but also their perceptions, attitudes and motivations. If they can gain their trust, they may be able to explore sensitive issues and assess emotional responses. Interviews which take place in homes or facilities will often allow valuable insights into relationships, processes and contexts simply by careful and prolonged observations. Of particular importance when there is limited knowledge at the start of a research activity as to which are the most relevant issues, qualitative studies can allow the gradual elaboration of concepts and theories as the research proceeds, delaying the often very difficult task of formulating precise definitions of variables and the expected relationships between them until the researcher has had an opportunity to experience the ground realities ([Kuznetsov et al. 2013](#)).

To some extent, the disadvantages associated with the archetypal qualitative study can be seen as the mirror image of the advantages. The flexibility which is so attractive to many researchers tends to place great weight on the regard in which the members of the research team are held by those whom they might wish to persuade of the value of their findings. The central issue is that of subjectivity, that given the extent of their control over the process of data collection it is likely that the research findings will be at least partly determined by the preconceptions of the researchers, i.e. they will tend, quite possibly unconsciously, to gather information that reinforces their personal perceptions as to how the world works. While it can reasonably be argued that quantitative research also has to contend with this issue, the use of predetermined instruments and procedures – questionnaires, manuals, sampling designs etc. – provides those who wish to determine the extent to which findings have been influenced by the decisions of the researchers with the documentary evidence they require. This indicates the way in which qualitative researchers can guard against their findings being dismissed as 'too subjective', by ensuring that every step in the data collection process is carefully documented, providing detailed descriptions not only of what was done but why. This should be a central component of an activity usually described as 'reflexivity' ([Finlay 2002](#); [Mruck and Breuer 2003](#)), ongoing assessments by each researcher of the extent to which their activities might be driven by personal factors and attempts to counteract that tendency.

One related common criticism of qualitative studies is that of sample selection biases, for example tending to gather information more from those who are in favour or those who are against the intervention, neither group being *representative* of the overall population. Researchers will usually try to avoid obvious potential biases, such as relying on local officials or 'community leaders' to determine their subjects, but it is easy to overlook other potential pitfalls, for example limited resources may result in a failure to seek out less accessible stakeholders, for example those who live in remote or less accessible areas. Sample sizes are often limited in qualitative studies. The essential need to use only capable, experienced researchers, because the quality of the findings is so dependent on their abilities, generally implies that the cost per respondent will be substantially higher than that for quantitative studies using enumerators to complete standardised questionnaires.

Small samples can raise difficult problems in terms of analysis and interpretation, given that we are often interested in the relationship between the diverse circumstances and characteristics of our respondents and their perceptions, attitudes, etc. We would often see it as essential to distinguish between respondents in terms of a range of attributes including gender, age-group, income/wealth, rural/urban, etc. even if we adopt a policy of stratification, such that we have respondents in each cell of the implied multi-way table, the numbers in each cell will be so small that we may be reluctant to infer that they can be extrapolated to other 'similar' individuals in the study population. One common challenge to qualitative findings is that they are *anecdotal*, interesting as descriptions of individual cases but *unrepresentative* and therefore of limited use in terms of reaching general conclusions and hence in terms of policy making. A similar complaint may arise with respect to comparisons between the various groups, for example differences in attitudes as between men and women. If there were relevant differences in the nature of the information gathering process between groups, for example different researchers choosing to vary the type or sequence of questions, or the use of male researchers to interview men and females to interview women, it might be argued that at least part of the observed differences may simply reflect inter-interviewer variation. A final, practical disadvantage, of qualitative studies is the sheer volume of information, mainly textual, that they almost always generate, posing substantial problems in terms of analysis and interpretation, even with the use of computer software packages ([ESRC undated](#)).

Potential advantages and disadvantages of quantitative studies

A well designed and implemented probability sample survey has the unique advantage of being able to provide reliable, bounded estimates of key population parameters, for example immunisation rates, illness prevalence rates, utilization of services, average length of stay in hospital, median cost of an outpatient visit, etc. Unlike any other methodology, it allows the researcher not only to generate such estimates but to specify how 'confident' they are that each estimate falls within a stated range (the 'precision' of the estimation). These estimates are derived using the area of mathematics known as statistical inference, which allows a researcher who can show that they have 'followed the rules' of probability surveys to present such estimates without the need for further justification. While other approaches, for example market research surveys or political opinion polls, may make similar claims, they are almost always not following the rules and therefore cannot legitimately use the language of statistical inference to support those claims.

This ability to generate reliable estimates to a given level of precision can be very attractive to policy-makers because it allows them to assess the potential quantitative impact of a given intervention. For example, China has recently started to introduce policies which provide improved health insurance coverage for the poorest members of rural populations. Such policies had been recommended by health researchers for many years but became much more acceptable to government when the costs of such changes could be reliably estimated from probability sample survey data. Again very useful from a policy perspective, the adoption of predetermined and standardised instruments for data collection in most quantitative studies enhances the credibility of making comparisons between different subgroups of the target population. Given that precisely the same questions are asked in what should be precisely the same manner to such subgroups, for example the heads of richer and poorer households, it will often seem plausible to directly compare their responses, for example in terms of the proportion of children under two vaccinated against polio. Quantitative studies generally try to minimise any variations in behaviour between those collecting data from different subgroups, which may be misinterpreted as between sub-group variation. Similar consideration apply to comparisons over time, for example estimation of trends in childhood malnutrition rates using DHS data for different years.

The desire to make comparisons between subgroups or over time is related to one of the main disadvantages of the typical quantitative approach, the difficulty of developing simple, *uniformly applicable* definitions of key concepts that are well understood and have a common interpretation across all subgroups of the population. For example, in one pilot exercise conducted by the author, a standard question as to whether anyone in a household had suffered an acute sickness in the previous two weeks produced incidence rates for those in the poorest rural area surveyed that were far too low to be believable. A follow up qualitative study found that fevers were so common that many people did not consider them worth reporting. Similar issues arise with respect to many of the covariates on which we often try to collect data in such surveys. The distinction between rural and urban areas, for example, is often problematic as is that between public, not-for-profit and for-profit facilities (if we are aware that they are all charging for services to a greater or lesser extent). Particular difficulties arise with studies that are concerned with equity. Measures of income, expenditure, wealth, indebtedness, vulnerability, etc. are notoriously difficult to define in ways which can be confidently expected to produce comparable findings across sub-groups ([World Bank 2003](#)).

The above indicates the need for a profound understanding of both the topics addressed and the population targeted at the design stage of any quantitative study. They should certainly not be used to explore issues about which the researchers have very limited understanding. That will almost always result in a substantial expenditure of resources to little purpose. In-depth knowledge is essential if the study is to be well designed and the design phase is often the key to a successful outcome. The implementation of a large-scale quantitative study is primarily an exercise in human resource management and logistics. Once launched it is very difficult to change course or rectify any major design defects that may become apparent. It is essential to ensure: (a) that the research team has the necessary management skills required and that those with these skills are willing to take a leadership role – along with the responsibility for ensuring that the exercise proceeds with as little divergence from the original intention as possible; and (b) that the resources are sufficient to allow for unexpected problems – bureaucratic delays, equipment failures, illness, bad weather, etc. – which will almost inevitably be encountered. Attempting to stretch an inadequate budget and ‘hoping for the best’ is a recipe for failure. Finally, it should be taken into account that those who most strongly favour quantitative studies often have a tendency to pay insufficient attention to likely data quality issues, preferring to make heroic assumptions as the reliability of the findings derived from this data, often substituting technical expertise for considered analysis and claiming general validity for what are typically very simplistic models of causality.

Combined Methods

In practice, it would be very unusual, and almost certainly a mistake, not to use both quantitative and qualitative approaches in any implementation research exercise. While there is a very long history of researchers combining quantitative and qualitative methodologies, the mid-1990s saw a more formal discussion of the opportunities and potential pitfalls of using ‘combined methods’ (sometimes described as mixed methods or qual/quant) ([Qualquant 2016](#), [Palinkas et al. 2011](#), [Barnett et al. 2016](#)). Most attention has focused on the potential advantages of using qualitative studies to complement and support large scale surveys ([Kanbur and Shaffer 2005](#)) they include:

- The use of qualitative studies to improve survey design
- The interpretation of counterintuitive or surprising findings from surveys
- Explaining the reasons behind observed survey outcomes
- Exploring the motivations underlying observed behaviour
- Suggesting the direction of causality
- Assessing the validity of quantitative results
- Understanding conceptual categories such as ill-health, household, etc.

- Interpreting local categories of social differentiation, e.g. poor/non-poor
- Provide a dynamic dimension to cross-sectional household survey data.

However, there are multiple pathways by which qualitative and quantitative studies might be linked. [Marsland et al. \(1998\)](#) categories these pathways under three broad headings:

A: *Swapping tools and attitudes: "Merging"*

1. Adopting standard sampling techniques in qualitative studies ([Barahona and Levy, 2002](#)).
2. Coding responses to open-ended questions using qualitative enquiries.
3. Using statistical techniques to analyse quantitative data obtained from qualitative studies, for example:
 - a. Creating frequency tables from coded responses to open-ended questions.
 - b. Constructing models based on binary and categorical data from ranking and scoring exercises.
4. Using participatory mapping to create sampling frames for questionnaire surveys.
5. Using findings from qualitative studies to reduce the non-sampling error (e.g. misunderstandings, offensive questions) in questionnaire surveys.

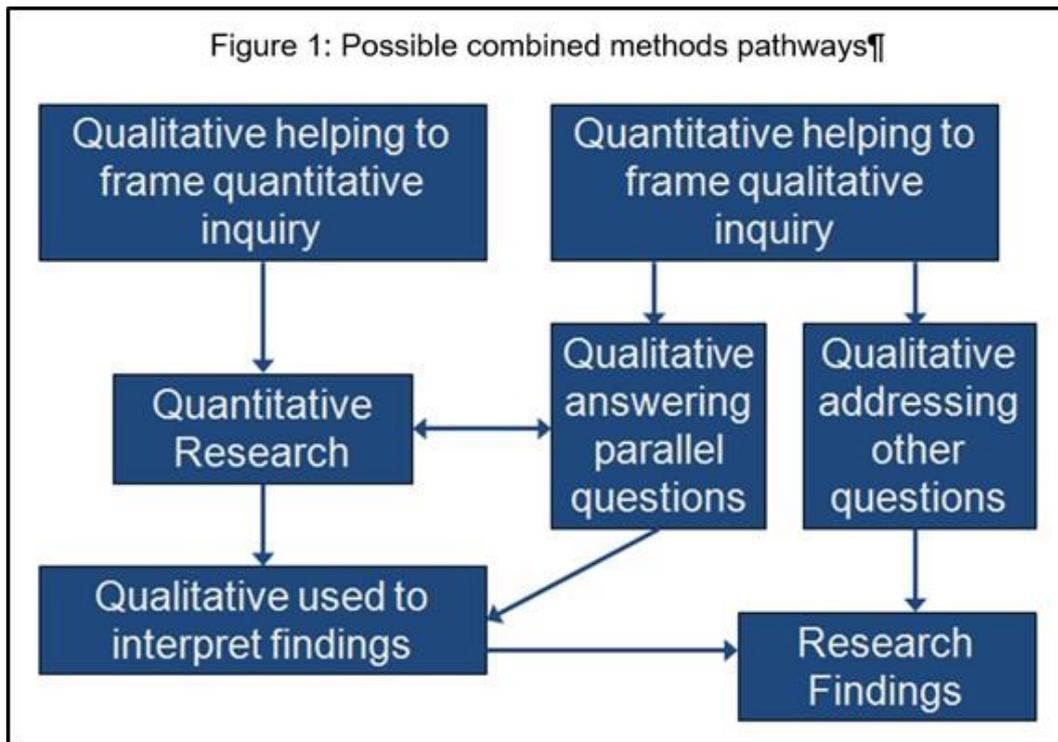
B: *"Sequencing"*

1. Using exploratory techniques to establish hypotheses which can be tested through questionnaire surveys.
2. Using a questionnaire survey to gather responses to a few key questions from a probability sample of respondents and then undertaking a qualitative follow-up study of respondents that appear to be of particular interest.

C: *"Concurrent use" of tools and methods from the different traditions*

1. Using a questionnaire survey to determine quantitative indicators (e.g. Likert scales ([Sullivan and Artino 2013](#))) on perceptions and attitudes relating to public and private health services.
alongside:
2. Qualitative exercises (key informant interviews, focus group discussions, participatory exercises) to address the same issues with the aim of gaining greater understanding.

These possibilities are reflected in figure 1 below.



For example, [Lucas et al. \(2009\)](#), used large-scale sample surveys in Cambodia, China and PDR Laos to identify households where at least one member had suffered from a serious illness over the course of the previous year. A limited number of geographical case studies, based on purposively selected counties in China and health districts in Cambodia and Lao PDR were undertaken. In each of these areas households affected by major illness were identified and studied using a two stage approach:

1. A rapid and reasonably large-scale household questionnaire survey was undertaken using cluster sampling of households within the selected study areas. This aimed to identify households substantially affected by different categories of serious health problems and to estimate the proportions of such households in the population.
2. The sampled households were analysed and classified into a number of strata based on the information provided by the questionnaire survey (the choice of stratification variables is indicated below). In-depth studies, typically requiring 1-2 person days, of a probability sample of the households in purposively selected strata were then undertaken by a team of social scientists.

4. Sampling in quantitative and qualitative studies

Implementation research aims to generate findings that are taken seriously not only by academics but much more importantly by the implementation team, policy makers, service providers and the general population. As indicated above, it will typically use a combination of quantitative and qualitative methods. However, whatever the methodology there may be advantages in selecting samples that can be seen as representative of some specified population, allowing findings to be more plausibly generalized, or alternatively selecting samples which exhibit a high degree of diversity, to demonstrate that all aspects of an issue are explored. Whatever the methods of data collection, it is at least worth taking time to consider how accusations of bias in the selection of research subjects, a very common tactic adopted by those who wish to discredit unpalatable research findings, might be addressed

([Barahona & Levy 2002](#)). A key point to remember is that your choice of data collection procedures does not dictate your choice of sample selection procedures. The only requirement is that these choices should jointly aim to deliver findings of a quality and nature that you believe: (a) justify their use to influence your target audiences; and (b) have the potential to be accepted by those audiences.

Why is sampling important?

Though we may tend to associate the concept of sampling with the formal procedures adopted, for example, in household surveys, quality testing of drug supplies, or audits of financial accounts, in practice it is central to our understanding of the world. Of necessity, almost all of our experiences are derived from samples: the people we have met, the organisations we have encountered, the documents we have read, the places we have visited, are a vanishingly small proportion of all those which might have influenced our perceptions and attitudes. In principle we know that information obtained from samples can be misleading, that it can be *biased*, to a greater or lesser extent, but it is often very difficult to act on that knowledge in a consistent manner. For example, even when engaged in our professional research activities, our impressions of people and places will often be strongly influenced by the small number of individuals that we meet and the observations that we make during a relatively short visit at a particular time of year. Even if we know or suspect that the doctors, nurses, local officials and community leaders that we encounter have been selected by stakeholders that have a strong incentive to ensure that we leave with a particular impression, it is often very easy to persuade ourselves that we are such skilful and experienced researchers that we can 'read between the lines' and make unbiased assessments of the true situation.

There are two problems with this approach. First, that we are almost certainly overestimating our abilities. Self-deception is a widespread human trait even among the most intelligent ([Lamba and Nityananda 2014](#)). Second, we are taking a serious risk in assuming that the target audiences for our research findings will accept our assurances that we have such abilities. As emphasised throughout this book, implementation research findings may be used to advocate major changes in health policy or major reallocations of health resources. It is appropriate therefore for those presented with such findings, particularly if they are in a position to make those changes or reallocations, to aggressively examine the methods we have employed. The simple assertion that you are convinced that the findings are not influenced by *sampling biases* is not, and should not be, a sufficient response.

Some definitions

The language used in discussing sampling procedures can be a cause of confusion. The following definitions can be applied to all forms of sampling:

- Population: A collection of entities – individuals, households, records, organisations – about which we wish to make qualitative or quantitative statements.
- Sample: The subset of entities on which we base those statements.
- Sampling Design: The procedure used to select that subset of entities.
- Sampling errors: Misleading findings arising from reliance on data from a sample.
- Non-sampling errors: Misleading findings arising for any other reason.

Note that some care is required in distinguishing between the sample we intend to obtain and the sample we do obtain. For example, if we make the elementary error of visiting a village at a time of day when most working-age adults are absent, we should be very cautious in assuming that the perceptions and attitudes of those available for interview reflect those of the community at large. Similarly, if any of the questions in a given study address sensitive issues,

we may well find that a substantial proportion of the members of our intended sample refuse to answer. Again, we would be foolish to assume that the responses can be considered representative of the overall population. One useful way to think about such issues is to decide which members of the population had at least some chance of being included for our sample. In technical terms, we usually describe such individuals as having a *probability greater than zero* of being included. This group is sometimes described as the *sampling population* because it is the population from which our sample is selected in practice rather than in theory.

We should always be extremely cautious in making inferences from a sample that relate to individuals who were not in the sampling population. For example, an utilisation study based on a sample of clinic returns available at the ministry of health will not necessarily provide reliable estimates at a national level if a substantial number of clinics in poorer, more remote, locations have failed to submit their returns. A study on healthcare costs using a sample of health insurance records cannot be assumed to allow inferences to a population that includes the uninsured. A study assessing satisfaction with health services by using mobile phone calls to a sample of patients following treatment must consider the possibility that the findings might have been very different if it had included those without access to a mobile phone.

[Chambers \(2006:28-32\)](#) uses the term 'rural development tourism' to explore the potential for sampling biases in assessments based on the type of short, infrequent visits to study sites that are common not only in research studies but which also play a major role in the routine activities of public officials, politicians, consultants, NGO/donor agency staff, etc. Such biases, which can dramatically limit the size and diversity of the sampling population include:

1. Spatial Bias: Staying close to urban areas, traveling only on tarmac roads and preferring to visit communities and individuals who live near to such roads.
2. Project Bias: Areas where activities are taking place, projects are in progress.
3. Person Bias: Elites, Males, Users (those who use services or are targeted by projects), Visible/Active Individuals, etc.
4. Seasonal Bias: Fieldwork done when travel is easy, avoiding floods, droughts, etc.
5. Diplomatic Bias: Reluctance to annoy elite members of the community by addressing sensitive issues, tendency to avoid potentially embarrassing encounters with the very poor, the very sick or community outsiders.
6. Professional Bias: Focusing on areas offering favourable opportunities for research. For example, where health facility records are well maintained.
7. Security Bias: Avoidance of areas with a risk of political unrest or violence.

There are many ways in which we can reduce the likelihood and/or extent of sampling bias. For example, in terms of seeking evidence from existing documents we often adopt a process called *systematic review* ([Hemingway and Brereton 2009](#)), which aims to ensure (a) that we clearly describe our approach to making best use of those documents and (b) that we have read at least what we can argue (on the basis of explicit criteria) to be the most important. In the situation described above, making a brief visit to a new study site, we might seek alternative ways of selecting our informants, for example asking to talk to male and female, older and younger community members or taking advantage of clinic visits to engage discretely with low level health workers who were not invited to the formal meetings. There are a wide variety of ways to select a sample. The basic principle is that we try to devise *and rigorously implement* sampling procedures that are less likely to result in samples that lead us to make incorrect inferences.

Example: a small-scale exploratory study in the Niger Delta, Nigeria, sought to discover barriers to use of bed nets ([Galvin et al. 2011](#)). Convenience samples of volunteers in a number of villages were asked about sleeping arrangements, perceptions of bed nets and barriers against their use. Because they tended to talk to the most easily accessible respondents, a substantial majority of those in the samples were adult males. The article

reports that the “gender bias of our sample” limited “the conclusions that can be drawn ... importance of structural and inconvenience factors, safety, and comfort may reflect issues of more concern to males”. A “more focused purposive sample” of “mothers of under-fives, and young people 12 to 19 years may illuminate some further barriers or motivations”.

Some of the most commonly used sampling methods by researchers, which vary considerably in terms of the extent to which they address the key issue of selection bias, depending on the context in which they are employed, are:

1. Convenience: selection of those who are easily available and willing to respond.
 - Snowball sampling is often used for hard-to-reach cases (for example those with a stigmatising health condition) and involves locating an initial respondent of interest and asking them to identify others, who can then each nominate further potential candidates ([Magnani et al. 2005](#)).
2. Chance: attempt to avoid accusations of bias by introducing a chance factor.
 - Quota sampling ([Scott et al. 2013](#)) involves selecting a pre-determined number of respondents in various categories (e.g. by age group and gender), usually to match population proportions. Selection typically takes place at a convenient location (e.g. interviewing individuals walking to market or waiting at a clinic).
 - Transect sampling ([Leslie et al. 2009](#)) is a process in which the researcher takes a central position in a community, selects a random direction (traditionally by spinning a bottle) and then selects respondents who live along that direction.
3. Purposive: sample determined by researcher using knowledge of the population and context to meet stated objectives ([Palinkas et al. 2013](#))
 - Typical (modal instance) sampling aims to select a sample that is reasonably representative of a given type of respondent.
 - Heterogeneity (diversity/maximum variance) sampling aims to include the full range of potential responses.
 - Stratified purposive sampling is similar to quota sampling, but with the specific respondents in each category determined by the researcher.

Probability sampling

One special group of sampling methods involves the use of ‘*probability sampling*’. This is the approach adopted in large scale sample surveys such as the Demographic and Health Surveys, World Health Surveys and the Multiple Indicator Cluster Surveys, and it is the reason that they are seen as providing *unbiased* estimates of key population parameters such as infant mortality rates. The following additional definitions apply:

- Probability Sample: A sample selected in such a way that the probability of selection could in theory be calculated for every member of the sample.
- Sampling frame: List of entities (or groups of entities) used to select the sample.
- Population Parameter: Quantitative information about the sampling population.
- Sample statistic: Quantitative information about the sample.
- Estimation: The use of a sample statistic to estimate a population parameter.

The first definition imposes a strict criterion for distinguishing between *probability samples* and *non-probability samples*. Because the former are widely regarded as least likely to be influenced by sampling errors, there is a tendency for researchers to proceed as if they analysing probability samples when they are not. For example, those conducting market research surveys or political opinion polls will typically adopt methods of analysis which are appropriate to probability samples even though they have not followed sampling procedures that meet the above criterion.

The nature of the criterion is perhaps most easily understood if we examine the distinction between a *chance* sample and a *random* sample. These words are often used interchangeably by those who have no knowledge of sampling theory but in fact have very different meanings. Consider a large meeting taking place in a conference hall. Researchers selecting a sample of ten individuals might stand by the door and interview the first ten people to emerge. If they were interested in gender differences, they might interview the first five women and the first five men. If they were concerned that those leaving first were likely to have different views than those less anxious to leave, they might sample every tenth person until they had reached their target. No matter what additional strategies the researchers introduce, they cannot claim that they have obtained a probability sample. There is no way that they can calculate the probability of selection for any given member of their sample. On the other hand this would be possible if, for example, the names of those attending were written on pieces of paper and placed in a bag, with ten names being selected from that bag. We would then say that every person attending the meeting had an equal probability of being selected, with that probability being equal to ten divided by the total number attending. This would be an example of random as opposed to chance sampling.

The key additional element is the use of a sampling frame. In the above, we can think of the names as being on a list which is then cut up and put into the bag, which simply serves as a mechanism for selecting names *at random* from the list. More traditionally, the sample would be selected from the list using a computer generated *table of random numbers* or, more recently, using a mobile phone app. In some cases such lists will already exist. For example, there may be a list of all licenced doctors or pharmacists practicing in a city or a list of all rural public health facilities in a given region. In the latter case, if we wished to obtain a sample of all the doctors working in such facilities we could take a random sample of facilities, ask each of the sampled facilities to prepare a list of their doctors and then randomly sample from those lists. This would be an example of *two-stage probability sampling*. Rural household surveys often follow a similar procedure, sampling villages at the first stage and households within the sampled villages at the second.

The most common application of probability sample surveys is *estimation of population parameters* using *sample statistics*. Most often this involves the estimation of population means (for example mean number of antenatal visits) or proportions (proportion experiencing a fever in previous two weeks), using the equivalent measures calculated using data from the sample. As indicated above, if it can be determined that the appropriate procedures have been followed in conducting the survey, such estimates are widely regarded as unbiased. A further advantage of probability sampling, which applies to no other approach, is that it is also possible to estimate mathematically the magnitude of the sampling error, the risk that the sample is unrepresentative of the population, from the survey data. This allows statements as to the confidence with which the value of the estimated parameter can be assumed to lie between lower and upper bounds. For example, it might be asserted the proportion experiencing a fever in the previous two weeks (P) may be assumed, with 95% confidence ([NCCMT](#)), to lie within the bounds $P - k_1 \times se$ and $P + k_2 \times se$. Where k_1 and k_2 are known constants and se is the sampling error calculated from the survey data. Such bounded estimates are not only of value to researchers, for example when they are trying to assess changes over time due to an intervention, but also to decision makers attempting to assess the potential costs and benefits of such an intervention when considering a possible change in policy. Perhaps surprisingly, relatively small probability samples can provide reasonably precise bounded estimates for very large populations – in fact the size of the sampling population does not significantly affect the precision of the estimate as long as it is much greater than the size of the sample.

It should however be noted that probability sampling need not be restricted to studies which seek estimates of population parameters. When sample sizes are very small (perhaps less than ten observations), it may well seem preferable to use some form of purposive sampling,

given that the risk of selecting an obviously inappropriate probability sample may be considerable. However, adopting a probability sampling approach can be very useful in larger scale qualitative studies that wish to argue that their findings can be generalised to a wider population. In this instance the primary advantage is that samples selected using this approach will be less open to challenges on the grounds of bias. As long as researchers can claim, and hopefully demonstrate, that they have followed the standard procedures, even critics should be willing to accept that there has been no attempt to subvert the study findings by the deliberate selection of a biased sample.

Against the above advantages must be set the potential costs associated with the need to obtain or construct a sampling frame. This may even raise difficult conceptual issues in terms of the entities to be included or excluded. For example, how should we define a private healthcare facility? Should we list only accredited facilities, even if the great majority of the population use unqualified providers who work from their own homes? Should we include traditional or religious healers? There will often also be considerable practical difficulties. How can we possibly identify all the unaccredited providers in a given area? How can we possibly construct a sampling frame that will allow us to study the healthcare needs of transient migrant workers who seek employment as day labourers? Existing lists, even of health facilities and communities are often outdated and incomplete. Construction of a new list may be difficult, time consuming and expensive, though researchers have found many ingenious ways to address this issue, often by adopting area sampling approaches ([Myatt et al. 2005](#); [Bennett et al. 1991](#)), a technique which has become much more readily available with the widespread use of mobile phones that can link to the Geographical Positioning System (GPS).

References

- Barnett, I., Yosellina, Sulisty, S., Befani, B., KariSari, K., Sharmin, S. and Dewi, D. (2016). *Mixed-Method Impact Evaluation of a Mobile Phone Application for Nutrition Monitoring in Indonesia*. IDS Evidence Report 200. Institute of Development Studies UK. <http://www.ids.ac.uk/publication/mixed-method-impact-evaluation-of-a-mobile-phone-application-for-nutrition-monitoring-in-indonesia>
- Barahona, C., & Sarah Levy, S. (2002). *How to generate statistics and influence policy using participatory methods in research*. Statistical Services Centre Working Paper: University of Reading. www.reading.ac.uk/ssc/n/TIP%20Archive%20CD/Methodology/How_to_generate_stats_and_influence_policy.pdf
- Bennett, Steve, Tony Woods, Winitha M. Liyanage and Duane L. Smith (1991). A Simplified General Method for Cluster-Sample Surveys of Health in Developing Countries. *World Health Statistics Quarterly* 44:98-106. http://apps.who.int/iris/bitstream/10665/47585/1/WHSQ_1991_44%283%29_98-106_eng.pdf?ua=1
- Bhuiya, A., Hanifi, S. M. A., Urni, F., & Mahmood, S. S. (2009). Three methods to monitor utilization of healthcare services by the poor. *International Journal for Equity in Health* 8.
- Bishai, David, Ligia Paina, Qingfeng Li, David H Peters and Adnan A Hyder (2014). Advancing the application of systems thinking in health: why cure crowds out prevention. *Health Research Policy and Systems* 12:28. www.biomedcentral.com/content/pdf/1478-4505-12-28.pdf
- Boslaugh, Sarah (2007). *Secondary Data Sources for Public Health: A Practical Guide*. Cambridge University Press. www.academia.edu/1630213/An_introduction_to_secondary_data_analysis
- Bowden, A, J A Fox-Rushby, L Nyandieka, and J Wanjau (2002). Methods for pre-testing and piloting survey questions: illustrations from the KENQOL survey of health-related quality of life. *Health Policy and Planning* 17(3):322-330.
- Bowen, Glenn A (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal* 9(2):27-40. www.academia.edu/7849680/Document_Analysis_as_a_Qualitative_Research_Method
- Carvalho, S. and White, H. (1997). *Combining the quantitative and qualitative approaches to poverty measurement and analysis*. World Bank Technical Paper 366. The World Bank, Washington D.C.
- Chambers, R. (2007). *Who Counts? The Quiet Revolution of Participation and Numbers*. Working Paper 296: Institute of Development Studies. <https://www.ids.ac.uk/files/Wp296.pdf>
- Chambers, Robert (2006). *Poverty Unperceived: Traps, Biases and Agenda*. Working Paper 270. Institute of Development Studies. UK. www.ids.ac.uk/files/Wp270.pdf
- Deitchler, Megan, Joseph J Valadez, Kari Egge, Soledad Fernandez and Mary Hennigan. 2007. A field test of three LQAS designs to assess the prevalence of acute malnutrition. *International Journal of Epidemiology* 36:858–864.

ESRC (undated), *ReStore: A Sustainable Web Resources Repository*. Economic and Social Research Council UK. <http://www.restore.ac.uk/lboro/research/software/caqdas.php>

Finlay, Linda (2002). Negotiating the swamp: the opportunity and challenge of reflexivity in research practice. *Qualitative Research* 2:209.
www.uk.sagepub.com/braunandclarke/study/SAGE_Journal_Articles/Ch_2_Finlay.pdf

Galvin, K. T., Petford, N., Ajose, F., & Davies, D. (2011). An exploratory qualitative study on perceptions about mosquito bed nets in the Niger Delta: what are the barriers to sustained use? *Journal of Multidisciplinary Healthcare* 4:73-84. <https://www.dovepress.com/an-exploratory-qualitative-study-on-perceptions-about-mosquito-bed-net-peer-reviewed-article-JMDH>

Greenhalgh. Trisha and Rod Taylor (1997). Papers that go beyond numbers (qualitative research). *BMJ* 315(20):740-743.

Hemingway, Pippa and Nic Brereton (2009). *What is a systematic review?* Hayward Group Ltd. (<http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/syst-review.pdf>)

Holland, J. (Ed.) (2007). *Tools for Institutional, Political, and Social Analysis of Policy Reform A Sourcebook for Development Practitioners*. Washington: World Bank.
http://siteresources.worldbank.org/EXTTOPPSISOU/Resources/1424002-1185304794278/TIPs_Sourcebook_English.pdf

Hoshaw-Woodard, Stacy. 2001. *Description and comparison of the methods of cluster sampling and lot quality assurance sampling to assess immunization coverage*. WHO.
www.who.int/vaccines-documents/DocsPDF01/www592.pdf

Hsieh, Hsiu-Fang and Sarah E. Shannon (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research* 15:1277-1288.
http://www.iisgcp.org/pdf/glssn/Supplemental_Reading_on_Coding_2.pdf

Kanbur, Ravi and Paul Shaffer (2005). *Epistemology, Normative Theory and Poverty Analysis: Implications for Q-Squared in Practice*. Q-Squared Working Paper No. 2. University of Toronto. <http://isiarticles.com/bundles/Article/pre/pdf/21648.pdf>

Kanbur, Ravi (2003). Q-squared? A commentary on qualitative and quantitative poverty appraisal. In *Q-squared: qualitative and quantitative poverty appraisal*, ed. Ravi Kanbur, pp. 1-21. Permanent Black, Delhi. <http://www.arts.cornell.edu/poverty/kanbur/qqz.pdf>

Kuznetsov, Vladimir N, Andrej M Grijbovski, Andrej O Mariandyshev, Eva Johansson, Donald A Enarson and Gunnar A Bjune. Hopelessness as a basis for tuberculosis diagnostic delay in the Arkhangelsk region: a grounded theory study. *BMC Public Health* 13:712.
www.biomedcentral.com/content/pdf/1471-2458-13-712.pdf

Lagarde, M. (2012). How to do (or not to do) . . . Assessing the impact of a policy change with routine longitudinal data. *Health Policy and Planning* 27:76-83.
<http://heapol.oxfordjournals.org/content/early/2011/01/27/heapol.czr004.full.pdf>

Lamba S, Nityananda V (2014). Self-Deceived Individuals Are Better at Deceiving Others. *PLoS ONE* 9(8).
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0104562#pone-0104562-g002>

Leslie, Toby, Dr Nader Mohammed, Hafiz Omar, Haroon Ur Rasheed, Eveline Klinkenberg, Frits van der Vorst (2009). Assessment of the Present Role of the Private Sector in Malaria Diagnosis, Treatment and Control in Afghanistan. *Afghanistan Annual Malaria Journal* 1:89-113. <http://moph.gov.af/Content/Media/Documents/AAMJ2812201015030182.pdf>

Lover, A. A., Sutton, B. A., Asy, A. J., & Wilder-Smith, A. (2011). An exploratory study of treated-bed nets in Timor-Leste: patterns of intended and alternative usage. *Malaria Journal* 10.

Lucas, Henry, Shijun Ding and Gerald Bloom (2009). What do we mean by 'major illness'? The need for new approaches to research on the impact of ill-health on poverty. In *Health and social protection: experiences from Cambodia, China and Lao PDR*. Bruno Meessen, Xiaomei Pei, Bart Criel and Gerald Bloom, eds. ITGPress, Antwerp. [http://www.eldis.org/vfile/upload/1/document/0809/What do we mean by major illness - The need for new approaches to research on the impact of ill-health on poverty.pdf](http://www.eldis.org/vfile/upload/1/document/0809/What%20do%20we%20mean%20by%20major%20illness%20-%20The%20need%20for%20new%20approaches%20to%20research%20on%20the%20impact%20of%20ill-health%20on%20poverty.pdf)

Magnani, Robert, Keith Sabin, Tobi Saidela and Douglas Heckathorn (2005). Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* 19 (suppl 2):S67–S72. http://journals.lww.com/aidsonline/Abstract/2005/05002/Review_of_sampling_hard_to_reach_and_hidden.9.aspx

Marsland, N., Wilson, I., Abeyasekera, S., and Kleih, U. (1998). *A Methodological Framework for Combining Quantitative and Qualitative Survey Methods - Background Paper: Types of Combinations*. Report written for DFID Research project R7033. <http://www.reading.ac.uk/ssc/resources/Docs/QQA/qqa.pdf>

Martineau, Tim, Tolib Mirzoev, Stephen Pearson, Bui Thi Thu Ha, Qian Xu, K V Ramani and Xiaoyun Liu (2015). Coherence between health policy and human resource strategy: lessons from maternal health in Vietnam, India and China. *Health Policy and Planning* 30:111–120.

Mruck, Katja; Breuer, Franz (2003). Subjectivity and Reflexivity in Qualitative Research—The FQS Issues. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 4(2). www.qualitative-research.net/index.php/fqs/article/view/696/1504

Murray, C. J. L. (2007). Towards good practice for health statistics: lessons from the Millennium Development Goal health indicators. *Lancet* 369:862-873.

Myatt, Mark, Teshome Feleke, Kate Sadler, and Steve Collins. 2005. Field trial of a survey method for estimating the coverage of selective feeding programmes. *Bulletin of the World Health Organisation* 83(1):20-26. www.who.int/bulletin/volumes/83/1/20.pdf

Neupane, Sunisha, Willem Odendaal, Irwin Friedman, Waasila Jassat, Helen Schneider and Tanya Doherty (2014). Comparing a paper based monitoring and evaluation system to a mHealth system to support the national community health worker programme, South Africa: an evaluation. *BMC Medical Informatics and Decision Making* 14:69. <http://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-14-69>

Nyandieka, Lilian N, Annabel Bowden, John Wanjau, and Julia A Fox-Rushby (2002). Managing a household survey: a practical example from the KENQOL survey. *Health Policy and Planning* 17(2):207-212

Nzinga, Jacintha, Patrick Mbindyo, Lairumbi Mbaabu, Ann Warira and Mike English. (2009). Documenting the experiences of health workers expected to implement guidelines during an intervention study in Kenyan hospitals. *Implementation Science* 4:44.
<http://www.implementationscience.com/content/4/1/44>

Palinkas, Lawrence A, Sarah M. Horwitz, Carla A. Green, Jennifer P. Wisdom, Naihua Duan and Kimberly Hoagwood (2013). Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Administration and Policy in Mental Health* 42(5).
https://www.researchgate.net/publication/258315317_Purposeful_Sampling_for_Qualitative_Data_Collection_and_Analysis_in_Mixed_Method_Implementation_Research

Palinkas, L. A., Aarons, G. A., Horwitz, S., Chamberlain, P., Hurlburt, M., & Landsverk, J. (2011). Mixed Method Designs in Implementation Research. *Administration and Policy in Mental Health and Mental Health Services Research* 38:44-53.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3025112/>

Peters, D. H., Ahmed, A., Noor, A., Singh, L. P., Kakar, F. K., Hansena, P. M., et al. (2007). A balanced scorecard for health services in Afghanistan. *Bulletin of the World Health Organization* 85(2):146-151. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2636277/>

Qualquant (2016). *Qualquant: Use the Right Tool*. National Science Foundation.
<http://qualquant.org/>

Rifkin, Susan B (1996). Rapid Rural Appraisal: Its use and value for health planners and managers. *Public Administration* 74:509-526.
http://www.academia.edu/3664632/RAPID_RURAL_APPRAISAL_ITS_USE_AND_VALUE_FOR_HEALTH_PLANNERS_AND_MANAGERS

Rog, Morgan, Bonnielin Swenor, Luis C Cajas-Monson, Wilson Mchiwe, Steven Kiboko, Harran Mkocho and Sheila West (2011). A cross-sectional survey of water and clean faces in trachoma endemic communities in Tanzania. *BMC Public Health* 11:495.
www.biomedcentral.com/content/pdf/1471-2458-11-495.pdf

Rowe, A. K. (2009). Potential of Integrated Continuous Surveys and Quality Management to Support Monitoring, Evaluation, and the Scale-Up of Health Interventions in Developing Countries. *American Journal of Tropical Medicine and Hygiene*, 80(6):971-979.
<http://www.ajtmh.org/content/80/6/971.long>

Scott, Jennifer, Sarah Averbach, Anna Merport Modest, Michele Hacker, Sarah Cornish, Danielle Spencer, Maureen Murphy and Parveen Parmar (2013). An assessment of attitudes toward gender inequitable sexual and reproductive health norms in South Sudan: a community-based participatory research approach. *Conflict and Health* 7(24).
www.biomedcentral.com/content/pdf/1752-1505-7-24.pdf

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education* 5(4), 541–542.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3886444/pdf/i1949-8357-5-4-541.pdf>

Uganda (2012). *mTrac: Information for Better Health*. Ugandan Ministry of Health.
<http://www.mtrac.ug/video-tags/uganda>

Wallerstein, N. B., & Duran, B. (2006). Using Community-Based Participatory Research to Address Health Disparities. *Health Promotion Practice* 7:312-323.

Walshe C, Ewing G, Griffiths J. (2012). Using observation as a data collection method to help understand patient and professional roles and actions in palliative care settings. *Palliative Medicine* 26(8):1048-1054.

https://www.researchgate.net/publication/51898162_Using_observation_as_a_data_collection_method_to_help_understand_patient_and_professional_roles_and_actions_in_palliative_care_settings

World Bank (2003). *Measuring Living Standards: Household Consumption and Wealth Indices*. Quantitative Techniques for Health Equity Analysis—Technical Note #4. World Bank. http://siteresources.worldbank.org/INTPAH/Resources/Publications/Quantitative-Techniques/health_eq_tn04.pdf