IMPLEMENTATION RESEARCH TOOLKIT

Research Methods and Data Management

Alison Krentel and Riris Andono Ahmad

TDR For research on diseases of poverty
UNICEF • UNDP • World Bank • WHO

World Health Organization

Research Methods and Data Management

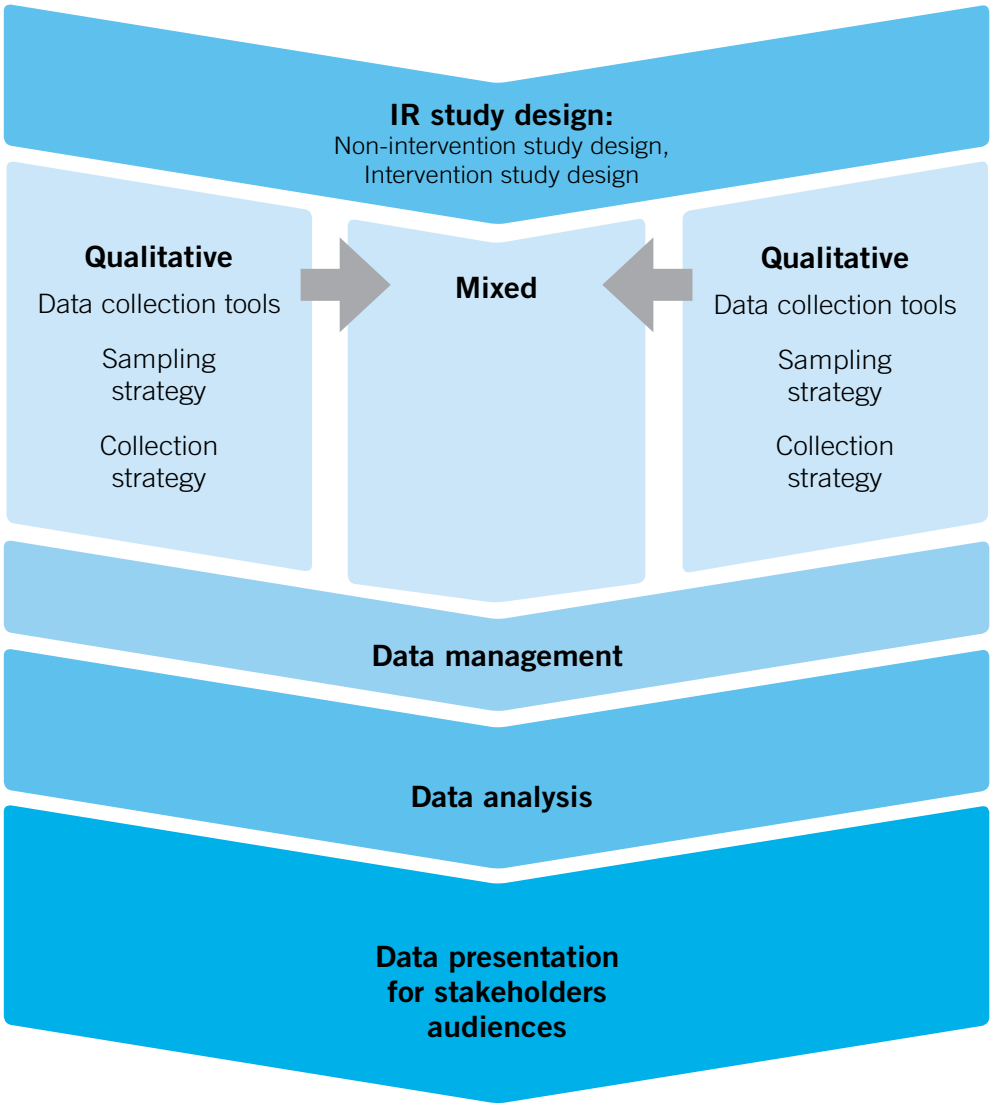
The purpose of this module is to describe the fundamentals of implementation research (IR) methodologies including study design, data collection methods, data analysis, presentation and interpretation of IR findings with the objective of enhancing their uptake and use by target audiences.

This module covers the concepts of: a. Research approach; b. Research designs c. Data collection methods; d. Data analysis; and, e. Data presentation. The overall pathway followed by the module is outlined in Figure 1. These concepts are illustrated through examples of completed IR projects. The module describes IR study design and research methodologies at a general level, and does not replace materials that specialize in research methodologies. For those interested in further information, many useful resources are available to supplement this module. Some of these key materials are signposted in the module.

After reviewing of this module, you should be able to:

- Describe the designs commonly used in IR projects.
- Identify the strengths of quantitative, qualitative and mixed methods approaches in IR data collection.
- Select the appropriate data collection approaches and tools for your IR project(s).
- Describe the sampling processes used for both quantitative and qualitative research tools.
- Highlight relevant ethical issues in data collection.
- Describe appropriate data analysis processes for both quantitative and qualitative data and for mixed methods approaches.
- Describe various formats of data presentation.

Figure 1: Outline of the research methods and data management module content



Study design for IR projects

Similar to other types of research, study designs used in IR can be interventional or observational. In an interventional research design, the researcher influences objects or situations and then measures the outcome of these manipulations. In an observational study design, the researcher observes and analyses researchable objects or situations without intervening. These non-intervention studies can be exploratory, descriptive and comparative (analytical) studies, while intervention studies can be experimental studies, quasi-experimental, before and after, cohort studies or randomized controlled trials. Below each of these study designs is briefly explained.

Non-intervention studies

Descriptive

Descriptive studies are used when you want to describe the implementation of health-related interventions and any problems or barriers within that context. Depending on your familiarity with the subject of the study, different study designs can be used to answer your research questions. If the subject is new and no prior knowledge exists, you can conduct an exploratory study using qualitative methods. The results from this qualitative study can then be used to develop subsequent research, using quantitative methods, to measure to what extent these problems occur. A descriptive study can also begin with quantitative methods (i.e. survey) to quantify the intervention barrier followed by qualitative methods to describe the context where the implementation problems exist. More details about methods are provided later in this module.

Most surveys used within descriptive studies use a cross-sectional design, a relatively simple and inexpensive design that is useful for investigating contexts with many variables to take into consideration. Data from repeated cross-sectional surveys provide useful indicators of trends, given that they have representative, independent and random samples as well as standardized definitions. Each survey should have a clear purpose. Valid surveys need well-designed questionnaires, an appropriate sample of sufficient size, a scientific sampling method and a good response rate.

Analytical

Analytical studies investigate and establish a causal relationship between the independent and dependent variables under study. Traditionally, cohort or case control study designs are used for non-intervention studies, in order to establish likely causal relationship. However, cohort study design is more commonly used for IR.

In a cohort study, the researcher recruits a group of people – who are free from disease, for example – and who are classified into sub-groups according to exposure status. Sub-groups are then followed up to see subsequent development of specific outcomes, such as specific health conditions. Cohort design can be used to measure typical IR-related outcomes over time (i.e. acceptability, adoption, appropriateness, feasibility, fidelity of interventions, implementation costs and cost-effectiveness, determinants of coverage and sustainability/maintenance). This design produces high quality, individual level data, enabling researchers to examine if better implementation outcomes are associated with exposures at the individual level, including the timing and direction of any effects.[1] A cohort design can also be used to assess the uptake and retention of patients in specific services, particularly for chronic illnesses such as the continuation of antiretroviral (ARV) therapy among people living with HIV or treatment adherence among people with multidrug-resistant TB (MDR-TB).

Analytical studies can have a cross-sectional study design. However, such study designs cannot establish causal relationships between independent and dependent variables, as the measurement of both variables is conducted simultaneously.

Intervention/Experimental studies

Experimental research is the only type of research that can establish cause and effect. The randomized controlled trial (RCT) study in particular, is known for establishing causal relationships due to its ability to control for confounders, and for ensuring that the only difference between different study arms is the intervention in question. In an experimental study, the researcher is interested in the effect of an independent variable (also known as the experimental or treatment variable) on one or more dependent variables (also known as the criterion or outcome variables). In effect, the researcher changes the independent variable and measures the dependent variable(s). There are usually two groups of subjects in experimental research: The experimental group, which receives an intervention (e.g. taught by a new teaching method, receives a new drug), and the control group, which receives no intervention (e.g. continues to be taught by the old method, receives a placebo). Sometimes, a comparison group will also be used in addition, or instead of a control group. The comparison group receives a different treatment from the experimental group. The control and/or comparison groups are critical in experimental research as they allow the researcher to determine whether the intervention had an effect or whether one intervention was more effective than another.

The following are different types of experimental studies.

Randomized control trial (RCT)

This is the 'gold standard' for efficacy studies in clinical trials. IR, on the other hand, focuses more on generalizability of results to different settings rather than the efficacy of a given intervention. For this reason, RCT is not a commonly used study design in IR. In RCT, the subjects should be randomly assigned to the treatment and control groups to ensure that all groups are homogenous before an intervention is applied, and that the intervention is the only difference between the groups. Randomization is used to ensure internal validity.

Quasi-experimental

This study design is similar to RCT but lacks the key characteristic of random assignment. The design is frequently used when it is not logistically feasible or ethical to conduct an RCT. The assignment to the treatment group uses criteria other than randomization e.g. matching by individual or matching by group of sociodemographic factors. Quasi-experimental design is suitable for IR by virtue of the fact that the design allows for real-life factors – such as cost, feasibility and political concerns – to be integral factors in the study.

Case study 1 — Community-directed education intervention: Quasi-experimental study in the malaria endemic areas of Sarpang District, Bhutan

Background: Malaria remains a public health problem in spite of the efficacious interventions such as long lasting insecticidal nets (LLINs) and artemisinin-based combination therapy. The Kingdom of Bhutan has achieved notable success in the prevention and control of malaria, and the country is moving towards the malaria elimination phase. For example, in 2011 only 194 malaria cases were registered compared to 5935 cases in 2000. To attain the elimination goal, current efforts need to be reinforced by community-directed interventions in order to empower the community to enhance their health-seeking and other preventive behaviours. Community-directed interventions have proved to be useful in the prevention and control of infectious diseases such as onchocerciasis. This study was conducted to elucidate the effectiveness of the community-directed educational intervention on malaria prevention and control in the malaria-endemic areas of Sarpang District, Bhutan. A quasi-experimental study design was adopted, using both qualitative and quantitative methods (Figure). The study district (Sarpang) was purposively selected from seven malaria endemic districts. The study basic health units (BHU) were Umling and Chuzerganga (intervention arm), and Jigmeling (control arm). These were purposively selected. These BHUs were similar in population size and other relevant contextual criteria. Baseline data was collected during the formative phase using in-depth interviews and focus group discussions (FGDs), household surveys and document/data review. The training tool was developed in collaboration with the BHU staff. Health workers and community action groups (CAGs) were trained on malaria transmission, care and in the use of LLINs, proper use of indoor residual spraying, control of mosquito breeding sites, and the importance of early diagnosis and treatment. The intervention package was implemented in addition to the regular programme activities in the intervention BHUs while in the control BHU, only regular programme activities were conducted. The effectiveness of the intervention was evaluated using household survey, FGDs, in-depth interviews and review meetings. Comparison of the pre- and post-intervention group, showed a significant improvement in knowledge, attitude and practice of the community intervention arm as compared to the control arm.

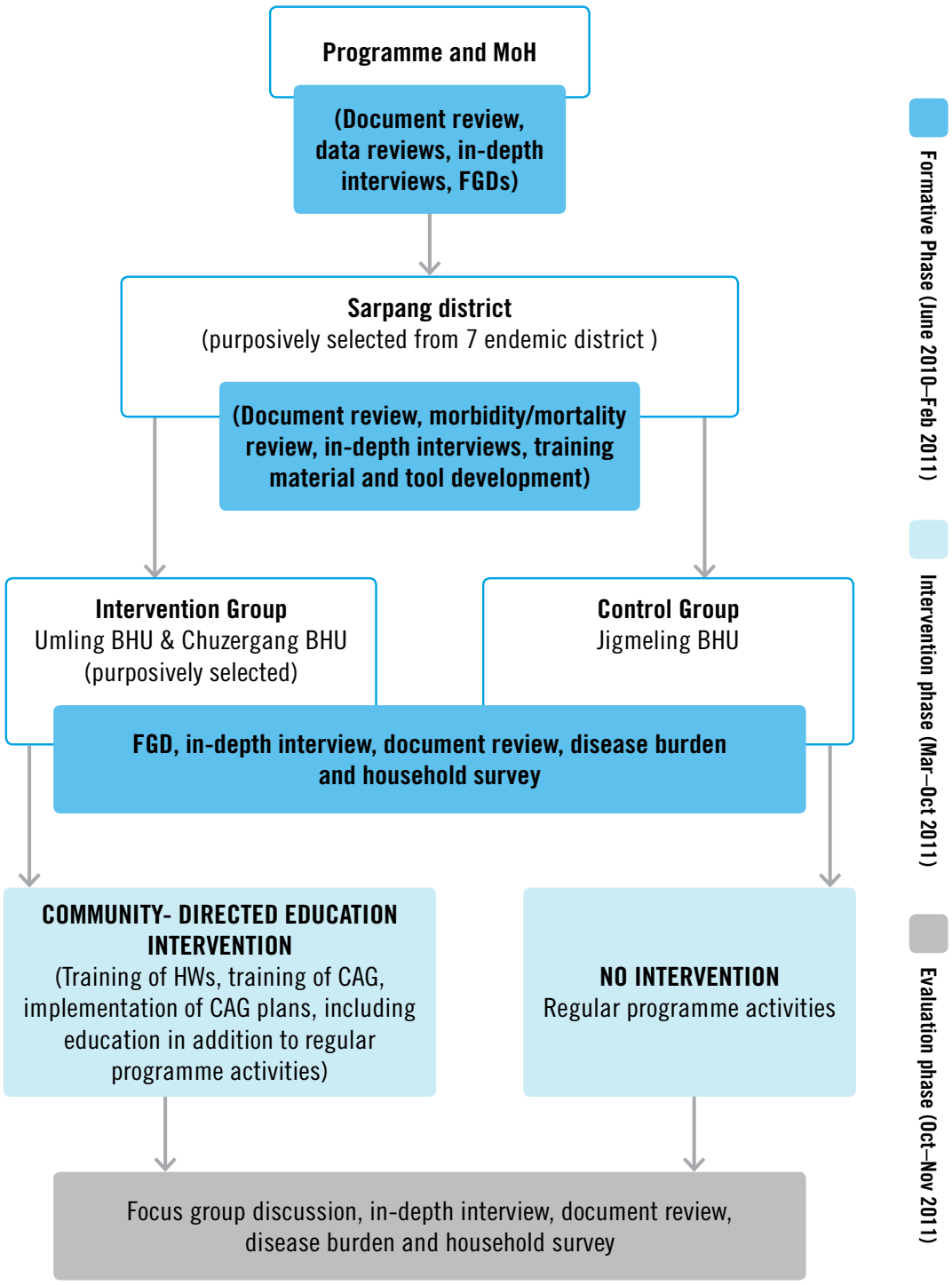
Conclusion: The quasi-experimental study design was able to elucidate the effectiveness of the community-directed educational intervention on malaria prevention and control in malaria-endemic areas.

Lessons: Quasi-experimental study design is an appropriate approach to establish the impact of a given intervention. However, to ensure reliable results, the intervention and control arms should be as similar as possible in terms of population characteristics and context. The only distinguishing variable should be the intervention in question.

Source: Tobgay T. et al. Community-directed educational intervention for malaria elimination in Bhutan: quasi-experimental study in malaria endemic areas of Sarpang district. Malaria Journal. 2013; 12(1):1.

Case study 1 Community-directed education intervention: Quasi-experimental study in the malaria endemic areas of Sarpang District, Bhutan

Figure. Schematic diagram of research activities



Pragmatic trials

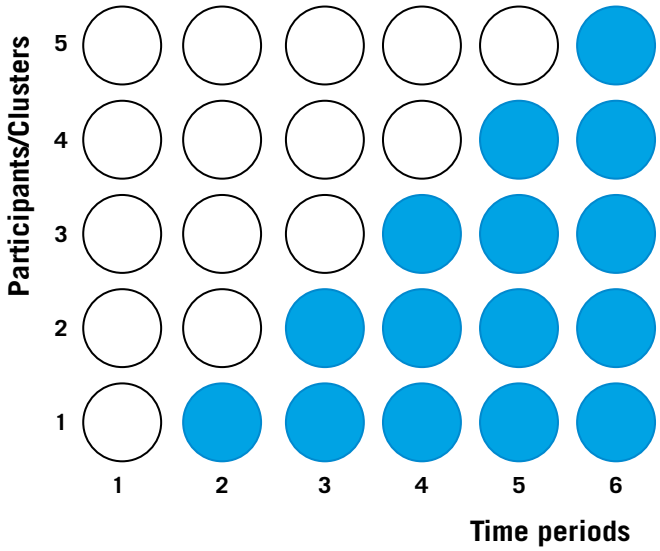
Pragmatic trials evaluate the effects of health service interventions under the human, financial and logistic constraints of typical, real-world situations. The aim of this study design is to measure effectiveness rather than efficacy.[2,3] *Contrary to the efficacy study, where participants are recruited from a homogeneous sub population (e.g. gender, age, ethnicity etc.) and randomly assigned to arms of the study, the design of the pragmatic trial presents higher degrees of variation in study participants. Participants are selected from within a real clinical or population setting to be representative of the population. To improve validity of pragmatic trials, randomization is conducted at the facility level (cluster randomization) rather than at the individual level.*

As the effectiveness of a treatment is influenced by the extent to which an intervention is acceptable to patients, a pragmatic trial not only measures treatment outcome but also evaluates measures designed to increase effectiveness. For example, while patients in both control and interventions arms receive identical treatment, the intervention group receives additional interventions to increase treatment acceptance or adherence (e.g. counselling, home visit or mobile reminder).

Stepped-wedge cluster randomized trial

This is a variant of cluster randomized trial design in which the selected clusters are randomly allocated at the time point when they receive the intervention. In this design, all clusters are assigned in both intervention and control arms (Figure 2). The cluster can be geographical areas, clinics or other types of facilities.[4] The advantage of this design is that each cluster can serve as a control for itself. The design also addresses any ethical issues where, for example, the randomization of patients to an intervention believed to be inferior or the withdrawal of an intervention believed to be superior, is considered unethical.

Figure 2: Diagrammatic illustration of stepped-wedge study design



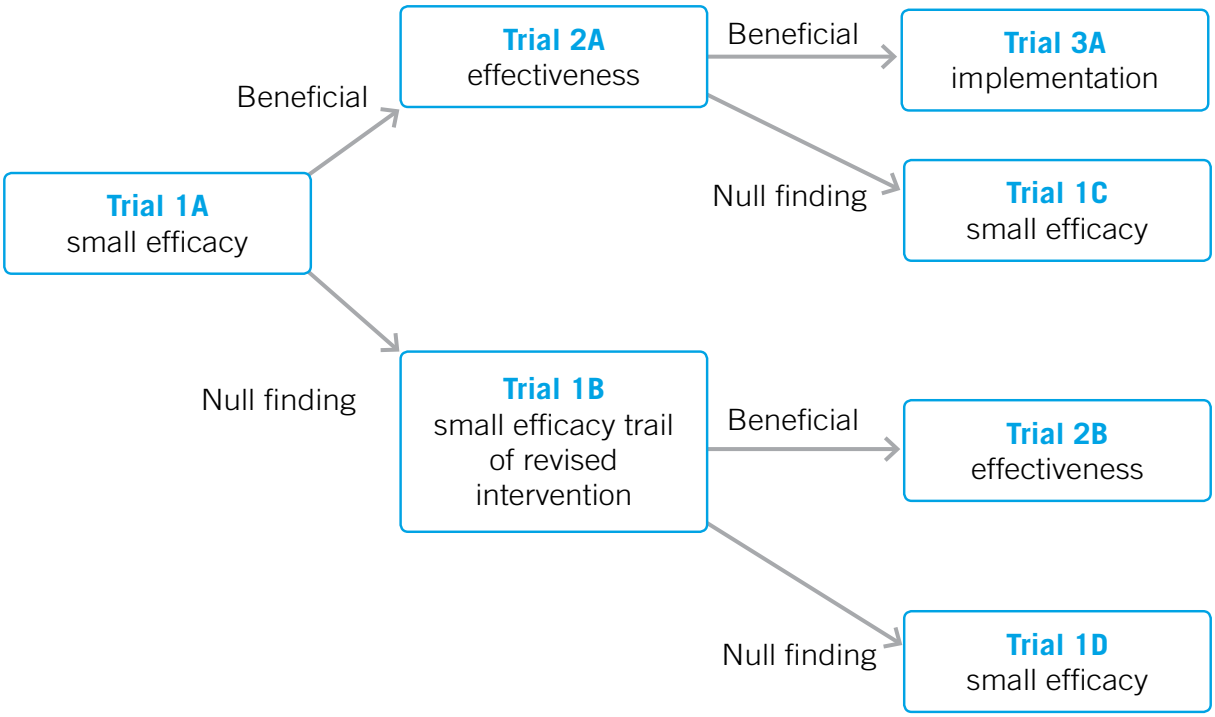
Shaded cells represent intervention periods
Blank cells represent control periods
Each cell represents a data collection point

Adaptive design trial

This variant of experimental design anticipates intentional changes to the trial plan. It is characterized by the idea that collected data will be used to make decisions regarding the trial while it is ongoing. The objective of an adaptive design is to maintain the validity of the study while maintaining the flexibility to identify optimal treatment. Researchers can modify both trial and statistical procedures. Adaptation of statistical procedures can include the sample size, randomization, study design, data monitoring and the analysis plan. Typical trial procedures can include eligibility criteria, enrolment design, study dose, treatment duration including early stopping, follow-up design, study endpoints or laboratory/diagnostic procedures.[5,6] Figure 3 provides a description of the adaptive sequencing of trials. The changes in the subsequent trials are conditional on the outcome of the previous trial and/or certain parameter values.

From an ethical standpoint, adaptive design is advantageous since the methods enable the researcher to detect outcome differences early and allow for changes to the intervention concurrently during the trial. However, this flexibility limits the measurement of treatment effect for each group.

Figure 3: A typical adaptive sequencing trial design[5]



The choice of the study design depends on:

- state of knowledge about the problem;
- nature of the problem and its environment;
- type of objectives;
- available resources;
- ingenuity and creativity of research team.

Table 1: Factors dictating the most appropriate choice of study design

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

Example

A mixed methods protocol to evaluate the effect and cost-effectiveness of an integrated electronic diagnosis approach (IeDA) for the management of childhood illnesses at primary health facilities in Burkina Faso.[7]

Background: Burkina Faso introduced the Integrated Management of Childhood Illnesses (IMCI) strategy in 2003. However, an evaluation conducted in 2013 found that only 28% of children were assessed for three danger signs as recommended by IMCI, and only 15% of children were correctly classified. About 30% of children were correctly prescribed with an antibiotic for suspected pneumonia or oral rehydration salts (ORS) for diarrhoea, and 40% were correctly referred. Recent advances in information and communication technologies (ICT) and the use of electronic clinical protocols hold the potential to transform health care delivery in low-income countries. However, no evidence is available on the effect of ICT on adherence to IMCI. A mixed methods study that aims to measure the effect of the IeDA innovation (an electronic IMCI protocol provided to nurses) is planned in two regions of Burkina Faso.

The study focuses on three key questions: (i) How does the effectiveness and the cost of the intervention vary by type of health worker and type of health centre? (ii) What is the impact of changes in the content, coverage and quality of the IeDA intervention on adherence and cost-effectiveness? (iii) What mechanisms of change (including costs) might explain the relationship between the IeDA intervention and adherence? In order to answer these questions, the study combines the following mixed methods: stepped-wedge trial, a realistic evaluation and an economic study in order to capture the effect of the innovation after its introduction on the level of adherence, cost and acceptability.

Table 2 provides a summary of study designs reviewed in this module, according to the stage of the intervention under study. This table has been adapted from Bowen et al (2009)[3] to reflect IR questions.

REFLECTION ACTIVITY

Considering your team's IR question, which study design is most appropriate for your work?

Table 2: Sample study design: Phases of intervention development by area of focus (adapted from Bowen at al 2009)[3]

	Intervention development phase		
	Can it work?	Does it work	Will it work
Area of focus	Is there some evidence that intervention X might work?	Is there some evidence that X might be efficacious under actual conditions, compared to whatever other practices might already be in place instead?	Will it be effective in real life contexts, settings and cultures/ populations that might adopt the intervention as practice?
Acceptability	Focus groups with target population participants to understand how this intervention would fit in with their daily-life activities.	An RCT to compare the satisfaction of the intervention group to that of a control group that does not receive the intervention.	A population based survey before, during and after implementation of a policy intervention. A cohort study comparing the actual use of facilities with and without the intervention over time.
Demand	Survey to determine whether people in the target population would use the intervention to guide their behavioural choices.	Pre–post intervention survey design to compare frequency of use and patterns of use across different populations.	Post-only design with multiple surveys over time to test reactions to the intervention in a new population.
Implementation	Pre–post design to evaluate whether the intervention can be deployed in any clinical or community context, using focus groups as the method of evaluation.	Different types of trial designs (pragmatic, stepped-wedge, adaptive design) to test whether the intervention can be implemented in different clinical settings or community contexts: Using surveys and observations to compare practices and outcomes before and after interventions; using focus groups and in-depth interviews to further explain what works during the intervention processes.	

	Intervention development phase		
	Can it work?	Does it work	Will it work
Practicality/cost	Small-scale demonstration study to examine the predicted cost, burden, and benefit because of appropriate and intensity, frequency, duration of the intervention, using key-informant interviews to gather data.	Cost-effectiveness analysis combined with in-depth interviews with community leaders or other stakeholders to determine how easily the intervention was used in the health system.	Cost analyses with in-depth interviews with providers to identify potential areas of concern during implementation.
Adaptation	FGD, key informant interviews to drive the adaption of the intervention. Quasi-experimental design using pre- and post-surveys to examine the effects of the adapted intervention in communities.	Experiment using adaptive design to examine whether an effective intervention continues to show evidence of efficacy once modified and implemented in a practice context.	Small-scale experiment testing appropriate intensity, frequency, and duration of the modified intervention, or intervention for the new target population.
Integration	Pre–post design to observe the extent to which people in the target setting are using the new intervention activities and with what costs and benefits to their other responsibilities.	Prospective longitudinal study to identify the sustainability of a recently tested package of intervention activities.	Annual monitoring of important systems to measure outcomes across years.
Expansion	Quasi-experimental, pre–post design using interviews with key informants to determine how well an expanded version of an intervention is perceived to work after implementation.	Uncontrolled pre–post study to test new, enhanced the version of a previously tested intervention.	Continued monitoring to identify any decay of intervention effects after implementation.

Selecting the research methods for your IR project

IR can use quantitative, qualitative research methods or a combination of the two. Quantitative and qualitative techniques can be said to offer a trade-off, between breadth and depth, and between generalizability and targeting to specific populations. Before you choose the most appropriate methods and design for your IR study, it is important to understand some of the principles behind both qualitative and quantitative research methods. Table 3 provides a summary of the characteristics of both methodologies.

Table 3: Summary characteristics of quantitative and qualitative research

	Quantitative	Qualitative
Purpose	Explanatory/causal.	Exploratory, descriptive.
Perspective regarding reality	Naturalist, positivist perspective common in natural science.	Interpretive perspective common in social science.
Research tradition	Naturalist, positivist perspective common in natural science.	Interpretive perspective common in social science.
Sample and sample size	Sample size is large using mostly probability sampling strategy.	Sample size is small using purposive sampling strategy.
Sample is representative of	Population being studied. Sample represents the existing variation in the population.	Phenomenon being studied. Sample with rich information regarding the phenomenon.
Methods	Structured/Semi-structured Surveys or observations.	In-depth interviews, focus groups, observations, etc.
Gathering Data	More efficient, tests specific hypotheses.	Time-consuming process; often real-world environments.
Administration	Researcher uses tools to gather data (requires less training).	Researcher is the instrument for gathering data (requires training).
Types of Questions	Closed, yes/no responses.	Probing, open-ended.
Forms of data	Numbers and statistics.	Words, stories and pictures.
Types of Analyses	Statistical, summarize results using numbers.	Interpretive, establishes themes or patterns.
Interpretation of study	Generalization of findings.	Context-specific findings.

The main difference between quantitative and qualitative approaches comes from the research traditions and the philosophy of how researchers in each research tradition see the nature of the world. Researchers from the natural science tradition developed quantitative research methods, where the philosophical

approach in creating knowledge is through positivism. Knowledge creation is characterized by empirical observation, the testing of theories and development of universal laws. Qualitative research methods on the other hand came from a social research tradition, where social phenomena (reality) are considered to be constructed through interaction among individuals in the community. A shared understanding or interpretation of its nature creates the meaning of the phenomena. These meanings are constructed within the context (for example cultural beliefs) where the phenomena exist. Therefore, the nature of reality is subjective and particular to the interpretation given to them. Click on each heading for details.

Strengths and limitations

Implementation research can employ both quantitative and qualitative methods. However, researchers need to be aware that each approach has its own strengths and limitations. Table 4 summarizes the strengths and limitations of quantitative and qualitative methods. In general, the strengths of one method can be seen as the weakness of the other. Therefore, combining both quantitative and qualitative methods can improve the strength of an IR project.

Table 4: Strengths and limitations of quantitative and qualitative research methods

	Strengths	Limitations
Quantitative methods	Provide wide coverage of a range of situations.	The methods can be inflexible and artificial (e.g. RCT).
	Can be fast and economical.	Not effective for understanding processes or the significance that people attach to actions.
	Statistics from large samples can provide considerable relevance for policy decisions.	They are not very helpful in generating theories.
Qualitative methods	Data gathering methods seen more as natural than artificial.	Data collection can be tedious and require more resources.
	Ability to look at change processes over time.	Analysis and interpretation of data may be more difficult.
	Ability to understand people's meaning.	Harder to control the pace, progress and end-points of the research process.
	Ability to adjust to new issues and ideas as they emerge.	Policy makers may give low credibility to result from qualitative approach.
	Captures a wide range of relevant themes through purposive sampling.	Lack of external validity/ generalizability.

(Adapted from Amaratunga D et al 2002)[8]

Assessing the quality of quantitative and qualitative studies

Quantitative and qualitative studies have fundamentally different criteria to assess the rigor of the study due to the paradigm used and the nature of the methods. The criteria are analogous but not interchangeable. Each has its own appropriate and no less rigorous standards. There are four analogous criteria that are comparable to assess the quality of both quantitative and qualitative studies i.e. truth-value, applicability, consistency and neutrality.[9]

Table 5: Comparable criteria in quantitative/qualitative methods, and questions addressed

Issues	Quantitative	Qualitative	Question
Truth value	Validity	Credibility	Do we measure what we are supposed to measure?
Applicability	Generalizability	Transferability	Could the research be repeated in different subjects or contexts?
Consistency	Reliability	Dependability	Could the research be repeated with the same result?
Neutrality	Objectivity	Conformability	How far does the personal interest of the researcher influence the result?

(Adapted from Krefting L, 1991)[9]

Truth Value

The quality of study depends on how effective the researcher is in measuring the concept being studied. In a quantitative method, this means the extent to which the measurement reaches the concept it intends to measure. Validity assumes correct operational measures for the concepts being studied. The validity of the study can be improved by ensuring that there are no selection and measurement biases (through the use of standardized instruments and procedures).

Credibility is the corresponding criterion of validity in qualitative research. It focuses on whether the investigator has achieved in seeing the truth according to the informants' eyes and in understanding of the context in which the research is conducted. Credibility can be accomplished by triangulation of informants, data collection methods or analysis; prolonged engagement with people; continual observation in the field; the utilization of peer researchers; researcher reflexivity; and participant checks, validation or co-analysis.

The strength of qualitative research lies in validity (closeness to the truth). Good qualitative research, using a selection of data collection methods, should touch the core of what is going on rather than just skimming the surface.[10]

Applicability

Applicability refers to how we can apply the research findings to a wider population, beyond that under study. In a quantitative study, it is known as the external validity or generalizability of the findings. Generalizability is a goal of quantitative studies. It is accomplished by selecting large enough random samples to minimize the probability of error and to be able to statistically represent the population from which the samples are drawn.

Transferability is the qualitative analogue to the concept of generalizability. Transferability is the degree to which the results of a study can be applied to other contexts and settings or other groups. It also means the level to which the audience will be able to generalize the study findings into his or her own context. The audience that requires to transfer the findings into different situations are responsible to assess the transferability of the study, rather than the researchers of the original study. Transferability can be achieved when the researcher gives adequate information about researchers' backgrounds, any prior knowledge and possible bias, as well as the research context, processes, members and researcher-participant connections so that the reader can decide to what extent the findings of the study is transferable to their own contexts.

Consistency

Consistency refers to whether the study conclusions would be similar if replicated with the same subject matter or in a similar context at a different time. In a quantitative study consistency refers to the reliability of measurement. When we measure variables under study, all measurements involve some degree of error. When the amount of error is low, the reliability of the measurement is high.

Consistency is defined as dependability in qualitative research. Dependability refers to the way researchers ensure that the study is conducted consistently across time, researchers and analysis techniques and that the procedures of the study are explicit and repeatable. This can be achieved by an audit trail, which is a process of keeping a detailed chronology and description of the research activities, including: an explanation of the choices and justification for the different research designs, data collection and analysis, emerging themes, and an analytic memo.

In summary, 'reliability' in a quantitative study is the repeatability and independence of findings from the specific researchers generating those findings. While in qualitative research, reliability implies that given the data collected, the results are dependable and consistent.

Neutrality

Neutrality implies that the researcher maintains objectivity, minimizing any possible bias due to the researcher's values or interests. In a quantitative study, objectivity can be achieved by avoiding selection bias (by randomization) and measurement bias (by standardized instruments, procedures and masking participants' status during measurement). In a qualitative study, however, these same strategies would be counterproductive. To be able to capture reality as accurately as possible according to participants' perspectives and experiences, the researcher needs to be inseparable from the study participants. Furthermore, the researcher can act as an instrument during data collection. As a result, the researcher cannot be fully objective. Objectivity (or conformability) is therefore a way of knowing that the researchers have maintained the distinction between their own personal values and those of the study participants. The readers should be able to see that the integrity of the study findings is based on the data, and not the researchers' beliefs or biases. Conformity can be achieved through the use of a reflexive diary.

Mixed methods: Combining quantitative and qualitative methods

After understanding the strengths and weaknesses of both quantitative and qualitative approaches to research, it is possible that your IR team will consider using a combination of these two approaches. In fact, many IR projects use mixed methods to provide a better understanding of the problem than either a quantitative or qualitative research approach can do alone. Before making this decision, it is important to review why you may want to combine the two kinds of research approaches. Table 6 (adapted from Bryman 2006[11] and Greene et al 1989[12]) can help to guide the decision-making process.

Table 6: A Guide to the decision-making process on whether to use mixed research methods

Question?	Explanation regarding your research design	Terminology
Do you wish to confirm the accuracy of your findings?	You want to see the convergence of results from different methods in order to confirm what you have found with one method is valid through the use of another method.	Using two research approaches to ask similar questions is called **triangulation**.
Do you want to elaborate the results from one approach with another approach?	It is important to elaborate, enhance, illustrate or clarify the results from one method with the results from another.	Using one research approach to further elaborate the results from a different approach is called **complementarity**.
Do you want to use the results from one research method to inform the development of additional data collection?	When results from one method help to develop the subsequent data collection method, by informing the sample as well as measurement decisions (e.g. questions to ask, scales to use).	Using one research method to sequentially inform the other is called **development**.
Have you have discovered or expect to discover something new or contradictory with one method, and you want to try and understand that new or contradictory thing better?	When one data collection method reveals results that are unexpected or contradictory to what is understood to be true, re-asking the same questions using a different methodological approach can lend more understanding.	Using one research method to further explore contradictory results from another research method is called **initiation**.
Do you want to maximize your understanding?	Extending the range of enquiry by using different methods for different components of enquiry.	This combination of approaches, called **expansion**, will allow you to extend your data collection methods more widely.

After considering how a mixed methods approach might contribute to your research, you will also need to justify the sequence and weight given to the two approaches. The four most common types of mixed methods research are: sequential explanatory; sequential exploratory; concurrent triangulation; and concurrent embedded (Table 7).

Table 7: Main mixed methods research approaches

Sequence	Description
Sequential explanatory	Collection and analysis of quantitative data in the first phase is followed by the collection and analysis of qualitative data that builds on the results of the first phase. Weight is typically given to the quantitative data. Mixing of the data occurs when the initial quantitative results are used to inform the secondary qualitative data collection. It can be especially useful when unexpected results arise from a quantitative study. The straightforward nature of the design is its strength and so it is easy to implement. The main weakness of the design is the time required to implement since it falls into two phases.
Sequential exploratory	Collection and analysis of qualitative data in the first phase is followed by the collection and analysis of quantitative data that builds on the results of the first phase. Weight is typically given to the qualitative data. This design tends to be used when the primary purpose is to explore a phenomenon (e.g. testing elements of an emergent theory or determining the distribution of a phenomenon in a given population). It is easy to implement but requires substantial time for data collection.
Concurrent triangulation	Quantitative and qualitative data are collected simultaneously and then the two datasets are compared to see if there is convergence, differences, or some combination of the two. Ideally, the weight given to the quantitative and qualitative findings is equal but in reality, more weight may be given to one methodology over another. Concurrent triangulation is one of the most popular types of mixed methods design. It can, however, be difficult to compare results, particularly if discrepancies arise. It also requires great effort and expertise on the part of the researcher to adequately study a phenomenon using two methods.
Concurrent embedded	Quantitative and qualitative data are collected simultaneously but there is a primary method that guides the approach. Either quantitative or qualitative data will be used to provide a supportive or supplementary role based on the primary data type. The researcher is able to collect two types of data during a single research phase. Often an embedded design is used to answer different research questions with a study.

When designing mixed methods research, the IR team will need to consider the following elements when planning data collection and analysis:

- Timing: Will quantitative and qualitative methods be used simultaneously (concurrent designs) or in two distinct phases (sequential designs)?
- Weighting: How much emphasis will be put on the quantitative or qualitative methods? Will they be weighted equally?
- Mixing: Data analysis needs to be matched to the design of the study. For example, in a concurrent design, one way of mixing the data is to provide a discussion about the emerging themes from the data and how they support or refute the statistical analysis. Another approach could be to combine the quantitative and qualitative data to arrive at new variables or new themes (Creswell 2009). In a sequential design, for example, a researcher might collect and analyse quantitative data in the first phase of the study and may then select some extreme cases to follow-up in a qualitative phase.
- Visual diagrams: An important mixed methods tool that incorporates a notation system and a flow chart of the research process.

If your research team decides to use mixed methods in your study, you will need to describe why you chose this approach, as discussed in Proposal Development Module.

SEE DEVELOPING AN IR PROPOSAL

Case study 2 **Use of mixed methods to explain malaria persistence in remote Central Viet Nam**

Background: Malaria remains a major global threat despite the availability of efficacious tools. Its effective control requires consistent action from both *health care systems* and *community* and an understanding of features that precipitate risk. The Viet Nam National Malaria Control Programme (NMCP) introduced in 1991 has controlled malaria through the provision of free anti-malarial drugs, impregnated bed-nets, bi-annual home insecticide spraying and early diagnosis and treatment. Overall, the number of clinical cases declined from 1.2 million and 4646 recorded deaths in 1991 to 185 529 clinical cases and 50 deaths in 2002. However, over 90% of severe cases and deaths occurred in mountainous, forested and largely ethnic minority areas of central Viet Nam, where populations are impoverished, poorly educated, culturally and linguistically distinct and living in dispersed, less accessible settlements. The researchers therefore considered it both instructive and timely to investigate persistent malaria in such settings.

Methods: Mixed methods (qualitative and quantitative) were used to collect data, in order to explore the complex interrelations between the various actors and system elements. Data was collected in two stages. The formative stage used mainly qualitative tools (e.g. community meetings, observation of bed-net use, and focus group discussions/semi-structured interviews) while health managers, providers and the community helped to define and expand thematic areas of enquiry. Outcomes informed the quantitative approaches (e.g. a provider quiz, structured surveys with community members and village health workers, and quality check of microscopy facilities and health records at district and commune levels). The table describes the methods that were used.

Conclusion: Use of the mixed methods informed researchers and the NMCP about the contextual factors that acted as bottlenecks to effective malaria control in the affected region.

Lessons: The complexity of contextual factors coupled with poverty, low education levels, cross-border mobility, and cultural diversity, made it appropriate to use mixed methods.

Case study 2 Use of mixed methods to explain malaria persistence in remote Central Viet Nam

Table. Summary of mixed methods used during the project

Formative stage		
Method	**Objective**	**Participants**
Community meetings	To explore beliefs, attitudes, awareness, care seeking/providing and circumstances relevant to malaria exposure and control	Malaria control officials, local government, mass organizations, hospitals
Focus group discussion		Provincial and district malaria control managers and Commune Health Station staff, village health workers, and community members
Semi-structured interviews		Provincial malaria control officials, district malaria control secretaries, district hospital staff, commune health staff, village health workers, community members
Informal group discussion		District hospital managers
Observation	• To identify antimalarial drugs on the market	Drug selling points
Observation	• To describe village environment/context	Villages/community
ASSESSMENT STAGE		
Tests/quiz	To obtain an impression of provider knowledge and guidelines adherence	District hospital staff
Observation checklists	To assess visibility and currency of malaria treatment guidelines Quality of microscopy	Health service points
	Bed-net quality during KAP survey home visits	Homes
Review of treatment records/logs		Malaria patient records
Structured questionnaire	To determine community knowledge, attitudes and practices(KAP)	Village health workers Community members

Source: Morrow M. et al. Pathways to malaria persistence in remote central Vietnam: a mixed-method study of health care and the community. BMC Public Health. 2009; 9(1):1.

REFLECTION ACTIVITY

In relation to your team's IR question, and the study design that you have chosen, consider and discuss/agree which research methodology (or combination of both methods) you will use in your research.

Research tools and techniques

Introduction

This section describes the tools and techniques that are used in quantitative and qualitative methods.

IR projects use a wide range and combinations of techniques and data collection tools.

Quantitative research tools

Quantitative methods involve the collection and analysis of objective data, often in numerical form. The research design is determined prior to the start of data collection and is not flexible. The research process, interventions and data collection tools (e.g. questionnaires) are standardized to minimize or control possible bias. Table 8 provides an overview of quantitative data collection strategies.

Table 8: Quantitative data collection tools

Type of instruments	Summary
Observation checklist	The researcher directly observes (watches and listens to) some phenomenon and then systematically records the resulting observations. Tool: Observation checklist is the instrument used for structured observation. The checklist consists of pre-determined specific categories of behaviours/arrangement/processes/procedures that will be observed.
Questionnaires	Survey instruments comprising a series of questions, designed to measure a given item or set of items. Tool: Questionnaires can be used for structured interviews, offline or online self-administered data collection, and telephone interviews. In a questionnaire, the subjects are required to respond to questions in writing or, more commonly, by marking an answer sheet. In the latter type of questionnaire, response options are often closed lists of responses.
Performance based instruments	Performance-based instruments are alternative forms of assessment used to demonstrate a skill or proficiency by having the participant create, produce or do something (e.g. write a paper, create a portfolio, do an athletic performance). Although popular in recent years, the use of the these approaches is fraught with technical difficulties. They are often time- consuming and may require equipment or other resources that are not readily available.

Type of instruments	Summary
Diary	A diary is a self-completed record of experiences during the study period (e.g. alcohol consumption, episode of sickness, or travel).
Electronic data capture	Electronic data capture is a method for collecting data entered directly into a computer or other electronic device (i.e. rather than paper forms). The instrument can be in web based, handheld/smartphone or computer format.

Qualitative research techniques and tools

Qualitative research is generally used to explore values, attitudes, opinions, feelings and behaviours of individuals and understand how these affect the individuals in question. Researchers using qualitative methods are concerned with individuals' perceptions of specific topics, issues or situations and the meanings they assign to their lives. This kind of research is important for generating theory, developing policy, improving educational practice, justifying change for a particular practice, and illuminating social issues. It may also be used to explain the results of a previous quantitative study or to prepare for the development of a quantitative study.

If your research team decides to use qualitative methods in your study, you will need to describe how qualitative methods will provide the information to help you address your research objectives and research question(s). For example, qualitative research may be appropriate because you intend to explore the values and behaviours of individuals in the study area in relation to a public health intervention, and to understand how these affect the phenomena in question. For example, why do some households have bed nets but do not use them? Or, why do individuals in a study area decline services from a specialized antenatal clinic? Qualitative methods can provide context, a deeper understanding of stakeholders' needs and participants' perspectives.

When collecting qualitative data, it is preferable to use more than one data collection method. Obtaining information on the same phenomena in a variety of ways allows the researcher to triangulate the data, adding rigour to the research. By nature, qualitative data collection is emergent and the design is intentionally flexible to enable the researcher investigate themes (findings) in more detail as they emerge.

Qualitative methods use data collection methodologies such as interviewing, observation, discussions and review of documents (e.g. diaries, historical documents). The results of qualitative research are descriptive or explanatory rather than predictive, and are typically time-consuming to collect and analyse. The following table may be helpful to you as you decide which qualitative tools and techniques are most appropriate for your IR project (Table 9).

Table 9: Qualitative data collection tools

	Summary and examples
Participant observation	The researcher participates in/observes the natural setting over an extended period of time: Systematic observation of verbal and non-verbal actual behaviour in which trained observers use a structured recording form. Data is collected by observing, interviewing, note taking and/or journaling. The researcher develops a relationship with the participants, which may affect the data collected. Tool: Participant observation checklist Example: Semi-structured direct observation will be carried out in selected facilities to assess and compare the behaviour of health staff towards patients who are not members of the revised schemes in at least two facilities in each study county, such as one township or commune health centre and one county or district general hospital. In this setting the observer can participate in the interaction between the health staff and the patients and can act as part of the health providers' team or as a client to the health providers.
Non-participant observation	The researcher does not participate in any activity in the natural setting. Data is collected by observing, note-taking and/or journaling. The researcher does not develop a relationship with the participants and therefore cannot explore further issues in relation to observations made unless this approach is complemented with a follow up. Tool: Participant observation checklist Example: The same study setting as the example above, but this time the observer does not participate in the interaction between health staff and the patients. He or she will independently observe the encounters.
Field observation during a 'transect walk'	Detailed descriptions of events, actions, behaviours, people and objects in a natural setting. Field observations are written in the form of field notes. Tool: Transect walk checklist Example: To understand the day-to-day activities, practices, and interactions in a village, a researcher walks through the village cross-sectionally and observes villagers activities, structures of houses, buildings, and interactions among villagers.

	Summary and examples
In-depth interviews	A purposeful conversation directed to the participant by the researcher. The researcher will typically develop an interview guide beforehand. The researcher encourages the participant to talk in-depth, prompting more detail whenever possible without leading the participant to specific answers. Interviews are often recorded and transcribed. The average length of an interview is one hour (or less). Tool: In-depth interview guide Example: In-depth individual interviews with: People suffering from 'catastrophic illnesses', including both members and non-members of revised schemes and those who have used and not used the services; health policy-makers at national and local levels; and rural health insurance scheme managers.
Review of documents and artefacts	Written or printed records of past events (e.g. letters, anecdotal notes, diaries). Material objects and symbols of a current or past event, groups, organizations, or a person that can reveal social processes, meaning, and value (e.g. diplomas, awards, papers, logos etc.). Tool: Checklist or other criteria to review documents Example: Analysis of printed posters, commercials etc. to understand values, messages and meanings for targeted audiences.
Video/film/photographs	Media that captures the daily life of an individual, group or event under study. Can be captured and viewed repeatedly to record behaviours. Tool: Checklist and/or criteria to review that media Example: Review photographs taken by community members showing the areas of public health need in their community.
Focus group discussion	A 1–2 hour discussion, guided by a trained moderator, in which 6 to 10 similar respondents (e.g. by age, gender, social status) focus on a list of defined topics. The discussion, designed to reveal beliefs, opinions and motives, should take place in an informal setting. Data collection may be enhanced by the interaction among participants. Tool: FGD topic guide Example: Focus group discussions using participatory techniques with: members and non-members of the revised schemes (including different age, gender and socioeconomic groups); and health service providers at county/district levels and below, including general practitioners/primary care providers, preventive service providers, and out-patient and in-patient providers.

Unlike quantitative data collection, qualitative data collection can be more flexible allowing the research to incorporate emerging themes in the ongoing data collection. This allows the researcher to test and validate findings as they collect the data. For example, perhaps in one in-depth interview, the researcher learns that people do not attend the lymphatic filariasis mass drug administration because they use traditional medicines and therefore feel that they are already under treatment. The researcher may then add a related question to subsequent in-depth interviews to see how prevalent this phenomenon is in the study population.

Table 10 describes situations when various qualitative data collection techniques can be used.

Table 10: When to use various qualitative data collection techniques

Data collection technique	Situation
Observation	• When the unit of analysis is individual or a group. • When verification is needed. • Anytime and in any situation where researchers want to understand first-hand phenomena under study.
In depth Interviews/ Key informant interviews	• At the beginning of the research as a stepping stone to FGDs. • When preliminary knowledge on a particular issue is needed. • When research interests are being defined. • When individuals or social settings are difficult to access. • To understand subjective experiences. • Where subject matter may be sensitive and people will not speak in FGD settings.
Focus Group Discussions	• When a single subject is being explored in depth. • When enough is known about the subject to develop a topic guide for discussion. • When the subject matter is not sensitive so that people will not mind talking in a group. • Quick results are needed but the research project has limited funding. • Acceptable number of people can be assembled to participate in a discussion group.

Case study 3 Data collection tools: Case of the NIGRAAN project

Background: Data collection tools enable a systematic collection of data about participants in any given study. The exact tool employed depends on the objective of the study. Due to the potentially complex nature of implementation research (IR), mixed methods – and hence different data collection tools – are often used as in the NIGRAAN project in rural Pakistan. The project was conducted by the department of community health sciences of the Aga Khan University (AKU) (Karachi) in collaboration with the Sindh Provincial Department of Health. Nigraan is an Urdu word meaning 'supervisor'. The two-year IR project sought to identify ways the structured and supportive supervision of lady health workers (LHWs) by lady health supervisors (LHSs) could be strengthened, and to improve community case management of pneumonia and diarrhoea in children under five years of age in Badin district, in Sindh. The study was conducted in three sequential phases. The study participants included LHWs, LHSs, community caregivers of children under the age of five and policy-makers. Quantitative data was collected using structured questionnaires, a knowledge assessment questionnaire and a skill assessment questionnaire (Table 1), while qualitative data was collected using in-depth interviews (IDs), focus group discussions (FGDs) and narrative interviews (Table 2).

Table 1. Quantitative data collection tools

Tool	Study participants	Purpose of the tool
Household survey questionnaires	Primary caregivers	To record socio-demographic information, caregiver practices regarding diarrhoea and pneumonia of the population under study, as well as to document the morbidity due to diarrhoea and pneumonia.
Knowledge assessment questionnaires	LHSs and LHWs	To assess the theoretical understanding and knowledge of LHSs and LHWs regarding community case management of diarrhoea and pneumonia.
Skills assessment scorecard 'A'	LHSs and LHWs	To assess the practical/clinical skills of LHSs and LHWs regarding community case management of diarrhoea and pneumonia.
Skills assessment scorecard 'B'	LHSs and LHWs	

Table 2. Qualitative data collection tools

Tool	Study participants	Purpose of the tool
Narrative interviews	Community caregivers	Explore caregiving practices and decision making for childhood diarrhoea and pneumonia.
FGDs and IDs	LHSs, LHWs	To record HWs' perspectives, knowledge and skills regarding community case management of childhood diarrhoea and pneumonia in rural Pakistan.
IDs	Policy-makers	Establish their opinions on the causes of the observed structural gaps.

Lessons: Data collection should be designed specifically, in accordance with the study population and objective.

Pre-testing

All study instruments (quantitative and qualitative) should be pre-tested to check the validity and reliability of data collection tools. Pre-testing allows the research team to check whether the research instructions and questions are clear, context specific, and that adequate time has been allowed to administer the questionnaire, etc. Pre-testing should be conducted from a comparable study population and environment. Since data management is critical to the success of the research, the data management team should be available during the discussion that follows the pre-test, in order to incorporate changes into the final design of the tool and facilitate the incorporation of appropriate checks into the data entry system. This stage includes designing the forms for recording measurements, developing programmes for data entry, management and analysis; and planning dummy tabulations to ensure the appropriate variables are collected.

Example

Table 11 summarises the range of research methods used in the different phases of an IR project in Bangladesh. It describes a cluster randomized controlled trial designed to test a home care and community health worker intervention in comparison to established neonatal care services.

Table 11: Research methods used in the different phases of an IR newborn care project in Bangladesh

Phase of the research	Methods	Objective
Pre- intervention phase	Quantitative household survey	Provided estimates of existing neonatal mortality and levels of skilled attendance.
	Formative qualitative research	Explored home care practices that put newborns at risk of death, and the barriers for safe delivery and postnatal care.
	Observation of newborn care	Demonstrated that community health workers could diagnose newborn illness.
Intervention phase	Household surveys and in-depth interviews	Demonstrated that the intervention was being implemented as planned.
	Surveys, observations and in-depth interviews	Established whether the newborn package was being implemented consistently ("implementation fidelity").
Post intervention phase	End-line household	Assessed both neonatal mortality and service coverage levels
	Qualitative research	Explained in detail how and why delivery and post-natal practices changed, largely because of the engagement of the local community in the programme, and the supportive supervision of the community health workers ("meaning enhancement").

Adapted from Baqui et al, 2008;[13] Baqui et al, 2009;[14] Choi et al, 2010;[15] Shah et al, 2010.[16]

REFLECTION ACTIVITY

In relation to your team's research problem and question(s), discuss and agree which data collection tools you will use in your IR project. As you reflect on this, you must take into account the selected study design, your IR questions, and also the time and budget available to your IR team.

Sampling

Now that you have chosen the most appropriate techniques and tools to collect your research data, it is important to know how many people you need to approach to participate in your research. This is called the 'sample size'. In general, when using quantitative research tools, you need to ensure that you recruit enough people to provide an accurate and reliable estimate of what you are studying. When using qualitative research tools, the aim is to reach enough individuals that you can represent the prevalent opinions, experiences and knowledge in the study population. In this section, we review the sampling designs used in both quantitative and qualitative research tools.

Sampling design in quantitative methods

Quantitative studies require a representative sample of the study population to be able to accurately portray the characteristics of the population and to yield maximum precision of such population parameters. The following criteria are critical when designing a sampling strategy: (1) What are the research objectives? (2) What are accurate estimates of sampling variability? (3) Is it feasible to apply the sampling strategy and obtain the calculated sample size? (4) Is it possible to mininize costs (or to achieve research objectives for minimum cost). As these criteria can conflict with each other, research teams must find a balance between them.

Sample size

A representative sample requires an adequate sample size, taking into account statistical power parameters. Power is the probability of rejecting the 'null' when an alternative hypothesis is true. In simple terms, this is the probability of actually detecting an effect under study. Different sample size calculations should be used for the various study design types. Sample size calculation formula and calculation procedures can be found in standard biostatistics reference materials.[17] Further discussion with a statistician will also help to confirm and calculate the appropriate sample size needed for various types of research methods.

Sampling strategy

As quantitative studies require a representative sample with regard to population characteristics, a 'probability' sampling is preferable. This enables every individual in the population to have a certain chance of being included in the sample. Probability sampling also allows estimates of sampling error to be calculated. There are several probability sampling strategies (Table 12).

Table 12: Probability sampling strategies

Sampling strategy	Summary
Simple random sampling	The ideal sampling strategy because each element of the population has equal probability of being included in the sample. The sampling procedure is to assign a number to each element in the sampling frame and use a random number to select elements from the sampling frame. Most statistical packages can generate random numbers.
Systematic random sampling	This sampling strategy uses a list of population elements. We assume that the elements are randomly listed. The first element included in the sample is randomly identified and the subsequent elements are selected using sampling interval. The sampling interval is calculated by dividing the desired sample size by the number of elements in the sampling frame.
Stratified sampling	Stratified sampling can be used in a population that consists of mutually exclusive sub-groups (e.g. school population with classes). A random sampling procedure is then used to select elements from each stratum/sub-group. Sample size can be selected proportionately to the stratum size.
Cluster sampling	Cluster sampling is commonly used when the population is very large or dispersed across a large geographical area. The goal of cluster sampling is to increase sampling efficiency. However, cluster sampling reduces the population variability in the sample since a group of individuals in the same geographical area is to some extent more homogenous and the probability of each element to be selected in the sample is not equal. To address this limitation, sample size calculation in a cluster sampling strategy needs to take into account design effect, which will increase sample size. Furthermore, the researcher can use the 'probability proportionate to size' procedure to correct the difference in cluster size and adjust the chance that clusters will be selected. A common example is the Expanded Programme for Immunisation (EPI) cluster sampling framework.

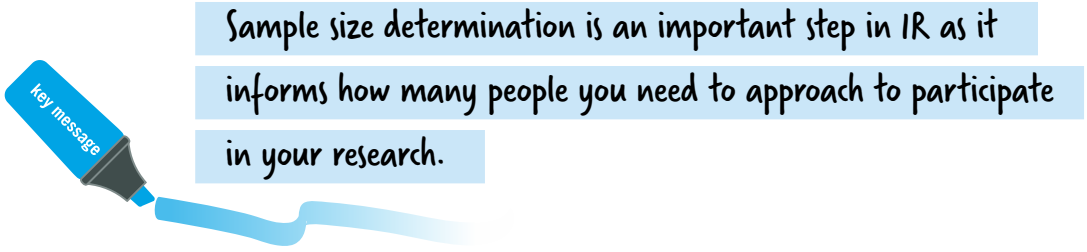
In some situations, random sampling is not the preferred option due to lack of specific resources (e.g. a list of the entire population), time, costs or ethical constraints. In other situations, the research requires some 'weighting' to the information being collected (e.g. a survey among experts). In this scenario, nonprobability sampling is preferable. There are several commonly used non-probability sampling strategies (Table 13).

Table 13: Non-probability sampling strategies

Sampling strategy	Summary
Availability or convenience sampling	Availability sampling refers to the technique in which the selection of sample is due to researcher accessibility. The limitation of this strategy is selection bias. An example of this strategy is sample from facility or institution where the researcher is employed.
Successive sampling	Successive sampling is when individuals are selected successively, for example, an exit interview with patients after an encounter with the health provider. All patients who just met with the doctor are offered the opportunity to participate in the study. If the study involves multiple sites, a combination with stratified sampling can be used. However, the patients are selected successively in each stratum.
Purposive sampling	Purposive sampling is used when the elements are selected based on the researcher's judgment regarding the desired information being collected. Participants are being selected on their knowledge of the topic being studied. The example of purposive sampling is a survey using a panel of experts.
Snowball sampling or respondent driven sampling	This type of sampling strategy is suitable to recruit participants who are members of a hidden population (e.g. victims of domestic violence, drug users). Snowball sampling is started when a researcher can identify the first participant that met selection criteria. The researcher then asks these participants to identify people with similar experiences or characteristics. To increase the variability of characteristics of the study participants, the researchers can ask the subsequent participants to find the next participants with the same experience but with different socio- demographic characteristics, for example with different gender, age group, socio-economic

Sampling in qualitative methods

Sampling strategy



Sampling in qualitative research uses quite different approaches to those used in quantitative studies. The aim in qualitative research is not to have a representative sample, but rather one that reflects the characteristics and richness of the context and/or study population. Whatever sampling method is used, the IR team will need to justify their sampling frame selection. Table 14 reviews the different kinds of sampling techniques used in qualitative research.

Table 14: Sampling techniques used in qualitative research[18]

Kinds of sampling	Explanation
Convenience	Studies the units that are available at the time of the research. It is more convenient than a random sample because the researcher uses what is available, rather than what is selected. There is however, a risk of measurement bias. If interviewing households in the morning is most convenient, which populations might be overrepresented (housewives, elderly) and which may be underrepresented in the sample (employed, men, students)?
Purposive	Used when the elements are selected based on the researcher's judgment regarding the desired information being collected. For example, researchers may decide to identify respondents according to their involvement in a particular health programme.
Maximum variation	Selects units that represent as wide a range of variation as possible (e.g. gender, socioeconomic status, population density, etc.).
Snowball	Identifies a few people who will be involved with the study and then asks them to identify more people who would be relevant to include in the research. Best to start with at least two individuals so as to reach different networks of individuals. This is the most common form of sampling in qualitative research methods.
Contrasting cases	Involves two or more units with distinct characteristics so that comparisons can be made when explaining problems and understanding the factors that influences them. For example, researchers may decide to study individuals living in a site where a health programme has been successful and another site where the programme has been less optimal.

REFLECTION ACTIVITY

Working together with the other members of your IR team, devise the sampling strategy for your study and the tools you will use. Think about the respondents you hope to recruit as you plan your study. How can they be reached? What time of day should you interview people? How much time will they have to participate in your study? Remember that IR takes place in real-life settings, so if you want to interview nurses in a health clinic, Monday morning may not be the best time!

Data collection

Now that you have a sampling framework, you can start to think about collecting your data. Before data collection begins (e.g. interviews, focus group discussions, survey, etc.), you must receive ethical approval from the bioethics committee in your country. In this process, you will have to develop an information sheet and informed consent form that will need to be read to each study participant. Before any data is collected, the participant must give his/her informed consent to the process. This information is outlined in detail in the module entitled Planning and Conducting an IR Project.

SEE IR-PLANNING & CONDUCTING IR MODULE

While planning your data collection, you will need to identify people who will actually carryout the data collection. A short training will need to be conducted prior to the start of the survey so that the enumerators understand the entire process, the data collection instruments as well as the sampling strategy that you will use. At this time, role plays can be carried out with the enumerators to ensure that they understand the research process and instruments.

Finally consider how you will supervise the data collection process. How will problems in the field be rectified? For example, what if people aren't home on the days that you start your surveys? What if key informants do not have the time for an in-depth discussion? How will you work with the local head of the health centre if he or she does not agree with using a performance-based instrument?

Regular meetings with the enumerators/field staff can help address some of these issues that may arise in the data collection process. These regular meetings will also provide an opportunity to amend topic guides if using qualitative data collection tools. For example, if you discover new and unexpected phenomena on the first day of FGDs, you will want to add the topic to your next planned FGDs to ascertain how common it is within your respondent population.

Data management

Most research projects generate a significant amount of data. This data should be of good quality because it is underpins the quality of the study. Therefore, good data management is fundamental for high quality research. Good practice in data management also helps researchers ensure that the required processes of data collection and analysis are organized, understandable, and transparent.

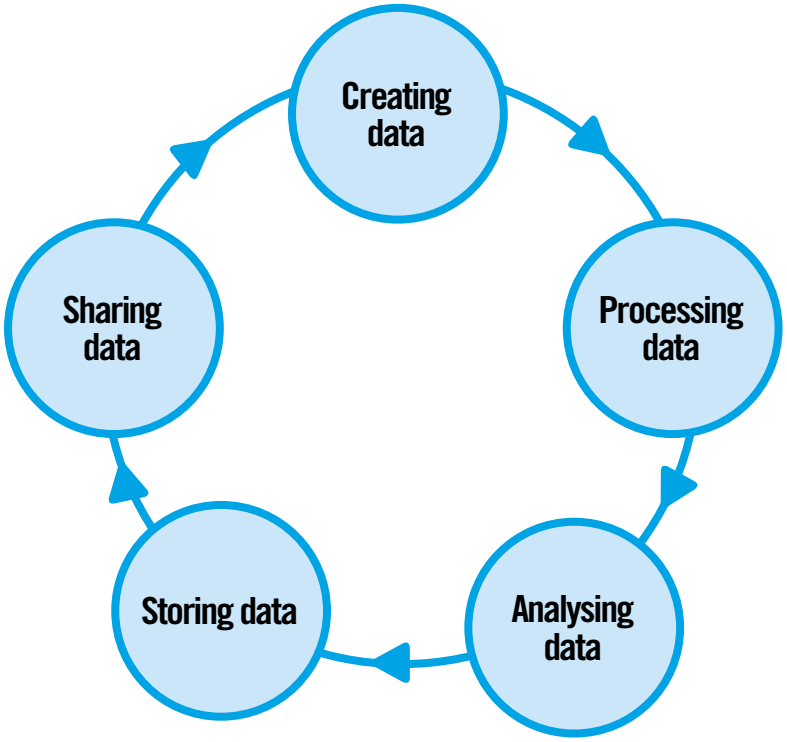
The main data management responsibilities include:

- Organizing and ensuring the collection of accurate data.
- Capturing the data on a database.
- Validating and correcting the data.
- Providing data in a form that will enable analysis.
- Storing and sharing the data.

It is important to remember that the confidentiality of the respondents' identities must be guaranteed at all times in the data management process. This is usually stipulated in the ethical approval you will receive (e.g. keeping files on a password-protected computer, locked cabinets, limited number of persons with access to any anonymized data). In addition, you will outline how long you will keep the data after the research has been completed. You must ensure that these ethical criteria are maintained throughout the course of your IR project.

Data management is a cyclical process (Figure 4). The data life cycle starts with creating data, followed by processing and using data for analysis. The last two stages of the cycle are storing and sharing the data.

Figure 4: Data management cyclical process



Adapted from Surkis & Read, 2015.[19]

Creating data

Creating data is the first step in the data management processes. In quantitative studies, this stage consists of defining what type of data will be collected, their format and the procedure to create them. The researcher must ensure that all the collected data reflect the realities, using standardized instruments, data collection procedures, checking error rates during data collection (e.g. checking the completeness and consistency of respondents' responses in the questionnaires, checking the validity of the responses by random re-interview process).

In a qualitative study this stage starts with defining different types of information the researcher intends to gather, different tools (e.g. interview or FGD guidelines) and data collection activities. The researcher needs to ensure that all recording devices are placed in a way that will best record the conversation or discussion, and that the space for interview or discussion creates a safe atmosphere for open discussion while ensuring privacy.

Processing data

This is the course of translating information from the rawest form to a form that is ready for analysis to the researcher. In quantitative study this means creating an electronic database that is appropriate to manage different types of data (e.g. multiple responses, numerical data, visual analogue scale data, etc.). It involves creation of file and coding structures that are understandable, the development of a codebook, making decisions on which data can be kept in the database which should be discarded. As data is entered, data entry errors should be prevented by double entry and checking the consistency of responses. In qualitative studies, this means that all the recorded data are transcribed verbatim and in some cases transcripts can be shared with respondents to verify content. It also entails the development of a codebook particularly when more than one researcher is conducting the analysis. All collected qualitative data should be saved in a qualitative data management application.

Analysing data

Data analysis in quantitative studies consists of identifying patterns through descriptive analysis, comparing data, hypothesis testing and finding relationship between variables. In qualitative studies, this process consists of identifying, understanding meaning and assigning code to the data, identifying patterns and emerging themes, and constructing framework to explain certain phenomena. This activity will be described in a subsequent section.

Storing data

Storing data involves activities not only during the study period, but also in the long term by archiving data in a repository or data centre. Presently, electronic data storage/repository is the medium of choice as it requires little space and is simple to back up. However, a data storage strategy is needed as digital storage media also have several limitations e.g. quality and life cycle of storage media, software interoperability, relevant data reading equipment and power supply. Data security is another issue in storing data. Security issues include physical data security (e.g. locked room or cabinet, access log book), and electronic data security (e.g. secure access using password, level of access, and data encryption for sharing and transmission). The WHO Good Clinical Practice Guideline recommends that data and essential documents should be stored for at least two years after the research project has ended.[20]

Sharing data

Sharing data is particularly important in collaborative multi-centre/country studies. Data sharing, together with data transfer, data storage and access for all collaborative partners or institutions can be challenging as it may involve different regulations. Cloud-based file sharing may be preferable, although it may not be suitable for all types of data, particularly identifiable, confidential data. Furthermore, researchers do not have control over where data is actually stored.

Data sharing is becoming mandatory in many fields as a way to ensure transparency, avoid duplication as well as plagiarism. Since IR may involve different institutions/ organizations, the guidelines for data sharing and ownership should be clearly spelt out at the beginning through agreements such as a memorandum of understanding. Data sharing should follow a clear process and can be carried out between two research institutions though not between two individuals. Please check your own institutional and national guidelines before designing any data sharing agreements.

Data quality management

Collection and storage/documentation of accurately recorded and retrievable results are essential for any research. Good data collection practices will ensure that data can be traced to their source and their original form (i.e. the raw data that constitutes the first recording of the observation). To ensure these characteristics, raw data must be recorded:

- **Promptly:** After a specific task is completed. Delaying data recording will reduce data quality as memory may fail or be inaccurate.
- **Accurately:** Inaccurate data recording will reduce the reliability of the data collected; Accuracy is therefore a critical part of the integrity of the study.
- **Legibly:** Hand-written data should be clearly written and electronic records should not be difficult to decipher.
- **Indelibly:** Handwritten raw data should be recorded in permanent ink. Any changes to the raw data should not obscure the previous entry. The date, reason for the change and signature of the person responsible for the change should be added.

Clear and regularly checked data flow prevents data loss. As IR collects different types of data (i.e. patient, organizational and surveillance-related data) from various sources (i.e. human subjects, medical records, health services and laboratory registers, surveillance systems, and administrative systems) a detailed chart should be made describing the critical pathway(s) to be used for the data collection process in handling questionnaires, coding, data entry, data verification, cleaning and storage of hard copies and back-up of data files.

Data quality is key to having authentic and scientific data and therefore should be taken seriously. Activities such as staff training, supportive supervision and data feedback can be used to enhance the quality of data. Refer to the planning of an IR project module for details.

SEE IR-PLANNING & CONDUCTING IR MODULE

Data analysis

Introduction

Depending on the research questions you want to answer and the type of data you have collected (i.e. quantitative or qualitative data), different types of analysis can be performed. Before we begin to analyse the data, we need to remember the different audiences to be reached with the results and recommendations of the IR project. What are their needs for information, and what is the best way to reach them? Click on each of the headings below for an explanation on each type of analysis.

Data analysis plan

To ensure that the analysis is undertaken in a systematic manner, an analysis plan should first be created. The analysis plan should contain a description of the research question and the various steps that will be followed in the research process. It is best practice to develop your data analysis plan at the start of your project, in order to capture the hypotheses you have about your research question. You may amend the data analysis plan as your research progresses.

Designing data analysis in an IR project is based on the premise that IR aims to: (i) understand the implementation processes for a given intervention, focusing on mechanisms that support or constrain those processes; and (ii) communicate that understanding of the implementation process to multiple stakeholders, who may consequently contribute to the integration of findings into current and/or future research, policy and/or programming.

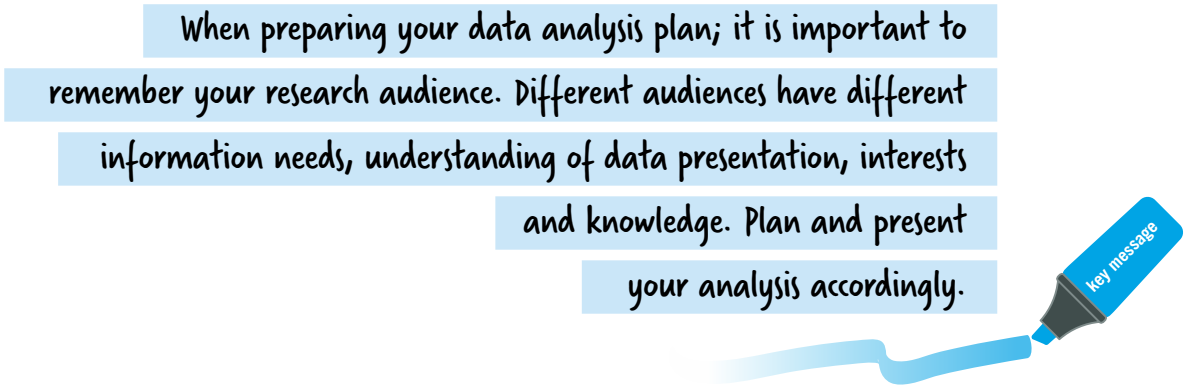
Most IR proposals use mixed methods in which quantitative and qualitative techniques are combined. Under many circumstances, mixed-methods approaches can provide a better understanding of the problem than either approach can achieve alone. However, few of the stakeholders in the IR project team are likely to have specialized knowledge of both quantitative and qualitative research methods. It is therefore essential that the analysis and most importantly, the presentation of findings, be carefully considered to avoid potential misinterpretations that could lead to inappropriate conclusions and/or responses. Emphasis should be placed on simplicity and interpretability because stakeholders need to both understand the information provided and also be able to interpret it correctly.[21] Data analysis should take place along with the data collection process. This continual data analysis process facilitates regular sharing and discussion of findings.

Designing analysis by purpose

An important preliminary consideration when designing your data analysis plan is to clearly define the primary objectives of the analysis by identifying the specific issues to be addressed. It is important to remember that data from IR is, by its nature, intended not only to simply describe an intervention but also to improve it.

For example, IR research may focus on:

- **Effectiveness:** Aims to modify implementation procedures in order to improve the generation of benefits.
- **Efficiency:** Attempts to assess the implications of possible modifications to the implementation process in order to increase the benefits in relation to resources.
- **Equity:** Focuses on distributional issues, i.e. how benefits and resource costs are distributed.
- **Sustainability:** Focuses on identifying essential inputs, potential constraints on their availability and other possible barriers to medium and long-term sustainability.



Quantitative data analysis

Before any statistical analysis is undertaken, some factors need to be taken into account in order to select the most appropriate statistical analysis approach. These are described briefly below.

Measurement scale and different statistical techniques

Measurement scale is a way to define and categorize variables. There are four different measurement scales (nominal, ordinal, continuous and ratio scale). Each measurement scale has different properties, which are required for different statistical analysis. Table 15 summarizes the properties for different measurement scales, described in detail below.

Table 15: Summary of measurement scale properties

Measurement	Category/ difference	Rank/order	Meaningful number	Meaningful number scale
Nominal	+	-	-	-
Ordinal	+	+	-	-
Continuous	+	+	+	-
Ratio	+	+	+	+

The nominal scale can only differentiate the category. We cannot say that one category is higher or better than the other category. An example of a nominal scale is gender. If we code Male as 1 and Female as 2 or vice versa (i.e. when we enter the variable into the computer), it does not mean that one gender is better than the other. The numbers 1 and 2 only represent categories of data.

Ordinal scales represent an ordered series of relationships or rank order. However, we cannot quantify the difference between the categories. We can only say that one category is better or higher than the other categories. An example of an ordinal scale is the level of a health facility (e.g. primary, secondary, tertiary).

Continuous scales represent a rank order with equal unit of quantity or measurement. However, in this scale, zero simply represents an additional point of measurement not the lowest value. An example of such a scale is a temperature scale in Celsius or Fahrenheit. In this scale, zero (0) is one point on the scale with numbers above and below it.

Ratio scale is similar to the continuous scale, in that it represents a rank order with equal unit of quantity or measurement. However, ratio scale has an absolute zero, in which zero is the lowest value. An example of ratio scale is body mass index (BMI) in which the lowest value (theoretically) is zero.

The continuous and ratio data are referred to as parametric as these types of data have certain parameters with regards to distribution of the population as a whole (assumption of normal distribution with mean as a measure of central tendency and variance as a measure of dispersion). Parametric also means that the data can be added, subtracted, multiplied and divided. The statistical analysis for these types of data is referred to as parametric test.

On the other hand, nominal and ordinal scales are referred to as non-parametric. Non-parametric data lacks the parameters that the parametric data have. Furthermore, it lacks quantifiable values and as such nonparametric data cannot be added, subtracted, multiplied or divided. Nominal and ordinal data are analysed using non-parametric tests.

A parametric test is considered to be more robust than a non-parametric test. Furthermore, there are more statistical options available for analysing parametric data. However, most parametric tests assume that the data is normally distributed.

Research questions

The way we formulate research questions also determines what kind of statistical techniques need to be used for analysis. Examples of IR questions include:

- Describing patterns/distributions of study variables in terms of "What, Who, Where, and When".
- Comparing differences between groups.
- Exploring possible associations/correlation between independent variables (exposures) and dependent variables (study outcomes).
- Exploring a possible causal relationship between independent variables (exposures) and dependent variables (study outcomes).

Descriptive statistics

Quantitative research generates large volumes of data that require organizing and summarizing. These operations facilitate a better understanding of how the data vary or relate to each other. The data reveals distributions of the values of study variables within a study population. For example:

- The number of children under five years in various households in a given population.
- Daily outpatient attendance in a health facility.
- The birth weights of children born in a particular health facility over a period of time.
- Educational levels of mothers of children born in a particular health facility.

Analysis of the type of data described above essentially involves the use of techniques to summarize these distributions and to estimate the extent to which they relate to other variables.

Example

In a sample of newborns we might summarize the distribution of birth weights by calculating the frequency of low, normal and high birth weights, classifying as normal those in some standard range. If we also calculated the frequency of different education levels for the mothers of those newborns, we could then estimate the strength of a possible relationship between these two variables.

The use of frequency distributions for this purpose has several advantages:

- Useful for all types of variables
- Easy to explain and interpret for audiences without specialist knowledge.
- Can be presented graphically and in different formats to aid interpretation (e.g. tables, bar charts, pie chart, graphs, etc.).

The different data presentation formats help to reach different target audiences. Tables are a useful presentation format when you want to communicate within the scientific community. Graphical data presentations help to communicate with a wider, less scientific, audience in the community or policy makers. You can read further about data presentation and how to present data to different audiences in the Advocacy and Communication module of this Toolkit.

COMMUNICATION & ADVOCACY MODULE SEE

Defining intervals for frequency distributions

A key decision in constructing a frequency distribution relates to the choice of intervals along the measuring scale. For example:

- Ordinal: Level of health facility (e.g. primary, secondary, tertiary).
- Continuous: Body temperature (e.g. below normal, normal, above normal).
- Ratio: Body mass index (BMI) (e.g. <25, 25–29, 30+).

There are two conflicting objectives when determining the number of intervals:

- Limiting the loss of information through the use of a relatively large number of intervals.
- Providing a simple, interpretable and useful summary through the use of a relatively small number of intervals.

Note: Distributions based on unequal intervals should be used with caution, as they can be easily misinterpreted, especially when distributions are presented graphically.

Summary statistics and frequency distribution

Careful examination of the frequency distribution of a variable is a crucial step and can be an extremely powerful and robust form of analysis. There can be a tendency to move too quickly to the calculation of simpler summary statistics (e.g. mean, variance) that are intended (but often fail) to capture the essential features of a distribution.

Summary statistics usually focus on deriving the measures indicating the overall tendency location of a distribution (e.g. how sick, poor or educated a study population is, on average) or on indicating the extent of variation within a population. However, the reasons for selecting a particular summary statistic should relate to the purpose for which it is intended.

Measure of central tendency

Example

To find out if a recently implemented intervention reduced the problem of malnutrition among five year-old children in a given village, a researcher may ask: "Which summary statistic is most appropriate?"

Change in mean or median daily calorie intake of all five year-olds in village?

Change in the proportion of five year-old children in village falling below the predetermined minimum calorie requirement?

The criteria for making such choices include:

- *Face validity (i.e. is the statistic relevant to the specific concern?).*
- *Whether stakeholders understand how the data was derived.*
- *Whether stakeholders are able to interpret the findings as intended.*

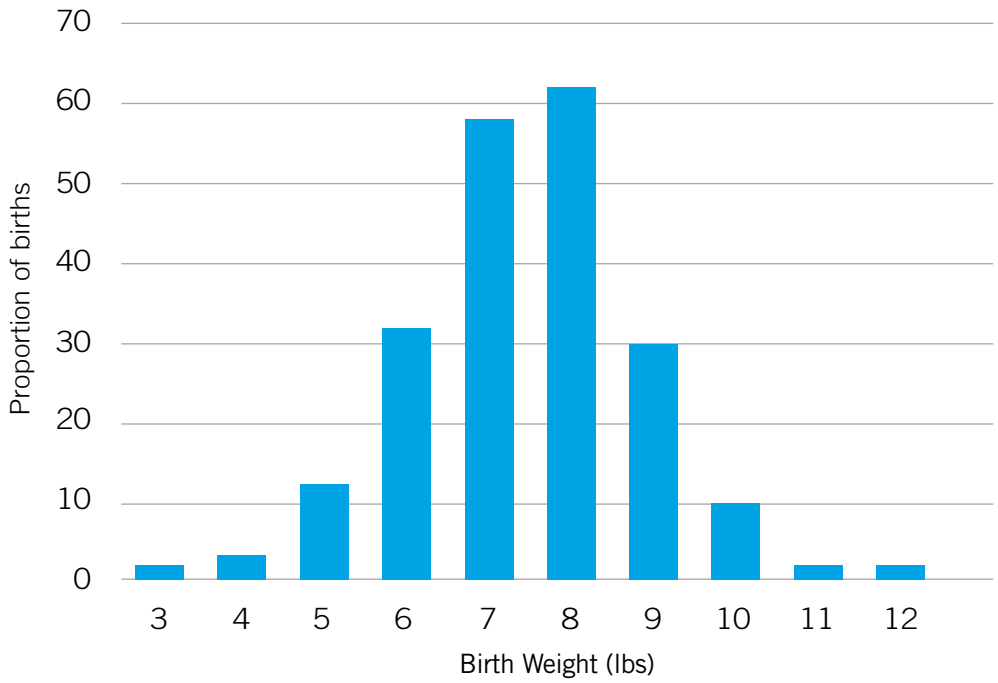
The central tendency measures the central location of a data distribution. The mean, or average, is the most commonly used parameter because the mean is simple to calculate and manipulate. For example, it is straightforward to combine the mean of sub-populations to calculate the overall population mean. However, the mean is often inappropriately used. It can also be misinterpreted as the typical value in a population.

Example

The GDP of a certain middle-income country was calculated as 3200US$. Interpreting this as the income of an 'average' person in that country does not reflect the reality (in fact, it was closer to 1200US$). The mean is often unrepresentative when the underlying distribution is skewed.

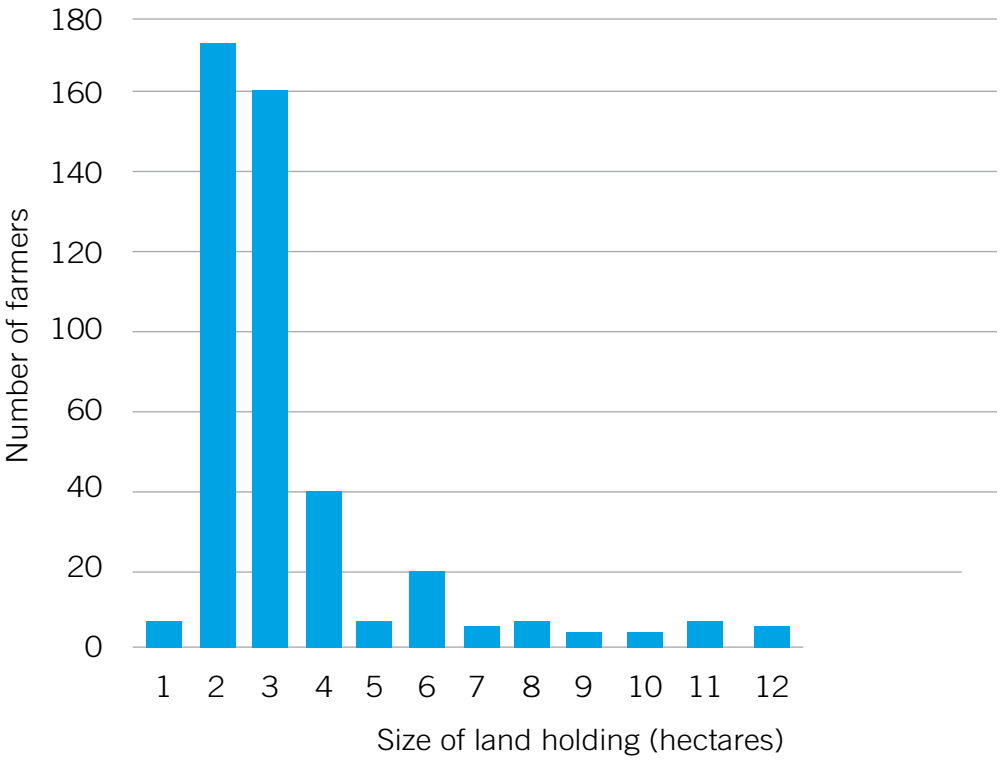
The median, defined as the middle value, is relatively easy to explain. The magnitudes of other values are irrelevant. For example, if the largest value in a given range increases or the smallest value decreases, the median remains unchanged. When a data set is not skewed (or when data is distributed 'normally'), the mean and the median are the same (Figure 5). It is therefore preferable to use median as a measure of central tendency when the data set is skewed as the value is independent to the shape of the data distribution.

Figure 5: Normal distribution: the mean is the measure of central location



In a skewed distribution, the mean is difficult to interpret (Figure 6).

Figure 6: Skewed distribution



Measure of dispersion

Measure of dispersion denotes how much variability occurs in a given population, as follows:

- **Low variability:** Measures of location can be seen as reasonably representative of the overall population; there is limited loss of information through aggregation.
- **High variability:** Measures of location are less useful; there is a substantial risk of losing information by aggregation unless the nature of the distribution is well understood.

Choice of measures

Variances, standard deviations and coefficients of variation are widely used in statistical analysis. As with the mean, this is not because they are always the best measures of variability (they can be easily interpreted for normally distributed variables but not for other distributions), but mainly because they can be readily calculated and manipulated.

For example, given the variances of two population sub-groups it is easy to combine them to calculate the overall population variance. However, while they may have technical advantages, these measures have serious limitations in terms of policy application.

Alternative measures

More readily interpreted measures include quartiles and percentiles. Quartiles: divide data into four quarters (Q1 to Q4), with 25% of available data in each:

- Q2 is the median.
- Q1 is the median of the data points below the median.
- Q3 is the median of the data points above the median.

Q3–Q1 is the inter-quartile range, comprising the middle 50% of a population. Percentiles divide the data into two parts:

- p percent have values less than the percentile.
- (100 – p) percent have greater values.
- 50th percentile = median; 25th percentile = first quartile.

Other common percentiles:

- 20th (which defines the first quintile group).
- 10th (which defines the first decile group).

Other descriptive statistics

Sub-group analysis

The outcomes of an intervention may vary substantially between different sub-groups of the target population. Sub-group analysis can be complex if the sub-groups are not pre-defined. Investigating a relationship within a sub-group simply because it appears interesting could bias the findings.

Data mining (i.e. exploring data sets to discover apparent relationships) is useful in formulating new hypotheses but requires great caution in IR. The context within which this sub-analysis is undertaken should be considered carefully, because relationships between inputs and outcomes may be mediated by contextual variables. For example, we might assume that it would be useful to undertake an analysis of chronic illness by age group and sex, as shown in Table 16. For meaningful interpretation of the results, the type of chronic illness and the background of the patients experiencing them will be important variables to consider.

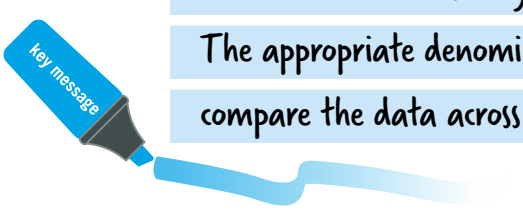
Table 16: Background variables of patients with chronic illness

Age group	Chronic illness prevalence	
	Males	Females
15-24	0.55	0.80
25-44	1.79	4.01
46-64	4.91	12.28
65	12.86	20.00
All	1.77	4.25

Measures of risk

Although measures of risk are widely used in health research, they are not always well understood. For example, risk and odds are often used interchangeably however they do not mean the same thing:

- Risk (P): number of people experiencing an event/population exposed to the event.
- Relative risk RR = (PA/PB): risk in group A compared to risk in group B.
- Odds: number of people experiencing an event versus number of people not experiencing the same event = P / (1-P)
- Odds ratio: OR = [PA/(1- PA)] / [PB/(1-PB)]



The denominator is very important in descriptive analysis. The appropriate denominator helps the audience understand and compare the data across different groups/characteristics.

Statistical tests

A statistical test is performed so that we can make inferences concerning some unknown aspects of a statistical population from the sample that we have collected from a study. There are different types of statistical tests that we can use depending on the research questions, type of measurement scale and assumptions about data distribution. A Simple univariate and bivariate analyses should be done before a sophisticated analysis such as the multivariate analysis, is undertaken.

Finding association/correlation

Association is a relationship between two variables which are statistically dependent. The two variables are equivalent; there are no independent and dependent variables. Correlation can be considered as one type of association where the relationship between variables is linear. There are several statistical tests to assess the correlation between variables (Table 17).

Table 17: Different statistical tests for finding associations according to existing assumptions

Measurement scale	Assumption of distribution	Analysis
Nominal or Ordinal	-	Chi square test
Continuous or Ratio	Normally distributed Avoid outliers data	Pearson correlation
	Not normally distributed	Spearman rank correlation

Finding causality: group comparison

Group comparison analysis is used to explore the statistically significant difference of study outcomes between groups. The groups can be categorized by exposures under study. When there is a significant difference between groups we assume that the difference is due to the exposures (Table 18).

Table 18: Different statistical tests for group comparison according to the existing assumptions

Measurement scale	Assumption of distribution	Type of group	Analysis
Nominal or Ordinal	-	Independent	Chi square test
	-	Paired (before-after)	Sign test
Continuous or Ratio	Normally distributed	Independent	Independent t test
		Paired (before-after)	Paired t test
	Not normally distributed	Independent	Mann Whitney
		Paired	Wilcoxon

Finding causality: prediction

Regression analysis is the type of analysis used to predict study outcome from a number of independent variables. If the outcome variable is on a continuous or ratio scale and has a normal distribution of data, we can use linear regression. If the outcome variable is dichotomous i.e the variable has only two possible values such as "yes" or "no", we can use logistic regression.

Qualitative data analysis

There are many traditions of qualitative research and it has been argued that there cannot and should not be a uniform approach to qualitative analysis methods (Bradley et al 2007).[22] Similarly, there are few 'agreed-on' canons for qualitative data analysis, in the sense of shared ground rules for drawing conclusions and verifying sturdiness.[23] Many qualitative studies adopt an iterative strategy: collect some data, construct initial concepts and hypotheses, test against new data, revise concepts and hypotheses. This approach implies that data collection and analysis are embedded in a single process and are undertaken by the same individuals. However, with the increasing use of qualitative research in health research, objectives are often pre-defined prior to the start of data collection, as opposed to being developed as information for the data collected emerges.

Researchers can also use several different computer qualitative data analysis (QDA) softwares to help them manage their data. The term "QDA software" is slightly misleading because the software does not actually analyse the data, but organizes them to make it easier to find and identify themes. Software can also be relatively expensive (up to around US$900 per single user). For these reasons, some researchers prefer analysing data manually. However, as the software improves, researchers are finding QDA increasingly useful in helping analyse data and saving time. Here are some of the more common QDA software names:

- AtlasTi (http://www.atlasti.com) deals with large data sets, unstructured coding, and mimics paper code and sort functions.

- MAXQDA (http://www.maxqda.com) provides powerful tools for analysing interviews, reports, tables, online surveys, videos, audio files, images and bibliographical data sets.
- QSR NVivo (http://www.qsrinternational.com) (previously called Nud*ist 6) caters for unstructured coding, finds patterns/relationships in codes.
- EZ-TEXT 3.06C (http://www.cdc.gov/hiv/topics/surveillance/resources/software/ez-text/index.htm).

Researchers should feel free to use whatever analysis method (with or without software) they are comfortable with. Whatever approach is used, all qualitative analyses involve making sense of large amounts of data, identifying significant patterns and communicating the essence of what the data reveal.

Qualitative data analysis consists of data management, data reduction and coding of data. In short, the goal is to identify patterns (themes) in the data and the links that exist between them. As mentioned, there is no set formula for analysing qualitative data, but there are three core requirements of qualitative analysis to adhere to:

A. Detailed description of techniques and methods used to select samples and generate data.

B. Carefully specified analysis, paying attention to issues of validity and reliability.

C. Triangulation with other data collection methods.

The following steps describe these three core components in more detail:

A. Detailed description of techniques and methods used to select samples and generate data

   - If conducting interviews or focus group discussions, all sessions are recorded (preferably with a recording device, although where this is not accepted by the participants, with hand written notes).
   - All recordings have to be transcribed verbatim (i.e. typed out in full, word-for-word).
   - If observation has been done, document the times, locations and important events (e.g. interruptions, significant events, etc.)
   - All background information about the participants should be appended to each transcript.

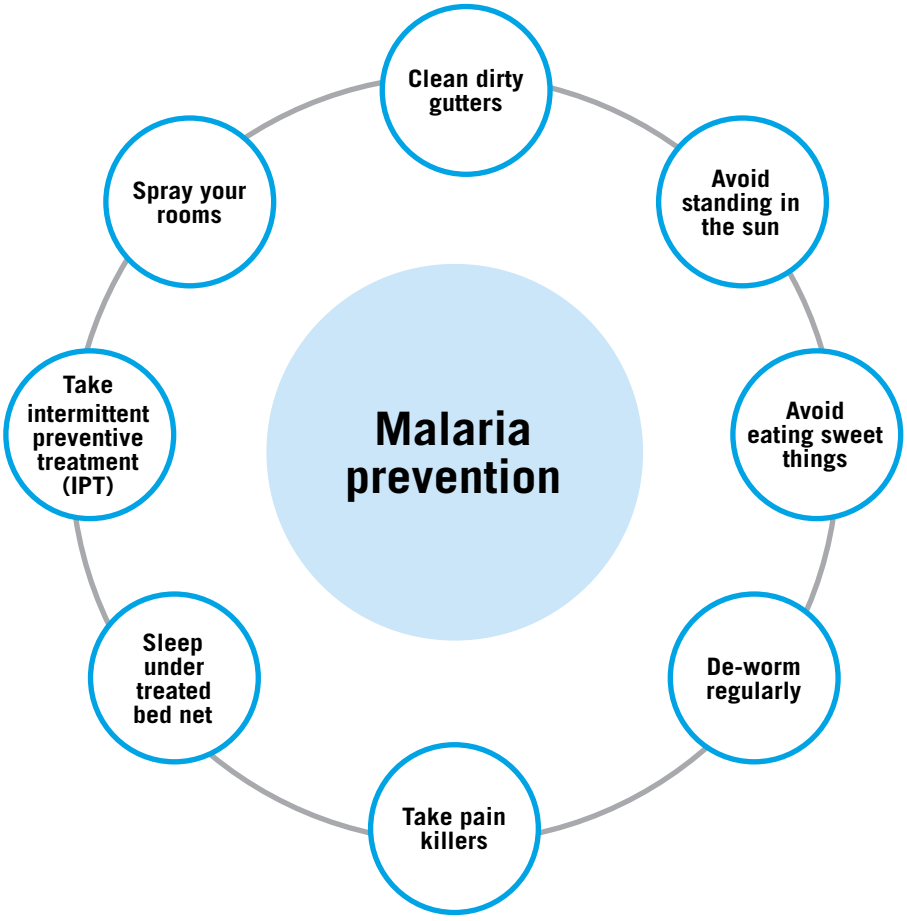
B. Carefully specified analysis, paying attention to issues of validity and reliability

   - In the initial step of the analysis, the researcher will read/re-read the first set of data and write notes, comments and observations in the margin, with regard to interesting data that is relevant to answering the research question(s).
   - While reading the data, the researchers should begin developing a preliminary list of emergent categories into which they will group the notes and comments. These categories are guided by the purpose of the study,

the researchers' knowledge and orientation, and the meanings made explicit by the participants.[24] A list of these categories is compiled and attached to the data.

- The next set of data collected is then carefully read and, with the previously constructed list of categories in mind, notes, comments and observations are once again recorded in the margin. This second data set is grouped into categories and a list of the categories is compiled. The two lists are then compared and merged to create a master list of categories. This list reflects the recurring regularities or patterns in the study.
- These categories are then given names. Category names may emerge from the researcher, from the participants or from the literature. According to Merriam (1998),[24] these categories should be: exhaustive; mutually exclusive; sensitive to what is in the data; conceptually congruent; and, in effect, the answers to the research questions. Category names or codes in data analysis can also be derived from the questions asked in the data collection tools based on the objectives of the study.
- Once the researchers are satisfied with the categories, the data is assigned to these categories. Taking a clean copy of the data, the researcher organizes the data into meaning units and assigns them to the relevant categories, writing the category code in the margin.
- The researchers then create separate files for each category and cut and paste the meaning units into the relevant category, creating a file containing all the relevant data. Care should be taken to avoid context stripping by carefully cross-referencing all units and coding them with the participants' pseudonym, the date of data collection, and the page number.[25]
- The researchers then try to link the categories in a meaningful way. Diagrams can be used to facilitate this process. For example, in a study to determine causes of malaria, a number of prevention themes emerged (Figure 7).

Figure 7: Diagram to describe reported methods of preventing malaria



C. Triangulation with other data collection methods

- Review your results against those collected using other data collection methods to determine the validity or truthfulness of your findings.
- Review if routine data sources confirm your findings.

Rigour in qualitative research

The research team must ensure scientific rigour in qualitative methods analysis. For example, will your study provide participants with a copy of their interview transcripts to give them an opportunity to verify and clarify their points of view? Will you use software to help manage your data and increase rigour? Will you conduct member checks (have more than one researcher analyse sections of the data to compare and verify results (called inter-rater reliability)? Will you triangulate the data to increase the rigour? Will you report disconfirming evidence?

Validity and reliability in analysing qualitative research

In quantitative studies, reliability means repeatability and independence of findings from the specific researchers generating those findings. In qualitative research, reliability implies that given the data collected, the results are dependable and consistent.[10] The strength of qualitative research lies in validity (closeness to

truth). Good qualitative research, using a selection of data collection methods, should touch the core of what is going on rather than just skimming the surface. When analysing your qualitative data, look for internal validity, where an in-depth understanding will allow you to counter alternative explanations for your findings.

Analysis of textual material

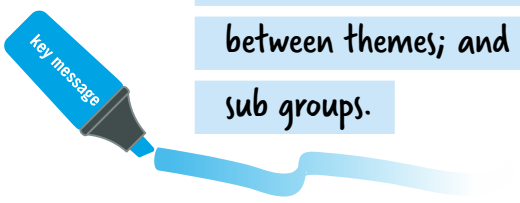
The basic process for the analysis of text derived from qualitative interviews or discussions is relatively straightforward and includes:

- Identification of similar phrases, themes and relationships between themes.
- Identification of similarities and differences between population sub-groups (e.g. men/women, rural/urban, young/old, richer/poorer, etc.).
- Initial attempts to generalize by identifying consistent patterns across or within sub-groups.
- Critical review and revision of generalizations, paying particular attention to contradictory evidence and outliers.

As far as possible, outputs of focus group discussions (FGDs) should be verbatim records. The notes taken by the recorder (a person) should be compared to a recording of the discussion. The recorder and moderator should agree on a final transcript. The transcripts (from multiple FGDs) should provide the material for systematic analysis.

FGD analysis will typically address a number of specific research topics and sub-topics, such as eliciting additional topics of local concern, which can be used to define the broad domains for analysis. These can be sub-divided further into themes, sub-themes, etc. and allocated systematic codes.

The initial descriptive analysis should also capture: (i) most common themes mentioned; (ii) less common themes; (iii) common associations between themes; and (iv) similarities and differences between sub groups.



Domain/theme analysis

One relatively simple approach is based on the identification of key topics, referred to as 'domains,' and the relationships between them.

There are four stages in domain/theme analysis:

A. Identify main issues raised by the interviewees – the domains /themes.

B. Group more detailed topics within each of these domains to construct a taxonomy of sub-categories.

C. Specify what was actually said and the components within each sub-category.

D. Explore the of interrelationships between the various domains.

A. Domain/theme identification

- Index texts, identifying topics line-by-line.
- Collate these topics across all interviews to identify a preliminary list.
- Some will recur more frequently than others and some of the latter can be classified as sub-topics.
- Systematically combine related topics to develop a list of just a few fairly broad domains.

After listing the domains, it is useful to start arranging the actual segments of text into the primary domains. This process groups actual phrases together and allows the sub-categories to emerge directly from the interviewees' own words.

Example[2]

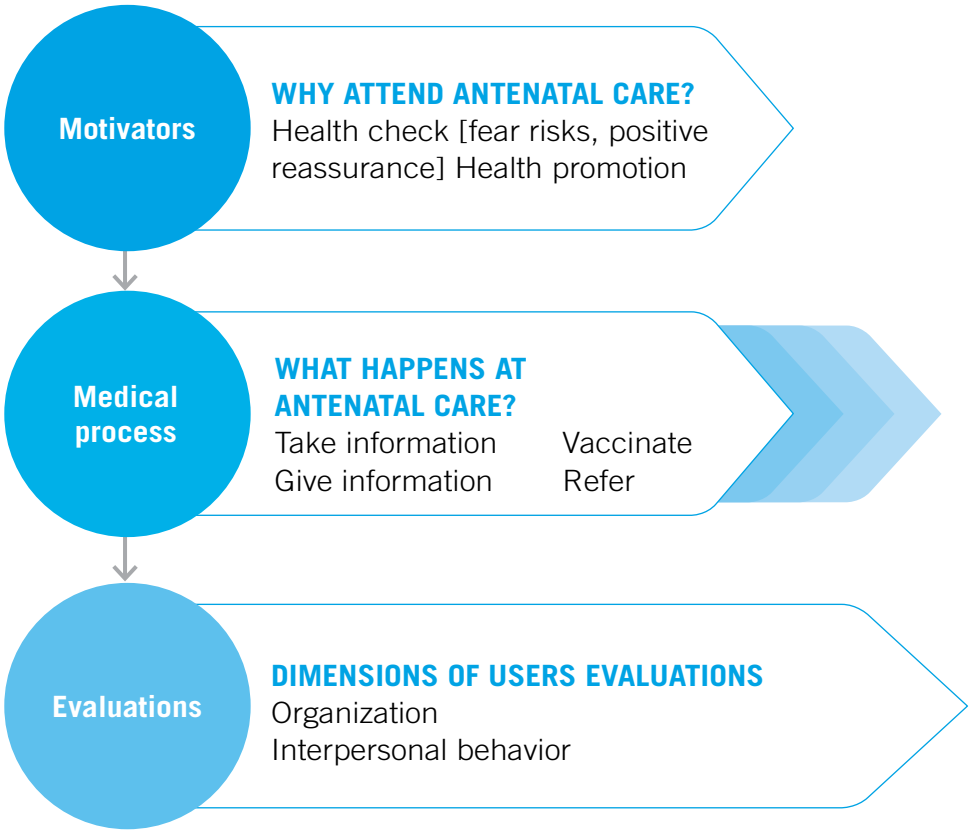
- ***Getting and being pregnant:*** *Signs of pregnancy, danger signs, physical problems.*
- ***Feelings during pregnancy:*** *Anxiety, anger/fright, worries, embarrassment, inconvenience, impressions.*
- ***Family planning:*** *Methods.*
- ***Advice/activities to promote health:*** *Exercise, activities, smoking, self-care, advice sources, information sources.*
- ***Birth and miscarriage:*** *Previous experiences, place, signs, caesarean/normal, birth weight.*
- ***Antenatal care:*** *Staff, place, experiences, meetings, tests, distance/cost, logistics, waiting time.*
- ***General background:*** Family, employment, geography.

From the above example, the following broad domains were identified:

- Motivations for antenatal care.
- Medical process (experiences of antenatal care and evaluation of that care).
- Risks during pregnancy.
- Reproductive histories.
- Socioeconomic background.

These domains are represented in Figure 8.

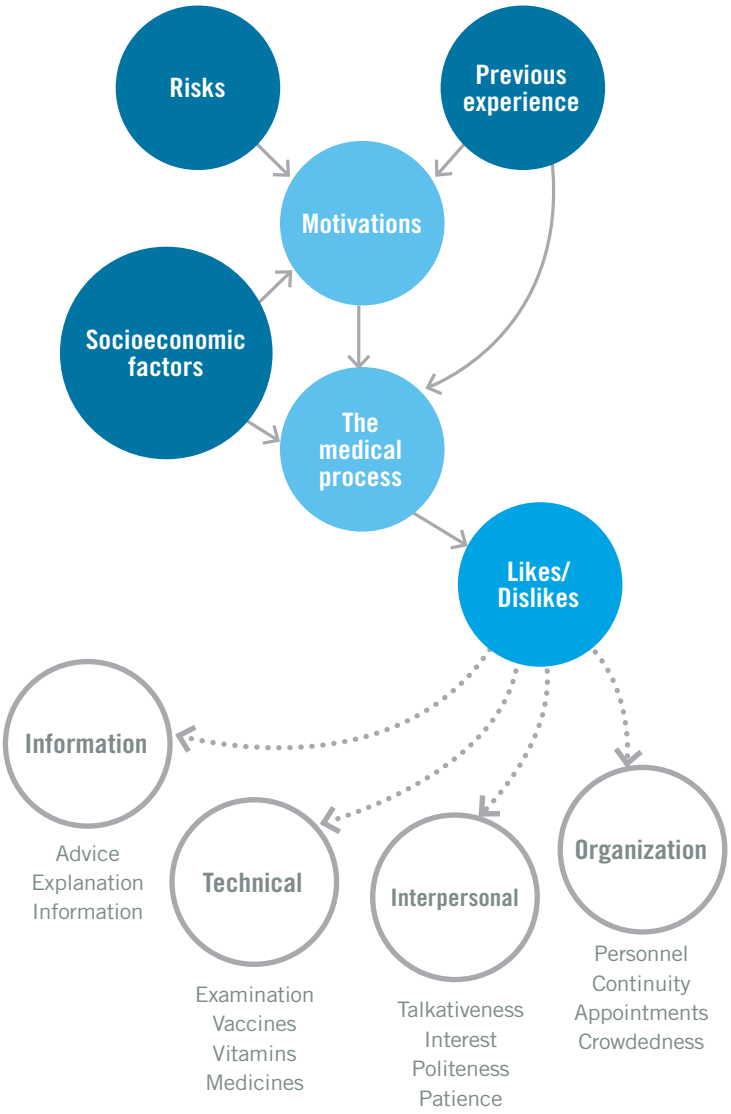
Figure 8: Taxonomy of sub-categories[2]



B. Relationships between domains/themes

This stage involves identifying relationships between the domains or topics to build up an overall picture. Within the collection of actual quotations from respondents, the researcher should identify statements that relate one topic to another. For example, in the study described above, researchers were able to establish associations between the domains that linked women's previous experiences, risk perceptions and socioeconomic situation and their evaluations of health services (Figure 9).

Figure 9: Relationship between domains[2]



Coding schemes

Following an initial analysis to gain an overall understanding of the main features of the data, many analysts apply a systematic coding procedure. The researchers determine the most appropriate way to conduct a systematic analysis, uncovering and documenting links between topics, themes and sub-themes.[23] These codes are then assigned to specific occurrences of words or phrases, highlighting patterns within the text while preserving their context, as in Table 19.

Table 19: Matrix of perceived causes and signs of malaria

Focus group discussion				
	Village A women	Village A men	Village B women	Village B men
Malaria signs	Hot body Yellow eyes White lips	Bloody stool Hot body Yellow Eyes	Hot body White lips Yellow eyes Bloody Stool	Hot body Yellow eyes White lips
Malaria causes	Mosquitoes Fresh mangoes	Mosquitoes Standing in the heat Fresh mangoes	Mosquitoes Standing in the heat	Eating fresh mangoes Mosquitoes

Description of qualitative data analysis

Transcripts from key informant interviews and group interviews will be coded and analysed according to emerging themes using Ethnograph software for qualitative analysis. Data will be reported in the form of narratives or frequency tables in addition to standard thick ethnographic descriptions.

Coding of focus group interviews, ethnographic field notes and interviews with health workers using Atlas-TI software will enable the analysis of emerging themes and presentation of data in the form of narratives or frequency tables.

Transcripts from life histories will be coded and analysed according to emerging themes (Ethnograph or Atlas-Ti software). Data will be reported in the form of narratives or frequency tables. In addition, videotaped recordings of patients will be used for national and international advocacy with the permission of interview subjects. Semi-structured, open-ended interviews with patients and family members of patients will be coded and reported as narratives or frequencies of coded responses to better understand the impact of the persistence of MDR-TB in this setting.

Mixed methods data analysis

In a mixed methods IR project, demonstrating how scientific rigour will be ensured throughout your study is critical. It is important to examine the validity (i.e. being able to draw meaningful inferences from a population) and reliability (i.e. stability of instrument scores over time) of the quantitative data.

To ensure qualitative validation, the researcher will use a number of strategies. First, opportunity will be provided for the participants to review the findings and then provide feedback as to whether the findings are an accurate reflection of their experience. Second, triangulation of the data will be used from various sources (transcripts and individual interviews) and from multiple participants. Finally, any 'disconfirming' evidence will be reported. This is to ensure that accounts provided by the participants are trustworthy.

Before beginning the analysis, consider how the mixed method study was designed. Refer to Table 7 on mixed methods approaches to review the order in which data was collected. This will guide the process indicating which data (qualitative or quantitative) should be analysed first.

One of the important aspects of mixed methods analysis is the capability in the presentation of these data to have the different methodologies 'speak' to each other. For example if the quantitative survey results show that 45% of mothers do not attend antenatal services, adding a direct quotation from a mother collected in a FGD will add a real-life and tangible element to this result.

Data presentation for your audience

When working through the analysis of the data collected in the IR project, it is important to remember who will receive the results of the research. This will determine how the research findings are presented. For example, if the results are disseminated in community meetings, it is important to use simple infographics and quotations or stories; in contrast during a workshop style meeting with high level policy-makers, more detailed information and numerical explanations will be required. This is dealt in more detail in the Communications and advocacy module of this Toolkit.

SEE COMMUNICATION & ADVOCACY MODULE

References

1. Bowen J.D. et al. How we design feasibility studies. American Journal of Preventive Medicine. 2009; 36(5):452–457.
2. Eldridge S. Pragmatic trials in primary health care: what, when and how? Family Practice. 2010; 27:591–592).
3. Roland M. and Torgerson D.J. Understanding controlled trials. What are pragmatic trials? BMJ. 1998; 316:24.
4. Mdege N.D. et al. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. Journal of Clinical Epidemiology. 2011; 64:936–948.
5. Brown C.H. et al. Adaptive designs for randomized trials in public health. Annual Review of Public Health. 2009; 30: 1–25.
6. Kilbourne A.M. et. al. Protocol: Adaptive Implementation of Effective Programs Trial (ADEPT): cluster randomized SMART trial comparing a standard versus enhanced implementation strategy to improve outcomes of a mood disorders program. Implementation Science. 2014; 9:132.
7. Blanchet K.B. et al. A mixed methods protocol to evaluate the effect and cost-effectiveness of an integrated electronic diagnosis approach (IeDA) for the management of childhood illnesses at primary health facilities in Burkina Faso. Implementation Science. 2016; 11:111.
8. Amaratunga A. et al. Quantitative and qualitative research in the built environment: application of "mixed" research approach. Work Study. 2002; 51(1): 17–31.
9. Krefting L. Rigor in qualitative research: the assessment of trustworthiness. The American Journal of Occupational Therapy. 1991; 45(3):214–22.
10. Lincoln Y.S. and Guba E.G. Naturalistic inquiry. Newbury Park: Sage Publications, Inc.
11. Bryman A. (2006). Integrating quantitative and qualitative research: how is it done?' Qualitative Research. 1985; 6:97–113.
12. Greene, J.C., Caracelli, V.J. and Graham, W.F. Toward a conceptual framework for mixed-method evaluation designs. Educational Evaluation and Policy Analysis. 1989; 11(3):255–74.
13. Baqui A.H.et al. Effect of community-based newborn-care intervention package implemented through two service-delivery strategies in Sylhet district, Bangladesh: a cluster-randomised controlled trial. Lancet. 2008; 7: 371(9628):1936–44.
14. Baqui A.H. et al. Effect of timing of first postnatal care home visit on neonatal mortality in Bangladesh: an observational cohort study. BMJ. 2009; 339:b2826
15. Choi Y. et al. Can mothers recognize neonatal illness correctly? comparison of maternal report and assessment by community health workers in rural Bangladesh. Tropical Medicine and International Health. 2010; 15: 743–753.

16. Shah R.et al. Community-based Health Workers Achieve High Coverage in Neonatal Intervention Trials: A Case Study from Sylhet, Bangladesh. Journal of Health, Population, and Nutrition. 2010; 28(6):610–618.

17. A useful list of biostatistics resources is available here: http://libguides.lib.msu.edu/biostatistics (accessed 17 October 2017).

18. Atkinson S. and Abu el Hap M. Domain analysis for qualitative public health data. Health Policy and Planning. 1996; 11(4): 438-42.

19. Surkis A. and Read K. Research data management. Journal of the Medical Library Association: Journal of the Medical Library Association. 2015; 103(3):154–156.

20. Guideline for good clinical practice. International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use (1996). (https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf, accessed 17 October 2017).

21. Walker N., Bryce J., and Black R.E. Interpreting health statistics for policymaking: the story behind the headlines. The Lancet. 2007; 369(9565): 956–963.

22. Bradley E.H., Curry L.A., and Devers K.J. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Services Research. 2007; 42(4):1758–1772.

23. Miles M.B., and Huberman A.M. Qualitative data analysis: An expanded sourcebook. Beverly Hills: Sage Publications, Inc.; 1994.

24. Merriam S.B. Qualitative research and case study applications in education (2nd ed). San Francisco: Jossey-Bass; 1998.

25. Maxwell, JA. Qualitative research design: an interactive approach (3rd Edition, Vol 41). Los Angeles: Sage Publications, Inc.; 2013.