

Session 2 Management Information Systems 60 minutes

Step 1: Re-introduce participants to the definition of monitoring and remind them of the earlier discussion that monitoring can be either a periodic or a day-to-day data collection of information about the programme. Explain to the participants that the information on monitoring activities is usually kept in a management information system (MIS) which is a part of the monitoring system. Ask participants to describe what a management information system is. Have the co-facilitator write their responses on a flip chart. Make sure the following points come out:

A Management Information System is a strategy and method for collecting, storing, organizing, and reviewing (using feedback) monitoring information for the purpose of management decision making.

An MIS is one tool that programme managers have at their disposal to enter, process, and review information about programme activities and they use it to make decisions on the programme in general. As such, an MIS is part of a monitoring system.

An MIS can exist at many sites and levels within a programme. For example, a given nutrition programme can collect, manage, and act on information at the community or district level. This information can also be analysed, reviewed, and used for decision-making purposes at higher levels. An effective MIS will also have mechanisms in place at all levels for feedback to be given and received on such decisions.

An MIS can help to improve the quality (adequacy and accuracy) of data collected. It is one way to develop a reporting system that identifies information flows and reporting levels, by increasing ease of access to information. It helps to institutionalize feedback.

Step 2: Ask participants to give examples of the MIS or monitoring data system from their own programmes and share how the data are used and how they think the system is important.

Step 3: Explain to participants that they are going to do a **SWOT** analysis. Point out that a **SWOT** analysis is a tool to help programme managers to think about the **strengths**, **weaknesses**, **opportunities** and **threats** of their MIS. Using a flipchart, ask participants to identify the strengths of their MIS and list these in the “Strengths” quadrant of a **SWOT** diagram. The following points should come out:

The MIS exists
Can collect and store data
Personnel is there
Different forms available to collect different types of information
Can be used to reach the lowest level

Then do the same for the Weaknesses. The following points should come out:

Information may not be specific/categorized
Too much upward, one-directional information flow
Reports are not valid
Gaps in information
Data not analysed and used
Data not timely
People who collect data don't understand the need to do so
No demand for the data/information
Too much data

Inform the participants that the methods used for collecting data determine how it is analysed. This will be covered in later sessions.

Step 4: Ask participants to complete the SWOT diagram by doing the same for Threats and Opportunities. Inform participants that weaknesses can be turned to strengths and threats to opportunities.

Step 5: Conclude this session by displaying **Transparency 4.2** on programme components and discussing the various activities at each level of a programme, and telling participants that when designing a management information system they need first to determine:

- What information will be most useful for programme management?
- What information will be most useful to programme implementors (at different levels)?
- What information will be readily available through programme implementation?
- What added information needs to be collected as part of programme monitoring?
- Which indicators can be realistically monitored by programme staff?

- How will the information be stored and retrieved?
- What will be the most appropriate/effective feedback channels?

Session 3 Population and Sampling 150 minutes

Step 1: Hang a card on the wall with the word “Population” written on it. Ask participants to define it.

A **population** is a group of individuals inhabiting a specific area or sharing specific characteristics.

Step 2: Ask participants what kind of population characteristics we are interested in for nutrition programme evaluations. Some examples are:

- height
- weight
- Hb or anaemia prevalence
- growth rate
- < 2 years - vegetable consumption
- income/expenditure

Step 3: Hang a second card on the wall with the word “Sample” written on it. Ask participants to define the term sample. The following should come out:

A **sample** is a subset of any category of stakeholders who represent the entire group of stakeholders or population.

Step 4: Ask participants what they think makes a good sample. The key concept to come from the discussion is **representativeness**.

Step 5: Ask participants if sampling is important for evaluation design, for monitoring, or both. Ask for examples of when sampling can be important for monitoring (e.g., special studies).

Step 6: Ask participants to say what determines how large a sample is necessary. Three concepts should emerge:

- **Representativeness:** the larger the sample, the more likely it is that the sample represents the population (small samples can, by chance, be unrepresentative).
- **Comparisons:** if making comparisons about populations, the sample needed from each will be larger than needed to estimate just one population.
- **Differences:** the larger the differences between populations, the smaller the sample needed to make conclusions about

the differences.

Step 7: Explain to participants that there are two mistakes that can be made when making conclusions about differences between populations. Point out that one error is to say that the populations are different when they are not and that the other possible error is that the populations are not different when they are.

Ask participants what types of populations we would want to compare in the evaluation of a nutrition programme.

Step 8: Point out that both of the two possible errors become less likely as the sample gets larger. Thus, we can be more confident about conclusions from larger samples.

Step 9: Inform participants that computer programmes such as Epi-Info are available that can determine the minimum necessary sample size based on how:

confident they feel they must be about conclusions concerning differences between populations;

big a difference they expect.

(If these computer programmes are available, offer to demonstrate how they work).

Step 10: Ask participants how they know the size of the difference to expect. Allow several to suggest possibilities, but ensure that the following points come out:

- other similar projects may have reported their effects;
- a review of the literature may help us guess how much effect we can expect from an intervention similar to the one we will evaluate;
- we also can determine what is a **practical** difference that we consider worthwhile based on the resources and effort planned for the programme.

Step 11: Divide participants into 4 groups and distribute **Handout 4. 4**. Assign each group one scenario. Provide each group with different types of beans which are appropriate for and depict the scenario. Ask each group to determine what type of sampling procedure they would use and to explain

why they would do so. Allow about 25 minutes for this activity and then share the groups findings in plenary.

Step 12: To conclude this session, distribute **Handout 4.5** on sampling procedures and ask participants to read through it. Answer any questions they may have about sampling and sampling procedures.

Session 4 Preparation for Implementation of Monitoring and Evaluation Activities 120 minutes

Step 1: Remind participants that monitoring and evaluation activities are an integral part of programme management. They need to be in-built in the design and plan of action. Inform the participants that the monitoring and evaluation team should be in place and indicators selected. In addition, there should be a management information system in place and the evaluation design should be selected.

Step 2: In groups of three, ask participants to brainstorm and identify all the tasks that are involved in the preparation for conducting monitoring and evaluation activities. Allow 15 minutes for this activity and then let the participants share their responses in plenary and write them on the flipchart. For monitoring, the following points should come out:

- setting up an administrative system for monitoring;
- planning of logistics;
- develop work plan/time line;
- mobilize the available financial resources;
- work out additional budget;
- mobilize the human resources required;
- mobilize all the needed material supplies, including computers, stationery, and equipment;
- arrange supervisory/field visits;
- organize transport;
- design a dissemination strategy for the findings;
- organize an effective feedback system.

For evaluation, the following should come out in addition to those that have been mentioned for monitoring:

formulate the terms of reference;

arrange for the introduction of the evaluation to the appropriate authorities.

Step 3: Explain that once all of the tasks have been identified, it is necessary to prepare a work plan. Show **Transparency 4.3** and walk participants through it. Point out that a work plan is a schedule, chart or graph that summarizes various components of monitoring and evaluation systems and how they fit together. It includes the tasks to be performed, when they will be performed, who is responsible for carrying them out and how much time will be spent on each task.

- Step 4:** Emphasize that a work plan should be realistic and that changes can be made when the need arises. It should cover preparation, training, implementation, data analysis, reporting and dissemination of results. The realities of local customs (holidays and festivals) and working hours should be considered.
- Step 5:** Explain that one important aspect of planning monitoring and evaluation activities involves the preparation of a budget. Let the participants brainstorm on the budget items and costing of the activity. Distribute **Handout 4.6** which provides information on the categories to be included in a budget.
- Step 6:** Divide participants into four groups and ask two groups to plan the logistics, draw up a work plan (start by listing the tasks to be performed in sequence) and draft a budget for a monitoring system of a hypothetical programme or their own programme. Ask the other two groups to do likewise for an evaluation system. Allow about 45 minutes for this and then share the groups' reports in plenary.
- Step 7:** Summarize this session by reminding participants that their work plans should be flexible and should be changed if the need arises.

THREATS TO VALIDITY THROUGH STUDY DESIGNS

History: The effects that are not part of the project/issue of intervention as they are not planned or anticipated events. They just happen and they produce an effect that influences the study results. It is a common threat to validity in intervention studies.

Selection: Where the units in the control group differ completely from those in the experimental group. Or the self-selection of groups in surveys; the selection of the day of the study + recall.

Testing effect: Common where a pre-test is given, and tends to have effect on the post-test. Or in longitudinal surveys, where the same people are repeatedly asked the same questions over a certain time period. After a while, people begin to remember the correct answers and their responses are due to their familiarity with the questioning tool and not their actual state.

Instrumentation effect: These are effects resulting from changes in the methodology, or equipment, or the way questions are asked, such that the changes in the way the information is asked or collected (as in the case of interviewers being more experienced) result in a threat to the validity of the findings; usage of the instrument (reading of meniscus + weighing scale [zeroing]).

Maturation effect: In longitudinal studies extending for a period, people (trainers and respondents) become older, bored, hungry, wiser, discouraged, resistant, or tired over time and this may cause the first findings to be different from latter ones.

Mortality/drop-out effect: Again in longitudinal/cohort studies, differences may be due to differential loss of cases between comparison groups.

OTHERS INCLUDE:

Accuracy of the instrument (design of the questionnaire; accuracy of the equipment; recall period).

Error on the part of the enumerator: recording problems, i.e., of the wrong code for a correct response.

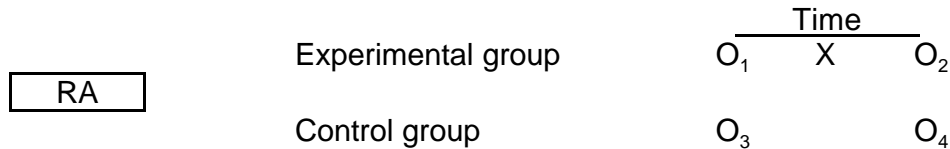
These threats must be carefully considered and controlled when designing research studies otherwise the effects may mar the conclusions.

TYPES OF STUDY DESIGNS

Three kinds of designs shall be discussed: experimental designs, non-experimental designs and quasi-experimental designs.

1. *Experimental Designs*

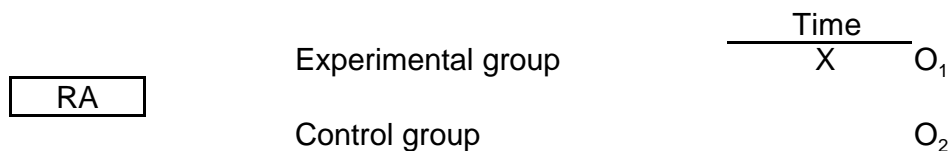
True Experimental Design



In this design, all subjects are randomly assigned (RA) from a single population to the experimental group. Both the experimental and control groups receive an initial measurement observation (the pretests O_1 and O_3). The experimental group then receives the intervention (X) but the control group does not receive the intervention. Finally, a second set of measurement observations are made (O_2 and O_4). We would expect that, since the experimental group received a special programme intervention (X), O_2 would be greater than O_4 . Also, since both the experimental and control cases were randomly assigned, we would expect that O_1 would be equal to O_3 on such key variables as age, sex, and education.

RA denotes randomization and ensures that the two study groups are equal on main baseline variables before the start of the intervention. Thus any difference observed between O_2 and O_4 are basically due to X (the intervention or the test variable). This is one of the strongest designs in controlling for the threats of validity. Randomization to intervention and control groups is sometimes impossible from a practical standard point since it may either be ethically or politically incorrect to deny one group the programme while giving it to another.

Post-test-Only Control Group Design



This is also a true experimental design, except there is no baseline (pre-test) measurement observation. Since cases have been assigned randomly to the experimental and control groups, these groups are assumed to be similar before the programme intervention. This design allows the investigator to measure the effect of a

programme intervention on the experimental group by comparing that group with the controls. But it does not allow the investigator to determine the extent or magnitude of the change within the experimental group because no baseline measurement was taken.

2. Non-Experimental Designs (NEDs)

There are several non-experimental designs used by evaluators. These designs are most appropriate for collecting descriptive information or for doing small case studies of a particular situation. They are not recommended for evaluation studies that attempt to determine the effect of a programme intervention, but they may be useful in diagnostic studies to determine the reasons why a problem or a success exists.

A Post-test-Only Design



An intervention X has already taken place for a certain duration after which measurement O_1 is made. Since a control group is not available or a pre-test measurement was not made there is no possibility of comparison. *All that measurement O_1 can do is provide descriptive information.* The threats to validity of history, maturation, selection, and mortality are not controlled and therefore are factors to consider. Multivariate data analysis techniques can be used if comparative analysis is desired.

Pre-test-Post-test Design



In this design, there is no control group but at least there is an earlier measurement that can be used to make comparison and examine changes over time. However, the pre-test-post-test design is subject to several threats to validity including history, testing, maturation, and instrumentation.

Static-Group Comparison



Unlike the other two design, this one adds a control or comparison group **BUT NO**

randomization in allocation of groups. Design can be used to compare patients in one clinic with those from another clinic on their knowledge on how to appropriately prepare weaning foods and feed children. Threats to validity are selection and mortality since there might be a difference in baseline variables and the fact that those given measurement O_1 are those who have remained in the group initially after having been given the intervention.

3. Quasi-Experimental Design

Time Series Design

These designs do not have the costly restrictions of random assignment but they tend to control for many threats to validity [a compromise between experimental and non-experimental designs].

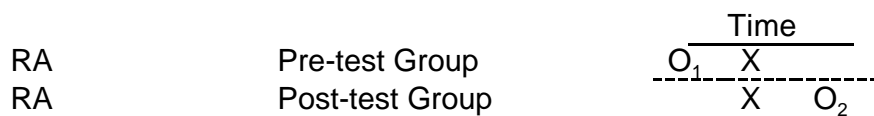
The time series design is similar to the non-experimental pre-test-post-test design, except that it has the advantage of repeated measurement observations before and after the programme intervention (X). The multiple measurements give us a trend. The best design is when you have access to regularly collected information, such as monthly service statistics. You can plot information on a graph and note the point at which the intervention occurred. A sudden change at the time of the intervention or some time after (Action Period) is likely to be associated with the intervention. The design does not control for history and possible instrumentation threats to validity. However, it allows for a more detailed analysis of data and program impact.

Non-equivalent Group Design



In this situation, we use two pre-tests (O_1 and O_3) to assess the extent to which the two groups were truly similar. Then we would compare the two post-tests (O_2 and O_4). We would expect that O_2 to be greater than O_4 because of the effect of the intervention. Useful in evaluation of programmes and for comparing villages and classes in school.

Separate Sample Pre-test-Post-test Design



The separate sample pre-test-post-test design is a frequently-used design in programming. It involves a baseline pre-test (O_1) with a randomly selected sample from the study population. Subsequently, a programme intervention (X) is introduced, and then a post-test (O_2) is made using a second selected sample from the same population. The design, however, does not control for history, maturation, mortality or possibly instrumentation effects.

Selecting an appropriate evaluation design can be tricky! It usually involves a careful consideration of ethical issues and a balancing of technical issues against practical and administrative issues.

Ethical issues. The first issues to consider in selecting an evaluation design are the ethical ones. If a particular evaluation design results in unethical procedures, a violation of people's rights and dignity, or a denial of services that otherwise would be available, then the design should be modified or abandoned regardless of the effect this may have on reliability, validity, time, funds, or available personnel. Indeed, if it is not possible to do an ethical study, then the study should not be done. There is no compromising on this point.

Practical and administrative issues. Most often, funds are in short supply, time is short, and personnel is few in numbers. These conditions obviously affect the choice of an evaluation design.

Technical issues. Some of the important technical issues you should keep in mind are:

- whenever possible, try to create experimental and control groups by assigning cases **randomly** from a single population study group;
- when random assignment is not possible, try to find a comparison group that is as **nearly equivalent** to the experimental group as possible;
- when neither a randomly assigned group nor a similar comparison group is available, try to use a **time series design** that can provide information on trends before and after a programme intervention;
- if a time series design cannot be used, as a minimum and before a programme starts, try to obtain **baseline (pre-test) information** that can be compared against post-programme information (a pre-test-post-test

design);

- always keep in mind the issues of validity. Are the measurements **true**? Do they do what they are supposed to do? Are there possible **threats to validity** (history, selecting, testing, maturation, mortality, or instrumentation) that might explain the results?

CHARACTERISTICS OF EXPERIMENTAL EVALUATION DESIGNS

Control groups:

Creation of a control group that shares the characteristics of the participant group permits the conclusion that any changes observed in the project group and not in the control group can be attributed to the project.

Control groups are important for demonstrating positive project effects in situations of deteriorating nutrition status, e.g., nutrition status declining even more during a drought in a control group compared to in a participating group.

Randomization to treatment:

Valid comparisons are possible when project participation is the only difference between participant and control groups and the only certain way to ensure that no differences exist between the groups is to randomly assign individuals to either participant or control groups.

Although random assignment is often an unrealistic (and sometimes unethical) choice for field-based nutrition projects, a valid comparison group can be found from a comparable area where the project has not yet begun activity. This, however, increases the possibility for error in the form of *bias* or *confounding*.

Pre/post analysis:

Baseline measurements determining the pre-intervention status for selected indicators are compared with *follow-up* measurements taken either during project implementation (for a mid-term evaluation) or upon project completion (for a summative evaluation).

Pre- and post-project information is necessary to demonstrate *if and to what extent change has occurred*. However, when measuring the magnitude of change, it cannot be assumed that pre-project information (a nutrition survey or needs assessment) can necessarily double as a baseline survey unless the data collected includes each of the relevant indicators, is geographically desegregated (project and control area are separate), and is followed immediately by the initiation of project services.

Group 1

As part of the programme to improve adolescent girls' micronutrient status through school-based interventions, a baseline will be performed to determine the causes of anaemia. The analysis costs for such an intensive study are high (e.g., blood analysis for malaria and hemoglobinopathy, stool analysis for parasites, dietary intake), and the funds available allow this comprehensive analysis among only 15 girls. One baseline will take place in one school, which has 60 adolescent girl students. What kind of sampling would you use to select the 15 girls for the anaemia aetiology study? Why?

Group 2

As part of the programme to improve adolescent girls' micronutrient status through workplace-based interventions, baseline prevalence of anaemia will be measured. Stakeholders are concerned that the intervention will not be effective among young girls (under 16 years), so assessment of the programme's effects among these young girls will be important. In the factories where the programme will intervene, 5% of the girl employees are younger than 16 years. What kind of sampling would you use to ensure that girls under 16 years are adequately represented in the analysis? Why?

Group 3

The programme intervention and comparison areas for the programme conducting growth monitoring and promotion in conjunction with food security improvements are very large, with rugged terrain and poor road communication. Additionally, population density is very low. Nonetheless, for proper evaluation of the programme's effects, a baseline assessment of household food security status is necessary from each area. How would you sample to obtain an estimate of food security status in the project and comparison areas? Why?

Group 4

In the programme addressing low birthweight, stakeholders will want to know if the reasons why women attend ante-natal care clinics change as a result of the behavioural change communication the programme will provide. A baseline assessment of women attending the clinics in the programme area and in the comparison area thus will be necessary, but survey costs will be high due to the amount of time needed to interview each woman to learn her reasons for attending. Thus it has been decided to conduct the baseline assessment on only 10% of women attending the clinics. How would you obtain a sample of 10% of women attending the ante-natal care clinics? Why?