Design, Monitoring and Evaluation Guidebook Supplement, Part 1

A resource to be used alongside
the original DM&E Guidebook.

MercyCorps

Table of Contents

Introduction

This handbook has been developed as a companion resource to the *Design, Monitoring and Evaluation (DM&E) Guidebook*, providing detailed information on how to implement a mix of common M&E tasks. The chapters in this handbook were written as stand-alone resources for the Mercy Corps online toolkit, DM&E-in-a-Box, and additional tip sheets and template tools can be found in this toolkit on the MC Digital Library. Please send any comments or suggestions on how to improve this set of resources to the DM&E Initiative at dme@mercycorps.org.

Acknowledgements

Key Steps to Effective M&E

1) Budget for M&E – It is recommended to allocate approximately 10% of the total project budget for M&E. This should include costs for staff, assessments, baselines, implementation of monitoring systems and evaluations.

2) Staff for M&E – Incorporate M&E skills in all of your project positions, look for M&E skills in all of your project staff, and consider project- or country-level M&E positions.

3) Incorporate M&E in project Workplan – The workplan should include critical M&E processes such as data collection, reflection sessions with project staff, dissemination activities and obtaining feedback from project beneficiaries and other stakeholders.

4) Conduct a DM&E Workshop at project start-up – Often project implementation occurs months after project design. Conducting a DM&E workshop at start up with project staff provides an opportunity to review and update the log frame to make sure it reflects current realities.

5) Develop an Indicator Plan – Indicator Plans help to identify the critical details on how staff will capture key output and objective-level data, which is the foundation of the project's M&E system.

6) Develop Data Collection and Management processes – Focus on simplicity and utilization, with clear staff roles and responsibilities.

7) Conduct regular meetings to Reflect on M&E data – Emphasize learning and build feedback loops into the project so that lessons learned from M&E data are used to improve implementation and ultimately project outcomes.

8) Make the Logframe a living document – Be sure to measure progress against targets and refine the logframe as the project evolves.

9) Report project results to Beneficiaries and other Stakeholders – We need to be accountable to the communities in which we work. Reporting shouldn't just happen upwards.

10) Conduct Baselines and Final Evaluations – Capture results, intended and unintended outcomes, and key programmatic lessons for organizational learning.

For more information and guidance on DM&E issues, consult the Mercy Corps DM&E Guidebook, DM&E in a Box on the Digital Library, the DM&E space on MC's Clearspace site (http://clearspace.mercycorps.org), or the DM&E Initiative at dme@mercycorps.org.

M&E During Proposal Development

Guide to M&E Sections in Proposals

This chapter is intended to help guide field and HQ teams in developing monitoring and evaluation (M&E) sections for project proposals. A description of key elements common in M&E sections can be found in the first few pages, with examples from Mercy Corps proposals presented at the end of the chapter.

M&E sections vary according to proposal context

The length and contents of the M&E section will depend largely on the donor and the proposal format. These can differ substantially; for example, some donors require a Performance Management Plan (PMP) or Indicator Plan while others do not, and some M&E sections will be just a brief paragraph or two while others will merit several pages.

> **TIPS FOR USING THIS GUIDE:**
> - *Adjust* these recommended contents to fit your particular circumstances, including donor format and proposal/section length.
> - *Do not* cut and paste the examples verbatim into new proposals – each project is unique and M&E sections should be adjusted accordingly.

With this in mind, please adjust the content to fit your specific proposal. While many of the examples cited here are from longer, more complex proposals, we can condense each major element covered below into one or two sentences each for shorter proposals.

Key elements and placement of M&E within proposals

It is often clear within specific donor formats where the M&E section should fit. Other times, however, we have to use our own discretion. Placing it towards the end of the technical narrative, in a subsequent stand-alone section, or in annex are all appropriate depending on proposal length, format and individual stylistic preferences.

> **KEY ELEMENTS OF AN M&E SECTION:**
> - Introduction to project M&E
> - Monitoring narrative
> - Baseline and evaluation narrative
> - Roles and responsibilities/M&E system
> - Mercy Corps DM&E capacity statement
> - PMP/Indicator Plan (if necessary)
> - Optional: in-depth research plans

The logframe[1] can also be found in various places, but it is usually placed outside of the M&E section, either in the technical narrative or attached in annex. This reflects the fact that the logframe is central to project design, not just M&E. Remember that Mercy Corps policy requires every project to be designed using a logframe. If it is not an explicit part of the donor proposal format, we can often still attach it in annex or develop it for internal use. Donors can have different terms for similar tables, such as USAID's Results Framework. There are also a variety of logframe formats – see Annex I in the Mercy Corps DM&E Guidebook for a table that deciphers the various donor terms for logframe elements.

Introduction to project M&E

In the first sentence or paragraph of the M&E section (depending on how long and detailed we want the M&E section to be), we want to give an overview of what monitoring and evaluation will mean to the project and why it's important. It is a good idea here to explicitly refer to the logframe, PMP (if relevant), workplan, and any other key tools that will be relevant to project M&E. Please see item A at the end of this chapter for ideas.

Monitoring narrative

We can provide a sentence or paragraph(s) detailing the monitoring system we plan to put in place to track ongoing progress on activities and outputs throughout the life of the project.

This may include:

- Listing the key **indicators** to be monitored on a monthly or quarterly basis;
- Describing staff and stakeholder **roles** for data collection, entry and analysis;
- Any particular monitoring **methods** that are worth mentioning; and
- Any specific software or other **data management systems** to be developed or used.

Please see item C at the end of this chapter for examples.

Baseline and evaluation narrative

Most proposals will benefit from a narrative describing the baseline study, mid-term (if necessary) and final evaluations, as well as any other needs assessments or research to be carried out over the life of the project. This could take a few sentences or a few paragraphs for each activity, again depending on proposal format and desired length of the M&E section. Please see item D for examples of baseline and evaluation narratives.

---

1 See the *Mercy Corps DM&E Guidebook* on the Digital Library for more on logframes and program design.

BASELINE/EVALUATION NARRATIVE TIPS:

- In describing a needs assessment, baseline or evaluation, we should first discuss what the **purpose** will be, i.e. to attain baseline and evaluation measurements for certain outcome (objective-level) indicators, acquire more in-depth knowledge of target communities and populations, conduct a training needs assessment, etc.

- For the **baseline**, it may help to list specific indicators that will be measured. See which indicators in the logframe or PMP/indicator plan require surveys, focus groups, or other in-depth undertakings (look for those indicating a % change – often a good tip-off!). After listing these, describe the methods used to collect the data, if they are known. Lastly, we can discuss roles and responsibilities in the baseline/evaluation, including staff and local stakeholder participation, as well as plans for analyzing the results and incorporating findings into implementation strategy.[2]

- **Mid-term evaluations** are generally planned in projects lasting two years or more; if the project timeframe is less than this, we should seriously reconsider whether a mid-term evaluation is worth the time and effort and what lessons or impact we'll be able to learn after less than a year of implementation. If there is a mid-term evaluation, we can describe it here as a separate sentence or paragraph.

- Finally, we should describe the **final evaluation** in a similar way. As with the baseline, we should start by stating the purpose, i.e. to gain lessons learned and document impact by measuring key outcome-level indicators. Then describe some of the indicators we'll be measuring, methods used, staff/stakeholder participation, analysis plans, etc.

- Consider **time implications** of baselines and evaluations and make sure this is reflected in the implementation plan and other areas of the proposal narrative.[3] A typical baseline or evaluation will take anywhere from two weeks to two months, including planning, implementation and report-writing. Surveys with large sample-size requirements and/or covering wide geographic areas are particularly time-intensive.

- Be sure to also note whether we're planning for an **external or internal evaluation**. This often depends on donor requirements and available budget; note that evaluation consultants have become very expensive, and we need to **budget adequately** for daily rates, airfare, in-country costs, etc.

---

2 See the *Baseline/Evaluation Scope of Work Template & Sample* on the Digital Library for a further description of the main elements to highlight.
3 See *DM&E Tip Sheet #1: Planning a Survey* on the Digital Library.

Roles & responsibilities and M&E system narrative

Discussing how M&E will be organized within the project and who will take on what responsibilities, as well as links to country-level DM&E Units or structures where relevant, is also a good idea. It is recommended for larger projects to have dedicated M&E Officers, and this is a good place to help justify the budgetary[4] costs by outlining what that person and others would do regarding training, M&E systems development, data collection, entry and analysis, etc.[5] We can also mention roles of M&E Focal Points and other staff with M&E responsibilities.

In addition, evidence shows that high levels of participation among local partners, beneficiaries and other stakeholders increases local ownership and the prospects for sustainability. We can take this into consideration and build in language and concepts around participatory M&E as it is appropriate to the project. For more ideas see *DM&E Tip Sheet #14: Participatory M&E* in DM&E-in-a-Box.

We may also want to mention any particular **software or other architecture** used to store and communicate data. This could include a description of MS Access or other types of databases, as well as plans for more sophisticated Web-based applications like MS SharePoint, if appropriate. Finally, some types of programming, such as Village Savings and Loan Association models or psycho-social programming (World Bank), may have pre-developed M&E systems associated with them which would be worth explaining.

Mercy Corps DM&E capacity statement

A Mercy Corps DM&E capacity statement can also be included, and should be tailored to the specific proposal. Items to highlight include:

- Mercy Corps' **global DM&E capacity**, resources and experience
- Any **in-country DM&E efforts** or experience such as existing M&E systems, units or staff, capacity-building efforts, etc.
- Any current or past **DM&E experience within Mercy Corps** that is relevant to the particular proposed project intervention, including sector, operating environment, etc.[6]

A boiler-plate paragraph on agency DM&E capacity, as well as an example highlighting in-country efforts, are presented in item E at the end of this chapter. However, it is important to adapt these to highlight specific aspects of DM&E that the donor may be looking for in the RFA (e.g., strong tracking systems, impact assessment, etc.), as well as to illustrate how our

capacity is relevant to the particular project and its operating environment. Citing specific experiences such as M&E systems or evaluations of note in other similar Mercy Corps projects or countries is recommended.

Performance Monitoring Plan (PMP) or Indicator Plan

The PMP is a major component of most USAID proposals[7]. For non-US government donors, a similar indicator plan based on the Mercy Corps format (see the DM&E Guidebook) can be developed and included in the M&E section. This is recommended if we want to provide in-depth detail on how we plan to measure our indicators. However, in some cases the indicator plan can also be developed at project start-up, and therefore is not included in the proposal. This is often the case for shorter proposals.

The PMP or indicator plan is sometimes annexed and simply referenced in the M&E narrative, while other times it is a central part of the M&E section itself. It is a good idea to briefly introduce and describe the PMP or indicator plan in narrative form, before presenting the table or referencing it in annex.

Generally, the PMP or indicator plan is a table that lists all project indicators in the left-hand column, then includes columns to note the indicator definition and measurement specifications (units, disaggregation, etc.); data collection sources and methods; frequency of collection; person responsible; and targets for each indicator. Please see item B for examples.

PMP/INDICATOR PLAN TIPS:

- We've found it helpful to develop a draft plan in **Excel** rather than Word to allow more flexibility in adding and arranging rows and columns. Once finished, the Excel version can then be imported (cut and pasted) into Word.

- For guidance on **potential indicators** to use, please see the various sector-based Indicator Menus in section three of DM&E-in-a-Box, *DM&E at Design/Proposal Development Stage*, on the Digital Library. You may also want to draw on indicators from recent proposals in that sector and indicators used in the Mission Metrics initiative[8].

- Think about the **budgetary and time implications** of tracking each indicator. We sometimes commit ourselves to extensive evaluation activities without analyzing whether we have adequate funding and staff capacity to carry them out! If there are many objective-level indicators requiring surveys, and/or a fairly complex monitoring system envisioned, it's a good idea to invest in a full-time M&E position.

---

4 See *DM&E Tip Sheet #9: Budgeting for M&E* on the DL for further information.
5 Example DM&E position descriptions are available in DM&E-in-a-Box on the Digital Library.
6 For ideas on specific sector or country examples please contact the DM&E Initiative at dme@mercycorps.org.

---

7 See *USAID Tips: Preparing a Performance Monitoring Plan* for more useful guidance on USAID PMPs, available in DM&E-in-a-Box on the DL.
8 Search "Mission Metrics" on the DL or Clearspace for more information on this initiative and associated indicators.

- In the *Definition of the indicator* column, we should make clear the **unit of analysis** (i.e., individual, household, association, etc.) as well as any **disaggregation**, or sub-groupings, needed for the indicator (e.g. sex, age, etc.). Also try not to simply restate the indicator; the definition should be more **detailed description** of the indicator and any elements or wording that might be vague or unclear in the indicator itself.

- For the *Data collection sources and methods* columns, describe the **original source** of the information (e.g., women-led cooperatives, heads of household, monitoring records, etc.), as well as **methods of inquiry** (e.g., survey, focus group, case study, etc.).

- We should pay particular attention to the *Frequency* column, as this will define what is **monitoring** (ongoing tracking) versus what is **evaluation** (discrete data collection at baseline, mid-term and end-of-project). We sometimes tend to wish for high measurement frequencies, such as monthly or quarterly, which can become challenging to implement. For example, in-depth quarterly surveys should generally be avoided. Also, we should try to **standardize frequencies** as much as possible so that multiple indicators can be rolled into the same data collection activities and field reports.

- Be careful to thoroughly think through planned **targets**, and double-check these with staff on the ground in-country as well as with sector specialists where appropriate to ensure they are realistic.

- Include an adequate level of detail for *Person(s) Responsible.* If needed, we can consider writing out a more detailed narrative or process flow detailing the **roles and responsibilities** of various staff, so that it is clear who is responsible when the indicator is collected, inputted, analyzed, reported on and disseminated.

- Where relevant and appropriate to the context, look for ways for **partners and beneficiaries** to be involved in tracking indicators. We may need to adjust contracts or MOUs to reflect these roles.

- Be sure that the PMP or indicator plan is **consistent with the workplan** or detailed implementation plan. We often underestimate the time it takes to develop M&E tools and systems, and implement major activities like baselines.

Optional: in-depth research plans

If there are any plans for in-depth research or advanced impact assessment beyond the normal project M&E activities, such as collaboration with an academic institution[9] or additional field studies to explore particular issues, then we should also mention this in the M&E section as a separate paragraph or as additional verbiage within the elements above.

Additional Resources

- The DM&E-in-a-Box toolkit on the Digital Library includes example logframes and indicator plans, as well as other guidelines and DM&E examples.

- The Proposal Toolkit on the Digital Library is the fundamental resource for planning and developing a proposal, accessible from the DL's home page.

---

9 Lessons on collaborating with academic institutions: https://clearspace.mercycorps.org/message/2148#2148.

Examples

A. EXAMPLE INTRODUCTIONS TO PROJECT M&E

i) Example introductory paragraphs from MC Kosovo proposal:

"Annex 4: Draft Performance Monitoring and Evaluation Plan: Project monitoring and evaluation (M&E) will be a continuous and collaborative process conducted by Mercy Corps staff, LNGO partners, municipalities, communities, contractors and other stakeholders. The purpose of these M&E activities will be to track progress against indicator targets and inform managerial and strategic decision-making in order to make ongoing adjustments to implementation strategies and ensure maximum impact.

"The Logical Framework and Performance Management plan (PMP) presented below will provide the basis of the WISE project's M&E efforts. While the former provides a basic overview of program strategy by identifying the key objectives and their corresponding activities, outputs and indicators of success, the latter further defines the way in which these indicators will be measured by listing targets and standards for measurement to ensure clarity among all users. Along with the project Workplan, these tools will be used as the primary project management tools, expected to assist staff in obtaining the most accurate and up-to-date information possible to inform on strategy and decision-making processes."

ii) Example introductory statements from MC DRC proposal:

"F.1. Monitoring and Evaluation Plan: Drawing upon its extensive Design, Monitoring and Evaluation (DM&E) experience across many countries, Mercy Corps will implement a thorough monitoring and evaluation plan over the course of the three-year NKNP program. Through these activities, the NKNP will be able to determine whether the program is on track to meet targets, ensure that beneficiary input is used to guide ongoing implementation and determine how the program has impacted the lives of vulnerable Congolese."

* Note that both of these examples happen to emphasize a high level of participation of local stakeholders in M&E. We should only include such strong language if we are truly intent on investing the time, budget and staff required to build local capacity and develop participatory M&E systems.

B. EXAMPLE PMPS/INDICATOR PLANS

i) Example PMP from MC DRC USAID-MYAP proposal:

Performance Monitoring Plan (PMP) Tracking Table

Performance Indicator		Data Acquisition				Analysis, Use & Reporting	
Performance Indicator	Indicator Definition and Unit of Measurement	Data Source	Method/ Approach of Data Collection and Data Entry (if indicated)	Schedule/ Frequency	Respon-sible Person(s) & Team	Method of Analysis and Data Use	Reporting deadline
Objective 1: Increase agricultural production and safe access to healthy food							
1. Average # months adequate food provisioning	Definition: self-reporting by household on # months with adequate food. Measures food access. Defined by FANTA	Population-based household survey	33 x 7 cluster method	Baseline Midline End of Project	M&E Manager	Analysis: disag-gregated by com-munity and beneficiary status	Y1Q1 Y2Q1 Y3Q4
2. Household dietary diversity	Definition: number out of 14 food groups consumed in household over one week. Measures food access. Defined by FANTA	Population-based household survey	33 x 7 cluster method	Baseline Midline End of Project	M&E Manager	Analysis: disag-gregated by com-munity and beneficiary status	Y1Q1 Y2Q1 Y3Q4
Intermediate Result 1: Increased agricultural productivity							
1. # ag-ricultural associations targeted	Definition: # agricultural associations that have participated in at least one capacity building activity	Program records	Monthly reports of agriculture team leader, drawn from team reports	Quarterly	Agri-culture Coordina-tor	Analysis: disag-gregated by community	End of each quarter

ii) Example PMP from MC Uganda USAID-CMM proposal:

Performance Monitoring Plan

Result	Increased community and institutional capacity to mitigate conflict and advance peace and reconciliation in Pader District, northern Uganda				
Indicator	**Indicator Definition & Disaggregation**	**Est Target**	**Data Source**	**Frequency**	**Person Responsible**
IR1 Establish and/or strengthen 39 parish, 18 sub-county and 1 district-level mechanisms for conflict mitigation, peace and reconciliation					
1.1 % increase in parish and sub-county survey respondents reporting "strong" local mechanisms for conflict mitigation, peace and reconciliation	"Strong" measured through ranking scale: • Accessibility of local CR mechanisms (very, somewhat, not at all) • Utilization of local CR mechanisms (frequent, sometimes, infrequent) • Local CR mechanisms' success at resolving conflicts (usually, sometimes, rarely) Disaggregated by age, ethnicity, gender and sub-county / parish.	Baseline + 75%; to be confirmed upon baseline completion	Random, in-person survey of district residents	Baseline and program end	Program team (CoP, Forum Coordinator, Pader Peace Forum)
1.2 Level of funds secured by the Pader Peace Forum in proportion to its total 3 – 5 year operating budget	Measures likelihood of forum sustainability. Final budget against which targets will be measured will be set by COP, FC, PPF at month 6.	20% - mo.8 50% - mo.13 80% - EOP	PPF admin. records	Quarterly	Forum Coordinator
1.3 # peacebuilding structures established or strengthened with USG assistance that engage conflict-affected citizens	Total number of parish Peace Committees created and sub-county Peace Committees strengthened (trained) through program. Disaggregated by sub-county / parish.	108	Project mgt & training attendance records	Ongoing	Forum Coordinator

C. EXAMPLE MONITORING NARRATIVE

Example monitoring narrative from MC Uganda USAID-CMM proposal:

"Monitoring: Mercy Corps has set a monitoring and evaluation plan to ensure effective follow up of the progress towards expected results on a regular basis. This will be further discussed and refined at a program team M&E kickoff meeting during program, designed to: 1) ensure in-depth understanding of program objectives, indicators and activities; 2) review individual roles and responsibilities relating to M&E activities; 3) begin developing the monitoring forms, surveys, and tools for performance management, and finally, 4) begin baseline planning and tools development.

"The Forum Coordinator and PPF Steering Committee members (for sub-counties) and sub-county PC members (for parishes) will lead quarterly monitoring and technical support visits to all parish and sub-county program sites, to collect quantitative and qualitative data on actual progress on the ground. All visits will entail community meetings to ensure full transparency and accountability to district residents. Monitoring teams will record qualitative feedback and observations for each site in a simple reporting format and review and collect quantitative information on program outputs from PC records, such as:

• Number of people trained in conflict mitigation/ resolution skills with USG assistance

• Number of peacebuilding structures established or strengthened with USG assistance that engage conflict-affected citizens

• Number of community-based reconciliation projects completed with USG assistance

"Upon conclusion of each series of monitoring visits, the Chief of Party, Forum Coordinator, and PPF Steering Committee will meet to review results, plan follow-up technical support and make strategic adjustments to future activities. The Forum Coordinator will enter all quantitative data into a comprehensive program database covering all sites. Qualitative reports and follow-up action plans will be kept on file and utilized as the basis for all quarterly and annual program reporting to USAID."

D. EXAMPLE BASELINE AND EVALUATION NARRATIVE

Example baseline/evaluation narrative from MC DRC MYAP proposal:

"Baseline and Endline. Mercy Corps will hire an external evaluator to lead the FSG baseline and endline evaluations. Both evaluations will perform a population-based survey in the seven targeted areas, gathering information on households' current health, nutrition and socioeconomic status. Indicators largely collect data related to households' food access and utilization. While some baseline information, based on provincial or national data, has been filled in on the Indicator Performance Tracking Table (IPTT) (See Appendix 18), the majority of the baseline figures will be provided once the baseline evaluation has been conducted. The external evaluator will be assisted by the FSG Program's M&E team, consisting of three individuals. The baseline will be conducted in the first quarter of program implementation.

"The endline survey will also be population-based and conducted with the same methodology as the baseline so that results can be compared. In addition to examining health, nutrition and socioeconomic aspects, the endline will examine the extent to which improved behaviors have been adopted. During the FSG Program, Mercy Corps will aim to improve practices related to hygiene and sanitation and safe water management. At the endline, the external evaluator and his/her team will report on the extent to which beneficiaries have incorporated such practices into their lives. In addition to measuring impact indicators, the endline will also measure whether the program has been successful in meeting its output targets. The endline will be conducted in the first quarter of the third program year.

"Both baseline and endline surveys will take place during the main hungry season, which usually occurs between October and December every year. By timing the two evaluations to take place during the same time of year, Mercy Corps will ensure that the results of each are comparable, genuinely reflecting the changes that occurred during the implementation of the FSG Program. Further, the hungry season is an optimal moment for evaluating this type of programming, since both challenges and positive effects stand in greater relief during this time of year. However, it is also true that this period is the main growing season, when staple crops are produced. Mercy Corps will organize its evaluations so as not to unduly disrupt the field work of participants while still ensuring that representative participation is achieved."

E. EXAMPLE DM&E CAPACITY STATEMENTS

i) Example Mercy Corps agency-wide DM&E Capacity Statement

Mercy Corps has made an agency-wide commitment to strengthening its capacity and systems to effectively design, monitor, and evaluate (DM&E) field programs. These investments include a network of more than 140 DM&E-focused field staff from Mercy Corps' programs worldwide linked together through a web-based DM&E Community of Practice that allows for field-to-field exchange of examples, experiences and technical assistance. Mercy Corps' global DM&E activities are supported by three HQ-based DM&E experts within the Technical Support Unit, who provide tailored assistance to individual projects and country programs, develop agency-wide capacity-building resources, and coordinate agency-level strategic initiatives. Furthermore, Mercy Corps has developed a comprehensive DM&E toolkit as a fundamental resource for implementing sustainable and meaningful DM&E activities. Additional key resources include the DM&E Guidebook, which provide staff with the foundation of basic concepts, terminology and tools; and Mercy Corps' online DM&E training available to all staff.

The strong investment in DM&E tools, resources, and training has profoundly impacted the capacity of Mercy Corps' in-country teams. Country programs in Afghanistan, Kosovo, Sri Lanka, Somalia, Indonesia, and Guatemala have developed integrated country program-level M&E systems, allowing for robust tracking of results and lessons learned that feed into dynamic, information-based strategic planning processes. Several country programs, including Iraq, Lebanon, Sudan, and Pakistan now have dedicated, centralized DM&E units with full-time staff to better ensure integration of solid DM&E practices across all projects.

Finally, Mercy Corps has invested in a concerted effort to better articulate the overall impact of the agency's work through a set of consolidated metrics that will serve to measure progress against key elements of our mission. The agency has also carried out several professional evaluations and strategic studies in recent years covering a range of sectors, allowing for documentation of intervention outcomes and cataloguing of lessons learned to improve future programs. These combined efforts contribute to Mercy Corps' ongoing development as a learning organization with the ultimate goal of designing and implementing programs that will improve the lives of our beneficiaries.

ii) Example of in-country DM&E capacity statement from MC Kosovo proposal:

"In the fall of 2007 Mercy Corps Kosovo initiated a country-level initiative to capture results on a country-wide basis, aggregating individual project data in order to measure objectives and indicators at the country program level to measure impact and more effectively inform management. This special initiative, spearheaded by the Mercy Corps Kosovo Design, Monitoring and Evaluation (DM&E) Coordinator with in-country and remote technical assistance provided from Mercy Corps' HQ-based DM&E Technical Support Unit, is a pioneering effort among Mercy Corps world-wide, demonstrating the commitment of Mercy Corps Kosovo to incorporate DM&E into their programming to the maximum extent possible. In addition to the recently established full-time position of DM&E Coordinator at the country-level, the effort also consists of a newly formed DM&E Support Unit which draws on the participation of specific DM&E Focal Points within each project. This team meets twice a month to build their DM&E capacity, revise M&E tools, assist project design, and work to mainstream the country-level DM&E system. Mercy Corps Kosovo's countrywide DM&E system is now fully operational, and preparations have already been made to immediately incorporate this draft Performance Monitoring and Evaluation Plan into this system, providing it with the necessary customization that will be required.

Budgeting for M&E Tip Sheet

This tip sheet is for field managers and other staff developing *project* budgets. Further pages provide guidance and greater detail on what to budget for monitoring and evaluation[10], including examples. This document outlines typical costs needing to be covered to adequately undertake monitoring and evaluation activities for the entire project, regardless of who ultimately will be responsible for implementing those activities (Mercy Corps or our partners, local or international).

Also, while this document does not provide details on budgeting for M&E at the country-level, those programs that aim to develop a country-level M&E system (with its own need for data collection processes and data management system) might consider building into the project-level budgets a small percentage that contributes to this country-level system (comparable to a small ICR to support the country-level M&E unit and its activities).

Why the focus on M&E budgets?

Monitoring and evaluation (M&E) is increasingly recognized in Mercy Corps field programs as an area of strategic importance. This is because it can be instrumental in ensuring quality programming, easing management tasks, remaining accountable to donors and beneficiaries, and documenting impact with an eye towards future funding.

Mercy Corps staff who have invested in M&E tell us[11] they have:

- Used concrete results information to secure follow-on funding to existing programs.

- Have successfully argued for innovative program approaches based on a solid evidence base.

- Used robust M&E data to attract new funding from non-traditional donors.

RETURNS ON M&E INVESTMENTS: MERCY CORPS INDONESIA
• **Rigorous baseline study** in Safe and Healthy Schools Program reveals anemia as major issue and provides rationale for shifting strategy.
• **M&E data convinces government** to change policies and adopt program; provides evidence for advocacy approaches.
• **Increased funding and geographical scope** of project by donors and government due to well-documented results.

10 Monitoring refers to the documentation of program work by routinely tracking data on project activities and outputs and comparing that information against targets as part of standard program management. Evaluation typically refers to measuring objective-level changes and is therefore done only periodically, typically at baseline, mid-term and end of project. Costs for both routine data collection and periodic evaluations need to be provided for in the project budget.

11 For the more information, check out, *Developing a Value Proposition for M&E: Experiences of Mercy Corps Staff Who've Invested in M&E*, available on the Digital Library.

Strong M&E performance requires adequate funding and planning. This tip sheet aims to give some guidance and basic parameters for funding *project* M&E activities, so that field programs can better incorporate M&E costs into new proposal budgets and other funding cycles.

How much should I allocate for effective project M&E?

A general rule of thumb is to allocate **about 5-10% of project budgets** toward M&E activities. This figure may increase or decrease according to various factors; however it is the range recommended by DfID's CSCF group (5%) and USAID (10%). This tip sheet explores some of the cost items typically needed to adequately implement M&E activities.

What do the items cost and how to calculate the 5 to 10%?

Please see the table, *Approximate cost ranges for common M&E items*, at the end of this tip sheet for detailed information on the potential cost ranges for each specific cost category.

To calculate the percentage of the budget supporting M&E activities, follow the table at the end of this chapter and add up all M&E-related expenses– whether it is a stand-alone line-item or a percentage of a larger line-item (such as a staff with part-time M&E responsibilities) – and see what percentage of the budget this equals. If it falls below 5%, review carefully the scope of the project and the project's indicators to confirm that you have in fact budgeted sufficiently for the baseline, mid-term evaluation (if needed), final evaluation and set-up and utilization of the monitoring system and the data it generates.

As budget templates vary by donor, sometimes significantly, we will not cover where these M&E cost categories should be included in the project budget. When available, utilize M&E specific sub-line items or note the relevant expense as M&E in the sub-line item title or "Notes" field. For example, while there may not be an M&E sub-line item in the Personnel and Benefits budget category, you might title one cost item, "M&E training for project staff".

What are some common M&E budget items?

Staff. Evidence shows that Mercy Corps projects and country programs perform better in M&E if they have someone focusing on it. Full-time M&E positions and units are increasingly common at the project and country levels. A local staff member dedicated to monitoring is a good investment for helping keep projects on track and preventing costly mistakes.

Country-level M&E positions can be funded by allocating percentages toward them in project budgets. These can be considered technical support, along the same lines as the technical assistance or management support provided by sector coordinators, program advisors, directors of programs, etc.

Whether or not the M&E position is local or expatriate staff depends largely upon the availability of human resources with M&E skills in country, and the size and profile of the project or country-level position. Some country offices have started with expatriate country-level M&E positions to help set up processes and systems and mentor other staff, before transitioning these positions to local staff.

DONOR AND AGENCY POLICIES:
- The **USAID** Administrator has recommended 7-10% of project budgets to be allocated toward M&E.
- **DfID**'s CSCF recommends that around 5% of budgets be spent on M&E.
- **Gates** and other foundations and corporate donors expect high-quality M&E and are increasingly looking to fund it, as **NIKE** did last year with the MC China GLOW project.
- **Mercy Corps'** field finance manual officially recommends 5-10% for M&E.

Sometimes, funding full-time M&E positions is difficult. We can also look for ways to incorporate M&E responsibilities into existing or other positions by allocating time percentages dedicated toward M&E in position descriptions. Recruiting for M&E skills in other important project positions will also increase performance on M&E.

But whenever possible, be sure to keep program staff on until the end of the project to allow for a strong final evaluation and reporting effort. Don't skimp on human resources when doing the final report as this will allow for a final report that surpasses donor expectations.

Consultants. Mercy Corps projects often use external consultants for evaluating projects, in order to bring objectivity and evaluation expertise. Consultants can also be useful, however, in ensuring a quality baseline study, conducting trainings, and helping to design M&E systems or processes.

Consultant rates have increased significantly over the past few years, with rates in the $350-$550 range now the norm for top-level evaluation consultants. When budgeting for consultants, remember to include items like international travel, lodging, per-diems, visas, and any other

BUDGETING # OF CONSULTANT DAYS:
- When calculating the number of days, don't forget about time for general preparation, training staff, analyzing data, and writing reports, as well as the actual field work.
- We often under-estimate consultant days; good evaluations and baselines can last three weeks to three months in total.

miscellaneous costs. If the external consultant may not speak the local language, than costs for an interpreter should also be included.

If we seem to be spending a lot on external consultants, such as for conducting baselines or developing systems, we may want to reconsider whether these investments might be better spent on a full-time M&E staff, or on IT or program staff with significant M&E responsibilities, rather than relying too heavily on outside support. Over-reliance on external consultants can sometimes cause sustainability concerns and be disempowering for regular program staff.

Baselines and evaluations. Both of these activities cost significant time and money. In spite of the importance of baselines in documenting changes in objective-level indicators – such as knowledge, attitudes, behaviors or conditions – baselines are often not funded sufficiently. A good baseline or evaluation can cost anywhere from $5,000 to $50,000, depending on quality and scope.

In both baselines and evaluations, we need to account for typical costs, including:

- consultant fees,
- transportation costs,
- field testing survey instruments,
- translating data collection forms,
- any local hires such as interviewers and interpreters,
- training local surveyors,
- data entry and analysis software,
- printing and copying, and
- possibly staff compensation for time spent away from their normal duties.

> **BUDGET FOR QUALITY BASELINES:**
>
> • M&E budgets are often skewed towards evaluations, with little money for baselines. Without quality baseline data, it is difficult to demonstrate impact in evaluations.
>
> • *Make sure you fund baselines adequately so that you have the data to assess impact!*

It is also important that we ensure funding for analysis and reflection workshops, in order to review findings of baselines and evaluations and incorporate reflections into new project strategies.

It can be very useful to have extra M&E funding for common activities like assessments in new sectors or areas of intervention, or field studies, after-action reviews, and case studies for particularly interesting results or experiences. This can also help provide cover in case other costs, such

as baselines or evaluations, stray over budget due to customary snags like weather delays, staffing, etc.

M&E systems development. We may need to install and/or train on new software applications as part of setting up monitoring and evaluation systems. This increasingly includes the help of specialized IT staff or consultants. However, be careful about using too much outside expertise in developing the system. It can be unsustainable if project staff do not have intricate knowledge of the system and are unable to manipulate it when changes need to be made. If external consultants are likely to be needed, will you be able to find them locally? If not, consider including additional costs for translation.

Once the system is set up, we also need to think about maintenance costs, such as software upgrades or replacements and training of new staff.

Training and cross-visits. Country offices with more enhanced M&E systems spend a lot of time regularly training staff on data collection processes and forms, data entry, maintaining data management systems, and analysis methods. This is in addition to basic DM&E trainings for new staff.

Increasingly, Mercy Corps country offices are using cross-visits as a way for staff to increase their M&E knowledge and skills. Budgeting for items like regional cross-visits or M&E-focused workshops can yield rewards as staff learn new ways to apply M&E methods to their projects. It can also make staff aware of potential pitfalls, and motivate them to ensure that adequate time is allocated to ensure the proper implementation of M&E systems.

External trainings can be a good way to enhance the M&E skills of promising staff, as well as to expose them to new ideas and methods for conducting M&E. These can cost anywhere from $2,000 to $5,000, including airfare. Scholarships are sometimes available, particularly if it's affiliated with a university.

In-depth impact assessments. We're sometimes interested in going beyond what is required on M&E by donors and our agency. For example, pilot projects or certain innovative interventions may merit a more rigorous impact analysis in order to provide evidence for scaling up. This may require

> **EXTRA M&E FUNDING BEYOND THE 5-10% MAY BE NECESSARY FOR:**
>
> • **Strong evidence** to scale-up pilot programs.
>
> • **Publishing** studies or reports externally.
>
> • **Impact assessments** with control groups and/or random assignment.
>
> • **Miscellaneous items** like assessments, field studies, after-action reviews, and case studies.

that a significantly higher percentage of the budget is spent on M&E, beyond the 5-10%. Likewise, if we're hoping to publish a study to an external audience, we will also likely need added resources.

Some types of impact assessments, particularly those using control groups, entail multi-year processes and need to account for rigorous baseline, monitoring and evaluation activities. These will undoubtedly entail higher costs.

Elaborate research initiatives that involve partnering with universities or think tanks often imply higher costs, even if the partner universities are working "for free". Be careful of hidden costs – like meetings, transport, field work, publishing, etc. – in these arrangements.

How can I determine appropriate M&E budget allocations for my project?

Look at the indicators in the logframe. Objective-level indicators, such as changes in knowledge, attitudes, behaviors and conditions, often require surveys and other high-cost data collection activities. The presence of these types of indicators is good reason to increase funding for line-items relating to M&E. Likewise, projects that need to monitor a large volume of outputs, such as small-grants projects, may want to invest in robust monitoring systems and processes.

Think about staff roles and responsibilities. Large projects can benefit by having at least one full-time M&E staff, or a small M&E team. It can also work well to have M&E "focal points" for each sector or area of intervention, who are part of the project team but have time dedicated (20%-50%) in their position descriptions to focus on M&E. Think about how the project team will be organized, and if the person(s) responsible for M&E will have enough time to give it the attention it deserves.

Consider the type of project. In any project, sufficient funding for M&E helps ensure basic good management practices and the ability to make strategic adjustments, while giving us the data with which to demonstrate success and advocate for new funding. Pilot projects, however, or those that are particularly high profile with donors, may merit more than the typical 5-10% to be spent on M&E.

Conversely, projects with particularly large budgets may require less in percentage terms. For example, in a $10 million program over 3 years, $300,000 (or only 3% of the budget) allocated for M&E may be sufficient.

Consider the length of the project. All of our projects should be designed with the overall objective or purpose in mind, in other words, we should know what larger changes we are trying to bring about with our project activities and deliverables. And ideally we will measure to what degree those larger changes are brought about by the activities in all of our projects. The reality, however, is that for shorter projects (typically nine to twelve months or less), it may not be feasible to measure much beyond outputs.

Consider the example of a nine month project with the primary intervention of distributing seeds. Ultimately we intend to increase net household income through this intervention, but because of growing cycles and the time it takes for product to be harvested and sold at market, it may not be feasible to include tracking changes in income in our M&E plan for this project. This means we may not need the full 5% of the project budget for M&E.

Understand the donor's information and evaluation requirements. While M&E systems should principally be designed to meet our program management and information needs, the requirements of the donor need to be factored in, including during budget preparation. Determine with the donor whether the evaluations will be internal or external, and if external, whether they will be organized by the donor (coming directly out of their budget) or by Mercy Corps (and therefore coming out of the project budget).

Additional Resources

- *Mercy Corps field finance manual*
- DM&E tip sheets on survey planning, data management, focus group discussions, and structuring country DM&E units, available in the DM&E-in-a-Box on the Mercy Corps Digital Library
- *Developing a Value Proposition for M&E: Experiences of Mercy Corps Staff Who've Invested in M&E* on the DL.
- *Evaluation Scope of Work Example & Template* available in the DM&E-in-a-Box on the Mercy Corps Digital Library

Approximate cost ranges for common M&E items

Cost Categories	Unit costs and range	Time period	Notes
DM&E Staff	Salary, local: $300-$2,000/mo.	Usually over the life of the project (1-5 years). Country-level positions may be continuous.	Percentages of country-level M&E positions can be funded through project budgets.
	Salary, int.: $30,000-$60,000/yr		
	Benefits/allow.: $5,000-$30,000/yr		
DM&E Consultants	Local: $50-$200/day	Consultant assignments can last three weeks to three months. Include time for training field team, data analysis & report writing.	International consultant rates are always increasing, with $350-$550/day rates for top consultants now common.
	Int.: $350-$550/day.		
	Int. travel: $1,000-$3,000		
	Per-diem: $20-$60/day		
	Lodging: $30-$150/day		
	Visa/documents: $60-200		
M&E systems development	IT staff/consultants: see above	Project start-up, as well as throughout the project for maintenance and training.	High-tech systems imply higher costs for set up, maintenance and training
	Database software: $0-$5,000		
	Web-based solutions: $1,000-$20,000		
	Training (IT, data entry, etc.): $500-$5,000		
Baselines & evaluations	$5,000-$50,000 for each M&E activity, to cover:	Assessments, baselines & evaluations can last three weeks to two months. Include time for analysis & report writing.	More data collection techniques within baselines and evaluations imply more time and higher costs. In-depth impact studies with control groups and/or random assignment can cost significantly more and are a 1-5 year commitment.
	Consultants: see above		
	Temp hires (interviewers, translators, etc.): $10-$100/day		
	Local travel: $20-$100/day		
	Printing: $50-$200		
	Data software: $70-$500		
	Analysis meetings: $50-$1,000		
	Miscellaneous costs: varies		
DM&E Training	By in-country staff: $500-$2,000	Project start-up, and continuously throughout program.	Due to high staff turnover, periodic M&E training should be part of ongoing program activities.
	By external consultant: $2,000-$5,000		
	International M&E trainings & cross-visits: $2,000-$5,000		
Meetings & Workshops	Strategic planning meetings; annual review meetings: $3,000 -$20,000	We should plan for strategic planning and regular analysis & reflection meetings on M&E data. Key for consensus building!	Level of costs depends on travel and set-up needs, number of participants and may be incorporated in other line-items.
	Standard, periodic monitoring results review with MC staff and partners: $300-$3,000		

Baselines and Evaluations

Baselines Tip Sheet

This tip sheet is for managers and staff implementing baseline studies. Further pages provide guidance and greater detail on baseline planning and implementation, including examples.

Introduction to baselines

A baseline is an M&E activity that should be undertaken at the beginning of every project. Its purpose is to collect key information that we need to track to be able to measure any changes realized during the life of a project. This is especially relevant for objective-level indicators, since they often require in-depth data collection activities like surveys or focus groups.[12] This tip sheet covers important factors to consider when conducting baselines.

Why should I conduct a baseline?

Without a baseline, it is very difficult to demonstrate the effects of our interventions. We may be able to show outputs, such as the number of microfinance loans disbursed or health clinics rehabilitated, but we will be challenged to demonstrate the changes to which these actions have contributed.

Objective-level indicators usually encompass changes in *knowledge, attitudes, behaviors or conditions*. For example, a microfinance project may wish to show increases in income, while a health project may contribute to increasing the percentage of mothers attending pre-natal consultations. The best way to demonstrate that these changes are taking place is to measure the indicators before the intervention and compare them to measurements at the mid-point or end of the project. Baselines also help us to refine our indicators and implementation strategy, and engage with the communities where we will be working.

What is the difference between baselines and assessments?

We sometimes confuse baselines with needs assessments, or believe that one can substitute for another. However, these activities differ in important ways, as illustrated in the following table:

---

12 See listing of Mercy Corps' DM&E Tip Sheets on surveys and focus groups in the Additional Resources section.

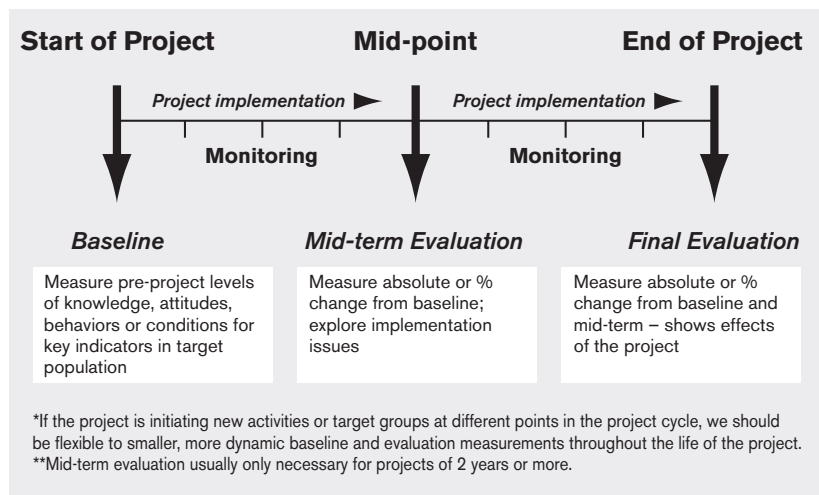|	Assessments	Baselines
Purpose	Explore issues with populations to inform project design	Collect data on the indicators chosen in our project design
Timing	Before project design or during start-up to finalize the workplan	After project design, prior to implementation of project activities
Target group	Wider population or broad sub-groups	Greater focus on planned project participants or beneficiaries
Data collection	Generally more qualitative, focusing on a range of issues	Generally more quantitative, focusing only on key indicators

Difficulties arise when we attempt to combine the vastly different functions of a baseline and an assessment. This happens a lot, usually because the initial assessment is out of date or lacking in detail, and we are seeking more information to tailor our implementation strategy.

THE PROBLEM IN MIXING ASSESSMENTS WITH BASELINES:

The type of information we must collect to assess the needs of the population is quite different from the information needed to measure indicators. It is hard to measure baseline indicators when the project's activities and target groups – and therefore the indicators that apply to them – are not yet clearly defined.

When should I implement a baseline?

Baselines should be implemented during the project start-up phase, after the proposal has been funded but before full-scale implementation begins. This concept is illustrated below:

Start of Project   **Mid-point**   **End of Project**

Project implementation ▶   *Project implementation* ▶

Monitoring   **Monitoring**

Baseline   *Mid-term Evaluation*   *Final Evaluation*

| Measure pre-project levels of knowledge, attitudes, behaviors or conditions for key indicators in target population | Measure absolute or % change from baseline; explore implementation issues | Measure absolute or % change from baseline and mid-term – shows effects of the project |

*If the project is initiating new activities or target groups at different points in the project cycle, we should be flexible to smaller, more dynamic baseline and evaluation measurements throughout the life of the project.
**Mid-term evaluation usually only necessary for projects of 2 years or more.

Recommendations for the exact timing of the baseline vary. Sometimes it is best to implement the baseline right away, in the first month. This often happens when we have a fairly straight-forward start-up process and a clear understanding of our informational needs.

In other cases, however, projects take longer to get off the ground. This can be the case when there are significant recruiting needs, when we are new to an area or sector, or when we still need to do an assessment during start-up to finalize the detailed implementation plans. In these instances, it may be preferable to wait a few months, when the project is more organized, in order to make sure the baseline covers all of the information we need to track. We should also be open to follow-on baseline measurements as needed, when new activities are introduced or strategic priorities shift. The timing of baselines may also have to shift according to seasonal factors, particularly for agricultural or economic development projects, to ensure that the most appropriate and relevant data is collected.

Remember also that baseline data may come in many different forms and from various sources, depending on the indicators. This includes monitoring data, secondary information, or smaller surveys targeting particular segments of the population or for a particular set of indicators. It may not necessarily require a large, comprehensive survey. The important point is that we have usable baseline data directly relevant to our indicators.

What if I do not have the time or resources for a baseline?

Simplify by reducing the scope of the baseline. Even in the most difficult of circumstances, such as in conflict environments or emergencies, responsible programming means tracking the effects of our interventions. There are ways to do this in reduced form when the situation dictates, by scaling down the amount of information collected and reducing the level of sophistication. When access is limited, remote third-party data collection or a phased approach may be possible. We can also talk to donors about the importance of baseline data and advocate for more time. The key is that we do not put off or cancel the baseline, but rather that we simplify and adapt it to make it feasible.

WAYS TO MAKE A BASELINE FEASIBLE:

1. Reduce the data collected, focusing only on the most important indicator(s).

2. Reduce the sample size.

3. Take a phased approach, conducting follow-on measurements as required.

4. Simplify the tools: shorter questionnaires, fewer methods, etc.

DO NOT eliminate the baseline due to time or resource pressures – instead, simplify!

How should I organize the baseline?

A baseline is just like an evaluation in terms of how we plan for it. Developing a good scope of work is essential.[13] Start with the *project logframe* to see which indicators we need to measure. This will help us to determine the objectives of our baseline. We should be careful to separate out the different methods that may be used to accomplish different objectives or measure certain indicators. For example, individual surveys may be best at obtaining the percentages or averages we need for our objective-level indicators, while focus groups may be an appropriate technique for exploring key qualitative indicators.

> **A SCOPE OF WORK FOR A BASELINE SHOULD INCLUDE:**
>
> - Purpose or objectives
> - Key baseline questions, tied to project indicators
> - Team composition and roles
> - Data collection methods
> - Data management plan
> - Budget and equipment needs
> - Workplan with timeline
> - Analysis and dissemination plan

A baseline scope of work can be developed during the project's kick-off workshop. Integrating DM&E planning into the project kick-off workshop is highly recommended and important to develop team consensus on the project's objectives and activities, as well as planning M&E tasks and defining roles and responsibilities. The workshop should ideally be held when all staff are on board and prior to the implementation of project activities.[14]

After sufficient planning through a scope of work, we can start to develop the tools needed for the baseline (see *DM&E Tip Sheet #2: Designing the Survey Tool*). Remember to pilot test the tools extensively, to ensure standardization in filling them out and minimize mistakes or misunderstandings. Data entry and management processes, as well as a plan for analyzing the data, should also be worked out.

We also might want to consider bringing in outside expertise to assist with the baseline. External consultants can provide objectivity and a higher level of rigor. The disadvantage is that results may be less likely to be internalized by the project team, if the baseline is not conducted in a participatory manner. In any event, baselines should be taken seriously and funded sufficiently, on par with evaluations, to ensure we have quality data upon which to measure impact over time.

13 See the *Baseline/Evaluation Scope of Work Template and Sample* on the Digital Library. *DM&E Tip Sheet #1: Survey Planning* also contains good guidance on planning for baselines and evaluations.

14 See Mercy Corps' *DM&E Tip Sheet #7: DM&E at Project Kick-off* for more guidance.

What are the key factors to consider in a baseline?

Planning for consistency and replicability. The key to effective baselines is the ability to compare the data collected to similar information gathered at the end of the project. This requires thinking through the methods used to ensure they will still be appropriate in a few years time. We should pay particular attention to sampling methods, so that we can reconstruct a comparable sample at project's end (see *DM&E Tip Sheet #4: Sampling* for more guidance). While we may not want to interview the exact same people, we do need to be able to sample the same target populations, using a comparable sampling strategy.

It is very important to clearly document the methodology we use for our baseline in reports or other project documents, since high staff turnover can erode institutional memory. Storing the data in a safe place known to all staff is also crucial, as baseline data often gets lost over many years.

Recognizing changes that can affect our baseline data, and re-measuring when necessary. As a general rule, Mercy Corps projects and the groups they work with change and evolve over time. These changes can render our original baseline data irrelevant by the end of the project.

It is important to plan ahead by conducting timely data collection around the events that can bring about these changes. This will allow us to capture important outcomes and keep our information up to date.

For example, a youth project might want to measure increases in capacity over time. However, each year, many youth graduate and leave their communities. In this case, we may be able to conduct end-line indicator measurements each year for the youth that are graduating, rather than waiting till the end of the project.

Assessing measurability and relevance of key indicators. The indicators chosen during the proposal phase are sometimes problematic. Part of the utility of conducting a baseline is to test the feasibility of collecting certain indicators, as well as assessing their ability to demonstrate results.

> **TYPES OF CHANGES DIMINISHING THE RELEVANCE OF ORIGINAL BASELINE DATA:**
>
> **Project strategy changes caused by:**
> - Shifting donor priorities
> - New assessment information
> - Changes in Mercy Corps strategy, capacity or staff
>
> **Target group changes caused by:**
> - Evolving needs
> - Shifting local priorities
> - Program 'graduation'
> - Migration
> - Turnover of local NGO staff or officials
>
> **Contextual changes caused by:**
> - Seasonal variations
> - Insecurity
> - Elections
> - Economic conditions
> - New laws

Based on the experience of collecting indicators during the baseline, the project may be able to make the case to the donor for altering the indicators to reflect current or more appropriate field realities. It is important that we identify poorly-designed indicators early on, so that we can make adjustments before it's too late.

> **KEY INDICATOR QUESTIONS:**
>
> 1. Is the indicator reasonable to measure?
> 2. Is it able to be measured in a consistent way over the life of the project?
> 3. Is it relevant for demonstrating the effects of the project?
> 4. Are the targets we've set realistic?

Planning for impact assessment through the use of control groups. The ability to rigorously assess impact at the end of the project can be enhanced by including a "control", or non-project participant, group in the baseline. This allows us to make comparisons of knowledge, attitudes, behaviors or conditions between groups that were involved in the project and those that were not, showing the added value, or impact, of the project.

> **CONTROL GROUPS:**
>
> - Should be reserved for projects requiring a high level of rigor or special interest in impact assessment.
> - Involve measuring indicators for groups or individuals who will not participate in or benefit from the project.
> - Should be as similar as possible to the project group (motivation, socio-economically, demographically, etc.).
> - Should be selected randomly in the same way as the project group.

Some donors are increasingly advocating for the use of control groups to assess impact. However, we need to recognize the additional time and resources this takes, and be selective about where and when it may be appropriate. Generally, these could include large, long-term projects, high profile pilot projects, or studies tied to long-term country program strategies.

In addition to selection issues, another primary challenge is how to sensitize and work with control groups when they are not receiving project benefits. A good strategy is to find some small programming element which they could participate in, or other token of appreciation for their participation in the study, that will not bias results. We will need to weigh the benefits of measuring impact through the use of control groups with very practical concerns about creating jealously or tensions, and the potential political implications.

We should also be aware of contamination issues, whereby the control group is still being somehow being affected by the project through, for example, contact with neighbors. This would decrease the differences we would expect to see.

Managing expectations and potential bias. Data collection involving the local population often creates heightened expectations. It is important that we attempt to manage these expectations by carefully explaining the objectives of our study and being realistic about the prospects of our working in a particular area or with a particular group. Sometimes, the desire to attract resources or interventions to the community can lead to bias on the part of respondents or local officials. We can counter-act this by attempting to cross-verify information and following up with direct observation where possible.

Analyzing and disseminating findings. As with any M&E activity, we should have a strategy for disseminating the results of our baseline for analysis among beneficiaries and other local stakeholders. Meetings should also be held with project staff and local partners to go over the results together and discuss their meaning for project implementation. For more on data analysis, please see *DM&E Tip Sheet #11: Data Analysis.*

> **AN ANALYSIS WORKSHOP WITH STAFF AND PARTNERS CAN DISCUSS THE FOLLOWING:**
>
> - Key findings from the baseline.
> - Possible adjustments to project activities.
> - Lessons learned in data collection.
> - Feasibility of indicators and targets.
> - Follow up data collection activities.
> - Plan of action for implementation.

Additional Resources

- Mercy Corps *DM&E Tip Sheet #1: Planning a Survey*
- Mercy Corps *DM&E Tip Sheet #2: Designing the Survey Tool*
- Mercy Corps *DM&E Tip Sheet #4: Focus Group Discussions*
- Mercy Corps *DM&E Tip Sheet #11: Data Analysis*
- Baseline guidance from MandENEWS listserv: https://clearspace.mercycorps.org/docs/DOC-1419
- Chapter on baselines from Search for Common Ground's *Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programs:* https://clearspace.mercycorps.org/docs/DOC-1444

Evaluations without Baselines Tip Sheet

This tip sheet is for managers, staff or consultants attempting to evaluate projects where there is no baseline data. Further pages provide detailed guidance on common issues and approaches, including examples.

How can we measure change without baseline data?

Some projects are unable to conduct a baseline during start-up. In addition, for various reasons, baseline data is sometimes no longer valid or relevant by the end of the project. Consequently, we sometimes face the challenge of evaluating a project that lacks relevant baseline data. This tip sheet explains strategies that have emerged for dealing with this challenge. A *Summary Table* illustrating the pros and cons of each approach is presented at the end of this chapter.

> **STRATEGIES FOR OVERCOMING THE ABSENCE OF BASELINE DATA INCLUDE:**
>
> - Project monitoring data
> - Secondary data sources
> - Comparison group of non-participants
> - Individual recall
> - Key informant interviews
> - Participatory rural appraisal (PRA)
>
> *Key point: **We should combine several techniques to cross-verify our information!***

Because some of these methodologies can be prone to error, we should be prudent when drawing conclusions based on them. It is recommended to consult HQ DM&E staff or other technical M&E specialists as needed when attempting to carry them out.

Using project monitoring data

Sometimes the data we routinely collect during project implementation can give us an idea of how participants have changed or evolved over the course of the project. The richness of this data is occasionally overlooked during an evaluation, and may be able to give a sense of prior conditions. This is another reason to ensure that our monitoring data is well-documented in project information systems.

Analyzing secondary data

Sometimes external data sources exist that contain information relevant to our project area. These could include data from central or local governments, UN agencies, other international NGOs, local organizations, or academia. Other times, our own data collected during an assessment or for other projects can help give us an idea of pre-intervention conditions.

Depending on the circumstances and relevance of the data to our project activities, this data can be used to follow trends within our target population. If the indicators or methodology used in secondary data are slightly different from ours, we may need to adjust our own indicators or collection methods slightly in order to be consistent.

Asking about the past

A rather obvious and often-used method is to ask people what conditions were like before the intervention, and have them compare this with current conditions. This is risky because people's views about the past are subjective and often imperfect, especially regarding quantitative information or past feelings.

> **QUESTIONS FOR SECONDARY DATA:**
>
> - Does it cover the same population and target groups as the project?
> - Is the time period appropriate?
> - How relevant is the data to project indicators?
> - What was the sampling methodology?
> - Who conducted the research? Is it possible to get the raw data?

We can address some of these concerns by sticking to broad, large-picture questions. For example, asking someone how many times they visited a clinic over the last three years is likely to lead to inaccuracies. A better question might be whether they visited the clinic more frequently or less frequently three years ago as compared to now, or whether their medical expenses have increased or decreased over the years.

Asking the question in multiple ways is a good way to double-check the accuracy of responses. We may also be able to access information at the village or district level, perhaps from local cooperatives, banks, or government authorities, to see whether our data makes sense. We can also talk to key informants, such as neighbors or local authorities. If major discrepancies exist, we can return to the respondent and attempt to reconcile the information.

> **TIPS FOR USING INDIVIDUAL RECALL:**
>
> - Group questions from the same time period together
> - Use major events to help mark time
> - Avoid too much quantitative or time-specific detail
> - Anticipate common biases
> - Ask the same question in different ways to cross-verify information

Conducting key informant interviews

These can be useful in getting a broad picture of past conditions, and can be less demanding than surveys. Key informants may include traditional or administrative authorities, teachers, doctors, religious figures, association members, etc. It is important to recognize that each person carries certain perspectives and biases. We should therefore seek to obtain a wide variety of informants, particularly from marginalized groups. Staff who were present during past periods can also be a great resource, particularly in getting started.

Using participatory techniques

NGOs sometimes use Participatory Rural Appraisal (PRA) techniques to help groups or communities reconstruct the history or conditions around a particular issue or series of events. They are particularly helpful in cross-checking information obtained from individuals, since the participatory exercises require group consensus. Two common techniques are outlined below[15]:

PARTICIPATORY TECHNIQUES USED TO ASSESS BASELINE CONDITIONS:

Historical timeline:
- Community members draw a timeline of significant events, denoting major changes that affected people's lives.
- The facilitator can ask them to fill in information about conditions in each particular period.

Seasonal calendar:
- Participants draw a chart depicting the months of the year, and denote key factors such as weather, harvesting, prices, etc., for each month.
- Typically used in agricultural livelihoods projects, but is applicable to other sectors.

Using comparison groups

Another strategy is to assemble a group of individuals or households that are similar to project beneficiaries, but did not participate in or benefit from the intervention. Comparing data between this group and our target beneficiaries, we can attempt to make the argument that differences among the groups related to our sectors of intervention are due in part to project activities. This is one way to illustrate our impact.

---

15 Further explanation and guidance on implementing PRA techniques can be found at: http://www.aidsalliance.org/graphics/secretariat/publications/Tools_Together_Now.pdf , http://www.ids.ac.uk/ids/particip/research/pra/pranotes02.pdf and in *DM&E Tip Sheet #14: Participatory M&E.*

COMPARISON GROUPS SHOULD BE SIMILAR TO PROJECT GROUPS IN:

- Demographic distribution (age, sex, ethnicity, etc.)
- Socio-economic status
- Geographical placement
- Access to resources and opportunities
- Motivation and initiative

In theory, the only underlying difference between project and comparison groups should be project participation or benefit. In practice, however, there are often non-random factors that we use to select project participants, such as targeting the most vulnerable or operating in places with reliable partners. A major challenge with the comparison method, therefore, is that those affected by the project often differ from non-beneficiaries in important ways. In addition to socio-economic and demographic differences, there are also less observable factors such as personal motivation, access to resources, and social capital. The comparison method is strengthened when we are able to show that the two groups share similar underlying characteristics, as demonstrated in the text box to the right.

One strategy is to use groups or individuals that are new to the project and compare them with groups or individuals that have been benefiting from the project for a longer period. We need to pay special attention, however, to how participants were selected in each phase, since selection itself is often based on important differences in pre-existing conditions.

EXAMPLE: MC AZERBAIJAN

Veterinarian Businesses		
	Control Group	Project Group
Avg. No. of clients	207	212
Avg. % of Clients with Small Farm	83%	81%
% with Access to Pricing Information (MC intervention)	3%	63%

First showing the similarities among groups adds credibility to any later analysis of differences caused by project.

Additional Resources

- *DM&E Tip Sheets* series (available on Mercy Corps Digital Library)
- Baseline guidance from MandENEWS listserv: https://clearspace.mercycorps.org/docs/DOC-1419
- Bamberger, Michael; Rugh, Jim; Mabry, Linda: RealWorld Evaluation, 2006.
- A good example of comparison and project group statistical tables is available on the DL: *Initial Phase Report for Azerbaijan: Mercy Corps Cluster Access to Business Services (CABS) Program in Rural Azerbaijan.*

SUMMARY TABLE OF APPROACHES FOR DEALING WITH THE LACK OF BASELINE DATA

Approach	Advantages	Disadvantages	Major issues
Project monitoring data	• Data has already been collected by the project • Multiple data points • Can show trends	• Depth of information may not suffice for objective-level indicators	• How detailed is the information? • Is it available for all time periods? • Do trends in behaviors or conditions emerge?
Secondary data	• Saves time and resources • Less susceptible to error or bias • Can be very comprehensive • Analysis sometimes included	• Difficulty in matching project population or indicators • Generally performed for different purposes • Methodology may be hard to replicate	• How relevant is the information, target groups, and time period to the project's activities? • How did they sample?
Individual recall	• Can ask specific questions related to the project • Can triangulate by asking questions in different ways, and consulting multiple sources • Good for broad, general welfare issues	• Hard to recollect details and timing of behaviors or conditions • Open to bias based on selective memory • Major issues in assessing prior attitudes or feelings.	• Is the information realistic to remember accurately and consistently? • Are there major events we can use to help mark time? • Is the data susceptible to biases based on cultural norms?
Key informant interviews	• Provides historical info quickly • Interviewees likely to be among the most informed • Can get official as well as common perspectives	• Can get skewed perspectives that do not represent views of wider population • Not optimal for obtaining individual-level quantitative info	• Are we interviewing a wide range of informants to obtain multiple viewpoints?
Participatory Rural Appraisal (PRA) techniques	• Group setting encourages consensus on key facts • Community members lead in describing and analyzing events	• Less relevant for individual-level quantitative info • Views can be dominated by strong personalities	• Can we use results to validate other data sources? • Are there biases in terms of how the group views past events?
Comparison groups	• Possible to demonstrate effects of participating in or benefiting from project • Able to show quantitative impacts	• Difficult to construct a robust comparison group • Almost always pre-existing differences in project and non-project groups, affecting findings	• Can we use random selection to help minimize differences? • Do we acknowledge factors that can cause differences that are not due to the project (such as geography, access to resources and opportunities, socio-economic status, motivation, etc.)?

Recommended Outline for Evaluation Reports

- **Cover Page, List of Acronyms**

- **Table of Contents** which identifies page numbers for the major content areas of the report.

- **Executive Summary** (2 to 3 pages) should be a clear and concise stand-alone document that gives readers the essential contents of the evaluation report in 2 or 3 pages, previewing the main points in order to enable readers to build a mental framework for organizing and understanding the detailed information within the report. In addition, the Executive Summary helps readers determine the key results and recommendations of the report. Thus, the Executive Summary should include: major lessons learned; maximum of two paragraphs describing the program, summary of targets and intended outcomes; areas of meaningful under or over achievement; and possibly a few lines describing the action plan developed to follow up on evaluation recommendations and how the evaluation report will be disseminated.

- **Methodology**: sampling method including strengths and weaknesses of method used, inclusion of stakeholders and staff, rough schedule of activities, description of any statistical analysis undertaken, including justification and software package used. The discussion of any random sampling used should include details on how the random respondents were identified and invited to participate. This section should also address constraints and limitations of the evaluation process and rigor. The methodology section should also include a detailed description of data collection techniques used throughout the evaluation.

- **Results**: Think about how best to organize this based on the evaluation questions. In some cases, it is helpful to organize the report against project/program objectives, but in other cases it may make more sense to organize the report against evaluation questions.

- **Synthesis, Recommendations and Lessons Learned**: This is space for the evaluation team to think about the data and results, and make concrete recommendations for current or future project improvements/changes, pull out organization lessons learned, and generally comment on data and results. Everything presented in this section must be directly linked back to the information presented in the Results section of the report. Ideally, items discussed here will not be completely new to the reader, but rather will refer to previous discussions. Recommendations that are not directly tied to Results can be included in an Evaluator Comments section for the report.

- **Annexes**: data collection instruments in English and translations; list of stakeholder groups with number and type of interactions; SOW, qualitative protocols developed and used, any data sets can be provided in electronic format, any required photos, participant profiles or other special documentation needed.

Data Collection

Survey Planning Tip Sheet

This tip sheet is for managers and staff planning to conduct a survey in the field. Subsequent pages provide greater detail on what to consider before, during and after implementing a survey, including examples.

Introduction to Surveys

A survey is a data collection method that we use in Mercy Corps to find out key information about a target population. It poses a standard list of questions to individuals or households, and can be oral or written. The key to generating effective survey results is good planning. This tip sheet covers some of the main issues we need to think about in order to organize and implement a successful survey.[16]

CHECKLIST FOR PLANNING A SURVEY:

- ☐ Purpose/informational needs defined
- ☐ Target audience & sampling strategy defined (see the Sampling tip sheet)
- ☐ Workplan & budget developed
- ☐ Survey team assembled/recruited
- ☐ Data entry & management plan developed
- ☐ Questions designed (see the Survey Tool Design tip sheet)
- ☐ Survey team trained
- ☐ Survey tool test-piloted & refined
- ☐ Survey implemented
- ☐ Data input & quality checks conducted
- ☐ Data analysis undertaken
- ☐ Dissemination & feedback carried out

When to use surveys? Surveys are most relevant for obtaining *quantitative data* (numbers, percentages, etc.) during a baseline, mid-term, or final evaluation. They are particularly useful in measuring objective-level indicators, such as changes in knowledge, attitudes and behaviors. For example, a survey may gauge the percentage of children using sanitary hand-washing practices, or attitudes towards community reintegration. A well designed and consistently applied survey can measure this information over time to illustrate a project's effects.

---

16 More resources for developing surveys are referenced in the *Additional Resources* section on the last page of this chapter. For more specific tips on writing the questionnaire, please see *DM&E Tip Sheet #2: Designing the Survey Tool.*

When not to use surveys? Most of our projects should avoid using surveys as part of their routine monitoring systems, as surveys can be quite costly and time-consuming. Likewise, mid-term evaluations in relatively short projects (1-2 years) may not necessitate a full-scale survey. In these instances, a more participatory method focused on improving implementation approaches may be more appropriate.

While surveys are good at answering "what happened" and "how many times it happened", they are generally not the best method for answering "why" or "how it happened". Focus groups, key informant interviews, and other rapid rural appraisal techniques are more appropriate methods for gathering *qualitative data*, which is more descriptive and allows for open-ended discussions, explanations and perspectives. Attempting to capture these types of information in a survey makes it difficult to organize and interpret results.

How long does it take to do a good survey? A survey can take anywhere from 3 weeks to 3 months or more, depending on the size and complexity. Some experts recommend at least 6 weeks for the survey process, including two weeks for design and preparation, two weeks for carrying it out, and two weeks for analyzing the data.

How should I plan for the survey?

We can think of the survey process as a mini-project, whose phases include preparation, implementation, and results analysis. A lack of attention devoted to the preparation and results analysis phases are the main reasons many surveys fail or lack credibility. Developing a scope of work document is a good way to plan appropriately. Common in baselines and evaluations, a scope of work generally contains the purpose, key questions, data collection methodology, sampling strategy, team composition, timeline, budget, and logistical issues of the M&E activity.

> **PLANNING A SURVEY:**
>
> Develop a **TIMELINE** of activities, and plan equally significant time for preparation, implementation and results analysis.
>
> **DO NOT** begin by writing questions – more thoughtful planning enhances efficiency and results.

Defining the purpose of the survey. A useful way to approach the survey planning process is to start with the end product and work backwards (similar to how we approach designing a project!). What do we want to state quantitatively in the final report? What is the purpose or role of the survey in the overall evaluation process? The "Questions" text boxes inserted throughout this document (see example at the top of the next page) will help remind us of these key issues.

We should begin the survey design process with a clearly defined informational need. Generally, this flows from the indicators in our project's *logframe and indicator plan* (see the *DM&E Guidebook*). The next step is to determine how we will obtain the information. Before committing to a survey, we should ensure that the specified data cannot be found elsewhere, such as in project records, secondary data, or through other methods. We should also look for areas where information from other sources can complement our survey data.

> **PLANNING QUESTION** ?
>
> *"Do we need this information, or would it simply be nice to know?"*
>
> Don't chase the information that might be nice to know. It can distract us from getting quality information that we NEED to know.

Developing a team, timeline of activities, and budget. Like any project, proper planning includes assembling the survey team and developing a work plan and timeline, as well as a budget. We should appoint an overall *survey supervisor* to be responsible for moving the process along, monitoring the quality of the data being collected, and serving as a focal point for any problems that arise. Interviewers will also have to be managed; a good rule of thumb is to ensure one survey supervisor for every 4 to 6 interviewers. We should also look at the sampling strategy and geographic scope when considering the management strategy and logistics.

Our team should be balanced appropriately - gender, local languages, ethnicity, age, etc. - according to the context and data needs. For example, one Mercy Corps project recruited mostly older, more experienced men to the survey team, due to experience factors and cultural norms. However, the project found out (too late!) that this team was not the most effective for discussing sensitive topics with women and youth. Results were incomplete or compromised. It is important to have balance to get a variety of perspectives, and to tailor our team demographics to the type of data needed and the audience.

> **ISSUES TO CONSIDER IN DEVELOPING THE BUDGET AND WORKPLAN:**
>
> - Interviewer/staff recruiting
> - Interviewer training
> - Designing and piloting the survey tool
> - Where/when/how to administer
> - Material resources & equipment
> - Access and transportation
> - Respondent availability
> - Data coding & management systems
> - Results analysis and interpretation
> - Translation needs

Designing the survey tool. We can get a head start on designing the survey instrument by analyzing project documents, monitoring data, and other previous M&E efforts. Likewise, it is very useful to do initial focus groups in the community, as well as staff or key informant interviews, to more clearly define the issues. As much as possible, we will want to conserve our limited data collection resources for the data that has clear uses and addresses information gaps. The tool should be as short and concise as possible, and consist mainly of closed questions that have coded responses. For further information on designing the survey tool, please see *DM&E Tip Sheet #2: Survey Tool Design.*

Planning for data coding and management. A plan for coding and processing our data is essential and should be considered early on in the planning process. Too often, we design survey questionnaires without sufficient forethought as to how the data will be inputted electronically and aggregated to show results. This creates difficulty in the analysis process. Consideration of coding issues will help us to determine what types of questions should be asked and how the interviewers will ask them. For example, our survey may ask, "What was your yearly household income last calendar year?" for which a wide range of numeric values are appropriate. Or, it could ask "Were you able to make enough money last year to adequately provide for your household's basic needs?" for which a yes or no answer (A or B), or a short list of logical responses (coded A-E), would be suitable. The most appropriate coding structure depends upon our survey's purpose and overall informational needs.

> **DATA MANAGEMENT QUESTIONS**
>
> - *What do we want the data to "say"?*
> - *How will we group individual responses?*
> - *How will the data need to be presented?*

Some common systems used to manage data in the field include *Microsoft Excel, Access* and *SPSS* software. Another option is *Epi Info 2002*[17], which was originally developed for public health projects but can be used effectively with most types of programming data. Whatever mechanism is used to enter and manage the data, before we design and undertake the survey we should clearly define who will enter the data and when. This person may have to be trained to do so. The team leader or another clearly identified individual should routinely cross-check the data as it is being entered to catch any possible mistakes or misinterpretations. See *DM&E Tip Sheet #8: Data Management* for more on this subject.

---

17 EPI Info 2002 data management software is available for free download at:
    http://www.cdc.gov/EpiInfo/epiinfo.htm

Training interviewers. Surveys in Mercy Corps projects are usually administered orally by interviewers, who ask questions to respondents and mark down answers. Interviewers will need to be adequately trained to perform these tasks. If the interviewers do not present and record questions in a standardized way, then responses will vary based on the personal style of the interviewer and our results will be hampered. Interviewers should also be prepared to consistently handle any confusion that may arise during the survey process. We can reduce this risk by investing in thorough training, and by developing clear guidelines for interviewers to refer to during the survey.

> **INTERVIEWER TRAINING:**
>
> Focus on consistency in:
>
> 1. Introducing the survey
> 2. Local language translations
> 3. Asking questions
> 4. Probing for further information
> 5. Recording responses

Pilot testing the instrument. Pilot testing, or pre-testing, our survey tool is critical in making sure that our survey is set to properly gather the information we want. Effective pilot testing will identify ways in which question wording, ordering and interpretation may *bias* (or inappropriately influence) our results. Pilot testing is also essential for verifying the relevance of pre-coded responses, detecting potential misunderstandings, and refining interviewing skills.

To pilot-test our survey tool, we should select a relatively small focus group (5-10 people) that is representative of the sample. Project staff are generally not ideal due to their heightened familiarity with the project – although initial role playing to get comfortable with the questions may be useful. The survey team can administer the survey to the focus group and then interview respondents to determine how the survey questions were interpreted. Our survey can also be piloted individually to different sets of respondents for further refining.

> **PILOTING QUESTIONS**
>
> - *Were questions interpreted correctly?*
> - *Were concepts or language unclear?*
> - *Were the response lists adequate?*
> - *Did any questions make respondents uncomfortable?*
> - *Did it take too long? Was there redundancy?*
> - *Is the resulting data useful and practical to input and analyze?*

It is often useful to analyze some example "mock" data from the pilot testing to ensure that all of the data being collected can be analyzed in a useful, strategic manner. This mock data analysis should include major stakeholders, such as project managers or country directors, who will be end-users of the

information. This way, we can eliminate unnecessary elements or re-arrange the survey tool or sampling methodology to ensure the information obtained is as relevant as possible.

Pilot testing can be incorporated into the interviewer training process. We should plan sufficient time for multiple stages of testing and re-designing of questions.

Determining the sample. Survey results are only useful in making statements about our target population if the survey sample has been constructed properly. A sample is a sub-set of individuals or households used to make wider observations about a target population. Our sample should strive to be representative of the population and randomly generated, meaning that respondents are chosen by chance, wherever possible. There are normally tradeoffs associated with the size and accuracy of the sampling methodology, and practical concerns such as access, time, and resource constraints. A proper sampling strategy seeks to balance these factors. Please see *DM&E Tip Sheet #3: Sampling* for further tips on how to select an appropriate sample.

Implementing the survey. We should carefully plan the data collection methodology in terms of logistics, survey team members' roles and responsibilities, and information sources. We should be realistic and plan ahead for potential problems. Surveys almost always experience delays, but usually we can predict these with advance planning and contingency analysis. Common causes of disruption include weather, transportation issues, miscommunication, low response rates, and insufficient grasp of the methodology by interviewers.

> **IMPLEMENTATION TIPS:**
> - Plan ahead to avoid delays
> - Invest in training and preparation
> - Develop contingency plans
> - Undertake routine data quality assurance checks
> - Be prudent in interpreting results

Entering data and analyzing results. An appointed team member should routinely check the data for errors as it is being entered and coded. Once the data is entered, the team should once again go through the data to identify mistakes and outliers. When analyzing results, it is important that we remember (and clearly state) the decisions we made about how to carry out the survey and who to include in the sample. We should be careful about making broad generalizations based on our results, as they may only be relevant for certain groups or lack statistical credibility due to sampling limitations. *DM&E Tip Sheet #11: Data Analysis* provides more insights and examples for ways in which to analyze data and present results.

Maintaining confidentiality. Keeping the data gathered in our survey secure and confidential is crucial for Mercy Corps' reputation and, more importantly, for the safety of respondents. Statistical tabulations need to be sufficiently broad so that individual respondents cannot be singled out. This is particularly relevant when we want to share results with local partners and communities (which should be part of the dissemination and feedback plan!). Another tip for ensuring confidentiality is to omit the names of survey respondents from electronic analysis files.

Additional Resources

- Mercy Corps' *DM&E Tip Sheet #2: Designing the Survey Tool*
- Mercy Corps' *DM&E Tip Sheet #3: Sampling*
- Mercy Corps' *DM&E Tip Sheet #5: Baselines*
- Mercy Corps' *DM&E Tip Sheet #8: Data Management*
- Mercy Corps' *DM&E Tip Sheet #11: Data Analysis*
- SPSS Survey Tips: http://www.spss.com/PDFs/STIPlr.pdf#search=%22spss%20survey%20tips%22
- "What is a Survey?" Fritz Scheuren. http://www.whatisasurvey.info/
- "The Survey System." www.surveysystem.com/sdesign.htm

Survey Tool Design Tip Sheet

This tip sheet is for managers and staff designing a survey tool or questionnaire. Further pages provide guidance on structuring the questionnaire and writing and coding questions, including examples.

Introduction to Survey Tools

A survey tool is the written questionnaire that we use to administer a survey. Using a few simple techniques when designing the survey tool can help us to get much better results. This tip sheet discusses the basic elements of survey tool design and how to create good survey questions.[18] For more detailed information, additional resources and references are listed at the end of this document.

What are the critical factors to consider?

A basic principle of survey design is known as **KISS: Keep It Short & Simple**. The most effective surveys are focused on gathering very specific information in the most concise way possible. Surveys that are too long can intimidate or tire out potential respondents and cause delays; about one hour should be the maximum time allotted. Overly complicated surveys can cause misunderstandings. These outcomes decrease the legitimacy and usefulness of our results. We should only include questions that link to a specific information need and are directly related to our survey's purpose.

EFFECTIVE SURVEY TOOLS:

- Keep It Short and Simple (KISS)
- Address key information gaps
- Are tailored to the audience
- Use close-ended questions

We also need to consider the underlying traits of the target population and context[19]. This will influence which approaches and topics are most appropriate. Survey team composition, including gender, age, ethnicity, and technical capacity, is also important in thinking about how questions are framed and which subjects are included. Language, access and logistical concerns are other factors that we need to bear in mind.

Conducting documentation review, interviews, and focus groups are some of the ways we can identify informational gaps and further refine the topics to be covered in our survey. As much as possible, we should ensure that the tool is efficient in gathering only that information which is critical and not found elsewhere. Pilot testing the survey tool in order to identify problem areas and refine questions is key to improving its effectiveness.

Developing the Survey Tool Components

Introduction. Whether written or oral, our survey should contain an introduction that explains the purpose of the study. This is important in building trust and confidentiality. The introduction should explain:

a. Who we (the interviewer) are, including our name and who we work for.

b. Why the survey is being conducted and how the information will be used.

c. How the information will be kept confidential (only if we can ensure confidentiality!).

d. How long the survey will take and what kinds of questions it will contain.

It often helps to have this introduction written out ahead of time on the survey guide for the interviewer to read from. If our survey is part of a needs assessment or other activity that may or may not result in an intervention, it is important that we manage expectations by frankly discussing the informational uses and what, if any, interventions may follow. At the end of the introduction we should ask if the respondent is willing to participate in the survey, and be sure to thank them for their time.

Socio-Demographic Questions. In order to identify any trends in results for different subsets of the population, we have to ask socio-demographic information. For example, it may be useful to analyze information by age group, education, income, gender, region, etc. Demographic questions are generally at the beginning of a survey, and should be coded. The kinds of socio-demographic information to be included are ultimately determined by how we envision analyzing the results and what the informational needs are.

Content Questions. These questions are the heart of our survey. They seek to capture the information needed to understand what is happening in the target population. While length can vary depending on the goal and the context, we should keep this section as short as possible. Some practitioners recommend a limit of 10 to 15 well-constructed questions. Questions should reflect only that information which is critically important, usually as dictated by the project *logframe and indicator plan*[20].

Wrap-up. In concluding, we should thank the respondent and explain once again the purpose of the survey, and how it will be used. Sometimes, a small token gift is appropriate to give as a measure of appreciation for their time and contribution to our study.

---

18 For more tips on survey planning in general, please see *DM&E Tip Sheet #1: Survey Planning*.

19 For guidance in sampling from the target population, please see *DM&E Tip Sheet # 3: Sampling*.

20 For more information on standard DM&E tools and methodology, please see the Mercy Corps *Design, Monitoring and Evaluation (DM&E) Guidebook*.

In addition to the survey tool itself, it's sometimes useful to develop an additional guide, or reference sheet, for the interviewer. This may contain alternative, agreed-upon ways for rephrasing questions if the respondent does not understand, or approaches for probing further if presented with a certain type of response.

Designing Good Questions

Good survey questions require a lot of forethought and pilot testing. An effective survey question is one where the question being asked is clear and easily understood. It is interpreted by all readers in the same way, and captures the needed data in the appropriate format. Key elements in designing effective survey questions are presented below, with examples showing common mistakes and how to improve upon them.

Vocabulary. We must choose our words carefully so that all respondents understand the question in the same way. If technical or more complicated terms must be used, we should define them in simple language within the question. We should use language that asks for the specific information needed and does not require conceptual interpretation.

QUESTION DESIGN TIPS – VOCABULARY

Example of What NOT to ask: **Have you learned anything new while working lately?**

Why: "Anything new" could be just that – anything. "While working" could include informal work, house work, helping out a friend, or anything else the respondent considers "work." "Lately" to some could mean last week; to others last year.

Better: **In the past two years, have you learned a new technique related to the work you do to earn an income?**

Question Ordering. Survey respondents need a little time to get comfortable with the survey process before they will divulge sensitive or personal information. Our questions should start with the easiest or most general and move towards the more sensitive and specific. Questions about a similar subject should be grouped together to make it easier for the respondent to focus on one issue at a time.

Avoiding double-barreled questions. The use of double-barreled questions is one of the most common survey mistakes. A good survey question will ask only one question. Double-barreled questions are those in which two questions are built into one, leading to different interpretations and problems with results analysis. Often, if a question contains the words "and" or "or," it risks being double-barreled. If the question we are asking is too complex to be covered in a single question, we need to use multiple questions.

QUESTION DESIGN TIPS – DOUBLE-BARRELED QUESTIONS

Example of what NOT to ask: **Has the project provided high-quality services that met all of your needs?**

Why: The respondent could feel that the program provides high quality services but that not all needs are met. Conversely, the interviewee could be of the opinion that all needs are met sufficiently even though service quality is just average.

Better: **In your view, did the project provide high quality services?**

Avoiding Bias. Eliminating potential sources of bias in our survey questions is another significant challenge. A question is considered biased when the wording or context influences responses in any way. In survey question design, bias can occur when subtle assumptions or assertions are made within the question. It often happens when judgmental or descriptive language is used, or when what we think are "common sense" assumptions are imbedded in a question's logic. We can help minimize bias by using consistent and neutral language to describe situations and ideas. Emotionally-charged language tends to provoke unintended reactions or sentiments. We can also limit questions that introduce new concepts.

QUESTION DESIGN TIPS – BIAS

Example of what NOT to ask: **Given the impacts on your child's health, do you think that exclusive breastfeeding is a good idea?**

Why: "Given impacts on your child's health" implies a judgment that the behavior is good or bad, and assumes that the respondent has the same assessment as you. "Good idea" is a value-laden term. Overall, the question seems to hint at the answer: "of course it is a good idea."

Better: **Do you practice exclusive breastfeeding during the first six months of your child's infancy?**

Closed vs. open-ended questions. Whenever possible, survey questions should be closed-ended. Closed-ended questions provide a set list of possible answers for respondents to choose from. These generally come in the form of multiple choice lists, yes/no options, or value ranges (agree-disagree, etc.). They facilitate later data analysis by making it easier to quantify responses. Any use of open-ended questions in a survey should be kept at a minimum. The range of possible responses to an open-ended question makes analysis extremely difficult. They can be quite time-consuming to go through, are difficult to group, and often invite multiple interpretations. This is why open-ended questions are generally better suited for focus groups or other interviewing methods.

QUESTION DESIGN TIPS – OPEN VS. CLOSE-ENDED

Example of an open-ended question: **What types of problems have you approached local authorities to help solve?** Respondents may cite a wide variety of problems in many different ways, which the survey team will be forced to analyze and group. This wastes time and resources, and causes the data to be less credible.

Example of the same question closed: **Have you approached local authorities to help solve a problem in the last year?** (Yes/No) **If yes, what type of problem was it?**

☐ personal ☐ community relations ☐ business ☐ land ☐ services ☐ legal

☐ security ☐ other (please specify) _____

The answer options we provide should be specific to the data we are interested in gathering. They should be tailored and adjusted through pilot testing. It is important that we are open-minded about the range of possible responses, and try not to introduce bias. When writing closed questions, we should always include a "Don't Know", "Not Applicable", and/or "No Response" option. It is also good practice to include an "Other" category, in case the listed responses are not sufficient. The key is not to leave the interviewer guessing on what to do when the respondent provides an answer that does not easily fit in one of the available response options. We should always strive to ensure that no questions are left unmarked!

Rating scales. Closed questions will often include rating scales as answer options. Rating scales ask the respondent to make an evaluation about a statement. Numeric values can be assigned to rating scales to make analysis easier. Rating scales should generally have five to seven points, with the middle point being a neutral option. After a long string of questions involving rating scales, a respondent will tend to become

COMMON RATING SCALES:

a) Excellent	**a)** Strongly Agree
b) Good	**b)** Agree
c) Average	**c)** Not sure
d) Mediocre	**d)** Disagree
e) Poor	**e)** Strongly Disagree

"habituated" and is more likely to think about the question less and answer with the same option repeatedly. Because of this, we should use rating scales sparingly and interspersed with other types of questions.

Time period. The time period being referred to in a question should be absolutely clear. A respondent may answer a question about "today" differently than they would about "this week." Questions should focus on current information that respondents can easily provide. Responses about the past (more than one week) and the planned future are notoriously unreliable. If a survey needs to ask questions about the past, these questions should reference the same time frame and be asked sequentially so that a respondent can concentrate on remembering a specific time frame.

QUESTION DESIGN TIPS – TIME PERIOD

Example of what NOT to ask: **How many times over the last three years have you been unable to work due to illness?**

Why: This question is virtually impossible for anyone to answer accurately, unless the number of times is extremely small or none. Additionally, does one "time" mean each day of illness, or does a week of illness count as one "time?"

Better: **Over the last month, how many days of work have you missed due to illness?**

Local language translations. Be sure to clearly discuss as a team, and write out if possible, exactly how the questions will be translated into local languages. If each interviewer is left to his or her own interpretation or translation, subtle differences in wording may lead to very different types of responses.

Additional Resources

- Mercy Corps *DM&E Tip Sheet #1: Planning a Survey*

- "The Survey System." www.surveysystem.com/sdesign.htm

- "What is a Survey?" Fritz Scheuren. Available at http://www.whatisasurvey.info/. Available in Arabic.

- *EZ Questionnaire* survey tips: http://www.ezquestionnaire.com/support/surveytips.asp

Sampling Tip Sheet

This tip sheet is for managers and staff developing sampling strategies for surveys and other M&E activities. Further pages detail key issues and methods, including examples.

Introduction to sampling

Sampling is what we do when we select a group of individuals to study in order to obtain information about a larger population. Sampling is very important in determining the credibility of our data. It can be a tricky process, but luckily for us there are some simple approaches we can use to make sampling easier. This tip sheet provides insights on how to use these methods in a field setting, with practical examples and a breakdown of various strategies provided at the end of the chapter. For more detailed information on sampling, please see the *Additional Resources* section at the end of this document.

> **EFFECTIVE SAMPLING:**
> - Strives for random selection at every level
> - Gives a sample representative of the target population
> - Balances time and resource constraints with sample size and methodology requirements

Sampling methods are useful for both quantitative and qualitative data collection, though they are particularly important for quantitative M&E activities, such as surveys. Sampling reduces the time and costs associated with having to interview the entire population, which is often not feasible. However, the results we obtain through sampling are only *estimates* of the true values for the population, since the sample includes some, but not all, of the people. Appropriate sampling techniques are therefore important to ensure that our results represent the characteristics inherent in our overall target population as much as possible.

Defining the unit of analysis

An informed sampling strategy begins by asking the question: what do we want to measure, and from whom? The project *logframe* is a good reference for determining this, as well as other key planning documents like the *indicator plan* and the evaluation *scope of work*.[21]

For example, evaluations often attempt to measure changes in knowledge, attitudes or behaviors as a result of an intervention. This could mean measuring the objective-level indicator, "Percentage of mothers aware of at least two pregnancy-related danger signs." In this case, the "what to measure" is the percentage of people displaying a certain level of awareness, and the "from whom", or unit of analysis, is mothers in the target area. The unit of analysis may pertain to individuals, households, associations, communities, etc.

21 See the *Design, Monitoring & Evaluation* (*DM&E*) *Guidebook* on the Digital Library.

Determining the sample size

Contrary to common belief, it is not recommended to select a valid sample size simply by calculating a certain percentage of the population. Instead, we recommend the following steps, with more detail in the sections below:

1. Think about the type of data we are collecting and how it will be used, in order to assess the level of precision or accuracy required.

2. Consider the types and number of sub-groups we'd like to credibly analyze and compare.

3. Determine how other factors may impact sample size, such as population size and sampling methodology.

4. Think realistically about tradeoffs between size and feasibility. Overly large samples can drain resources and take time away from other important tasks.

Taking into account all of these factors can be difficult, especially given that some elements, such as variation, are not generally known. A good rule of thumb in the field, therefore, is to first ensure that numbers are large enough to cover the minimum required for key sub-groups (about 25-30; see the section, "Ensuring sufficient size for key sub-groups" below for more information), then increase the sample size in line with what is feasible or appropriate given time and resource constraints.

Sample size and level of precision. *Precision* has to do with the level of certainty with which we are seeking to make statements about the target population.[22] A larger sample size increases precision. For example, conclusions drawn from a sample of 300 are much more likely to be statistically credible than similar data drawn from a sample of only 50. Every increase in sample size increases the likelihood that our results will be an accurate assessment of the population.

However, the ability to improve precision by increasing sample size is often only particularly strong for the first few hundred or so units. After that, it may not be worth the added costs and effort that comes with large increases in sample size. The *quantity* of the data is less important than the *quality* of the data.

There are sample size calculators online that can help ensure a statistically robust sample size. A good one is at: http://www.raosoft.com/samplesize. html. Keep in mind, however, that sample size calculators often recommend larger samples than are necessary or feasibly practical in the field. If using an online calculator, be sure to keep a record of the calculation for documentation and reporting purposes.

22 See *DM&E Tip Sheet #11: Data Analysis* for more on statistical significance testing and margin of error, as well as the additional resources listed there as well as at the end of this tip sheet.

Level of precision required varies according to survey needs. Usually in Mercy Corps projects, we use basic descriptive statistics such as numbers, averages and percentages to make comparisons across time and groupings, or comparisons of means in standard two-by-two tables. In these cases, the simple rule of thumb for sampling based on sub-group size requirements will suffice (described in the next text box). Indeed, larger than necessary samples can be a burden, lowering the quality of the data collection process, causing delays, draining budgets and reducing the time available for important data management and analysis tasks.

There are a few times, however, when a greater degree of statistical rigor is needed. For example, a comprehensive survey measuring child malnutrition may require a high level of precision, due to industry standards and the importance of small percentage changes.

Likewise, if we are seeking to publish our results, or if we are piloting a new technique, there may also be a need for increased precision. For more technical surveys, we may wish to consult Mercy Corps DM&E staff or an external statistician.

Ensuring sufficient size for key sub-groups. One of the most important steps in sampling is taking into account any sub-groupings that we might be interested in describing in the final report. For example, if our unit of analysis is individuals, we may want to be able to analyze results by gender, age, or other demographics. If households are our unit of analysis, we may be interested in characteristics such as household income level or area of origin. Other key sub-groups may include users and non-users of a certain service, or different groups participating in different aspects or phases of a project. We want to make sure we have large enough numbers of people or households within each of these groups in our sample so that our results, when broken down by sub-groups, remain statistically viable.

SUB-GROUPS & SAMPLING:

- We should ensure adequate minimum numbers for each sub-group that we need to make comparisons on or draw conclusions about.
- The rule of thumb for ensuring meaningful results is 25-30 units in each sub-group.
- We can then make increases to our sample size to improve the precision of our results, depending on what is feasible and appropriate.
- Sub-groups can include those based on community or region, gender, income, education, ethnicity, household type, age, etc.

For example, let's say we want to be sure to capture women who regularly use a health clinic within our overall sample, and we know that this group represents about 20% of the female population. Our overall sample size should be sufficiently large so that this 20% contains enough people (at least 25-30) to make statistically valid conclusions about them. Using simple or systematic random sampling techniques (as described in subsequent sections), this would require a total sample size of at least 125 to 150 women, since 20% of 150 = 30. It may also be worth adding an extra 10% or so to the sample in case the random selection does not result in the expected distribution (for example, we may get far less than 20%), and to account for possible non-responses or lack of availability. Example 1 at the end of this chapter provides another example of how to sample based on sub-groups.

It may be, however, that applying the full survey to 150 women just to ensure adequate numbers for one sub-group is not entirely feasible or necessary. An alternative approach would be to use stratified sampling techniques, which treat the sub-group as a different population to be sampled specifically and separate from the general population. This approach can help prevent our overall sample size from becoming too large. For example, it may be possible to randomly sample 30 health clinic attendees, then to generate a separate random sample of 30 or more women who do not attend the clinic and compare them. More information on using stratified sampling for sub-group analysis is described below.

Population size is less important than people generally assume. In general, the larger the population size the bigger our sample has to be to sustain the same level of statistical accuracy. However, this is less relevant once the population size reaches past a few hundred. For example, if the population size is 10,000, the needed sample for a certain level of precision is 223. When the population size increases to 200,000, the corresponding suggested sample size is only 228 (according to the online calculator). Unless the population is under a few hundred, large changes in the population size do not affect sample size requirements by all that much.

Tradeoffs: size vs. feasibility. There are generally tradeoffs associated with obtaining an ideal sample size versus doing what is feasible given time, resource, and logistical constraints. Overly ambitious sampling efforts can detract resources from other important M&E efforts and lead to rushed surveys or delays, decreasing the quality of our overall effort. Likewise, too little attention to sample size requirements may jeopardize our ability to make statements about important sub-groups. The optimal solution is to remain practical and focus on what is logically feasible, while at the same time taking into account minimum sample size requirements to ensure meaningful results.

Selecting the sample

In Mercy Corps, we employ many different approaches for selecting our sample, but the key principle to remember is *random sampling*. This means that once we set the overall population upon which our sample will be based, the units we select within that population are chosen entirely by chance. This is the best way to ensure that our sample is representative of the population we are trying to analyze. In general, we should always strive to randomly select the people we survey. Even if we are targeting specific groups, people or households within those groups should be sampled using random methods.

> **KEY DEFINITIONS:**
>
> - *Random Sampling:* The deliberate process of selecting a sample so that each individual or household chosen is selected entirely by chance.
> - *Sources of Bias:* Any factors introduced during the sampling process, either intentionally or unintentionally, that may influence our results, causing them to differ from the true characteristics of the population.

The purest form of random sampling is called *simple random sampling*. This implies that all of the units (individuals, households, etc.) are chosen individually and directly through a random process and that each unit has an equal chance of being selected into the sample. To do this, we need to have a list of all the possible respondents. For example, if we have a list of all of the participants in a business skills training project, we would simply take this list and randomly choose people from it to be included in the sample. One of the easiest ways to randomly select units from a list is to use Excel functions that generate random numbers, which correspond to the numbered individuals or households in our list. An illustration of how to do this is presented in Example 3 at the end of the chapter.

A similar approach commonly used is called *systematic random sampling*. This approach first calculates a *sampling interval;* which is the overall population size divided by the desired sample size. If we have a list of sampling units, then a starting point on the list within this sampling interval range is chosen randomly, and additional selections are chosen by adding the sampling interval to the numbers already chosen. This can be useful to select from a list that is ordered in some way (alphabetically, geographically, by age, business type, etc.) to ensure good representation across the list. A detailed example of this approach is described in Example 2 at the end of this chapter.

Systematic random sampling **can be conducted in the field without a list.** Often, we do not have the luxury of having a complete listing of households or individuals that represent the population we are sampling. In these cases, we use the same sampling interval principles as we would with a list, but instead apply it in-person on the ground. This can be done by interviewing every 7th, 10th, or 15th household in a given community, depending on what the sampling interval may be. The key to this approach is ensuring that every household in the community has a chance of being selected.

To do this, we first need to determine a feasible sampling interval, based on both the minimum numbers required as well as logistical factors. The same principles used above and in the examples at the end of this chapter can be used to determine our sample size. To implement the sampling interval approach in a community, we need to sketch out a rough mapping of where people live. Then, we can determine a route that our interviewers will follow. This route should encompass all households or individuals in the community, so that we obtain an accurate representation. We then follow the same procedure for the other communities chosen.

> **KEY DEFINITIONS:**
>
> - *Simple and systematic random sampling:* When each unit (person, household, community, etc.) of our sample is selected at random (by chance), and each unit has the same chance of being selected.
> - *Stratified sampling:* When the population is broken down into different sub-groups, based on demographic profile, type of experience with the project, or other characteristics. Units are randomly selected within each group, ensuring sufficient size to be able to analyze and compare.
> - *Cluster sampling:* When the population is broken down into geographic areas, or clusters, which can then be used as sampling units from which to select a smaller sub-sample. It is most relevant when we do not have a listing of all the units, or are seeking to save time and limit costs in a population that is geographically dispersed.

Stratified sampling **can help ensure representation by particular sub-groups (or strata) of interest.** Generating a simple or systematic random sample large enough to include adequate numbers for key sub-groups can sometimes result in sample sizes that are too large to implement. An alternative approach is stratified sampling, which breaks the population of interest down into different sub-groups, usually based on demographic data or exposure to the program being evaluated. Then, individual units are randomly sampled within each sub-group or strata (using simple or systematic random sampling). We need to have adequate numbers in these

sub-groupings to ensure that we can analyze the data to look for trends in sub-groups of interest. We should be careful here about generalizing results to the larger population, since sub-groups may be over- or under-represented proportionate to their representation in the overall population.

For example, in our health clinic example above, using a stratified sampling approach we might randomly sample 30 regular health clinic attendees and compare this data with a randomly selected group of 30 or so non-attendees. While this may be useful for the sake of sub-group comparison, we cannot simply aggregate the results and claim that they represent the characteristics of the overall population. This is because regular clinic attendees comprise only about 20% of the actual overall population, while in our aggregated sample they would comprise 50% and be over-represented.

If comparing sub-groups is the priority and time is limited, a stratified sample with two groups of at least 30, as described above, is a good approach. However, if want to credibly analyze sub-groups *and* the overall population, then it is better to go with one general sample that is large enough to ensure adequate numbers of key sub-groups based on their proportional representation in the population. Examples of doing this were discussed previously, as well as at the end of the chapter. This is a good example of the kinds of tradeoffs that need to be made when choosing a sample; in this case, tradeoffs based on size, feasibility, and different kinds of analysis preferences (sub-groups vs. general population).

Cluster sampling **can save time and resources.** For example, we may wish to analyze access to clean water among the population of a certain province. It might be too much work to try to construct a listing of all the households in the province. Even if we do have a listing, it may be too costly and geographically difficult to interview each widely-dispersed household that is selected through a process of simple or systematic random sampling.

Cluster sampling breaks the population down into different physically-defined areas, called clusters. We can define clusters using a map, a list of communities, or some other form of local knowledge. Clusters need to be well-defined physically or administratively so that their boundaries are easily identifiable in the field. Regions, communities, or neighborhoods, or

THOROUGHLY DOCUMENT THE SAMPLING METHODOLOGY:

Sampling strategies and the rationale for making key decisions should be described in detail in the final report.

This will increase credibility and transparency about how our results were obtained.

logical geographical separations created by physical boundaries such as roads, rivers, and mountains, are generally good ways to organize clusters. Clusters may also emerge due to the operational considerations of the project and how it is administered, such as groupings of communities located geographically close to one another.

STEPS IN CLUSTER SAMPLING:

1. Define clusters, or groupings of communities or households.
2. Randomly select a sub-set of these clusters to survey.
3. Interview either all or randomly selected units within the clusters.

We can randomly select the clusters we will include in our sample after developing a listing of all the different clusters in our project area. To randomly select our clusters from a list, we can use an approach similar to those described above using Excel or systematic random sampling. We might want to incorporate the population size, or number of project participants, of each cluster into our selection process. This will make our sample more representative.

Often, two-stage and multi-stage cluster sampling may be required. Once we have chosen the clusters, we can either interview all subjects within those clusters or randomly sample individual units within each cluster selected (depending on the size of the cluster). Randomly sampling individual units within the selected clusters is called *two-stage cluster sampling*. Other times, we will need to randomly select units within the cluster that fall in particular sub-groups or strata. In these cases, once the clusters are selected we then break those clusters down into the sub-groups of interest and randomly select among them. This is a process called *multi-stage cluster sampling*.

DON'T OVERLOOK MARGINALIZED GROUPS:

Some sampling strategies can unintentionally under-represent poor or marginalized segments of the population. For example, these individuals may not appear on official lists like voter registration rolls or school lists, or may live in areas with limited access.

If these groups are among the targeted beneficiaries of an intervention or if the survey is intended to be generalized to the entire community, we should be extra careful to ensure they are included proportionately in the sample and that we have not unintentionally biased our sample towards more affluent or accessible populations.

In two-stage or multi-stage cluster sampling, we may wish to limit the number of communities we select, and interview more people in each community, so that the sampling interval remains practical. For example, interviewing only 1 in every 40 households across several communities would consume a lot of time and energy in

traveling from house to house. It may be more practical to interview 1 in every 6 households in a smaller number of communities to come up with the same overall sample size. While including fewer communities may risk under-representing certain geographically-based groups, it may be far more feasible given the circumstances.

One potential disadvantage of cluster sampling is that we may miss out on key segments of the population, particularly if they are organized geographically. For example, if ethnic minority communities are concentrated primarily in two out of fifteen cluster areas, and these clusters are not randomly selected in our sample of clusters, we may miss out on key data. In this case, we may want to combine a purposive sampling approach with the cluster approach to ensure that certain communities or sub-groups are included.

***Bias* occurs when the sample chosen is influenced in some way by the selection process.** This decreases the ability of the sample to mirror the characteristics of the population. Every time we make a sampling choice that is not random, we can introduce bias. For example, for logistical reasons, we may prefer to interview people who are easiest to reach. However, our results would be biased towards the populations that are most accessible and would not accurately represent the entire population. Indeed, those that are more accessible are likely to be better off, biasing our results upward. Bias often occurs in subtle ways that we are not fully aware of, and can very easily decrease the validity of our data. For this reason, the best rule of thumb is to strive for random selection at every level.
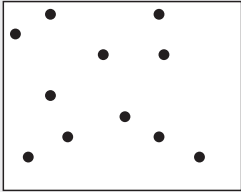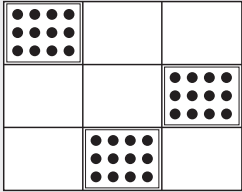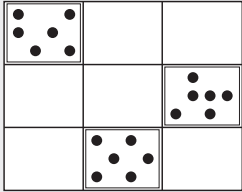
Different sample selection methods require different sample sizes to retain accuracy. Any technique beyond simple random sampling may require increasing the sample size. In doing so we are introducing potential sources of *bias,* or outside influences, which can alter the representative nature of the sample. Some practitioners suggest doubling the sample size for each successive grouping in the selection process. A high rate of non-response also calls for increasing the sample size.

Additional Resources

For more detailed assistance on sampling, consult the references below or contact Mercy Corps' Headquarters-based DM&E staff.

- Magnani, Robert. *Sampling Guide*. Top recommended technical guide on sampling in development projects. http://www.fantaproject.org/downloads/pdfs/sampling.pdf

- Online sample size calculator: http://www.raosoft.com/samplesize.html

- The Statistics Home Page -- http://www.statsoft.com/textbook/stathome.html

- Russ Lenth's Power and Sample Size page: http://www.cs.uiowa.edu/~rlenth/Power/

- JHU on-line stats lecture: http://ocw.jhsph.edu/courses/StatisticalReasoning1/PDFs/Lecture1.pdf

- Margin of Error and Confidence Intervals Made Simple: http://www.isixsigma.com/library/content/c040607a.asp

- EPI Info 2002 data management software: http://www.cdc.gov/EpiInfo/epiinfo.htm

Comparison of Sampling Methods

Simple random sampling	Cluster sampling	Two-stage or muti-stage cluster sampling
Units are selected randomly from the population at large, generally using a list.	Clusters are selected randomly, and all units within each cluster are selected.	Clusters are selected randomly, and units within each cluster are selected either randomly using lists or on-ground sampling intervals (two-stage cluster sampling), or based on sub-groups or strata of interest (multi-stage cluster sampling).

Common Sampling Methods

Sampling method	When to use	Advantages	Disadvantages
Simple random sampling and systematic random sampling. Households selected at random from a list, or using sampling interval (i.e., every n^{th} number on list, or n^{th} household in a community).	• To maximize the representative nature of the sample with regard to the target population. • The target population is clearly defined.	• Simple to construct if we have a reliable list. • Minimizes chance of sampling error. • Relatively easy and practical way to ensure representative sample.	• Costly/time-consuming if sample is widely dispersed. • Requires extensive and complete population list. • May not ensure adequate numbers for sub-groups.
Stratified sampling. Select specific sub-groups and randomly sample units within those subgroups.	• When we want to ensure that key sub-groups are represented in sufficient numbers, regardless of their size relative to the overall population.	• Ensures that we'll be able to credibly analyze and compare results for key subgroups.	• Harder to aggregate results and generalize to the larger population, since sampling sub-groups are not proportionately representative.
Cluster sampling (multiple forms). Designate clusters according to areas of project activity or geographical boundaries. Randomly select a few of the clusters, and interview all units or a random selection of units in each.	• Time and budget is limited. • Population is geographically disperse. • Sampling frame, or list, is incomplete or unreliable.	• Saves time and resources.	• Can miss key sub-groups if they are organized geographically. • Results could vary widely depending on clusters selected.

SAMPLING EXAMPLES

DETERMINING SAMPLE SIZE BASED ON SUB-GROUP ANALYSIS

Example 1: Suppose we wanted to analyze the results of an intervention by income level across three communities. The two sub-groupings of interest are communities and income level. Using our minimum size requirement of 30, the sample size breakdown for each sub-group could be as follows:

- Community 1, income level 1 = 30
- Community 1, income level 2 = 30
- Community 2, income level 1 = 30
- Community 2, income level 2 = 30
- Community 3, income level 1 = 30
- Community 3, income level 2 = 30

Total minimum sample size required = 180

When multiple sub-groups are added, one useful rule of thumb is to ensure that we have the minimum sample size requirement for the least common sub-group. Another approach is to pick two or three sub-groups which are of most importance to the project, and sample based on the minimum size requirements for these. It is important to remember that if we purposely target or over-represent key sub-groups, we cannot draw conclusions about the wider population based on our data. If we wish to do so, we will need to readjust so that the sample is consistent with the larger population.

SELECTING A SAMPLE USING SYSTEMATIC RANDOM SAMPLING

Example 2: The table to the right shows the names of participants in a business training project. To select a random sample from this list:

1. Put all units into a numbered list, as shown.

2. Decide the desired sample size – let's say 30 for this example.

3. Divide the total population size (150) by the desired sample size (30). This gives us our *sampling interval,* which is 150/30 = 5.
   A sampling interval is defined as the interval between two successive sampling points.

4. Pick a random number within the sampling interval (between 1 and 5). This number becomes our random start. To do this, you can draw from a hat, or use the RAND function in Excel.

5. Let's say the number you randomly picked is 2. This means that our *random start* is the second number on the list, which will be incorporated into our sample.

6. We then count down our list by the sampling interval (5). 5 spots down on the list takes us to number 7, which is then selected.

7. We continue to go down the list, repeating this process until we have selected 30 units from the list.

1	Abdallah	
2	Arif	X
3	Azimi	
4	Blasevic	
5	Bolotkan	
6	Cevallos	
7	Chikalanga	X
8	Cye	
9	Dalimare	
10	Demas	
11	Dieuaide	
12	Ekedahl	X
13	Evanitka	
14	Faku	
15	Floyd	
…	….	…
150	Zacuto	

RANDOMLY SELECTING SAMPLING UNITS FROM A LIST IN EXCEL

Example 3:

1. Put all units into a numbered list, as shown on the right (could be individuals or households).

2. Enter this formula into a cell: =RAND()*(150-1)+1, where 150 is the number of the last unit, and 1 is the number of the first unit. This will generate a random number between 1 and 150. In our example, as shown in the results column, 148 was the first random number generated. Therefore, the first unit in our sample will be the name corresponding to number 148 on our list, which in our case is Yeley.

3. The second random number generated is 3. Therefore, in our case, Azimi is added to the sample.

1	Abdallah	
2	Arif	
3	Azimi	X
4	Blasevic	
…	…	…
148	Yeley	X
149	Zacuto	
150	Zuluku	
	Random calculation	Results from calculation:
1	=RAND()*(150-1)+1	148
2	=RAND()*(150-1)+1	3
3	=RAND()*(150-1)+1	65
4	=RAND()*(150-1)+1	18
5	=RAND()*(150-1)+1	98
…	…	…
30	=RAND()*(150-1)+1	115

4. Repeat this process 30 times to obtain the 30 units that we will use for our sample.

5. If the same number is generated twice, run the calculation again to get a different number.

Note: Another way to easily generate random numbers, with no repeats, is to go to the Web site: *http://www.mdani.demon.co.uk/para/random.htm* and enter the appropriate numbers.

Participatory M&E Tip Sheet

This section is for managers and staff aiming to incorporate more participatory approaches into their M&E activities. Please note that the field of participatory M&E is vast, and this tip sheet does not purport to summarize all possible approaches. Please consult the Additional Resources at the end of this section for further information.

Why participatory M&E?

A central purpose of monitoring and evaluation is to provide information to people to make decisions. The more project participants and other stakeholders are able to lead the process of information gathering and analysis, the more they will be empowered to act on it to improve programs. Evidence shows that increasing local stakeholder participation in M&E efforts helps build ownership and enhances sustainability, leading to improved impact.

How can we make M&E more participatory?

There are a range of ideas about what is meant by participatory monitoring and evaluation. On one end of the spectrum, there is the ideal scenario where communities and partners are empowered to lead the entire process. This could include project design, establishing indicators, leading data collection and analysis and making program decisions. On the other end of the spectrum, some believe that simply including beneficiary viewpoints through surveys or focus groups represents participatory M&E (it does not).

> **EXAMPLES OF PARTICIPATORY M&E:**
>
> - Stakeholders lead in designing program indicators and/or adapting locally-relevant indicators.
> - Lead role for participants in collecting and analyzing baseline & evaluation data.
> - Monitoring of certain indicators becomes responsibility of partners/ beneficiaries.
> - Use of participatory evaluation methods such as PRA, self or peer evaluation, participatory impact assessment, photo voice, most significant change, etc.
> - Heightened stakeholder involvement in analysis and dissemination workshops.

The most feasible option for many projects lies somewhere in between. While increasing participation in M&E can be worthwhile, it can also be time-consuming and often requires significant capacity building efforts.

The central question therefore becomes: How can we begin to incorporate more participatory elements into our M&E activities? The text box above and the remaining pages of this tip sheet provide insights and suggestions for doing this.

When to build in participatory M&E?

M&E activities can be made more participatory at any point during the project life-cycle. This includes:

- **In the assessment/design/proposal development phase.** This could include local stakeholder contributions to program design and project indicators. We can also build in language and objectives or activities relating to participatory M&E in the proposal narrative and logframe, with corresponding time and budget allocations.

- **During the project start-up phase.** This can be an excellent time to organize M&E systems with staff, partners, beneficiaries and other stakeholders, and see where and how various actors can contribute to M&E. These roles can be built into the indicator plan and workplan, as well as position descriptions and partner agreements or MOUs.

- **During ongoing implementation.** At any point during the project, we may recognize that certain aspects of ongoing monitoring can be transferred to local partners or participants, particularly as relationships evolve and capacity grows over time.

- **At project mid-term and final evaluations.** Evaluations or field studies also provide opportunities for participation in research design, data collection and analysis. There are a range of approaches for building participatory methods into evaluations, as will be explored later in the tip sheet.

When to scale back participatory M&E efforts?

There are some cases where it might be less appropriate to incorporate fully participatory M&E mechanisms. Barriers may include a lack of time or budget for adequate local capacity building, or instances when we only need certain information for our own or our donor's benefit. For example, emergency response projects or assessments, or strict impact assessments for donor accountability, may be less suitable to full-scale PM&E techniques. We should be realistic in recognizing the time and cost investments involved, and strive for a scope of participation in M&E that is appropriate and likely to work well. Note that this tip sheet only relates to participation in M&E; it is still good practice to incorporate other forms of beneficiary and stakeholder participation throughout ongoing project implementation.

Participatory M&E during project design and start-up

Discussing and developing participatory M&E elements during project design and/or start-up in an inclusive and consultative manner with local stakeholders is the best way to get them on board with M&E as well as the overall project approach. There are many ways to incorporate participatory M&E planning and methods into a project design and start up:

- **Participation in assessments.** Pre-project assessments that are part of new project design processes can be opportunities to build the participatory M&E experience of existing or future partners and beneficiaries. Facilitating potential partners and beneficiaries to take an active role in designing and carrying out the assessment can help build local ownership of the design process from the very beginning. PRA (Participatory Rural Appraisal) methods can be particularly effective in exploring a range of issues with communities.

- **Participatory design processes and locally-relevant objectives and indicators.** Once an assessment has been undertaken, participatory design workshops or focus groups with key stakeholders can help give local perspective and make project design more relevant to local needs and conditions. These may also be used to develop locally-driven indicators and objectives for our project design. For instance, participants can define what "success" means to them and propose ways that information could be collected and measured. It also allows us to stress the idea of Mercy Corps working together with local actors towards a common goal. Jointly defining objectives and indicators early in the process helps build ownership and makes it easier to implement a participatory M&E system once the project is underway.[23]

**EXAMPLE: MC SERBIA
INCORPORATING LOCALLY-RELEVANT INDICATORS**

Interviews and focus groups conducted for the Community Revitalization through Democratic Action (CRDA) program in southern Serbia revealed that many residents felt a key indicator for the revitalization of their communities would be the "# of community organized cultural and sporting events." These activities had previously been a valued part of community life in the region, but had decreased due to the effects of the war. Their renewal, residents argued, would be as important an indicator of "revitalization" as the donor-mandated indicators such as "# of new social services" or "% increase in employment."

- **Participatory M&E in proposals.** Emphasizing participatory M&E approaches in new proposals, where appropriate, is a great way to plan for more participatory M&E systems on the ground. In proposals with elements of participatory M&E, we should also ensure that time and money are budgeted appropriately to ensure that it is feasible within the project scope. A good way to do this is to make M&E capacity-building of stakeholders an explicit project objective or activity. This gives it a priority on par with other project activities in terms of level of importance and helps ensure that it will be planned for adequately. If

possible, be specific in outlining the role of partners or beneficiaries in monitoring and/or evaluating specific aspects of the project.

- **Clarifying roles and responsibilities at project kick-off.** During the start-up period, there is usually a workshop or series of meetings to plan activities and procedures. Project start-up offers a great opportunity to build in participatory M&E approaches, which may include:

→ Involving stakeholders in project "re-design", such as further defining vague objectives or indicators in the logframe, clarifying activities in a workplan and agreeing on targets.

→ Developing concrete roles and responsibilities for stakeholders in monitoring and/or evaluating a particular sub-set of indicators. Adjust key documents such as position descriptions for staff and/or contracts or MOUs with partners to reflect changes.

- **Participation in the baseline study.** Consider having stakeholders participate in the design of the baseline so that they fully understand the indicators and measurement mechanisms. Then train them to carry out surveys, focus groups and other methodologies, and/or take on other roles such as data entry or management. Finally, and most importantly, facilitate them to lead or be included in the analysis of results, through a participatory analysis workshop or results dissemination meeting.

**EXAMPLE: MC NIGER
INVOLVING PARTNERS IN A START-UP WORKSHOP AND BASELINE**

The SKYE project in Mercy Corps Niger, focused on youth empowerment, held a 5-day M&E workshop with staff to revise the logframe and organize M&E processes. On the 6th day, they invited government partners and youth association members to discuss their roles in monitoring the peer education objective and participating in baseline data collection.

It was decided that youth peer educators would administer a before/after survey for each training module they carried out, as part of evaluating the indicator related to increased knowledge of basic life skills. They would also assist in collecting baseline data related to income and youth well-being, and analyze results with project staff and partners.

Participatory monitoring

Monitoring is usually about tracking ongoing activities and outputs, such as the number of community meetings or the progress of an irrigation construction project. In considering ways to increase the level of participation, ask the question:

---

23 For an excellent guide to facilitating local indicators, see *Participatory Impact Assessment* on Clearspace.

➜ How can we empower or organize partner or beneficiary groups to track these activities and outputs themselves and collectively analyze progress?

While investments in capacity building may be significant at first, in the end increased monitoring by local partners, beneficiaries, local government and/or other stakeholders can have the dual benefit of increasing ownership and reflective analysis for improvement amongst key participants while also decreasing the monitoring burden of Mercy Corps staff.

A few ideas for developing participatory monitoring approaches:

- **Organize a workshop with partners and beneficiaries** to analyze how they could be more involved in program monitoring, and continue with progress review meetings.

- **Think about activities and capacities of local stakeholders, and play to their strengths.** What information are beneficiary groups already collecting, in what format, and how does it relate to program monitoring? For example, farmer associations often keep internal records of members, attendance, trainings, collective purchases, loans or sales, etc. Health committees, local governments, youth associations and other groups also likely have some experience in record-keeping.

  The more we can build upon existing experience and capacities and tailor our own monitoring approaches so they are similar to already-established processes, the more successful we are likely to be. Conversely, the more we make them adapt to processes and tools that are not as familiar, the greater the capacity-building efforts required.

- **Be open to different methods, tools or technologies – and simplify.** For example, some local groups may be more accustomed to paper-based or oral methods for data collection and analysis. As long as it is done in a systematic way, these can also serve as viable monitoring systems. We should think creatively about how these could be incorporated into our M&E systems – such as MC staff recording

---

EXAMPLE: MC AFGHANISTAN LOCAL MONITORING & SUSTAINABILITY

Mercy Corps trained Veterinary Field Units (VFUs) to regularly monitor key indicators to track progress towards financial sustainability.

Indicators tracked by VFUs included:
- Income (using income/expense sheets)
- Lists of common diseases or sicknesses
- Medication distribution data
- Vaccination data

From a program evaluation report: *"This form of monitoring was very responsive to the local context and was a vital tool in helping establish the sustainability of the VFUs."*

---

or converting local records to electronic format for reports. Local languages should also be used where possible to reduce complication and translations issues for local stakeholders.

Requiring IT or other expertise produces barriers to participation, when instead we should be looking for ways to "de-mystify" M&E so that non-experts can participate. Think about the other skills – such as language ability, unique understanding of the context, and ability to act on information – that local stakeholders will add to the process which make up for lack of expertise in other areas.

- **Participatory analysis is equally important as data collection.** The benefit of participatory M&E systems is that local stakeholders are able to internalize information and make improvements; this cannot happen unless they routinely analyze and reflect on the monitoring information. Simply collecting monitoring data for reporting purposes without sufficient collective analysis defeats the purpose.

  It is therefore critical to systematically plan for regular analysis sessions, ideally led by the local stakeholders involved in collecting the data. Even if data is collected by Mercy Corps staff, we can still have participation of stakeholders in analysis meetings, so that they are aware of program progress or obstacles and can help be part of the solution.

LESSONS FROM MERCY CORPS EVALUATIONS: HOW TO INCREASE PARTICIPATION IN M&E

A <u>food security</u> project evaluation in MC Indonesia noted *"the program should use simple measures that can be collected by communities and used for mobilization and competition among communities."*

"All indicators used should be developed with Government and community partners. Joint MC/Government learning teams should be established to modify approaches together as required."

A field study on <u>infrastructure</u> sustainability in Indonesia learned: *"Mercy Corps should provide capacity building in monitoring & evaluation for project committee members and other key players in the community."*

"If the project is built by the community, Mercy Corps can create a reporting mechanism via community facilitators, project committee, village leaders or sub-district offices."

Increasing participation in baselines and evaluations

Those who learn the most from baselines or evaluations are generally those who participated in carrying them out. Therefore, in order for our staff, partners and beneficiaries to truly internalize lessons learned from evaluations, and to use those lessons to improve future interventions, it is important that we try to build in as much participation as possible. This includes all phases of evaluation, not just data collection:

- **Designing the methodology, tools and SOW.** As with any project, partners, beneficiaries and other stakeholders will feel more ownership of the findings (and be less skeptical of results) if they are involved in evaluation planning stages. This may include jointly developing evaluation objectives and methodology; helping to construct, pilot-test and provide feedback on data collection tools; and planning the community outreach strategy, survey teams and other logistics. Simplify methodology to ease implementation, such as shorter surveys, smaller sample sizes, etc.

> **EXAMPLE: MC UGANDA INVOLVING PARTNERS & COMMUNITIES IN BASELINE & EVALUATION**
>
> MC Uganda's Pader Peace Program mobilized local stakeholders at baseline and mid-term.
>
> *"At each site prior to carrying out the baseline survey, Mercy Corps staff and Pader Peace Forum representatives trained a team of local volunteers from the community's peace committee on how to administer the survey. These small teams of five to ten volunteers then went to administer the survey."*
>
> This participation helped local actors understand how the program was influencing indicators such as access to information, perceptions of conflict resolution mechanisms and overall confidence in the peace process.

- **Collecting the data.** This will require training, but again partners and community members are much more likely to understand and use the findings if they are involved in data collection. They also may be more skilled in accessing various sub-groups or discussing sensitive issues due to their local knowledge and standing in the community.

  Training for data collection should include an in-depth review of the data collection tool, clarification of the meaning of questions and translation issues, and a role play and pilot-testing of the tool, coming back together to analyze how it went and make adjustments.

- **Analyzing findings and producing lessons learned & recommendations.** Active participation in data analysis enables a real understanding of the results. This can happen through a facilitated data analysis workshop with staff, partners, beneficiaries and other stakeholders.

The more that participants are able to play with the data and actively analyze various aspects of it, coming up with their own findings and conclusions, the more they will understand and appreciate the results. For example, an analysis workshop could break down into small groups to analyze results from a particular set of questions in-depth, then come back together to discuss findings and conclusions with the larger group.

Also note that sometimes we need different approaches for different settings. Community-level feedback sessions, for example, may want to look at more simplified graphs or data on flipcharts, whereas an analysis meeting with partners in an office-type setting may enable more technical analysis using the data management software.

> **EXAMPLE: MC SRI LANKA PARTICIPATORY ANALYSIS & FEEDBACK**
>
> Each year, MC Sri Lanka conducted an Annual Results Review, analyzing survey results, including capacity index data, with a wide selection of staff, partners and community representatives. Small groups analyzed various outcomes data and discussed what it means to them, before sharing with the larger group. Together they discussed how it illustrates achievements or areas for improvement, and implications for current and future programming.
>
> NGO partners and community representatives are also responsible for holding feedback sessions in communities. Slightly different methods, such as flip-charts, are used, but the goal is similar - to analyze results, discuss what it means locally and plan next steps.

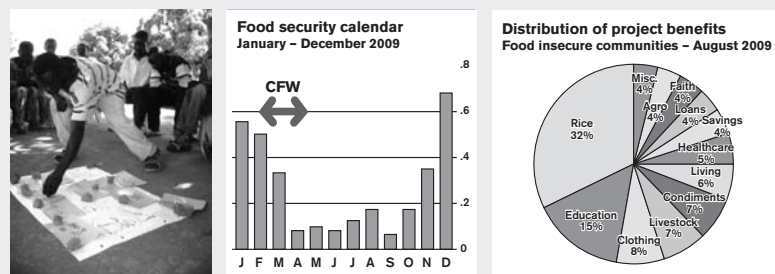CBO Index by Dimension – Ampara CAG Annual 2007 – Bokotuwa
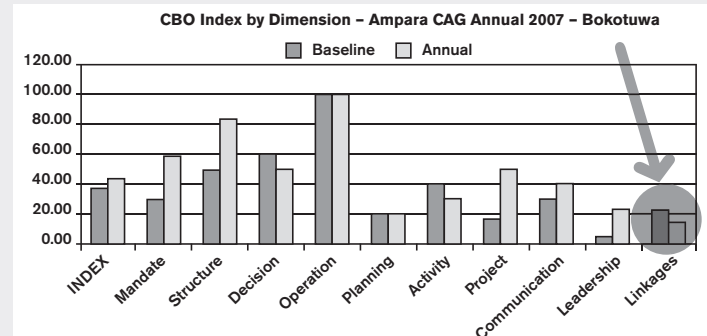
Participatory evaluation methods

Specific techniques geared especially for high participation in evaluations include:

- **Participatory Impact Assessment (PIA) and other PRA-type techniques.** The *Participatory Impact Assessment Guide*[24], developed by five international NGOs over a two year research and field-testing period, is available on Clearspace and provides an excellent guide on how to use a range of Participatory Rural Appraisal-type techniques for evaluation. They include mapping, ranking, charting, proportional piling, observation and other techniques where community participants lead in the analysis. PIA techniques can be used to analyze issues such as the relative benefits from or importance of different interventions, food security, disease mapping, ranking of needs or priorities, and analyzing before/after situations, and are useful in that they can provide both robust quantitative data and rich qualitative data from the perspective of beneficiaries. These methods have been increasingly recognized in recent years by donors and practitioners as legitimate and robust.[25]

**EXAMPLE: MC CAR
PARTICIPATORY IMPACT ASSESSMENT (PIA)**

**Food security calendar
January – December 2009**

**Distribution of project benefits
Food insecure communities – August 2009**

The Gates-funded Global Food Crisis Response project used PIA for their primary evaluation method across five countries. Here, MC CAR used a locally-developed food security calendar to analyze timing vis-à-vis periods of relative food insecurity, and proportional piling methods to analyze how Cash-for-Work payments were spent.

- **Self-evaluations.** An example of this could be an evaluation carried out by the communities or beneficiary groups themselves on projects they've initiated or helped to implement. This helps develop critical analysis tools and a heightened understanding of achievements and lessons learned among local actors. It can be especially useful in a process-oriented evaluation of an ongoing project, such as in a mid-term evaluation.

- **Peer evaluations.** In this methodology, communities or beneficiary groups evaluate projects of peer communities or beneficiary groups. This gives groups exposure to other similar programs and can generate learning about the sector as well as reflections about the groups' own projects through the evaluation of peers. It may also be less open to positive bias as self-evaluation methods.

- **Most Significant Change.** This method essentially consists of collecting a series of personal stories from beneficiaries, by asking them an open question along the lines of: *"What was the most significant change that occurred in your life as a result of your participation in the project?"* The collected stories are then analyzed in successive rounds by different stakeholder groups in order to emerge with the most significant or meaningful examples of changes brought about during the program, and why these were deemed significant. These stories can also be grouped by theme to show the relative importance of various types of impacts in the eyes of beneficiaries. This method may not represent participatory M&E per se, since it is not necessarily participant-led, but it does give a unique, beneficiary-centric perspective on project impact (as opposed to donor or MC-driven indicators of success) and so is included here. The official MSC guide gives the full details of how to apply the technique.[26]

- **Photo Voice.** This method is a good way to gain beneficiaries' perspective of program impact, and results can present a powerful visual and narrative. Steps include:

  1. Distribute cheap disposable or film cameras to participants and do a brief training on how to use them. Instruct them to take pictures according to certain themes. For example, it could be what represents the most significant change in their lives desired or brought about by their participation in the project, or other, more specific themes. A week or two later, collect the cameras and develop the photos.

  2. Then host a participatory analysis workshop whereby:

     a) Each participant is given their full set of photos and selects two or three that best illustrate the impact of the program for them.

     b) In small groups, participants explain why they chose the photos, and they collectively vote on the three best or most representative photos/stories.

---

24 https://clearspace.mercycorps.org/docs/DOC-7685
25 Also see *DM&E Tip Sheet #13: PRA in Evaluation on* the MC Digital Library

26 See Most Significant Change guide at: http://www.mande.co.uk/docs/MSCGuide.pdf

c) The plenary group then analyzes all selected photos and explanations from the small groups, and votes on the top two or three.

d) Winners are announced, and participants invited to explain why they voted for them. Don't forget to record the explanations and the process!

3. The winning photographs and accompanying personal stories can be presented in ceremonies, meetings, conferences, etc., as well as in the form of accompanying qualitative data in reports, presentations, videos, etc. One idea is to host a gallery walk with participants, donors and/or other stakeholders, whereby each participant has a poster of his or her selected photos and explains how they represent impact.[27]

**EXAMPLE: MC INDONESIA
PHOTO VOICE TECHNIQUE**

The HP3/Lestari program, focusing on urban sanitation, wanted a photo baseline to accompany their traditional quantitative baseline. After dividing participants into groups of youth, fathers and mothers, each group was given a basic film camera and trained on how to use it. Participants were then told to photograph key attitudes and behaviors on solid waste management, both positive and negative, important for changing sanitary conditions.

Staff then collected the photos along with participants' explanations, and organized a community panel of judges to select the winners in a fun, participatory event. This resulted in a clearer depiction of the key attitudes and behaviors the community wanted to change through the project.

- **Participatory focus groups using elements of Appreciative Inquiry.** It is also possible to stimulate a higher level of participation and beneficiary-driven perspectives using Appreciative Inquiry principles in a focus group setting. Appreciative Inquiry is a broad overall approach to evaluation that begins with the underlying philosophy that what we focus on becomes our reality. Therefore, an AI approach to evaluation

seeks to shift the focus from problem identification to discovering successes in order to learn from them and enhance factors for increased success.[28] It can be applied to DM&E in a wide variety of ways, many of which are not necessarily participatory in nature. The following example is a very specific use of AI ideas that can be used to generate high participation and beneficiary-led analysis in focus group settings.

Focus group participants split into pairs and spend 5-10 minutes each relating a personal story to the other about a positive experience they had within the project or relating to a specific theme. Guiding questions for facilitating work in pairs are:

1. Share stories of most outstanding experiences related to the topic. What happened? Who was involved? What made it so outstanding?

2. How did this affect the overall situation?

3. How could more situations like this be possible? If you could grant three wishes for making more experiences like this possible, what would they be?

Participants then share examples in plenary, as facilitators record and group the various stories and wishes by theme. As commonalities emerge, discussions can take place about what the results generally seem to indicate regarding people's experiences in the project and implications for future implementation strategies.

This technique could be used in an assessment, baseline, mid-term or final evaluation, and usually results in lively conversations among participants. Most importantly, it provides managers with sets of recommendations from beneficiaries themselves on how to generate more positive experiences, and therefore improve overall impact.

- **Outcome Mapping.** Outcome mapping is a highly participatory method developed by the Canadian-based International Development Research Centre (IDRC) and is often used with IDRC and CIDA-funded projects, though it is gaining popularity among other agencies as well. Its aim is to focus the M&E of social change programs explicitly on behavior change among local actors.

It usually begins with a participatory workshop, where the team works towards expressing the long-term, downstream impacts that it is working towards. These desired impacts will provide reference points to guide strategy formulation and action plans. Progress markers used to track performance are then developed for each "boundary partner", or stakeholder, to identify changes the program hopes to influence.

27 The following website provides useful tips on photo voice:
http://www.photovoice.com/method/index.html.

28 For more on overall Appreciative Inquiry approach, see
https://clearspace.mercycorps.org/docs/DOC-5387.

Ongoing monitoring under the outcome mapping approach is centered around three data collection tools: an outcome journal to monitor boundary partner actions and relationships, a strategy journal to monitor strategies and activities, and a performance journal to monitor the organizational practices that keep the project or program relevant and viable. Finally, participants plan and carry out an evaluation to examine particular issues and outcomes more precisely.[29] To learn more about this technique, please consult the outcome mapping websites listed in the Additional Resources section.

Implementation challenges to participatory M&E and potential solutions

Implementing participatory M&E is not without challenges. Some common challenges are listed below, as well as potential solutions for overcoming them:

- **Ensuring accuracy of data and limiting bias.**

  Potential solutions:

  → Plan adequate time in workplans for training and mentoring to improve familiarity with the system and accurate data collection and reporting. Consider having project staff accompany local teams for surveys and collection of monitoring data.

  → Build in data quality checks to look for errors and work with partners and/or participants to correct mistakes. Quality checks may include regular "spot-check" surveys or monitoring visits by staff to ensure data matches what is being submitted via beneficiary record-keeping or partner tracking.

  → Identify particular problem areas and simplify the nature of the data or the process of collecting it to reduce the likelihood of errors.

  → Think about the various incentives that may cause partners or beneficiaries to inflate results or under-report negative findings. Will they be negatively affected by bad results? If so, try to change the incentive structures so that accuracy of information is rewarded, and not necessarily the end-results.

  → Also consider which stakeholders have inherent interests in monitoring or evaluating the work. For example, community action groups have an interest in monitoring construction of community infrastructure. However, stakeholders monitoring their own work could be less objective and more open to bias. In this sense, use of peer evaluation methods or other third-party stakeholders may be a good idea.

- **Lack of time, staff or budget to implement participatory M&E.**

  Potential solutions:

  → Plan realistic budgets and staff time and responsibilities, reflected in project proposal narratives, logframes and budgets, staffing plans, Position Descriptions, workplans, country annual plans, etc.[30] Some can be adjusted even once the project is already underway.

  → Make building local M&E capacity an explicit project objective and/or activity. This will help ensure it is resourced adequately in subsequent workplans, etc.

  → Work with donors to emphasize the importance of participatory M&E methods for sustainability and lobby for more time and/or resources to implement it.

  → Make adjustments to the scope or level of sophistication of M&E activities to make it easier and less time-intensive to implement with local partners.

- **Low partner, beneficiary or staff capacity for M&E.**

  Potential solutions:

  → Try to develop M&E systems that build on local skills and knowledge, with indicators and processes that are familiar and relevant to them, rather than trying to train them to work within Mercy Corps' or our donors' systems.

  > **BUILDING M&E CAPACITY AS AN EXPLICIT PROGRAM OBJECTIVE:**
  >
  > Focusing an objective or activity explicitly on building capacity for M&E can ensure it gets prioritized. For example, one objective of a recent MC Tajikistan annual plan seeks to "increase community outreach and participation in planning, implementing and evaluating health services."

  → Consider reducing the scope or sophistication of certain M&E activities to make them easier. For example, smaller sample sizes, shorter and more basic surveys, less use of computers and IT skills, and greater use of local languages can all help make participatory M&E more feasible to implement.

  → Focus a specific area for participatory M&E, choosing only a few key indicators or activities that can be easily monitored and reported on by local partners.

---

29 http://www.adb.org/Documents/Information/Knowledge-Solutions/Outcome-Mapping.pdf

30 Tip Sheets *on Budgeting for M&E and Structuring the DM&E Unit*, as well as example PDs, in DM&E-in-a-Box.

Participatory M&E in specific sectors

The above techniques and examples can generally be applied to different types of projects. However, some approaches work particularly well in certain sector areas. Note that many of the techniques or tools referenced below are available on the Digital Library.[31]

- **Economic development and agricultural livelihoods.** Participatory tracking of indicators such as sales, value of production, income and profit with businesses, farmers and local NGO partners has worked well when building capacity in these areas is part of regular program activities. Training and ongoing support for record-keeping, budgeting, income/expense tracking, etc., can allow us to get information directly from program participants rather than conducting our own surveys (though some programs choose to do both in order to ensure accuracy). As seen in the Afghanistan VFU example, local tracking of business operations and economic performance indicators can help build critical business skills and financial literacy among partners and beneficiaries. This increases the prospects for sustainability and financial solvency.

- **Community mobilization and infrastructure.** There is wide scope for heightened local participation in M&E of community mobilization programs. For example, community groups can help monitor the progress, quality, maintenance and use of community infrastructure projects. Peer and self-evaluations can also work well in

  EXAMPLE: MC DRC LOCAL MONITORING

  Members of local environmental associations in IDP camps in DRC help monitor the quality of fuel-efficient stoves installed by the program. Data analysis allows them to identify problem areas and take corrective action, leading to improved quality and impact.

  these projects, in addition to PRA and other participatory evaluation techniques.

Capacity indices that are participatory in nature and include self-reporting and analysis are also ways to build local M&E capacity as well as provide greater focus for strengthening key areas. Finally, tools like the community-government linkage card provide a way for stakeholders to jointly assess responsiveness and accountability.

- **Civil society and partnerships.** As mentioned previously, building the capacity of partners to monitor and evaluate their work can be made into an explicit program objective or activity. Tools such as capacity indices and NGO monitoring and reporting formats can help this process. We can also encourage them to take a lead role in data collection and analysis during assessments, baselines and evaluations.

- **Conflict management.** Participatory assessment of changes in relationships, such as through PRA-type techniques involving mapping of group relationships, social interactions, etc., can help local stakeholders gain greater understanding of underlying conflict dynamics and how to improve them. Likewise, local monitoring of impact indicators such as changes in number of violent incidences or number of conflicts resolved can help increase buy-in for conflict management initiatives and provide critical information for early warning and response systems, advocacy efforts, etc.

- **Food security.** There is much literature within this sector regarding the use of community-based food security monitoring and early warning systems, including lessons learned from Mercy Corps' own experiences. Likewise, some food security surveys or assessments use a household food log to track dietary diversity and other indicators. This has the add-on effect of increasing beneficiaries' awareness of their own food security issues and patterns so that they can take appropriate action.

- **Youth empowerment.** The example from the SKYE program in Niger shows youth can help monitor indicators related to their activities, increasing ownership and self-reliance.

---

31 Data collection tools for specific sectors, including participatory techniques, are in the general Design, Monitoring and Evaluation section of the DL.

Additional Resources

GENERAL

- Search the "Participatory M&E" tag in DM&E Clearspace space (https://clearspace.mercycorps.org/dme).

- Participatory M&E website portal: http://portals.wi.wur.nl/ppme/

- Participatory M&E toolbox (scroll down to M&E section): http://www.livelihoods.org/info/info_toolbox.html

- Participatory Impact Monitoring: http://www2.gtz.de/dokumente/bib/96-2007_IV.pdf

SPECIFIC METHODS

- Participatory Impact Assessment Guide for NGOs: https://clearspace.mercycorps.org/docs/DOC-7685

- Photo voice website: http://www.photovoice.com/method/index.html.

- Most Significant Change guide: http://www.mande.co.uk/docs/MSCGuide.pdf

- Appreciative Inquiry: https://clearspace.mercycorps.org/docs/DOC-5387

- Outcome mapping learning community: www.outcomemapping.ca; general overview of method: http://www.adb.org/Documents/Information/Knowledge-Solutions/Outcome-Mapping.pdf.

- Empowerment evaluation web portal: http://homepage.mac.com/profdavidf/empowermentevaluation.htm

Dealing with Data

Data Management Tip Sheet

This tip sheet is for managers and staff involved in developing data management systems for monitoring and evaluation information. Subsequent pages include discussion of key issues and examples from Mercy Corps projects.

Introduction to data management

In this tip sheet, data management refers to the approaches, tools and information technology applications we use to store project or program-level monitoring and evaluation information.[32] These can also be referred to as Management Information Systems (MIS), or database solutions. While various strategies have evolved across Mercy Corps field programs for managing data, many still feel they have to "reinvent the wheel" when developing project or country-specific systems.

The purpose of this tip sheet, therefore, is to draw on lessons learned and provide a general reference guide for setting up data management systems for project information. More detailed guidance, including examples and case studies of specific applications, can be found by following the links throughout this tip sheet and at the end, as well as in the DM&E-in-a-Box toolkit.

DATA MANAGEMENT CHECKLIST

Systems should:

☐ **Be strategic,** collecting only the most useful or required information.

☐ **Reflect the technology** environment of the project or country.

☐ **Stay utility-focused,** facilitating management and reporting tasks.

☐ **Strive for simplicity** so staff can learn & use easily, even if high turnover.

☐ **Ensure timely and accurate data entry,** with multiple quality checks.

☐ **Avoid duplication & ease analysis** with an integrated solution.

☐ **Clearly define roles & responsibilities** for data entry, storage and analysis.

☐ **Adapt easily** to programming changes, like new indicators or sectors.

---

32 Monitoring information refers to the documentation of program work by routinely tracking data on project activities and outputs and comparing that information against targets as part of standard program management. Evaluation information typically refers to data that measures objective-level changes and is therefore collected only periodically, typically at baseline, mid-term and final evaluations.

Key questions to ask before considering IT solutions

What are our key data needs and uses? Data management systems generally work best when they track key information that is used strategically for project management. The project logframe and indicator plans should be our starting point for determining this.[33] Be careful of a common tendency to collect too much data, which can over-complicate our systems. Instead, try to simplify and focus only what is most important or relevant.

We should also look at any existing M&E forms or reporting formats to get a better sense of what information we're looking to manage and see how we can build on processes already in place for collecting and managing data.

Next, think about how the various stakeholders – project managers, program staff, country leadership, partners, M&E staff – will want to use the information (management, reporting, etc.). The optimal solution will manage information in a format that is usable for these tasks. This could include automatically generating tables or charts that facilitate visual analysis, or organizing data monthly or quarterly to reflect reporting timeframes.

What are the major components or desired features of the system? We should think about how answers to the first question above will affect our system requirements. Some factors to consider include:

- **Integration and compatibility:** Think about which data sets can be stored in one integrated place, and whether they should be compatible with or linked to other data management solutions for standardization, aggregation, or analysis/reporting purposes. For example, project data may require a certain standardized format to feed into larger country-wide systems.

- **Quantity and organization of data:** The amount of information we will be storing, as well as the number of dimensions to the data (i.e., regional, sector, etc.), will help determine the best solution. Some projects find it useful to organize the data by objective, as this can ease reporting to donors. In other cases, a geographical approach organizing data by location can be useful for discussing program strategy, resource allocation and management. Data with many dimensions or units may require a more advanced database such as MS Access, while MS Excel may be appropriate for more limited data sets.

- **Data transfer:** Who needs to access the information, from where, and how often is a major issue. Sometimes, data can be stored in one central field office and transferred only occasionally by email. Other times, we might prefer information to be updated and shared in real-time from multiple locations, and a web-based solution may

be desired. There are other technical solutions as well, such as using Access replicas in field offices or establishing virtual private networks (VPNs).

- **Adaptability:** The simpler the system, the easier it can adapt to changes. We can test this by assessing how easy it is to add new objectives or sectors. Back-end linkages, such as look-up tables and auto calculations, ease data entry but can also make the system more complex and less flexible.

EXAMPLE: MC IRAQ

Data management problems:

- Time-consuming system was difficult to maintain and teach to new staff.
- Repetitive data entry and tedious calculations led to inefficiency and mistakes.
- Did not fully address information needs.

Sample spreadsheet and table from MC Northern Iraq

Data management solutions:

- Simplified tool integrates project indicators and other info into one Excel database.
- Auto-calculations in Excel limit errors.
- Easily aggregates data and links to summary tables and charts for analysis.

What are our main constraints? In thinking about what is feasible for our project or country office, we should keep in mind some practical issues that affect design as well as implementation and sustainability of the data management system over time.

- **Budget:** The costs of implementing and maintaining data management solutions are often under-estimated. It is common for programs to develop cost estimates for system development, which can include consultant fees and software or equipment upgrades. The costs of *maintaining* the system, however, are often over-looked and can be

---

33 For more information on standard DM&E tools and methodology, please see the Mercy Corps *Design, Monitoring and Evaluation (DM&E) Guidebook*.

far more significant. These include employing knowledgeable staff, training, adjusting to changes, and updating or replacing technologies if they break down or become obsolete.

- **Time:** Another important, often over-looked factor is the time needed to design and implement the system. The more complex the system, the more it will take staff away from other activities and cause delays in actually generating data. Systems with higher-end technology or a large number of data points generally require more staff time to design, implement, and maintain.

- **Technology environment:** The level of technology we use should be consistent with the local project or country environment, in terms of IT sophistication, Internet connectivity, etc. This is important for sustainability of the system in the face of staff turnover or programming changes. Also, program staff feel more comfortable entering and analyzing data in software programs that they are more familiar with.

- **Staff capacity:** Data management structures should be developed in line with the capacity of program staff to implement these systems. There can be difficulty in maintaining the system or making changes if the level of technological sophistication outstrips the knowledge or capacity of Mercy Corps staff. Over-reliance on outside help, or even on one or two tech-savvy staff members, can lead to time delays and threaten overall sustainability, particularly in cases of high staff turnover.

TRADEOFFS: TECHNOLOGY VS. COST/IT EXPERTISE
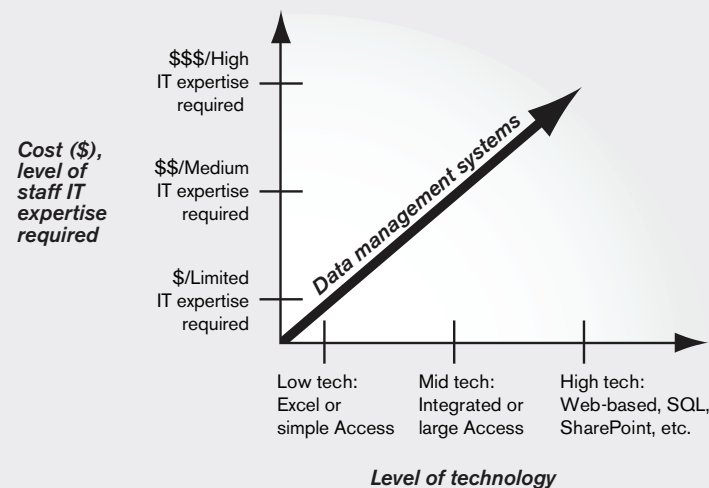
Analyzing software options

MS Excel and MS Access. The table on the next page discusses two very common software programs for data management and analysis, Microsoft's Excel and Access.

The table illustrates how MS Excel may be preferable in a project with a relatively simple data set and fairly basic IT capacity. MS Access, on the other hand, could be advantageous for multi-faceted programs that have the expertise to maintain it. Staff tend to be more familiar with Excel, and it requires less back and forth database developers to design and maintain. However, Access offers more options for integrated management of complex data sets.

OUTPUTS VS. OBJECTIVE-LEVEL DATA:

It can be useful to distinguish between monitoring data tracked routinely, and data on objective-level indicators that are part of one-off activities like baselines and evaluations.

The former requires a mainstreamed system geared towards daily use, while the latter may only be used by the specific survey team.

Pros & Cons of MS Excel and MS Access

	Excel		Access	
	Advantages	**Disadvantages**	**Advantages**	**Disadvantages**
Data entry	• Open interface to see entries & note mistakes • Minimal training required	• Data entry form not as user-friendly • Less automatic standardization of entries to guard against error	• Simple, user-friendly forms easy to enter data • Separate form for each data record • Can use data validation to guard against errors	• More difficult to set up and adjust to changes • Harder to readily see errors
Data storage	• Easy to maintain due to staff familiarity, especially if high turnover	• Hard to store many different types of data (flat structure) • Not as effective as an integrated data solution • Difficult to manage multiple versions	• Can store many types of data and high volume • Allows for multiple databases with one interface for integration	• Requires advanced technical staff to maintain; staff turnover a major issue • Can cause delays if relying on database experts
Data analysis	• Easier for program staff to manipulate • Can readily convert to pivot tables, charts, graphs, etc. (with training)	• Does not automatically generate formatted reports as in Access • Harder to integrate data from multiple sources	• Can generate auto-formatted reports based on data • One report integrates multiple data sources • Multiple analysis tools include Data Access pages, pivot charts, etc.	• Program staff not as familiar with software • Requires IT/M&E staff to format reports and make changes

Some M&E systems use a mixture of both Excel (for entering data and analyzing outputs) and Access (for storing all data in central location).

An intermediary option could be using Access as the main IT solution, but having data entry tables linked to Excel and enabling users to export the Access data into Excel for analysis. The need for linking to Excel may diminish as staff capacity and familiarity with Access increases. However, starting with data analysis in Excel can help encourage program staff to use the data more readily and tailor it to their unique needs or preferences. It can also help staff begin to understand the linkages among different data sources in Access and how a report is generated through the use of queries.

Other data management software. Database options could also include use of a Microsoft SQL server. This is a back-end[34] data management tool robust in storing and synchronizing large quantities of data. Some Mercy Corps programs have found it useful to back up more user-friendly front-end solutions such as Access, Excel or InfoPath. Staff capacity to manage an SQL server can be an issue, however.

There is also a range of more sophisticated software developed by private companies specifically for managing M&E data. While these may prove handy for some projects, they can be unsustainable due to the need to repeatedly train staff and the lack of in-house expertise.

Setting up the process

Team approach to system design. The end-users of the system, including program managers and staff, should be involved in the design of the system as much as possible. M&E, IT and program staff need to work together to ensure that data management processes are able to be implemented and strategic in terms of programmatic information needs and uses.

Training staff. Continual staff training, particularly in data entry procedures, can help reduce errors and ensure standardization. It is important to be detailed and specific, with written instructions for how to fill in each data point (e.g., does one enter "5" or "5%", or leave blank vs. write "NA").

Training of program staff on key Excel data analysis functions, like pivot tables and creating charts, can also greatly increase use of M&E data. More sophisticated software solutions may require more even extensive training to increase comfort levels. Training takes time but is a worthwhile investment: the more staff with intricate knowledge of the system, the better guarded we are against sustainability threats such as staff turnover and programming changes.

Defining roles and responsibilities. Clearly defining who will do what, and when, is critical. Responsibilities should be clearly detailed in

indicator plans and staff position descriptions. It usually works best when M&E responsibilities are shared among the project team, with focal points coordinating overall M&E activity and specific staff are assigned data entry roles. Supervisors or others should conduct routine data quality checks to ensure accuracy and consistency.

KEY DATA MANAGEMENT ROLES AND RESPONSIBILITIES:

- Designing the database
- Developing data collection forms
- Training on data collection forms
- Collecting the data
- Entering the data into a database
- Maintaining the database
- Cleaning/sorting data for analysis
- Producing analysis formats
- Analyzing the data
- Writing up findings in reports

Other data management considerations

Data entry. As a general rule, we should configure our data collection forms and systems so that data-entry staff enter exactly what is on the form into the spreadsheet or data-entry form. This helps to reduce error.

Many Mercy Corps programs find that setting up auto-calculations in Excel or Access, as opposed to tasking staff to make these manually, reduces mistakes. Look-up tables are also useful for ensuring standardization. Be careful of the tradeoffs, however; they can make the IT set-up more complex and less flexible to changes.

Data storage. Sometimes events like staff turnover or computer break-downs cause projects to lose important data. Keeping the data in one central location, and making sure it is backed up in multiple formats (hard-drives, servers, CDs, etc.), is crucial to sustainability.

It is also recommended to keep an updated manual-type document describing the data management system, so that later staff can understand the set-up details and the rationale behind decisions made.

Data security. We need to ensure that data is kept safe from corruption and that access to it is suitably controlled. This requires protecting databases and networks from unauthorized users through passwords and other encryptions.

Data sharing and access. Facilitating timely access to data is a common challenge for M&E systems. In projects with more limited geographic scope, data can be entered, stored and accessed on a central intranet or stand-alone server. Larger programs, however, have to work creatively to update and access information. A few options for this are listed below:

---

34   The term *back-end* refers to how the data is stored within the IT system, while *front-end* refers to the user interface that staff interact with to enter and analyze data. Systems can maintain different software solutions for each task.

EXAMPLE: MC INDONESIA

Data management issues:

- Geographically dispersed program.
- Program staff used to Excel.
- Range of programming sectors.

Data management solutions:

- Web-based application, MS SharePoint, provides access to data entry and analysis from anywhere.
- Data entered and stored in Access, then exported to Excel for analysis.
- Games and prizes to increase use of Web-site; M&E staff show visual examples of data to program staff.
- IT staff with M&E responsibilities help maintain the system.

Mercy Corps Indonesia Data Management System

Web-based MS SharePoint for common interface & real-time data integration

MS SQL Server for back-end database

Office Intranet Server

Data Entry in MS Access or InfoPath

Analysis & Reporting with Excel Pivot Tables & Charts
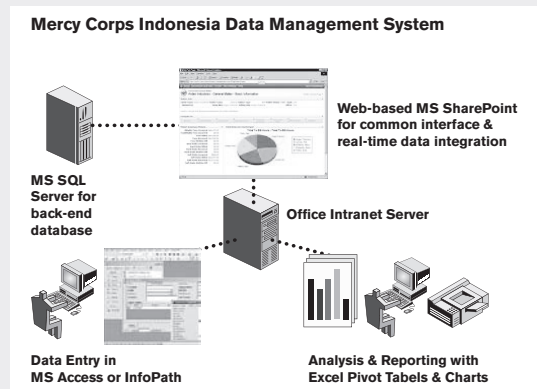
- *Low-tech options* include having field officers send their data sets electronically by email or flash-disk. These can then be aggregated manually by someone in a central location and shared. This limits over-reliance on Internet access or IT systems. However, it will not be possible to have real-time data, and manual updates and aggregations can be prone to error and cause confusion about different versions of data sets.

- There are also a variety of *Web-based* options that can facilitate access to databases through the Internet or VPN. These include Microsoft's SharePoint[35] and Groove, both of which provide an Internet website, or portal, in order to link many users in different sites to common databases in Access, SQL or other applications. Central servers accessed through the remote desktop function are another option. These solutions have proven effective where the

technology environment is appropriate, though some may require significant IT expertise to maintain.

Refining and maintaining the system

Start small and field test. Setting up a large, complex system in one go can be overwhelming for staff, and make it more difficult to correct mistakes. It can be useful to start with only one project, or one component or objective of a project, to develop and test IT solutions. Lessons learned on a smaller scale can then be incorporated into wider roll-out.

Continually adjust and refine. Major events like annual surveys, evaluations, workshops or reporting cycles can be useful times to step back and analyze the system. At each instance we should validate the utility of each set of data, and decide what is most relevant and what elements could be eliminated. We should also double-check to ensure that data collection and documentation responsibilities are able to be carried out by staff, in a quality manner and without over-burdening other activities. Routinely assessing quality and identifying areas that are error-prone is also important.

DATA ANALYSIS MEETINGS SHOULD:

- Clearly compare expected and actual results.
- Identify reasons for lower than expected results (if applicable).
- Outline a plan of action in response to results.
- Communicate info to stakeholders (partners, beneficiaries and donors) and solicit feedback.

Conduct data analysis meetings. M&E data can be useful for project management, learning, reporting and accountability. It is important that our data management solutions have the ability to aggregate and display data in tables, charts and graphs that are easy to understand. Regular analysis meetings including all project or program team members should be an integral part of the work-plan.

Consider the role of partners. Think about which elements of the system would be appropriate for local implementing partners to manage or input directly into. In the push for more participatory M&E, we should be wary of overly-complex M&E systems and IT solutions which may alienate partners or add to the perception of M&E as requiring very specialized expertise. Instead, we can look to simplify basic processes and encourage partners to play an increasing role in collecting, managing, analyzing and reporting on project information.

---

35 MS SharePoint is a common web-based software facilitating real-time data entry and analysis from multiple locations. See *SharePoint: Pros and Cons (https://clearspace.mercycorps.org/docs/DOC-1563).*

Additional Resources

Mercy Corps programs across the world continue to innovate and adapt to data management challenges. Check the DM&E-in-a-Box toolkit on the Mercy Corps Digital Library, the DM&E space on the MC Clearspace website (https://clearspace.mercycorps.org/community/cops/dme - search the "data management" tag) , and the resources listed below, or email the DM&E Initiative at dme@mercycorps.org, for more examples and advice on setting up a data management system.

All tip sheets listed below are available in DM&E-in-a-Box on the Digital Library, and specific country examples and reports are provided in further detail in the DM&E Community of Practice space on Clearspace.

- *DM&E Tip Sheet #12: Monitoring Systems*
- *DM&E Tip Sheet #7: DM&E at Project Kick-off*
- *DM&E Tip Sheet #10: Structuring the DM&E Unit*
- *DM&E Cross-visit report – MC Indonesia, Sri Lanka, & Afghanistan, March 2007*
- *MC Northern Iraq data management example*
- *MC Sri Lanka database, annual review summary, and AEA presentation*
- *Cross visit: MC Indonesia country slides*
- *MS Sharepoint: Pros and Cons*
- *MC Sri Lanka, Aceh, Eritrea and Afghanistan data management systems reports (Sherif Rushdy)*

Data Analysis Tip Sheet

This tip sheet is intended for Mercy Corps staff involved in organizing and analyzing monitoring and evaluation[36] data for strategic use. It gives an overview of basic data analysis methods, displaying several Mercy Corps examples and highlighting key issues and calculations.

Introduction to data analysis

Immediately following monitoring and evaluation activities, we often face the question: Now that we've collected all of this data, what do we do with it?

Mass amounts of data can be difficult to understand without proper organization and analysis. Data analysis is the process of converting raw data into meaningful statistics and themes. It helps us to draw out significant findings and use data to tell the story of how our programs are impacting beneficiaries. Knowing how we would like to analyze data also helps us plan our M&E activities more strategically.

This tip sheet provides some basic methods for organizing and analyzing both quantitative and qualitative data. It displays several Mercy Corps examples and highlights key issues and calculations to enhance data analysis skills and empower use of M&E information.

> **BASIC STEPS TO GOOD DATA ANALYSIS:**
>
> 1. **Clean & organize the data.** Check for errors or gaps, & identify outliers.
> 2. **Generate summary statistics** using sums and averages.
> 3. **Analyze & compare quantitative data** in tables, charts, and graphs. Experiment with different groupings and cross-analyses, looking for patterns & insights. Also compare to targets, especially for monitoring data.
> 4. **Analyze qualitative data** by identifying and tabulating common themes or experiences.
> 5. **Draw conclusions** by reviewing and cross-verifying the evidence in the data.
> 6. If needed, perform **advanced statistical calculations** for further level of rigor.

---

36 Monitoring information refers to the documentation of program work by routinely tracking data on project activities and outputs and comparing that information against targets as part of standard program management. Evaluation information typically refers to data that measures objective-level changes and is therefore collected only periodically, typically at baseline, mid-term and final evaluations.

QUANTITATIVE VS. QUALITATIVE DATA ANALYSIS:

- **Quantitative data analysis** involves performing statistical calculations, such as averages, percentages, and percentage change, based on numerical data. This data can come from surveys, technical assessments, observation, or monitoring data and reports. It is particularly aided by close-ended survey or interview questions with response categories.

- **Qualitative data analysis** deals with non-numerical data, such as open-ended written responses to interviews and focus group discussions. It often involves identifying common themes and trends in the text to draw out conclusions and make summary statements. Qualitative statements are more robust when we can add numbers to quantify them, such as by calculating the frequency of certain types of broad responses or themes mentioned.

Step 1: Clean the data

Data cleaning means reviewing the data to check for any possible errors or gaps. These can occur through lapses in data entry, misunderstanding of coding or response marking processes by interviewers, or even misunderstanding of the questions by respondents. If left unchecked, these errors can vastly skew our results and cause time delays.

A first step in organizing data, therefore, is to scan the data entry table question by question to look for anything that seems missing or out of place. In fact, we should also do this manually before the data is entered or while it is being entered, by scanning the data collection forms for abnormalities. Once the data is entered electronically into the database, some methods to aid the process of data cleaning include:

- Developing an acceptable or likely range for each set of data to check for valid values and investigate any outliers[37]. Numbers that fall outside the range should be analyzed to ensure they are accurate before including them in the data set. For example, if we're interviewing young people aged 15-25 for a youth project, anything lower than 15 or higher than 25 in that data set should be cause for further investigation.

- Cross-verifying values with other data points. For example, age and pregnancy columns should not allow a 70-year old woman to be marked as pregnant. The *filter* function in Excel can help compare values for different sets of data.

Step 2: Generate descriptive or summary statistics

To get a better initial hold on our data findings, we can reduce it by summing

---

37 An outlier is a data point that lies far outside the normal range for the data set. Because outliers can greatly affect averages, it is important to verify whether they are correct or due to errors in data collection or data entry.

or averaging the rows and columns containing quantitative data in our data sheets. This can also help to identify errors. After providing some summary statistics,e'll eventually want to look at sums and averages for different sub-groups, and cross-analyze these findings with results from other data points.

First, it's usually good to get an overview of the sample through a few basic descriptive statistics before digging into further analysis. This includes sample size and some of the demographic information commonly found at the beginning of surveys like age, income grouping, education, gender, etc.

We can break this information down according to which sub-groups or sets of information are most important for our analysis.

Basic calculations for data reduction and generating descriptive statistics

Sum. We're often interested in summing, or aggregating, data to show overall numbers or outputs for our projects. Examples of sums include total numbers of seed bags distributed, wells constructed, NFI kits distributed, people trained, Km of road rehabilitated, etc.

CALCULATING "SUM" AND "AVERAGE" USING FORMULAS IN EXCEL:

fx =SUM(B3:B7)		
	A	B
1	**NFIs Distributed**	
2	**Community**	**Kits**
3	Ak Sai	35
4	Beshariq	102
5	Bo'ston	58
6	Dombrachi	69
7	Jailgan	25
8	**Total**	**289**

fx =AVERAGE(B2:B7)		
	A	B
1	**Yearly income**	
2	Abdallah	$ 835
3	Arif	$ 1,360
4	Azimi	$ 968
5	Blasevic	$ 1,050
6	Bolotkan	$ 1,193
7	Cevallos	$ 750
8	**Average**	**$ 1,026**

Mean (or average). A mean is simply the statistical term for average. The most common way to summarize and display a data set is to show a typical, or average, value for that data set. The equation for calculating an average is as follows: Average = sum of all values / # of units. Averages are often more insightful when we analyze them according to different sub-groups or char-

BE CAREFUL WHEN AVERAGING AVERAGES!

A common mistake is to look at average columns or rows for different data sets, and then average these numbers to obtain an overall total average. This method gives the wrong result when sample sizes are not equal.

fx =((45*400)+(12*150))/(45+12)		
	Sample size	Avg income
Community X	45	$400
Community Y	12	$150
Total avg income		$347

As shown in the formula at left, it is better to calculate total average based on the sample sizes.

acteristics of the population. For example, we may see very different situations for male and female participants by averaging a few key indicators.

Median & mode. Median is the statistical term for the middle value in a series of numbers. It indicates that 50% of the numbers lie above this value and 50% lie below it. The median is useful since it's less sensitive to outliers – or unusual values – than the mean.

For example, when collecting the age of a series of 5 individuals [12, 10, 13, 7, 9], arrange the numbers in numerical order [7, 9, 10, 12, 13] and the middle number [10] is the median. In the case of an even set of numbers, add the middle 2 numbers and divide by 2 in order to find the median value. Mode is defined as the number occurring with most frequency in the data set. It can be helpful but is generally used less often than the other summary statistics listed above.

Range and maximum/minimum values. The range of a data set shows the difference between the largest and smallest number in the set. This allows us to see how dispersed, or far apart, the values are. To find the range, simply subtract the minimum value from the maximum value in the set.

When the range value is paired with the maximum and minimum value, as well as the mean and the median, it is easy for those reading the data to have a clear idea of the data set without having to examine all the data.

Presenting summary data. A common mistake many people make is to present a large volume of uncategorized data for the reader's consumption, particularly when presenting summary statistics. In general, for presentation purposes, it is best to keep tables and charts as simple as possible to best illustrate or summarize the point at hand. Tables or graphics with too many components can cause confusion and distract from key points we are trying to emphasize.

SUMMARY DATA TABLE BY TYPE OF PROJECT AND LOCATION:

Frequency distributions are a good way to display percentages for outputs or response types. The example below shows the distribution of types of infrastructure projects implemented:

Figure 2: Distribution of CAIP Infrastructure Projects by Type of Projects
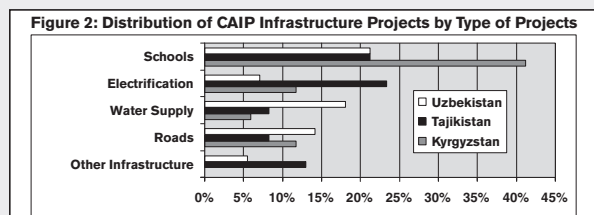
Step 3: Analyze and compare quantitative data

Percentages. Percentages are a great way to make data more understandable and comparable. They standardize data by converting it to a number out of 100. For example, if 13 youth members out of 56 interviewed expressed interest in vocational training, this means that 13/56, or 23%, show this interest.

The equation for calculating a percentage is therefore:

Percentage (%) = number of units fitting a particular value range / total units in the relevant population

It is often advisable to display both percentages and numbers for data, to give the reader an understanding of both relative and absolute size.

Cross-tabulating data. We can gain further insights by cross-analyzing data with another set of data points or sub-groupings. This gives more depth to the analysis and provides additional information about why certain groups or elements might be experiencing the certain types of results.

For example, let's say an assessment survey shows that 38% of males are unemployed. It may be interesting to see which vocational training skills were prioritized by these unemployed men, versus those vocational skills prioritized by men who are employed, or by employed or unemployed women. If answers differ largely among these groups, we may want to focus our project design on different skills for the different groups.

It is usually well worth the time to experiment with different groupings and cross-tabulations to see which might provide the most interesting insights or best illustrate the impact, results or experiences of people in the program.

Use of pivot tables. A pivot table is a convenient cross-analysis tool found in spreadsheet programs such as Excel and Access. They are especially effective in displaying data according to different sub-categories. Pivot tables

PIVOT TABLE:

The pivot table below uses three data fields – province, schools, and water source (yes/no) – to show how many schools do or do not have water systems installed, by province.

Count of school_Name	Water_so ▼		
Province ▼	No/Tidak	Yes/Ya	Grand Total
Bengkulu	3	45	48
Lampung	14	28	42
Riau	1	18	19
West Sumatra	6	30	36
Grand Total	24	121	145

We can also perform follow-on calculations based on the data in the pivot table to present the info in different ways, as shown below.

Province	% schools with water source
Bengkulu	94%
Lampung	67%
Riau	95%
West Sumatra	83%
Total	83%

generate summary data by looking up a range of data in a spreadsheet and automatically counting frequencies or summing or averaging the contents, according to the type of data we'd like to analyze.

In Excel, the pivot table function is accessed by selecting "PivotTable" from the Data menu. Using the "Pivot Table Wizard", the first step is selecting the full range of data in the set, including column titles. The wizard will then create an empty pivot table, and list a set of data fields (which were generated according to the column headings in our original spreadsheet) in a dialogue box. We choose which data fields we'd like to analyze by dragging them from the list and into the pivot table. It is worth experimenting in Excel to get a feel for how pivot tables work.

In the example in the text box on the previous page, the pivot table counts the number of schools according to province and whether or not they have a water source. It is possible to include more data fields – such as village, or type of water source - to further filter the data. Or, we could reduce the table to only two data fields for simplification.

Comparing averages and percentages. One of the most common analysis techniques is to compare averages or percentages for various sub-groups, time periods or response types. When comparing data it is necessary to have the same units of analysis and sampling strategies, in order to make the comparison valid.

EXAMPLE: MC SRI LANKA
COMPARING BASELINE AND ANNUAL SURVEY DATA

The example below is from MC Sri Lanka's annual review workshop. It analyzes average results from the various elements of a capacity index for community-based organizations from baseline to year one. At bottom are the detailed components of one of the elements of the capacity index ("Linkages").

Color-coded bar graphs are a simple and easy way to analyze differences among groups or to compare data from different time periods.
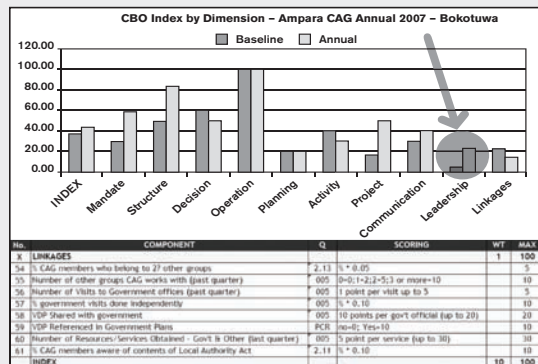
When comparing data that use response categories, we should try to look for the best way to group them so that the most important distinctions are highlighted. The example in the text box at right gives an idea of ways to do this.

A common mistake is to compare data from different groups or across different time periods based on very low sampling numbers for one of the groups. This is especially problematic when the sample size for one of the groups is very small, such as under 10 units. Percentages based on such a small sample size are generally not valid due to the large potential margin of error. We usually prefer to have at least 25 to 30 units in each group or time period to make a valid comparison, though this minimum threshold depends on the context and the applied sampling strategy.

It is also essential when comparing data that the sampling strategy for each data set is the same. This means following the same sampling strategy that was used in the baseline during the mid-term or final evaluation, to the extent possible. Otherwise, different sampling strategies may result in different target audiences and findings will be biased.

GROUPING AND COMPARING RESPONSE CATEGORIES:

Consider the data table below showing project sustainability ratings by region:

	Region 1	Region 2
Excellent	25%	45%
Satisfactory	38%	23%
Unsatisfactory	21%	22%
Poor	16%	10%
Total	100%	100%

From the above table, it is difficult to figure out which region had the better results. Look for ways to re-group or simplify the data to better highlight key findings.

For example, one way to do this might be to group the first two categories, as below at left.

	Region 1	Region 2
Satisfactory or above	63%	68%

	Region 1	Region 2
Excellent	25%	45%

Or, to show a starker difference among regions, display only the difference in "Excellent" projects, as above at right. Play around with different data groupings and arrangements to see which offer the most insight or display findings most clearly.

In past/present comparisons, we also need to check whether there were any other outstanding circumstances during the time period, or external factors that happened along the way, which could have greatly influenced differences in the data.

Control groups. When using a control group, as when comparing present data with past data, it is extremely important that we verify the methodology of how each group was sampled. Any differences should be duly noted, and a judgment call made to determine whether the comparison is still valid.

Control and comparison groups should be as similar to the participant group in every way possible, with the only major difference being the Mercy Corps intervention. This helps to convince the reader that any differences in project indicators for the groups – such as increases in income or access to pricing information – are a direct result of participation in the project rather than due to random factors.

Calculating percentage change. Percentage change is a useful calculation to show increases or decreases over time. It is often used to demonstrate objective-level indicators, such as those measuring changes in attitudes, knowledge, behaviors or conditions. When presenting this data, the reader often wants to know the context of that percentage change and the starting point (percentage found at baseline). Be sure to fully explain changes in percentages and provide the raw numbers and/or the starting and ending percentages. For example, instead of simply stating that school attendance increased by 10 percentage points, be sure to also say that it rose *from 65% to 75%* (or whatever the case may be). The below examples demonstrate how to calculate percentage change.

CALCULATING PERCENTAGE CHANGE:

Example 1: The following example demonstrates how to calculate percentage change:

- In year one there were 20 participants in the community action group
- In year two there were 35 participants in the community action group

% increase in participation = difference/original # = (35-20)/20 = 15/20 = 75%.

Example 2: People sometimes get confused when the indicator itself is a percentage:

- Baseline study shows 52% of children practicing hand-washing behaviors
- Final evaluation shows 72% of children practicing hand-washing behaviors

The *absolute change* in the indicator is 20 percentage points, but the *percentage change* is the difference divided by the original: (72-52)/52 = 20/52 = a 38% increase in the indicator.

Analyzing monitoring data and identifying trends. Monitoring data generally consists of routine tracking of outputs and activities on a monthly or quarterly basis. For this data, we're often interested in analyzing actual versus expected results as well as identifying any upwards or downwards trends in the data over time. This is especially important for project management, to ensure that we are on track with regard to our targets and that we are reaching our intended beneficiaries. It also allows us to make strategic adjustments to better maximize our impact. We can use the tips in this document to analyze and present monitoring information in a compelling fashion.

Organizing monitoring data systematically and displaying it visually in graphs and charts can give us a good idea of how the project is going. We can then make adjustments in the intervention or attempt to adjust targets based on the evidence.

To make the best use of monitoring data, we should ensure some kind of systematic analysis meetings or workshops on a weekly, monthly or quarterly basis, depending on the project. Monitoring data can also get overlooked in evaluations. Analyzing outputs or activities data can give important clues as to how, why, and when certain results or situations happened the way that they did.

EXAMPLE: MC IRAQ OUTPUTS DATA ANALYSIS

The table below from MC N. Iraq is a common example of how to analyze output monitoring data. It compares expected targets versus actual results.

Indicator	Current #	Targets	Difference	Progress
WATER & SANITATION	13 projects			
Pumps/Compact Units	7	7	+0	100%
Meters of Pipe	2,111	8,000	-5889	26%
Families receiving health training	410	1,500	-1090	27%
Villages w/improved wells or tanks	10	7	+3	100%

It is also possible to graph this data, and to show monthly or quarterly trends.

Step 4: Analyze qualitative data

Qualitative data generally stems from activities like focus groups, open-ended interviews, observation, story-telling, and participatory rural appraisal techniques (PRA). Responses are usually contained in written text format, but can also include drawings, audio, photos, video, etc. Because it can be time-consuming to collect and analyze, qualitative data collection is usually reserved for major M&E events like baselines, mid-terms, and final evaluations. However, more limited qualitative data collection and analysis can be part of regular project monitoring if processes are sufficiently simplified and streamlined.

We can use qualitative information to effectively complement and provide depth to quantitative outcomes. While quantitative data is good at telling us "what" happened, qualitative information can provide important insights on "why" or "how" it happened.

ISSUES TO CONSIDER IN QUALITATIVE ANALYSIS:

- What **patterns or common themes** emerge across responses? How can we broadly group and **quantify** these themes? How can they be used to complement other data?

- Are there any **deviations** from these patterns? If so, what are the prevailing factors that might explain these atypical responses?

- What **interesting stories or experiences** emerge from the responses? How can we further develop these case studies to illuminate broader findings?

- Do any of these patterns or findings suggest that we should analyze other data in new ways, or **collect additional data** for follow-up?

- Are patterns **consistent** with the findings of any corresponding qualitative or quantitative analyses that have been conducted? If not, what might explain these discrepancies? If so, is it sufficiently triangulated to support conclusion statements?

Culling through many sets of written responses takes a lot of time and effort. Building response categories or themes checklists into qualitative data collection tools helps us to sift through the information later in the analysis stage. This requires careful training and coordination of the data collectors to ensure they all share a common understanding of the themes and categories and how to assess what responses fall into them, as well as extensive field testing to ensure that response categories are the most relevant. Of course, we can still leave room for open-ended comments, stories and insights.

Quantifying the qualitative. Much of qualitative analysis involves looking for ways to quantify the information to strengthen claims or summary statements.

STRENGTHENING QUALITATIVE ANALYSIS BY QUANTIFYING RESULTS:

Consider the following statement:

- "Many community groups mentioned skill-building as an important benefit of participating in the program."

A more convincing way of analyzing and presenting these results would be:

- "65% of community groups alluded to skill-building as one of the most important benefits of program participation." Or,

- "The top three benefits mentioned by community groups were skill-building (15), access to resources (12), and a greater participation in decision-making (11)."

One way to quantify text data is to tabulate or list the frequency of common types of responses, as in the text box above. We can use similar techniques in other qualitative formats. For example, in analyzing results from a mapping exercise, we can attempt to quantify the number of maps denoting the local partner as an important resource, showing a key income-generating crop, or identifying positive relationships with local government. As with data collection, it is important to ensure that all data analyzers are consistent with their understandings of how to allocate responses into the relevant themes and categories.

Computer-assisted qualitative data analysis software, also referred to as CAQDAS, can automatically sort and quantify qualitative information from interviews and focus groups. These can be purchased online (see *Additional Resources* section) and may be useful for analyzing and sorting responses in written text format. Of course, we can also attempt to do word-searches ourselves in Word or Excel to help rapidly group answers into response categories.

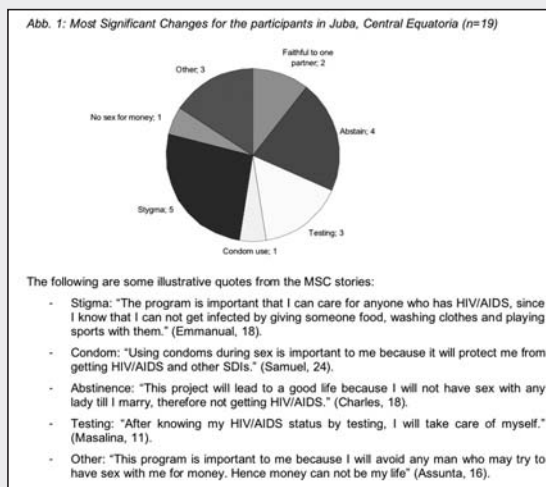**Using qualitative data to complement quantitative results.** Wherever possible, we should use the information provided from qualitative techniques to support or enhance quantitative data. Let's take an example where participation in a community group increased by 75% (from 20 members to 35 members) over the first year. Perhaps there are clues in our focus group and interview responses that can better tell us why this participation increased.

A nice complementing chart, therefore, would be to list the top 5-10 reasons that people listed for the increasing participation. We could also complement this with a few insights from groups where participation increased greatly, along with findings from groups whose participation decreased or was below average, to get a better idea of contributing factors.

One way to analyze a range of distinctly different experiences is to review the quantitative data for ways to categorize people or groups according to their experience with the project. Once the data is broken down in this way, we can then look for common themes or purposively select case studies from among these different groupings to learn more detail about how or why different experiences occurred.

Abb. 1: Most Significant Changes for the participants in Juba, Central Equatoria (n=19)

The following are some illustrative quotes from the MSC stories:

- Stigma: "The program is important that I can care for anyone who has HIV/AIDS, since I know that I can not get infected by giving someone food, washing clothes and playing sports with them." (Emmanual, 18).
- Condom: "Using condoms during sex is important to me because it will protect me from getting HIV/AIDS and other SDIs." (Samuel, 24).
- Abstinence: "This project will lead to a good life because I will not have sex with any lady till I marry, therefore not getting HIV/AIDS." (Charles, 18).
- Testing: "After knowing my HIV/AIDS status by testing, I will take care of myself." (Masalina, 11).
- Other: "This program is important to me because I will avoid any man who may try to have sex with me for money. Hence money can not be my life" (Assunta, 16).

The richness of individual quotes and stories in qualitative data can be powerful. While grouping and quantifying themes is important, the best analyses also utilize direct quotes, case studies, and individual experiences to shed light on more quantifiable outcomes.

Step 5: Verify results and draw conclusions

Conclusion drawing involves stepping back to consider what the analyzed data means and to assess the implications for the key questions or indicators at hand. Triangulation refers to the process of looking at different data points, sources and methods to cross-check or cross-verify any emergent themes or conclusions. Clear triangulation of results allows us to make more confident and credible conclusions based on our findings.

We should be very clear about what our data means and to what group it is applicable. Particularly with highly specific data that is focused on a key group(s), we should be sure to state that the information is relevant only for that group(s) and not the general population as a whole. This will help minimize misinterpretations. It is also important to clearly address any limitations in the methodology and avoid overly sweeping statements.

An issue that often arises is how to weigh different evidence drawn from different data points. Hopefully, the triangulation process allows us to confirm certain results or happenings. However, at times the results do not necessarily corroborate one another, and may even conflict. In these cases, the conclusions we draw require a judgment call to be made about which evidence or data is most convincing. It may also require more

nuanced conclusion statements. A detailed description of the rationale behind these conclusions, as well as the methodology used to arrive at them, is also very important!

Step 6: Perform advanced statistical calculations (if necessary)

Though generally beyond the scope or needs of most Mercy Corps M&E activities, sometimes we're required to perform a more robust statistical analysis such as calculating variance, confidence intervals, margin of error, etc. This may depend on sector norms; for example, health and nutrition studies often require such rigor. Other types of in-depth impact studies or assessments, particularly those that we're looking to publish to an external audience, might also demand more robust statistical calculations. Below is a very basic overview of some of these calculations, with links to further resources if needed.

Variance. Calculating the variance for a range of data can be useful to give us an idea of how much the results differ from one another and the general distribution of the data. Knowing the variance also helps determine margin of error and the statistical significance of findings.

Standard deviation. This is just the square root of the variance; it is basically the average of how much each data point differs from the mean. It is useful for calculating the statistical significance of findings.

Margin of error. The margin of error (MOE) expresses the amount of random sampling error in a survey's results. It is what we mean when we say that we found a result, "plus or minus X%." The larger the margin of error, the less confidence one can have that the results are close to the "true" situation.

The margin of error generally decreases with larger sample sizes but is also dependent on the amount of variability in the sample and the degree of precision, or confidence interval, deemed acceptable for our study. Online sample size calculators can also assist us with MOE calculations – see Additional Resources section for details.

Testing for statistical significance. In more rigorous studies, such as those aiming to be published externally in journals and other such publications, one way to increase legitimacy and demonstrate that our findings are valid and can be generalized to our broader population of interest is to test for statistical significance. The term "statistically significant" simply means that there is statistical evidence to support the fact that a difference has been observed in the sample and is likely to be found in the broader population, and unlikely to be erroneous or due to chance.

Steps in testing for statistical significance include:

1. *State the Research Hypothesis.* For example, skills levels increased.

2. *State the Null Hypothesis.* This is the opposite; it is what we are looking to disprove. For example, skills levels did not increase.

3. *Select a significance level.* This is the confidence we can have that our results are correct and not due to chance. We can usually accept lower levels of confidence (such as 90%) in exchange for a more feasible sample size.

4. *Select and compute the test for statistical significance.* The most common are T-tests and Chi-squared tests. Formulas are provided in the links below.

5. *Interpret the results.* The significance tests will produce a value that is compared with values in standardized tables to determine whether the finding is significant, and at what level. In general, for a one-sided T-test, a p value $< .10$ is significant at the 90% confidence level, while $p < .05$ is significant at the 95% confidence level.

Detailed instruction on calculating and interpreting the various significance tests is beyond the scope of this particular tip sheet. However, if you are interested or your study merits this level of sophistication, please consult the online resources listed below and/or contact the DM&E Initiative at dme@mercycorps.org.

Additional Resources (DM&E Guidebook and tip sheets in the DM&E-in-a-Box toolkit on the Digital Library.)

- *Mercy Corps DM&E Guidebook* for standard DM&E tools and concepts.

- *Mercy Corps' DM&E Tip Sheets on Planning a Survey, Designing the Survey Tool, Sampling, Focus Group Discussions, and Data Management.*

- *PinkMonkey.com:* Online stats textbook http://www.pinkmonkey.com/studyguides/subjects/stats/contents.asp

- *A Painless Guide to Statistics:* http://abacus.bates.edu/~ganderso/biology/resources/statistics.html

- The Statistics Home Page -- http://www.statsoft.com/textbook/stathome.html

- MOE calculators: http://www.raosoft.com/samplesize.html;http://www.americanresearchgroup.com/moe.html

- MOE overview: http://www.isixsigma.com/library/content/c040607a.asp; http://www.westgroupresearch.com/research/margin.html; http://www.cs.uiowa.edu/~rlenth/Power/

- Magnani, Robert. *Sampling Guide.* Top recommended technical guide on sampling in development projects. http://www.fantaproject.org/downloads/pdfs/sampling.pdf

- Hypothesis tests: http://www.csulb.edu/~msaintg/ppa696/696stsig.htm; http://en.wikipedia.org/wiki/T-test.

- EPI Info 2002 data management software: http://www.cdc.gov/EpiInfo/epiinfo.htm

- Computer-Assisted Qualitative Data Analysis Software (CAQDAS) available for purchase at: http://www.qualisresearch.com/default.htm; http://www.qualproinc.com; http://home.satx.rr.com/hyperqual

MercyCorps