The Rapid Evaluation, Action, and Learning (REAL) Approach

A toolkit to measure and refine changes and interventions in health campaigns
Primary users & target audience

This toolkit is geared towards readers who have some familiarity with operations research and monitoring and evaluation tools and approaches, including:

- **Staff of the Bill & Melinda Gates Foundation**, including country offices and program support teams who fund or provide technical assistance to campaigns as well as staff and leadership who make investment decisions about whether to scale up or transfer campaign effectiveness investments.

- **Campaign implementers and key stakeholders**, including Ministry of Health program managers and technical assistance providers/partners, with local contextual knowledge and decision-making influence, who are involved in processes that contribute to transforming results along the causal chain, thereby improving campaign effectiveness.

- **Campaign effectiveness grantees**, responsible for implementing campaign effectiveness interventions.

Learning objectives

By the end of this toolkit, users should be able to:

- Identify what information is needed to make a decision to scale up, scale down, or revise a change to a campaign.

- Describe the trade-offs and criteria for selecting the most feasible and appropriate (“fit-for-purpose”) approach to monitoring, evaluating, or operations research of campaign effectiveness changes.

- Describe the steps and components of the REAL approach to measure changes in the processes, outputs, outcomes, and impacts of interventions and strategies to improve campaign effectiveness.

- Articulate how to iteratively and continuously adapt interventions—and the implementation of interventions—to optimize outputs and outcomes.

- Identify what types of intervention they want to test and the most important characteristics of the intervention to test.
Before you begin

We recommend you use the companion toolkit “Identifying critical campaign challenges and diagnosing bottlenecks” before you embark on designing and testing a change.

Introduction

The goal of this toolkit is to improve the timeliness and relevance of evaluation and learning to accelerate improvements in the effectiveness of health campaigns.

The toolkit’s rationale is driven by growing global interest in “fit-for-purpose,” rapid-testing, adaptive learning approaches to evaluation, and the need for a culture shift towards iterative adaptation and improvement that integrates measurement and evidence-informed decision-making into daily practice (Kainz & Metz, 2015). This toolkit supports moving toward an ethos of “failing fast,” understanding what works (or not), and removing interventions that are less promising.

This toolkit describes a five-step approach to rapid evaluation, action, and learning (REAL) and how to implement it for a range of changes (i.e., solutions or interventions implemented to improve a campaign’s effectiveness). This approach improves efficiency by acknowledging that not every link in a logic model’s causal chain requires monitoring or testing. It improves transfer and scale-up of effective changes by involving diverse stakeholders to question program or intervention logic and then understand how and why an intervention works, as well as the role of context.

This approach can help to answer a range of questions, and this toolkit will help to identify which information is needed to make operational, strategic, or policy decisions.

Glossary of terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive management</td>
<td>An intentional approach to making decisions and adjustments in response to new information and changes in context.</td>
</tr>
<tr>
<td>Change</td>
<td>An intervention, activity, or solution that is the focus of the REAL approach.</td>
</tr>
<tr>
<td>Causal chain</td>
<td>Describes the process by which an initial change or input is transformed into a series of results.</td>
</tr>
<tr>
<td>Indicator</td>
<td>A specific, observable, and measurable piece of information used to track progress along a logic model.</td>
</tr>
<tr>
<td>Link</td>
<td>A single logical relationship between an input and its result.</td>
</tr>
<tr>
<td>Logic model</td>
<td>An illustration of the logical relationships between inputs and results (outputs or outcomes).</td>
</tr>
<tr>
<td>Result</td>
<td>What is produced by a link in a causal chain. A logic model can have multiple results along the causal chain.</td>
</tr>
<tr>
<td>Strategic question</td>
<td>A higher-level question whose answer can inform a management or strategic decision. For example: “Should we scale up the change?”</td>
</tr>
<tr>
<td>Testing question</td>
<td>An operational question that can be feasibly answered but which will also contribute to answering a strategic question.</td>
</tr>
</tbody>
</table>
Key questions answered through this toolkit include:

- Does the change produce its intended effect?
- Why or how does the change produce its intended effect?
- How can we further improve the process?
- What does the change cost?
- How could the change be scaled up?
- How could the change be transferred to another setting?

The last two questions are of particular importance to funders as they consider whether to recommend and/or invest in a given intervention or package.

Principles of the REAL approach:

- **Right-sized:** Aims to balance rigor with efficiency.
- **Strategic:** Prioritizes the production of information that is strategic for decision-making.
- **Context matters:** Aims to identify how a change is working, not just whether it is working.
- **Participatory:** Draws on multiple stakeholder views to strengthen the validity of logic models.
- **Timely:** Emphasizes rapid and right-sized assessment to better align with policy windows.
- **Adaptable:** Can be adapted to various types of change (simple, complicated, complex) or to changes in complex systems.
- **Ongoing:** Integrates design, measurement, and adaptive management to rapidly test interventions and enable continuous learning and improvement.

A hybrid approach: Agile science meets causal link monitoring

The approach presented here is heavily influenced by Heather Britt, Richard Hummelbrunner, and Jacqueline Greene’s 2017 white paper “Causal Link Monitoring.” We have adapted their approach and examples for campaign effectiveness interventions, drawing from other approaches, where appropriate, based on our review of dozens of protocols, publications, and monitoring and evaluation toolkits, including: agile science, causal link monitoring, continuous quality improvement, developmental evaluation, participatory action research, problem driven iterative adaptation, process evaluation, rapid cycle evaluation, and realist evaluation.

Key concepts of REAL:

- Not every link in a causal chain requires monitoring or testing!
- Transfer and scale-up can be improved by engaging diverse stakeholders to critique program logic.
Five Steps of REAL

REAL includes five steps. You should familiarize yourself with all of them before starting, as the REAL process is not always linear. Before you begin, consult the companion toolkit on identifying root causes of campaign challenges.

01. Identify a change to address challenges
02. Build a logic model for the change
03. Prioritize what to measure
04. Collect data
05. Use data for adaptive management

Figure 1. Rapid evaluation, action, and learning (REAL) approach.

Before you begin: identify the campaign’s main challenge and its root causes (see companion toolkit)

Step 1. Identify a change

Step 2. Build a logic model for the change

Step 3. Prioritize what to measure

Step 4. Collect data

Step 5. Use data for adaptive management

INTRODUCING THE TWO CASE STUDIES

This toolkit includes two hypothetical case studies to help the reader apply the toolkit to real-world campaign changes. One case study applies the REAL approach to a relatively simple change from India’s Intensified Mission Indradhanush (IMI) vaccination campaign and is used throughout this toolkit as an example. The characters in this case study include the District Immunization Officer (DIO), the WHO Surveillance Medical Officer (SMO), the Bill and Melinda Gates Foundation (BMGF) Program Officer (PO), and the IMI Monitoring & Evaluation (M&E) officer.

The second case study applies REAL to test a package of more complex changes from a human African trypanosomiasis (HAT) elimination initiative in the Democratic Republic of the Congo (DRC). This case is described in full in Annex 1.
STEP 01

Identify a change to address challenges

You have a campaign challenge to solve. In the companion toolkit on root cause analysis (RCA), you identified a challenge or bottleneck that constrained the effectiveness of a campaign. You learned to use root cause analysis to identify the root causes of the challenge, and assessed the actionability and potential impact of addressing each root cause. Based on the actionable causes you identified, you now need to identify a change to make and test.

Changes can be simple process refinements or a new intervention. The change may be obvious from the root cause analysis you did during the first toolkit, or it may require brainstorming and rapid prototyping to come up with a solution. Here are some suggestions to help identify an actionable and impactful idea for a change:

Adapted from The Field Guide to Human-Centered Design (IDEO.org, 2015)

→ Write the question or prompt somewhere where everyone can see (e.g., “How can we increase access to reliable transportation for campaign staff?”). Make sure the team understands the question before starting.

→ Brainstorm as many ideas as possible. We recommend using sticky notes. Encourage the team to build on others’ ideas and to not criticize any idea.

→ Once you have dozens of ideas on sticky notes, bundle them into groups of similar ideas. If you have a diagram of the RCA you can map them to the root causes and causal factors.

→ Choose which bundle to implement and test, or even multiple bundles of ideas. Continue considering feasibility and acceptability, how the solution will address the identified challenges, and what evidence we have from elsewhere on the solution’s effectiveness (particularly if it is an expensive or resource intensive solution).

Involve diverse stakeholders in this step, including those who can speak to the feasibility and acceptability of ideas, others that can speak to evidence of what works and technical considerations, and those who will ultimately implement the change (Seth & Menon. 2019).

The Field Guide to Human-Centered Design provides excellent guidance and examples for these steps and others. Because design has been described extensively elsewhere, this toolkit focuses on testing the changes. Here is some additional recommended reading to help you design a change in Step 1:

→ The Field Guide to Human-Centered Design
→ Design for Health website
→ Six Steps in Quality Intervention Development (6SQuID)
In this step you will learn how to build a logic model for your campaign change. According to CDC, a logic model is:

*a graphic depiction (road map) that presents the shared relationships among the resources, activities, outputs, and outcomes/impacts for your program. It depicts the relationship between your program’s activities and its intended effects, in an implicit ‘if-then’ relationship among the program elements — if I do this activity, then I expect this outcome. Among other things, a logic model helps clarify the boundary between ‘what’ the program is doing and ‘so what’—the changes that are intended to result from strong implementation of the “what.”* (CDC, 2017)

We often tend to oversimplify the ‘if-then’ relationships between a proposed change and its intended results. The added-value of REAL is that we unpack those assumptions to not only design more effective changes, but to identify what we should measure.

**This step is critically IMPORTANT because:**

- This step establishes the theory underlying how a change will contribute to achieving the intended outcomes, through one or more causal pathways.
- An accurate logic model will help you identify what to measure in Step 3.
- The more complicated a change, the more important this step is to explicitly articulate how the change(s) will produce the intended outcomes.

**2.1 Describe the change**

Describe the campaign ‘change’ as an input or activity. For example:

- “Mobile phone-based micro-planning”
- “Payment of campaign staff using mobile money”
- “Transportation vouchers for campaign staff”

Describing the change will build from the (hopefully) participatory approach you took in Step 1 to identifying or designing the change. Your description will be based on the perspectives and feedback from diverse but knowledgeable stakeholders who will be able to give different perspectives of what needs to be done.
2.2 Describe the intended outputs and outcomes of the change

List all the intended results of the change. It is often easier to first list them out and then order them sequentially. The campaign change should be written on the left with lines leading from left to right showing each result that is produced. List as many results in your chain as you can—don’t worry, you won’t have to measure all of them, but it is important to write out the full results chain.

Simple changes will produce relatively simple results chains (Figure 2). More complicated changes may produce more complicated results chains, with a single box producing multiple results (see HAT example in Annex 1).

For complicated and complex interventions, you should consider:

- Did we get the sequencing right? Are there any steps that need to occur first for other steps to occur?
- Does the change, or any result, lead to multiple results?
- Do we expect any feedback loops, such that the achievement of one result will influence an earlier result?

The team in India developed the following logic model in a few minutes (Figure 2).

2.3 Unpack assumptions in the causal links

This step unpacks the assumptions we make between the change and the result we think it will produce (Britt, Hummelbrunner, & Greene, 2017). Often, we tend to oversimplify results chains and logic models. By unpacking the assumptions between two boxes:

- Causal leaps become visible.
- Missed or forgotten steps are identified.
- Optimal sequencing is uncovered (for complicated changes).
- Measurement priorities become clearer (Step 3).

The team in India discussed their original logic model (Figure 3) and agreed that they had missed some key assumptions between the change and the first result. They were able to do this in a team meeting, but then vetted these assumptions with other stakeholders.

![Figure 2. Example logic model describing a campaign change and its results.](image)

![Figure 3. Logic model with added assumptions.](image)
2.4 Seek other perspectives and contextual factors

Vetting the logic model with diverse stakeholders using participatory approaches will ensure its validity and will improve the likelihood that findings will be used later. You can do this during regularly planned meetings or schedule in-person or telephone interviews.

Annex 2 provides a facilitation guide adapted from CDC for conversations about the logic model.

The team in India shared the logic model with the larger Intensified Mission Indradhanush team during a planned meeting. The meeting stakeholders were already familiar with the proposed change, so the team used the facilitation guide to vet the proposed logic model. One staff supervisor pointed out that there were many transportation options in the area, but that female campaign staff would not feel comfortable traveling alone in an autorickshaw or on a motorcycle; female staff would only ride in mini-buses, and even then, the decision to take a bus would depend on other factors such as how busy it was, what route it took, how the driver was behaving, etc. This conversation spurred the team to refine and add assumptions about transportation availability (Figure 4). ‘Acceptability of transportation (including gender)’ was added as a contextual factor to the logic model, noted as a grey chevron.
How to incorporate context in REAL

To answer questions about why and how the change works, and about potential for scale and transfer, a rich understanding of context is needed. Context can include local environment, decision-making structures, health systems organization, stakeholder norms and preferences—anything that might influence the achievement of a causal link (see examples of contextual factors in Figure 5). When these contextual factors differ across settings, the causal links may behave differently, producing different or unexpected results. Annex 3 provides a “context tracker” tool to help note and describe important contextual factors. In the case of India, the BMGF PO spent 10 minutes describing the local transportation infrastructure based on her initial observations, along with questions to follow up on during monitoring. She attempted to flag aspects which seemed locally-specific and which might not be found in other Indian states, or other countries.
Keep iterating your change

The logic model will help you identify how you need to iterate your change to make it even more likely to succeed. As you update your logic model and gain new insights, you will need to update your change. Consider whether you have the right staff or people involved; whether you have the right implementation strategy, training plan, or communications plan; and whether the physical components of your change (e.g., the transportation vouchers themselves) are easy to use. See The Field Guide for Human-Centered Design for more ideas, and Step 5.4 for an example from the IMI case.

2.5 Iterative updating of the logic model

Vetting with diverse stakeholders may identify the need to revise or adjust the change activity itself. Campaign stakeholders in India discussed whether they should introduce an intervention to mobilize bus drivers or schedule buses in advance. They opted to test the initial change for a few days and quickly assess whether it was sufficient or needed additional changes.

Your turn

What does the logic model look like for your change? You can use pen and paper, sticky notes, PowerPoint, Visio, etc. to create and update your logic model.
In this step, you will work with diverse stakeholders to identify which causal links are the most strategic to measure. The added value of the REAL approach is that not all results or relationships between changes and results need to be monitored or measured. Rather, the choice of relationships to monitor should be prioritized based on which are most strategic for success (Britt et al. 2017). In most cases, “success” is defined as the effective implementation and/or scale-up of interventions. As a general rule of thumb, focus on causal links where uncertainty exists or where there is disagreement between decision-makers, while keeping in mind the choice of areas of observation can change over the project lifetime (Britt et al. 2017). Another way to prioritize what to measure is based on the costs or other risks of making the wrong decision. If a process change has any potential to negatively affect the health or well-being of beneficiaries, for example, it should be rigorously evaluated.

You might already know quite a bit about some causal links from existing research evidence or evaluation. Or, some may be low-risk enough that you can count on expert opinion or strong theoretical plausibility of effect. Others may be dependent on earlier steps, and we don’t want to waste resources measuring a distal effect that is unlikely to occur because of uncertainty in more proximal causal chain. The goal of this step is to identify how to measure what will be needed to make a strategic decision to scale up, scale down, transfer, or not transfer the campaign change.

This step is critically IMPORTANT because:

- Not all causal links are necessary to measure.
- Choosing which to measure should be based on considerations of strategic importance:
  - What information do decision-makers need to scale up or down?
  - How much, and what type of information will they need?
  - Where is there the greatest uncertainty?

You may identify different criteria to guide your approach of deciding what to measure, and at what level of rigor, in your setting. The best practice is to include decision-makers in this stage so that you are clear about their decision-making needs and how they evaluate evidence.

For some decision-makers, it may be important to use an experimental design such as a randomized controlled trial to produce estimates of impact with a high level of confidence; however, experimental designs are typically more expensive and take longer, and your job is to weigh the trade-offs between measurement options (see an excellent overview of experimental designs) (Collins et al. 2011).

*Note that steps 3.1 and 3.2 can occur in any order, or in parallel.*
Scale vs. transfer and information needed

We define ‘scale-up’ as the expansion of a change, typically within an administrative jurisdiction such as a health district, province, or country. Replication, diffusion, or transfer refers to the process of taking a change or intervention from one context and setting and implementing it in another. Different types of information are needed to inform scale vs. transfer decisions (see Table 1 below). When considering whether a change can be replicated in or transferred to another setting, it is essential to understand how the campaign context interacted with the change mechanisms to produce the observed results. While many checklists exist for considering the transferability of interventions (Burchett et al. 2018), and all are useful in different ways, we will walk you through steps you can take to collect the right information to inform decisions to replicate, diffuse, transfer, and invest elsewhere.

3.1 Agree on what information is needed for ‘success’

A limitation of traditional impact evaluations is that while they can very precisely estimate the impact of an intervention in the study setting, they do not often produce information telling us how the intervention worked or whether it would work in another setting. For BMGF and other global stakeholders, questions such as ‘how and why did the intervention produce its results here, and could it produce similar results elsewhere?’ are as important as ‘was the intervention effective?’.

Common questions include:

- Does the change produce its intended effect?
- Why or how does the change produce its intended effect?
- How can we further improve the process?
- What does the change cost?
- How could the change be scaled-up?
- How could the change be transferred to another setting?

In this step, identify the key decision-makers of the REAL findings. This could be BMGF staff deciding whether to support the transfer of the intervention to other settings. It could be national or subnational policymakers and program managers (e.g., the Expanded Program on Immunization [EPI] manager) who makes key decisions about the design and implementation of the campaign. Discuss which questions are the greatest priority for each of them. While you are welcome to answer all these questions through rapid cycle testing, most testers prefer to prioritize a few for the sake of time and effort.
Table 1. Comparison of information priorities for each type of key decision-maker.

<table>
<thead>
<tr>
<th>Question</th>
<th>BMGF India Country Office, Program Officer</th>
<th>District Immunization Officer</th>
<th>WHO Surveillance Medical Officer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the change produce its intended effect?</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Why or how does the change produce its intended effect?</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>How can we further improve the process?</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>What does the change cost?</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>How could the change be scaled up?</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>How could the change be transferred to another setting?</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

In the IMI case, the team discussed whether to try to answer all questions or focus only on some. Local decision-makers were less interested in scale-up and transfer than was the BMGF PO. They all agreed that, at a minimum, they needed to know whether the change was working and how/why it was working to inform transfer and scale.

Strategic questions resulting from participatory discussions with decision-makers:

- Are transportation vouchers effective in improving the on-time arrival of campaign staff?
- Under what circumstances and for whom do transportation vouchers work to improve on-time arrival?

Your turn

Enter your key decision-makers in the columns, adding columns as needed. For each decision-maker, rate each question as ‘high,’ ‘medium,’ or ‘low’ priority to them. The facilitation guide in Annex 2 provides discussion questions for helping to establish their information priorities during meetings/interviews with them.
3.2 Identify causal links with greatest uncertainty

This is another step that is ideally done during meetings or interviews with key stakeholders and decision-makers. The goal is to identify the causal links with the greatest uncertainty; a central feature of rapid cycle testing is that it targets the testing and measurement to a few priority areas instead of having to measure everything.

In the testing team, consider which existing data are available for any of the assumptions.

During the stakeholder engagement for the logic model, ask the question: “How certain are you that this step will occur?” It can be helpful to walk through from beginning to end.

In India, after discussion and feedback from stakeholders, the team agreed that the greatest uncertainty existed for the assumptions and results highlighted below. They were confident that the provision of the vouchers (change) would occur because it was under their direct control. The team had no data on the frequency and timeliness of public transportation. One team member spent a morning observing bus flows in the main town and asked drivers about their routes. During meetings with stakeholders, there was much doubt whether all villages were covered by the bus network and whether campaign staff would be able to navigate the networks. There were additional concerns about female staff’s safety in the event they needed to change from one bus to another in a remote locale and some stakeholders questioned whether the drivers would even know to accept the vouchers. Basically, the team were collecting more questions they needed to answer and continuing to refine the logic model (Figure 6).

It is not uncommon to be equally uncertain for multiple assumptions/results. All else being equal, measurement should first focus on earlier links in the process which are necessary for subsequent results. In India, the team agreed they needed to simultaneously measure box 2 (available) and box 3 (on time).

![Figure 6. Refined logic model assessed for uncertainty.](image-url)
### 3.3 Combining the two criteria to clarify data to collect

In this step, we combine the priority strategic questions from 3.1 with the logic model's causal links with the most uncertainty to develop more operational, actionable testing questions and indicators.

In India, the team reworked the first causal link into testing questions that would help answer their strategic questions (i.e., whether, why, and how it works).

Ex. Testing questions for causal link “Appropriate transportation is available”:

- What proportion of female staff report being able to find a bus within 30 minutes?
- What proportion of vaccination sites are served by a bus route?
- New: Do drivers always accept the voucher?

The final articulation of questions or indicators should be informed by available data sources. For example, the team knew it could be easy to do an informal, verbal group survey during the weekly campaign staff meeting. They also knew it might be possible to build additional questions into the supervisors’ monitoring checklists, to enable real-time data collection of these new indicators. They also updated microplans to include the question of whether the vaccination site was on a bus route. Considering these feasible data sources, they updated the questions as follows:

<table>
<thead>
<tr>
<th>Original question</th>
<th>New question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What proportion of female staff report being able to find a bus within 30 minutes?</strong></td>
<td>No changes, plan to ask in staff meetings</td>
</tr>
<tr>
<td><strong>Do drivers always accept the voucher?</strong></td>
<td>No changes, plan to ask in staff meetings</td>
</tr>
<tr>
<td><strong>What proportion of vaccination sites are served by a bus route?</strong></td>
<td>Revised: Is this vaccination site served by a bus route? (to add to microplan and supervisor monitoring form)</td>
</tr>
</tbody>
</table>

The team realized, however, that in asking the verbal survey questions to staff during a meeting, they would not be able to link transportation challenges (or improvements) with site-level outcomes, which was important to demonstrate the effect of the intervention. The most feasible way of making this link would be through supervisors’ monitoring forms, so they added three simple questions which could help assess whether transportation remained a challenge:

- Vaccinator was over 30 minutes late for scheduled session (yes/no): ____
- Vaccinator reported challenges in finding transportation (yes/no): ____
- Vaccinator reported delays once on the bus (e.g. additional stops, inefficient route, etc.) (yes/no): ____

---

i. Be sure to follow all local or institutional human subjects research and research ethics policies.
Your turn

Using the blank decision-maker mapping tool from 3.1, list your top 3 strategic questions here:

1. 

2. 

3. 

Write the causal links with the greatest uncertainty here:

1. 

2. 

3. 

What are the actionable rapid cycle testing questions you will set out to answer?

1. 

2. 

3. 

Step 4 will provide tools and resources for identifying the most appropriate data sources and data collection approaches to answer these questions during a campaign.
At the end of Step 3, you prioritized which information you would need to be successful, articulated as actionable testing questions. In this step, you will develop a data collection and analysis plan that is fit-for-purpose (e.g., timely, feasible, affordable) and produces information required to make strategic decisions. Like the other steps, this step’s success depends in part on knowledge of the testing context and feedback from diverse perspectives.

Who should be involved? In addition to the existing recommended team, this step should most certainly include staff who are responsible for designing or implementing monitoring or data collection systems, or those involved in the analysis and interpretation of data. While this step doesn't include the analysis (see Step 5), effective analysis and interpretation require the right data, so it is important to include those analysts as stakeholders or team members.

**This step is critically IMPORTANT because:**

- Strategic decisions should be based on data or evidence (in its many forms) and this step helps to identify the most strategic data to collect/measure.
- Many readers will come to this toolkit with extensive background in measurement, data collection, analysis, evaluation, etc. This step reminds us to think somewhat differently about which data are needed and how to collect them, keeping in mind the principles of REAL.
4.1 Develop a data collection plan with indicators and data sources

In this step, we will convert the actionable testing questions into measurable indicators. Aligning these with existing indicators, for example those collected through routine campaign monitoring, can save time and avoid the need for new questions or data collection tools.

Indicators should be feasibly measured, whether through existing systems (e.g., routine campaign monitoring systems) or simple primary data collection (e.g., focus group discussions). Other innovations exist—depending on the context—such as robocalls, observation, meeting voting, etc.

The timing of measurement should align when the result is expected to occur. In the India case, whether a driver accepts a voucher can be measured immediately (or when is first convenient to do so) and does not need to wait. In the DRC case, the causal chain involving the new staff payment system was expected to take longer to produce results and there was acknowledgement that ongoing monitoring would be necessary to determine whether positive results persisted (Annex 1).

In India, the team fairly easily converted the actionable testing questions into measurable indicators (see Table 2 for illustrative data collection plan). In Step 3, they had already refined some to be more easily measured and had based their decisions on knowledge of data sources. The team agreed to collect and look at data as frequently as possible. These were the types of changes that should produce immediate outcomes, and if they weren’t, the team would refine the change.

It is important not to lose track of contextual factors. The team agreed that the member who volunteered to own the context tracker (see Annex 3) should continue to update it and take notes. She also offered to use that tool to record informal conversations with bus drivers or campaign staff about the new program and how it was working.
### Table 2. Illustrative data collection plan.

<table>
<thead>
<tr>
<th>New question</th>
<th>Indicator</th>
<th>Definition</th>
<th>Data source</th>
<th>Frequency of data collection</th>
</tr>
</thead>
</table>
| What proportion of female staff report being able to find a bus within 30 minutes? | Proportion of female staff able to find a bus within 30 minutes.           | **Numerator:** # female staff who raised hand in agreement when asked if they had tried to find a bus, and found one within 30 minutes that morning  
***Denominator:** # female staff who raised hand in agreement when asked if they tried to find a bus that morning | Staff meeting voting | Bi-weekly (each meeting)    |
| Do drivers always accept the voucher?                                        | Proportion of voucher attempts that are successful.                       | **Numerator:** # female staff who raised hand in agreement when asked if they had tried to pay a driver with a voucher that morning, and he accepted it  
***Denominator:** # female staff who raised hand in agreement when asked if they tried to pay a driver with a voucher that morning | Staff meeting voting | Bi-weekly (each meeting)    |
| Revised: Is this vaccination site served by a bus route? (to add to microplans and supervisor monitoring form) | Vaccination site served by a bus route.                                   | Yes/No                                                                                                                                                                                                     | Supervisor forms | At each site visit           |
| Do campaign staff arrive on time?                                            | Proportion of campaign staff that arrive on time.                         | **Numerator:** # staff who arrive on time in a given day  
***Denominator:** # staff assigned to a campaign post in a given day                                                                                                                                  | Supervisor forms | At each site visit           |
4.2 Work with systems owners to integrate indicators/fields into data collection tools and processes

Early in the process, it is important for the team to note all potential data collection systems and tools in use and whether opportunities exist to alter or add data collection fields. Not only is it important to be able to record the data collected, it is also important to be able to view or export the data for analysis. During this step, teams should think carefully about how they plan to look at the data and which analyses they may be performing. Teams should explore the ease of running reports or exporting data from routine systems, whether they need someone’s support to do this, and how frequently this could occur.

For planned changes occurring prior to the launch of a campaign, the team should quickly engage with the individuals or organizations responsible for the data collection systems and tools.

Qualitative data

As noted earlier, teams should be aware of and align with local Institutional Review Board (IRB) policies. In most cases, this type of work is considered routine program monitoring, and thus non-human subjects research. However, the team must think very carefully about whether any of the data could put campaign staff at personal or professional risk (e.g., through highlighting poor performance) or whether even casual conversations with campaign staff could introduce power imbalances and risks.

In India, the team included the IMI monitoring and evaluation (M&E) lead, who had full editor control of the campaign monitoring data collection application. She was able to add a new field to the digital supervisor form to capture data for the third indicator. To facilitate data collection and recording at the staff meeting (via voting), one team member marked themselves responsible for creating and managing a spreadsheet to keep track of numerator and denominator for each indicator at each staff meeting. He asked another team member to help with the counting at each meeting.

4.3 Implement data collection and look for trends

In REAL, data collection should begin as soon as is feasible but with attention paid to being efficient. For example, not all indicators may need to be collected immediately—timing may depend on the staging of expected results in the causal chain.

Training or supervision checks may be necessary before launching data collection. If routine reporting is modified, campaign data collectors (including staff and supervisors – whoever is using the modified tools) should be trained on how to enter new data and why they are being asked to do so. If additional support is used to implement qualitative data collection, training should focus on being flexible to answer the testing and strategic questions. Most qualitative data collectors will appreciate reading this toolkit to understand the larger context for this approach.
Once data collection has started, teams should look at the data to ensure there are no quality issues related to data collection or entry. Common data quality checks include (CDC, nd):

- **Completeness**: proportion of stored data against the potential of 100% complete.
- **Timeliness**: the proportion of stored data submitted on time.
- **Validity**: data are valid if it conforms to the definition for that data.
- **Accuracy**: the degree to which data correctly describes the “real world” event.
- **Usefulness**: for qualitative data, is the data collection useful in answering the testing questions and strategic questions? If not, data collectors may need additional coaching on how to frame questions, how to follow up with probes, etc.

Depending on the change, and the data and recording frequency, teams should begin to look for trends immediately. In the India case, the M&E lead managed to update the supervisors’ monitoring app within a few days to include the new field which was timed with a planned supervisors’ meeting where he explained the change. The team member responsible for recording contextual factors decided to go out on the first day of the change to try herself to use vouchers to see whether they were accepted. Three of four were immediately accepted, and the fourth driver accepted it as soon as she explained the scheme. During the first staff meeting, the team asked the survey questions and discussed the results immediately after the meeting.

### 4.4 Adapt and iterate data collection

Ongoing adaptation of questions, indicators, and data collection tools is encouraged. Adaptation should be based on:

- Experiences implementing the data collection plan.
- The need to respond to new information.

### Implementation lessons

Constant reflection and discussion should help identify whether the data collection plan is producing the information needed. If not, tweak it. This is not a research study, and thus there is no harm changing an actionable testing question mid-course to better be able to respond to strategic questions.

### Need to respond to new information

The logic model is dynamic, and as the change is implemented, the team may identify new assumptions and contextual factors or observe other causal links or unexpected consequences. The best practice is to continue to iterate on the logic model and cycle through the entire rapid cycle testing process to update what needs to be measured and how.
**Your turn** Based on the illustrative data collection plan in Table 2, propose a data collection plan for your campaign change.

<table>
<thead>
<tr>
<th>New question</th>
<th>Indicator</th>
<th>Definition</th>
<th>Data source</th>
<th>Frequency of data collection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In Step 4, you collected data for a carefully selected set of indicators. In this step, you will use the data to answer the initial strategic questions posed in Step 2 so that you or other decision-makers can make the necessary decisions to scale up, scale down, or continue to refine the change.

Your choice of analytic approach will depend on your strategic questions and decision-making needs. Upon reading this section, you may be inspired to go back to Steps 2 or 3 to refine your strategic questions, testing questions, or data collection plan and indicators.

As with all the other steps, Step 5 should be happening continuously until you have the information needed to take a strategic decision about the change. And although you are able to make a strategic decision, it doesn’t necessarily mean you are done collecting and analyzing information. At the end of this step, you will learn how to transition from a testing mentality to a change management mentality.

This step is critically IMPORTANT because:

- Many methods and approaches exist to analyze data and produce actionable, synthesized evidence from that data. This step will help to identify which approach is appropriate to support strategic decisions related to your change.

- In public health, epidemiology, and other fields related to campaigns, we tend to produce wonderful evidence that is not followed by the appropriate actions and decisions. This step will help you turn evidence into action—not only in terms of refinements to the campaign change but also in terms of broader evidence-informed decision-making for improving campaigns.

- Ongoing analysis and interpretation of data will support the ethos of adaptive management and continuous quality improvement.
5.1 Identify appropriate analytic methods for each strategic question

There are many different analytic approaches you can use to turn your quantitative or qualitative data into actionable insights. As you decide on approaches, keep in mind the strategic question you are trying to answer as well as decision-makers’ evidence needs and leverage existing meetings with key decision-makers to vet your approach (see examples, Table 3).

This step is one that should be done together with other question development and data collection steps as it may identify new data that are needed.

Table 3. Examples of analytic approaches for various strategic questions.

<table>
<thead>
<tr>
<th>Question</th>
<th>Quantitative</th>
<th>Qualitative</th>
</tr>
</thead>
</table>
| Does the change produce its intended effect? | Basic descriptive statistics (means, proportions, trends) Tests of differences across time or groups (e.g., if controls exist) | Case study approaches:  
  → Content analysis / thematic analysis |
| Why or how does the change produce its intended effect? | Associations between level of change and outcomes Regression analysis to disaggregate effect of the change | Case study approaches:  
  → Middle-range theory  
  → Thematic analysis  
  → Context mapping |
| How can we further improve the process? | RCA (companion toolkit) and human-centered design approaches for iterating |  |
| What does the change cost?            | Costing, budget impact analysis                                             |  |

The India team knew they would need to use a mix of quantitative and qualitative analyses to answer their strategic questions. They started by discussing how to best measure whether the change led to the intended result: an increased proportion of staff who arrived on time. They decided to plot that indicator over time from the daily monitoring data. They prepared an Excel line graph to do so and planned to use a Z test (a statistical test of differences in proportions) to test whether the change from pre-intervention to post-intervention (one and two weeks following implementation) was statistically significant.

The team knew it would also be important to visualize the entire causal chain, particularly if they did not observe a change in the results, in order to identify the 'blocked' link in the causal chain. They prepared an Excel graph of cascading bars to help with this analysis.

Finally, they planned to draw from the observations captured in their context tracker to explain why and how the intervention worked and key contextual factors that influenced outcomes.

5.2 Carry out the analysis

Plan for who will be responsible for the analysis and whether the team includes the right skills. Depending on the analytic approach, this might happen continuously (see Step 5.3).

ii. See for example Figure 3 in Mazzucca et al., (2018) for a helpful decision tree to choose among experimental and quasi-experimental study designs.
5.3 Take action based on information

During the first two days of the change, the India team realized that the change was largely successful but that a small proportion of villages were simply not served by existing transportation networks. Unfortunately, the microplans had not collected this information, so it was difficult to predict which villages required a private taxi service for the staff to get there. This resulted in the following refinement:

- A field added to the microplan template for future campaigns asking whether the village was served by an acceptable bus route.
- The logistics coordinator worked with Auxiliary Nurse Midwives (ANMs) to identify which villages were not on bus routes and thus would require a taxi during this campaign phase.

5.4 Summarize findings for strategic questions to key decision-makers

You involved key decision-makers in the choice and articulation of strategic questions, and now it is time to provide the answers to those questions. For each question or sub-question, you will want to provide the following information:

In the case of India, the team prepared a short slide presentation for key decision-makers for each strategic question. For example, they framed their content for the strategic question “Are transportation vouchers effective in improving the on-time arrival of campaign staff?” as follows:

- Providing campaign staff with transportation vouchers for public buses increased on-time arrival from 40% to 90% in one week.
- 70% of staff used the vouchers, at a cost of INR20 per staff per day. We anticipate the vouchers will be cost-neutral or saving over time as they will shorten campaigns and supervision costs.
- We recommend:
  - Provide vouchers for public buses.
  - Map bus routes as part of microplanning and flag off-route sites.
5.5 Transition to a change management mentality

Congratulations! You have collected, analyzed, and shared information to make a strategic decision about a campaign change. But the measurement and analysis are never over, nor is the pursuit of improvement.

<table>
<thead>
<tr>
<th>If...</th>
<th>...Then</th>
</tr>
</thead>
<tbody>
<tr>
<td>You found the change was not producing the intended effect, but interest remains in solving the bottlenecks.</td>
<td>Re-assess the causes of the challenge, particularly now having observed the context and issues for more time. Do you need to bring in different/new stakeholders or ideas? Revisit the logic model and its assumptions; go back to brainstorming ideas and solutions.</td>
</tr>
<tr>
<td>You found the change was not producing the intended effect and decision-makers opted to scale it down or halt it.</td>
<td>Iterate on the change</td>
</tr>
<tr>
<td>You found the change produced the intended effect and decision-makers opted to continue it.</td>
<td>Record and disseminate learnings</td>
</tr>
<tr>
<td>You found the change produced the intended effect and decision-makers opted to scale up or replicate it elsewhere.</td>
<td>Shift to a change management approach</td>
</tr>
<tr>
<td>You have the core tools and information needed to test the change anywhere: a logic model, strategic questions, and testing questions. However, these will all need to be refined for a new context and applied again. If the new context is not very different, or you found the change was not highly context-sensitive, you will need to make fewer refinements. If you cannot be involved in testing in the new context, package your learnings and tools for colleagues in the new context.</td>
<td>Revise the testing approach for a new context</td>
</tr>
</tbody>
</table>
Your turn
For each strategic question, discuss each heading as a group.
Conclusion

This toolkit describes an approach to rapid evaluation, action, and learning (REAL). REAL aims to improve the timeliness and relevance of information collected through monitoring and evaluation information to help inform strategic decisions about whether to scale up, scale down, or transfer a change. While this toolkit was written primarily to assess changes made to health campaigns, the principles can be transferred to other health or non-health changes.

The overall message we hope to convey is that timely and fit-for-purpose evaluation is feasible and should be done as part of a culture of continuous measurement and learning. Evaluation need not be onerous and time-consuming, but can generate real-time and useful information to act on. In applying the five steps of REAL, we hope that users will not only have more and better information on whether, why, and how their change works, but will also be able to improve the change as it is implemented, leading to higher impact public health interventions and changes.

References

Annex 1. DRC HAT case study

Case Example Overview: DRC’s Human African Trypanosomiasis Elimination Initiative

The Democratic Republic of the Congo (DRC) Human African Trypanosomiasis Elimination Initiative example was introduced in the RCA toolkit. Here we extend the DRC case to provide an end-to-end “how to” example for operationalizing the five-step approach to rapid evaluation, action, and learning (REAL).

Human African trypanosomiasis (HAT), also known as sleeping sickness, is a vector-borne disease transmitted by the bite of tsetse flies carrying Trypanosoma brucei protozoan parasites— if left untreated, HAT is often fatal. Although HAT is endemic to 36 countries in Sub-Saharan Africa, over 98 percent of cases occur in West and Central Africa from the T. b. gambiense parasite sub-species, and over 70 percent of reported cases are found in the DRC, where an estimated 53 percent of the population lives in areas considered at-risk for HAT infection.\(^1\)–\(^3\) The World Health Organization (WHO) is leading global efforts to eliminate HAT by 2020, with an elimination strategy that hinges on interrupting the disease transmission cycle through broad testing and treatment of people living in areas of risk—an approach supported by the development of new rapid diagnostics for detecting HAT\(^4\) and registration of Fexinidazole, the first all-oral drug that treats all stages of the disease.

In line with DRC’s commitment to the London Declaration on Neglected Tropical Diseases, the Ministry of Health and the National Control Program for Human African Trypanosomiasis (PNLTHA) developed a National Strategic Plan to eliminate HAT by 2020. The strategy draws upon new tools and technologies, including: rapid diagnostic tests, insecticide-treated traps for tsetse fly vector control, an awareness-raising advocacy campaign, digital technology for finding and confirming new cases, and mini-mobile teams to provide door-to-door HAT screening in high-risk areas of remote provinces.\(^5\) The new approaches appear to be working to bring the country closer to elimination: the latest data indicates the number of HAT cases in DRC declined from 1,200 in 2017 to 650 in 2018.\(^6\)

However, aspects of the HAT elimination initiative in DRC have faced numerous challenges that have required problem solving and subsequent implementation tweaks. In addition, a consortium of international partners has supported the development and introduction of innovative approaches to improve the effectiveness of the HAT elimination strategy.

DISCLAIMER: While DRC did not utilize the five-step approach to rapid evaluation, action, and learning described in this toolkit, we adapted this case as a hypothetical scenario for illustrating the application of each step. In Toolkit 1, root cause analysis was used to determine why the HAT campaign was not on track to meet the elimination targets. Here we focus on two of the root causes and how the DRC team could have implemented the five-step REAL approach to test solutions.
Toolkit 1 summary: Identifying critical campaign challenges and diagnosing bottlenecks

In Toolkit 1, the DRC team identified several root causes of why the HAT elimination initiative was not on track to meet targets. They assessed the actionability and potential impact of each root cause to determine which bottleneck to prioritize and address first. As depicted in the root cause analysis (RCA) pathway below, a key challenge is that HAT campaigns are missing opportunities for diagnosis and treatment, which is in part due to reduced effectiveness of the mobile screening teams (Figure A1). The effectiveness of the mobile screening teams has been compromised by two key factors:

01. Low motivation among HAT campaign staff due to delayed and incorrect payment.
02. Inefficient mobile route planning and targeting which misses at-risk villages due to reliance on hand-drawn maps of village locations. These maps are not drawn to scale. Additionally, an estimated 20 percent of villages in DRC are unknown/unmapped due to the difficult geographical terrain.

Figure A1. RCA diagram to understand why DRC is not on track to meet their HAT elimination target.

Figure A1 depicts the two primary root causes of DRC’s inability to meet HAT targets, that is:

» The suboptimal payment system.
» The difficult terrain, some of which is unknown/unmapped, and therefore not targeted by HAT mobile teams.
Analysis of the outcome of the RCA

These two primary root causes were analyzed to determine whether they could be resolved, using “actionability” and “impact” as criteria. Table 4 shows the outcome of this analysis:

Table A1: Assessment of potential actionability and impact of addressing root causes.

<table>
<thead>
<tr>
<th>Primary root cause</th>
<th>Actionability</th>
<th>Impact</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suboptimal payment system</td>
<td>Moderate</td>
<td>High</td>
<td>Payment delays and incorrect payments are large contributors to low worker motivation, but this is challenge is difficult to address because it requires escalation to the national level to introduce system-wide changes. In addition, some stakeholders have a vested interested in maintaining the status quo.</td>
</tr>
<tr>
<td>Difficult terrain (up to 20% of the country unknown and/or unmapped)</td>
<td>High</td>
<td>High</td>
<td>Technical solutions to address the bottleneck exist at a moderate cost (e.g., digitalization of geographic locations); political will and interest in digital health solutions exists; and addressing the root cause will be high impact given it is a country-wide barrier and a major contributor to inefficient HAT mobile team planning.</td>
</tr>
</tbody>
</table>

DRC Team resolution, action, and outcome

The DRC team decides to first focus on addressing the payment system by developing and launching the mobile money payment system for HAT mobile screening team staff.

In parallel, two technical partners collaborate to support DRC in undertaking rapid digitalization of priority geographic locations targeted for HAT mobile screening initiatives. The HAT mobile screening teams were accordingly provided with the new digital maps. However, the teams were still having issues with efficient route planning.

To assess whether the mobile payment system and digital maps are resulting in the desired outcomes, the DRC team decides to apply the rapid evaluation, action, and learning approach.
Rapid Evaluation, Action, and Learning (REAL)

Here we present a hypothetical scenario for how the DRC team could apply the five steps of rapid evaluation, action, and learning to address both the suboptimal payment system and the use and usefulness of digital maps for efficient route planning by mobile screening teams.

**Step 1. Identify a change to address challenges**

At the national level, the DRC team convened a diverse set of stakeholders to help brainstorm what change(s) could be implemented, keeping in mind cost, feasibility, and acceptability considerations. Stakeholders included partners, government officials, implementers of the HAT mobile screening strategies, and other health system actors with technical knowledge of the government payment processes for health care workers. Through discussion and debate, the stakeholders reach consensus around implementing several changes:

- Mobile money payment system to ensure prompt payment.
- Refresher training on digital map reading skills for all mobile screening team members to ensure their optimization of the digital tools.
- Monthly meeting between the HAT health zone-level supervisor and mobile screening teams to coordinate route planning for the 21-day missions.

While not all stakeholders are on board with these changes, there was enough consensus and buy-in from key actors to move forward with the rapid evaluation, action, and learning approach on a smaller scale (a few health zones) to assess whether the changes effectively address the root causes and should be scaled nationally to support the HAT elimination initiative.

**Key takeaways from Step 1:**

- Engage diverse and knowledgeable stakeholders to select intervention(s).
- Consider COST, FEASIBILITY, and ACCEPTABILITY.
- NOT every stakeholder will agree; aim at getting the view of the majority and building consensus.
Step 2. Build a logic model for the change

The second step of rapid evaluation, action, and learning is composed of five components:

2.1 **Describe the change(s):**
   - Payment of HAT mobile screening staff using mobile money system.
   - Refresher training on digital map reading skills for HAT mobile screening staff.
   - Monthly health zone coordination meeting for mobile screening team route planning.

2.2 **Describe the intended outputs and outcomes of the change:**

The DRC team spent a few minutes brainstorming intended results stemming from the changes:

- **Payment of HAT mobile screening staff using mobile money:**
  - HAT mobile screening team staff receive payment on time.
  - HAT mobile screening team staff receive correct amount of payment.
  - Reduced absenteeism.
  - Improved motivation and job satisfaction.
  - Improved number of clients screened in target villages.

- **Refresher training on digital map reading skills for HAT mobile screening staff:**
  - Improved digital map reading skills.
  - Improved ability to execute mobile screening routes/plans.

- **Monthly health zone coordination meeting for mobile screening team route planning:**
  - More efficient routes are planned.
  - Greater number of target villages reached with mobile screening.

Next, they developed a results chain to show the linkages between the change and the intended results (intermediate and longer-term), including any anticipated feedback loops where achieving one result may influence an earlier result. The process of developing the result chain required many rounds of reordering and iteration using sticky notes before arriving at the following logic model (Figure A2).

![Draft logic model highlighting linkages between intervention changes and results.](image-url)
2.3 Unpack assumptions in the causal links

In this step, the DRC team unpacked the assumptions between the change and the result(s) it is intended to produce. The team reviewed their original logic model above and determined that a few key assumptions were missing from the results chain. Through discussion they agreed on which key assumptions to include and developed a revised logic model (Figure A3). By going through this process, the team realized that each change depended on one or more assumptions before a result would occur. They also noted critical feedback loops whereby the achievement of a result could reinforce earlier results along the chain. For example, meeting daily screening targets for mobile teams would likely reinforce motivation and job satisfaction among the screening staff. In addition, the DRC team noted that sometimes the timing of one assumption or result depended on earlier ones. This is evidenced by the length of the results chains, where the team determined up to four intermediate results were needed before achieving the main target result of increasing identification and treatment of HAT cases. The team concluded that if they had implemented the changes without thinking them through, they may not have realized their intended results.
2.4 Seek other perspectives and contextual factors

The team convened a meeting to get feedback on the draft logic model and key assumptions. They invited multiple stakeholders that had been involved in the original change discussions, representing different levels of the health system and diverse perspectives from implementers, planners, partners, and the HAT mobile screening teams. The stakeholders raised additional points for consideration:

» The link between improved motivation/job satisfaction and staff absenteeism is likely more complex: at the health zone-level, there may be other contextual factors related to absenteeism (e.g., concerns about Ebola) that are important to consider.

» What aspects are important to ensuring the mobile money system works properly? For example, are there enough cash points conveniently located for mobile team staff to easily access payments? Do the cashpoints consistently carry enough money to pay mobile team staff? How will the mobile money payment system be maintained to ensure optimal performance?

» The monthly health zone coordination meetings need to ensure route planning is embedded within microplanning and considers key contextual factors (local environment, norms, and preferences). For example, a key assumption is that digital map data can be used to support development of realistic microplans. For villages that have never been visited before (e.g., previously unknown/unmapped), extra time should be estimated for community sensitization.

» Route planning should use the digital maps in conjunction with local knowledge on how best to sequence the villages (considering accessibility, roads, time), and considering any available data on the location of incident HAT cases.

Based on these discussions, the logic model was further refined to include reference to these additional considerations and contextual factors.
2.5 **Iterative updating of the logic model**

Through vetting the logic model with a diverse stakeholder group, the DRC team realized small revisions and adjustments may be needed in terms of how the change interventions are implemented. For example, the classroom-based refresher training on digital map reading skills is a necessary change but may not be enough, and the team is considering whether additional training modalities may be necessary (e.g., on the job supervision) to reinforce map reading skills. They opt to first try out the refresher training and assess whether it is producing the desired outcomes, or whether additional modalities are required.

---

**Key takeaways from Step 2:**

- Describe in clear terms the changes that the team wants to make.
- Define the intended outputs and outcomes of the change(s), and develop a results chain to show the link between the changes and the intermediate and longer term results; this gives you the first draft logic model.
- Review the draft logic model, critically analyzing assumptions in the causal link and revise the logic model accordingly.
- Conduct a stakeholder review of the draft logic model and incorporate feedback.
- Development of a robust logic model is an iterative process!
Step 3. Prioritize what to measure

3.1 Agree on what information is needed for success

Keeping in mind that not all causal links are necessary to measure, the goal of this step is to identify how to measure what will be needed to make a strategic decision to scale up, scale down, transfer, or not transfer the change.

The DRC team consider three likely decision-makers that will be interested in the rapid evaluation findings. Through knowledge of these stakeholders and informal interviews to understand their information priorities, the team mapped out level of interest in each of the questions below (Table A2). Based on this assessment, they decide to focus on answering whether the changes produce the intended effect (Q#1) for all three of the changes. They also decide to focus on answering how much the change costs for the mobile money payment system (Q#4). Given interest by international donors in scale and transferability, the team deliberates on whether to prioritize Q#2 and Q#6, but ultimately due to resource constraints decide to focus on collecting some limited information on how and why the intervention is working, which may help in understanding scale transferability later.

<table>
<thead>
<tr>
<th>Question</th>
<th>BMGF Program Officer</th>
<th>Provincial Health Officer</th>
<th>WHO Surveillance Medical Officer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the change produce its intended effect?</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Why or how does the change produce its intended effect?</td>
<td>Med</td>
<td>Low</td>
<td>Med</td>
</tr>
<tr>
<td>How can we further improve the process?</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>What does the change cost?</td>
<td>Low</td>
<td>High</td>
<td>Med</td>
</tr>
<tr>
<td>How could the change be scaled up?</td>
<td>High</td>
<td>Low</td>
<td>Med</td>
</tr>
<tr>
<td>How could the change be transferred to another setting?</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

Table A2. Comparison of information priorities for each type of key decision-maker
3.2 Identify causal links with greatest uncertainty

A central feature of REAL is targeting the testing and measurement to a few priority areas instead of having to measure everything. Thus, the DRC team returns to their logic model to begin identifying the causal links with greatest uncertainty (Figure A4). Based on stakeholder buy-in for a mobile money feasibility pilot, they were confident the mobile money system would be developed and rolled out but were less confident that mobile screening staff would receive the correct amount of payment and receive the payment on time, so they decided to focus measurement on this proximal indicator of success of the mobile money payment system.

Figure A4. Refined logic model pathways assessed for uncertainty.
3.3 Combining the two criteria to clarify data to collect

In this step, the DRC team combines the strategic questions from step 3.1 above with the logic model’s causal links with the most uncertainty to develop more operational, actionable testing questions and indicators. The team reworked the first causal link into actionable questions that would help answer their strategic questions (i.e. whether it works and how much it costs).

» Ex. Causal link: HAT mobile screening teams receive payment on time and in the correct amount:
  • What proportion of HAT mobile screening staff receive payment on time?
  • What proportion of HAT mobile screening staff receive payment in the correct amount?
  • When HAT staff do not receive payment on time or in the correct amount, why not?
  • What are the fixed startup costs to introduce the mobile payment system?
  • How much does the mobile payment system cost to maintain over time?

» Ex. Causal link: HAT mobile screening teams can correctly execute routes/plans using digital maps:
  • What proportion of HAT mobile screening staff can:
    — Correctly navigate digital maps following a refresher training course; and
    — Report feeling confident in their digital map reading skills
  • What proportion of target villages are reached for mobile HAT screening (per micro-plans)?
  • Why are some target villages still not being reached?

» Ex. Causal link: Monthly health zone coordination meetings, in conjunction with digital maps, support efficient route planning:
  • To what extent are the digital maps supporting identification of villages that have not previously been visited by the HAT mobile screening teams?
  • How is local knowledge of terrain and accessibility being combined with digital maps and data on incident HAT to determine mobile screening routes?

Key takeaways from Step 3:

» Only collect the information you require to make a strategic decision to scale up, scale down, transfer, or not transfer the change.

» To identify what data you need to collect, consider your strategic questions AND causal links with the most uncertainty in the logic model.
**DRC CASE EXAMPLE** RAPID EVALUATION, ACTION, AND LEARNING (REAL)

**Step 4. Collect Data**
To support this step, the DRC team invites colleagues involved in monitoring, evaluation, analysis, and interpretation of data to think through the design of the data collection plan and approach. In Steps 4 and 5, we focus on only one of the intervention solutions introduced: the mobile payment system.

### 4.1 Develop a data collection plan with indicators and data sources

In this step, the DRC team converts the actionable testing questions into measurable indicators (Table A3). Since the mobile payment system was newly being introduced, they could not rely on existing indicators; however, by considering priorities for measurement, the monitoring and evaluation staff provided input on what indicators to build into the routine data monitoring of the mobile payment system.

#### Table A3. Proposed data collection plan.

<table>
<thead>
<tr>
<th>New question</th>
<th>Indicator</th>
<th>Definition</th>
<th>Data source</th>
<th>Frequency of data collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>What proportion of HAT mobile screening staff receive payment on time?</td>
<td>Proportion of HAT mobile screening staff that receive payment on time.</td>
<td>Numerator: # HAT mobile screening staff who received payment on time following the mobile screening event Denominator: # HAT mobile screening staff who participated in delivering the mobile screening event</td>
<td>Mobile money system: Compare date HAT screening event ended with date payment sent to accounts of HAT mobile screening staff and retrieved from cash points.</td>
<td>Monthly, following each round of HAT mobile screening events.</td>
</tr>
<tr>
<td>What proportion of HAT mobile screening staff receive payment in the correct amount?</td>
<td>Proportion of HAT mobile screening staff that receive payment in the correct amount.</td>
<td>Numerator: # HAT mobile screening staff who received payment in the correct amount Denominator: # HAT mobile screening staff who participated in delivering the mobile screening event</td>
<td>Mobile money system: Compare the amount of payment sent to accounts of HAT mobile screening staff to the standard payment amount.</td>
<td>Monthly, following each round of HAT mobile screening events.</td>
</tr>
<tr>
<td>When HAT mobile screening staff do not receive correct and/ or on-time payment, what are the reasons?</td>
<td>Not applicable</td>
<td>Not applicable</td>
<td>Qualitative interviews to understand when and why payments not received on time or in correct amount.</td>
<td>Monthly for 3 months following initial rollout of mobile money payment system.</td>
</tr>
<tr>
<td>What are the fixed startup costs to introduce the mobile payment system?</td>
<td>Total cost to introduce the mobile money payment system.</td>
<td>Not applicable</td>
<td>Project records</td>
<td>Once</td>
</tr>
<tr>
<td>How much does the mobile payment system cost to maintain over time?</td>
<td>Total recurrent cost to maintain the mobile money payment system over time.</td>
<td>Not applicable</td>
<td>Project records</td>
<td>Once</td>
</tr>
</tbody>
</table>
4.2 Work with systems owners to integrate indicators/fields into data collection tools and processes

The DRC team worked with the mobile money payment programmers to ensure the prioritized indicators were incorporated into the system design and would be easy to view and export for routine analysis. By getting involved early in the design phase, the programming team provided initial prototypes of the data collection database, including snapshots of data exports and routine reports, for the DRC team to review. Through this process, the DRC team provided input on small tweaks to the data collection set up through the mobile money payment system. In addition, given the team's interest in the cost of this intervention, they were careful to collect and record all start up cost input information required to set up the mobile money payment system in the pilot area, as this evidence could help inform decisions around potential for scale-up.

4.3 Implement data collection and look for trends

To establish a relevant baseline comparison, while the mobile money payment system was being programmed and developed, the DRC team collected a 3-month baseline to determine the proportion of HAT mobile screening staff that received payment on time and in the correct amount. The DRC team knew this information would be very difficult and time consuming to collect because there was no routine system with the relevant indicators; while issues with the payment system were widely known and acknowledged, there was no systematic quantitative data that documented the extent of the problem. The DRC team therefore decided to conduct a short phone web-based survey among all HAT mobile team screening staff to gather information to provide a snapshot about the timeliness and accuracy of payment for the last two campaigns that the staff member had participated in (by shortening the duration of baseline, the team hoped to also limit recall bias). This approach required limited resources to set up and design yet provided enough data with appropriate rigor.

Introduction of the mobile payment system was expected to have an immediate impact on reducing delays in payments and ensuring HAT mobile screening staff received the correct payment amount. Therefore, as soon as the mobile money payment system was launched, the DRC team immediately began looking at the trends in the data through monitoring the routinely collected indicators in the payment system, including data quality checks of data completeness, timeliness, and accuracy.

4.4 Adapt and iterate data collection

Through routine data analysis, the DRC team determined that the mobile money payment system was increasing the proportion of HAT mobile screening staff receiving the correct payment amount. However, in the first two months following the payment system launch, it did not seem to have the intended effect on timely receipt of payment. To explore why, the DRC team decided to collect qualitative and process data about why and how delays in mobile payments occurred. They developed a short, semi-structured questionnaire guide on possible reasons for payment delays and devised a targeted approach that would be responsive to reported delays in payment receipt. They enlisted provincial-level HAT elimination staff to help in carrying out the qualitative data collection in areas where payment delays were longest and in areas where no payment delays occurred. Through comparing these two settings of higher and lower performance, the team aimed to determine the likely drivers of the delays. Collecting qualitative data can help to explain trends in qualitative data and can also be useful for documenting key contextual information that may relate to implementation success of a solution or intervention change.
Step 5. Use data for adaptive management

5.1 Identify appropriate analytic methods for each strategic question

The DRC team knew they would need to use a mix of quantitative and qualitative analyses to answer their strategic question: whether the mobile money payment system led to the intended result (increase in proportion of staff paid on time and in correct amount)? They developed an analysis plan to plot both indicators over time to visually compare the changes in trend lines—both overall at the national level and sub-analyses by provincial level. The provincial level analyses were prioritized as important given that qualitative data was revealing specific barriers and issues within specific provinces.

5.2 Carry out the analysis

The DRC team’s M&E officer conducted a Z test to compare whether there was a statistically significant increase in the proportion of staff paid on time and the proportion of staff paid in the correct amount, relative to the baseline. The quantitative analysis revealed certain provincial geographies that were underperforming. Based on the qualitative and process data, the DRC team began synthesizing information to help explain reasons for the underperformance. Through the analytical process, they determined several issues with HAT mobile screening staff being able to access payments from cashpoints. Although payments were sent to staff accounts on time and in the correct amount, in some areas, cashpoints were not conveniently located and/or did not have enough cash on hand to be able to pay the mobile screening staff, especially if multiple staff were trying to use the same cashpoint.

5.3 Take action based on information

Based on the evidence, the DRC team decided they needed more information from decision-makers on how to address the challenges faced with accessing money from cashpoints. Therefore, they organized a meeting with key stakeholders and decision-makers to share their early findings and to elicit brainstorm ideas on how best to address issues related to cashpoint placement and the availability of funds at cashpoints. After reaching consensus around some small tweaks to implementation, the DRC team resumes monitoring the key indicators for another two months.
5.4 Summarize findings for strategic questions to key decision-makers

The team feels confident in their data/findings and prepares a short slide presentation to share findings with decision-makers about whether the mobile payment system is effective in improving on time payments, and payments in the correct amount, to HAT mobile screening staff. The team presents their key findings, the evidence in support of the findings, the related costs of establishing the mobile payment system, and recommended actions to take.

5.5 Transition to a change management mentality

After collecting, analyzing, and sharing information with decision-makers, the measurement process is not over! The team continues to use the routinely collected indicators to monitor performance. Based on their finding that the change produced the intended effect, in several provinces, the decision-makers have opted to continue the payment system but work to refine aspects of it before moving to national scale-up.

References

Annex 2. Facilitation tool

This facilitation tool is adapted from the “CDC Program Evaluation Framework Checklist for Step 1” but has been adapted for REAL and health campaigns

The first step in the CDC Framework approach to program evaluation is to engage the stakeholders. Stakeholders are people or organizations that are invested in the campaign or the change, are interested in the results of the evaluation, and/or have a stake in what will be done with the results of the evaluation. Representing their needs and interests throughout the process is fundamental to good program evaluation. A program may have just a few or many stakeholders, and each of those stakeholders may seek to be involved in some steps or all six steps. This checklist helps identify stakeholders and understand their involvement in the evaluation.

Although “Engaging Stakeholders” is the first of the 6 steps, the first three steps of the CDC Framework are iterative and can happen in any sequence. For instance, identifying the right stakeholders may make more sense to do for your evaluation after drafting the purpose, user, and use of the evaluation that happens in Step 3. That said, this checklist will help you think through the key points in identifying and engaging stakeholders throughout your evaluation.

**Brainstorm potential stakeholders.** These may include, among others:
- People affected by your program, campaign, or proposed change.
- People involved in implementing the program or conducting the evaluation.
- People who will use the results of the evaluation. These may include internal staff, partners, program participants, community members, and other organizations, among others.

**In brainstorming the list be sure to think broadly,** including in your list:
- People in the above categories who share your priorities, and people who don't.
- People in the above categories who are critics as well as supporters.

**Especially if the list is very long, try to extract the subset of most important stakeholders.** Some helpful criteria for identifying whether a person or organization is a key stakeholder include that they:
- Increase the credibility of your program or your evaluation.
- Are responsible for day-to-day implementation of the program activities that are being evaluated and will need to implement any changes.
- Can advocate for the changes to the program that the evaluation may recommend, OR actively oppose the recommended changes.
- Fund or authorize the continuation or expansion of the program.

1  [https://www.cdc.gov/eval/steps/step1/Step-1-Checklist-Final.pdf](https://www.cdc.gov/eval/steps/step1/Step-1-Checklist-Final.pdf)
Discuss with key stakeholders individually the best way to engage them—in person, phone, email etc. Regardless of chosen medium, in the engagement discussions get clarity on the following questions: [NOTE: If a preliminary logic model for the change (Step 2) has been completed, then use it to help frame and target the questions.]

» What do you see as the main objectives of the proposed change?
» Which objectives are most important to you? That is, to retain your involvement and support, which objectives or results must be achieved?
» What is needed to implement and ensure the success of the change?
» What do you see as the most important evaluation questions at this time? [If outcomes are included]
» How rigorous must the design be?
» Do you have preferences regarding the types of data that are collected (e.g., quantitative, qualitative)?
» What resources (e.g., time, funds, evaluation expertise, access to respondents, and access to policymakers) might you contribute to this evaluation effort?
» In what parts or steps of this evaluation would you want to be involved?
» All or just some specific ones? How would you like to be kept apprised of this evaluation?
» How best to engage you in the steps in which you want to be involved?
» (How) will you use the results of this evaluation?

Examine the results of the stakeholder discussion for insights related to development/refinement of the logic model (Step 2), questions you will measure (Step 3), and analysis plan (Step 4).

Especially if there are many stakeholders, summarize the results of the engagement discussions with a [simple or detailed as you prefer] plan for stakeholder involvement, including which stakeholders will participate/provide input during the major stages of the project and what their roles and responsibilities will be for each step.
## Annex 3. Context Tracker Tool

Instructions: Use this tool to track how contextual factors influence the realization of specific causal links, or of the overall effectiveness of the change. You can enter your own contextual factors. See Figure 5 for a list of contextual factors.

<table>
<thead>
<tr>
<th>Contextual factor</th>
<th>Key insights: How might this factor influence the realization of the causal link?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socioeconomic</td>
<td></td>
</tr>
<tr>
<td>Stakeholders</td>
<td></td>
</tr>
<tr>
<td>Health system</td>
<td></td>
</tr>
<tr>
<td>Institutional</td>
<td></td>
</tr>
<tr>
<td>Cultural</td>
<td></td>
</tr>
<tr>
<td>Political</td>
<td></td>
</tr>
<tr>
<td>Physical/ environmental</td>
<td></td>
</tr>
<tr>
<td>Others:</td>
<td></td>
</tr>
</tbody>
</table>