Guide to evaluating behaviourally and culturally informed health interventions in complex settings
Abstract

This framework proposes a stagewise model for evaluating the effectiveness and sustainability of behaviourally and culturally informed interventions in complex settings, with detailed guidance and accompanying tools. It presents the theoretical background, addresses the challenges of assessing causality during times of change and of influencing factors, and provides a method for measuring the unintended positive and negative effects of interventions on well-being, trust and social cohesion.
Contents

Acknowledgements iv

Introduction 1
Exploring causality and contribution 1
Exploring intended and unintended effects 1
Who is this guide for? 1

Key concepts 2
Contribution analysis 3
Theory of change 3
Main evaluation questions 3
Unintended impacts on health equity, well-being, trust and social cohesion 3

How-to guide 5
Stakeholders and project coordination 5
An iterative process 5
Guiding principles 5
Managing expectations 5

Stages of the evaluation framework 6
STAGE 1. Design an evaluation plan 7
STAGE 2. Develop a theory of change 13
STAGE 3. Create an indicator framework 17
STAGE 4. Gather and analyse evidence 20
STAGE 5. Summarize and assess findings 23
STAGE 6. Share and follow up 26

Tools 28
TOOL 1. Indicators of well-being, trust and social cohesion 29
TOOL 2. Main evaluation questions 32
TOOL 3. Evaluation report template 33
TOOL 4. Terms of reference for the evaluation team 35
TOOL 5. Team communication plan 37
TOOL 6. Summary timeline 38
TOOL 7. Evaluation matrix 39
TOOL 8. Influence on decision-making 40
TOOL 9. Overview of evaluation methods 41
TOOL 10. Health equity indicators for rapid socioeconomic pre-assessment 43
TOOL 11. Theory of change visualization 47
TOOL 12. Indicator framework 48
TOOL 13. Evidence table 49
TOOL 14. Influencing factor rating 50
TOOL 15. Evidence rating 51
TOOL 16. Dissemination reach 52
TOOL 17. Exercises 53

Theoretical background 57
A transparent methodology 57
Understanding evaluation 57
Methodology 58
Evaluating interventions during a public health crisis 61

Glossary 66

References 68
Acknowledgements

This document was developed by the Behavioural and Cultural Insights Unit of the WHO Regional Office for Europe. The lead author was Nadia Aleyna Scott and co-authors were Katrine Bach Habersaat, Tania Aase Dræbel and Nils Fietje, with valuable input from Felicity Thomas, Katey Warran, Sideeka Narayan, Brett Janson Craig, and Martha Scherzer.

Early versions of the framework were used to evaluate three separate BCI interventions in Romania, Portugal, and Greece. The insights from these pilot evaluations were invaluable for refining the framework. The BCI Unit would like to thank the following individuals in connection with these pilot evaluations: Virgil Musta, Mihai Petcovici, Sorana Mocanu, Cassandra Butu, Andreea Popescu, Oana Blaga, Ágata Salvador, Leonor Costa, Shivani Atul, Ana Beato, Miguel Telo de Arriaga, Lydia Asoniti, Zoi Tsimtsiou, Apostolia Karampatea, Christos Hadjichristodoulou.
Introduction

Evaluation is a critical element of any intervention. The evaluation process involves exploring and documenting the effects of the intervention, what works well and what could be improved, and then using this information to continually improve or to scale up or replicate successful interventions. Using a theory-based method helps to explore all relevant factors in a systematic manner and reach reliable findings.

Public health interventions are often introduced in a context of continual change and high complexity, not least when the focus is on changing health behaviours and addressing health inequity. This evaluation framework introduces a stagewise method for evaluating interventions that take place in complex settings.

Exploring causality and contribution
Sometimes the direct effects of an intervention can be documented and it is possible to say that the intervention definitely caused the health outcome. In more complex situations, the health outcome may be caused by a number of factors, either within or external to the intervention. This evaluation framework can be used to uncover the contribution of the intervention to the health outcome.

Exploring intended and unintended effects
An intervention may successfully reach its targets and yet have unintended negative effects; alternatively, it may fail to reach its targets and yet have unforeseen positive effects. This evaluation framework takes account of both the intended and unintended effects (positive and negative) of an intervention. Well-being, trust and social cohesion in the target population and other populations have been identified as key factors that must always be explored alongside the targets set for the intervention.

Who is this guide for?
This guide is intended for health authorities and other organizations involved in evaluations of behaviourally and culturally informed interventions in complex settings. The best evaluations are created through an interplay between those who commission an evaluation and those who conduct it and produce the results. Therefore, this document is intended as a tool and reference for both parties.
Key concepts

This section introduces a few key concepts which will be referred to throughout the document.
Contribution analysis
When evaluating an intervention conducted in a controlled environment, it is often possible to assess the direct effect of the intervention, that is, the direct causality. The focus will be on the attribution of the intervention: did the intervention cause the expected results to occur?

When evaluating interventions in complex settings, the investigation aims to uncover the contribution of the intervention: did the intervention play a role in bringing about the observed results?

This evaluation framework builds on a contribution analysis (1) and provides a method to retrospectively map the steps between the intervention and the observed results in order to uncover and assess the contribution of the intervention.

This framework does not cover methods used to determine causal impact of interventions in real-world settings, such as randomized controlled trials (RCTs) or quasi-experimental methods.

Theory of change
Interventions in complex settings require a careful consideration of the contextual dynamics and of relationships between the various steps in the intervention. A theory of change provides a structure to describe such considerations before, both during and after the intervention.

TOOL 11 provides a template for the theory of change illustration.

A theory of change is a comprehensive description and illustration of how and why a desired change is expected to happen in a particular context. It accounts for all steps from the challenge to the goal and maps out the underlying assumptions and risks, as well as influencing factors.

Ideally, the theory of change is created prior to the intervention and is continually refined throughout the course of the intervention, as well as during the evaluation process. If a theory of change was not developed prior to the intervention, then it can be reconstructed as part of the evaluation process. This framework provides a method to reconstruct the theory of change (STAGE 2).

Main evaluation questions
All evaluations conducted using this framework should be based on the set of main evaluation questions. This provides a narrow focus for the evaluation and ensures that the implications for well-being, trust, and social cohesion are accounted for. To prepare the evaluation, the evaluation team will first tailor the main evaluation questions to the specific context and then develop subquestions to guide the data collection. The main evaluation questions cover three themes.

1. How has the intervention contributed to health outcomes?
2. How has the intervention contributed to well-being, trust, and social cohesion in the target group?
3. How has the intervention impacted people outside the target group – particularly those experiencing disadvantage?

TOOL 2 lists the main evaluation questions.

Unintended impacts on health equity, well-being, trust and social cohesion
There are many examples of interventions that failed to produce the expected results or that produced unintentional negative effects despite reaching their goals, as follows. If scare tactics are used, people may be frightened into changing their behaviours in the short term, but negative impacts on their well-being, trust and social cohesion may reduce their long-term motivation to maintain these behaviours. Alternatively, a well-meant campaign targeting a specific population group may unintentionally create stigma against the group.

For this reason, it is crucial to understand the environment of the intervention and to uncover whether changes in the environment could trigger opposing reactions or whether the intervention might have unintended impacts on participants. These unintended effects can be positive, negative or both. Accounting for unintended impacts is an essential part of the methodology laid out in this evaluation framework (Fig. 1).

Assessing health equity
Health equity is associated with education, living standards, gender equity, distribution of power, policy frameworks and social values (2). In developing this evaluation framework, special emphasis was placed on health equity and on exploring how interventions affect those experiencing disadvantage and who is being left behind. This is done by completing a socioeconomic assessment and a gender equality checklist as part of the planning phase (included in TOOL 10).

Assessing well-being, trust and social cohesion
To account for unintended impacts on the target population or on other populations, this framework
is based on assessing three factors that are critical to achieving long-term impacts on behaviours and cultural context and reaching public health goals: well-being, trust and social cohesion (3–6). These factors were selected because they are drivers of behavioural change, are sensitive enough to reflect the current situation in the population, and can be assessed both quantitatively through surveys and experiments and qualitatively through focus groups and interviews. TOOL 1 includes a recommended minimum questionnaire.

**Well-being** is about how we are doing as individuals, communities and nations. Well-being exists in two dimensions: subjective and objective. It comprises an individual’s experience of their life and a comparison of their life circumstances with social norms and values. Although people have an intuitive understanding of what it means to be well, what exactly creates well-being can be very different from person to person (7).

**Trust** is key to understanding the drivers of social outcomes (8). However, once lost, it is difficult to re-establish (9,10). Factors that affect trust in authorities include perceptions of their competence, objectivity, fairness, consistency, sincerity and faith (11,12). Therefore, any intervention or policy that changes people’s perceptions of authorities are also likely to affect their trust in these authorities – negatively or positively. Trusting relationships with **stakeholders** help to avoid confusion, distrust and misconceptions. Psychological research shows that the human mind is more trusting, positive and receptive when messages are clear and simple (13). Therefore, if an intervention is confusing for people to understand, their trust is likely to decrease.

**Social cohesion** covers multiple dimensions (including a sense of belonging and active participation) and is interrelated with the concepts of social capital (14). The Organisation for Economic Co-operation and Development recommends using social cohesion as a guiding principle for all policy design (14); it is also one of the five pillars of the United Nations socioeconomic response to the COVID-19 pandemic (15).
How-to guide

This section includes a guide to conducting evaluations of behaviourally and culturally informed health interventions. Each stage will be explained in more detail in the STAGES section and you will find a number of support tools in the TOOLS section.

Stakeholders and project coordination
It is suggested that all stages are led by a lead evaluator, who will document findings and decisions in a progress report and share this with an evaluation steering committee for input. The proposed actions in this guide should be conducted collaboratively by the steering committee. Other stakeholders can also be consulted for input. Step 1.1 includes more information about the recommended roles and responsibilities.

An iterative process
The framework ensures a theory-driven and methodologically rigorous process; however, users do not need to follow all steps exactly. Throughout the process, users are encouraged to develop rigorous thinking using peer-to-peer review.

Users are recommended to revisit steps and update evaluation goals as they collect more evidence and get a clearer picture of the role of the intervention. Reflective cycles throughout implementation of the evaluation will help you to identify the main concerns and problems, as well as the things that work and those that do not. These reflections should be documented in the form of notes and incorporated into the final report under gaps and challenges. See TOOL 3 for an evaluation report template.

Guiding principles
When conducting the actions in this guide, make sure to work as a team and to follow the guiding principles.

- Be respectful of diversity of opinion. Questioning the intervention logic only serves to make the process stronger.
- Encourage skepticism of the theory of change and the intervention as a whole. Recognize that skepticism has cultural contexts and that these also apply to you, not just to others.
- Don’t strive for consensus as there can and should be multiple competing hypotheses for how the intervention is supposed to address the problem.

Managing expectations
Conclusions based on the approach used in this framework will not provide definitive proof, but will rather highlight the factors contributing to an effect. No matter how much one tries to account for all of the dynamics and implications of implementing an intervention in a complex setting, it will never be possible to fully document its contribution. The theory of change will help to account for behavioural nuances and influencing factors, but will still be a simplification.

The framework aims to highlight both individual experiences and contextual factors. In addition, a mixed methods approach including both quantitative and qualitative research methods produce a combination of numerical and generalizable data that generates a deeper insight than either method alone. Quantitative research is based on a structured collection and analysis of numerical data and can provide information on the frequency of certain behaviours, beliefs and knowledge. If data are collected from a representative sample, it is possible to generalize the results to a larger population. Qualitative research is conducted to gain an understanding of a target group’s perspectives and experiences and can explore the reasons why people make certain choices and adopt specific behaviours.
Stages of the evaluation framework

STAGE 1
Design an evaluation plan
1.1. Build the team
1.2. Estimate the budget and time
1.3. Determine the purpose and evaluation questions
1.4. Identify users and applications
1.5. Plan for dissemination of findings
1.6. Select sources and methods
1.7. Collect evidence
1.8. Conduct a socioeconomic assessment

STAGE 2
Develop a theory of change
2.1. Visualize the theory of change
2.2. Identify the influencing factors
2.3. Identify the assumptions
2.4. Identify the risks

STAGE 3
Create an indicator framework
3.1. Revisit the evaluation matrix
3.2. Choose the indicators
3.3. Determine the baseline
3.4. Detail the methodology
3.5. Rate the evidence
3.6. Update the evaluation plan

STAGE 4
Gather and analyse evidence
4.1. Test the theory of change
4.2. Review the influencing factors
4.3. Identify supporting evidence
4.4. Consider additional factors
4.5. Rank the influencing factors

STAGE 5
Summarize and assess findings
5.1. Write the contribution claim
5.2. Rate the evidence
5.3. Rate the theory of change
5.4. Test the contribution claim
5.5. Write up the evaluation findings

STAGE 6
Share and follow up
6.1. Conduct the final team session
6.2. Disseminate the findings
6.3. Follow up
STAGE 1

Design an evaluation plan

1.1. Build the team
1.2. Estimate the budget and time
1.3. Determine the purpose and evaluation questions
1.4. Identify users and applications
1.5. Plan for dissemination of findings
1.6. Select sources and methods
1.7. Collect evidence
1.8. Conduct a socioeconomic assessment

Objectives
The objectives of STAGE 1 are to prepare the evaluation in detail, assess the organizational readiness and plan accordingly.

Steps
Although it is tempting to jump straight into the evaluation, investing in this first stage will make every next stage easier and increase the value of the findings. In the following subsections, you will find a breakdown of the steps in STAGE 1.

Tools
In this section you will encounter references to the following tools:

TOOL 2. Main evaluation questions
TOOL 4. Terms of reference for the evaluation team
TOOL 5. Team communication plan
TOOL 6. Summary timeline
TOOL 7. Evaluation matrix
TOOL 8. Influence on decision making
TOOL 9. Overview of evaluation methods
TOOL 10. Health equity indicators for rapid socioeconomic pre-assessment.
1. **Build the team**

**Team and steering committee.** The evaluation team can be an existing group or a group established for the purpose. It may be useful to establish a steering committee. While the steering committee members do not have to be experienced evaluators, they should contribute with an awareness of behavioural, cultural, social, organizational, historical, public health, medical, or other related perspectives.

**Map resources.** Negotiate the time resources for the evaluation and state these decisions clearly in the evaluation plan. Determine what financial resources are required and also state these in the evaluation plan. Be sure to account for financial costs that might be incurred beyond analysis and reporting.

**Define responsibilities.** Define team responsibilities and reporting structure by referring to the terms of reference in TOOL 4. An important part of this work is to decide who will conduct the actual evaluation and who will manage the evaluation internally.

**Plan communications.** Ensure that all team members can participate in (i) providing input into the evaluation questions, (ii) receiving regular updates on progress, and (iii) participating in the interpretation of evaluation results and implications for action. Create a communication plan to determine and maintain effective flows of information. The communication plan should cover stakeholder groups, communication goals, and the content and frequency of communication. See TOOL 5 for the team communication plan.

1.2. **Estimate the budget and time**

Set out a rough budget estimate based on the decision in Step 1.1. Then determine when the evaluation should aim to be completed, bearing in mind what it seeks to achieve. It is essential that the findings be completed within a time frame that best aligns with the decision-making requirements. Use the summary timeline in TOOL 6.

1.3. **Determine the purpose and evaluation questions**

The four main categories of evaluation purposes are described below. Each category includes a number of facilitation questions to support the evaluation team.

1. **To make a judgement about the worth or merit of a programme, policy or activity:**
   - Is the intervention meeting – or has it already met – its objectives?
   - Should it continue to be funded, or is there another intervention that is more **effective**?
   - Should this intervention be **scaled up** within the country/region or promoted for **replication** or **adaptation** in other countries/regions?

2. **To improve a programme:**
   - How could the current intervention be **improved**?
   - How could these improvements be **incorporated** into the intervention process?
3. To help support the design and development of a programme/organization:
   • what is the best strategy to achieve the objective of increasing protective measures while minimizing unintended consequences?

4. To create new knowledge:
   • does the intervention work as theorized?
   • what can be learned about the barriers and drivers of public perceptions and behaviours in order to increase adherence among the public?

When you have determined the purpose of the evaluation, you are ready to tailor the main evaluation questions (see TOOL 2) and to brainstorm on the subquestions in your evaluation. These should be recorded – and updated if necessary – in an evaluation matrix, which is an organizing tool to help to plan for the data collection, analysis and report writing phases (see TOOL 7).

1.4. Identify the primary intended users and applications
You need to take into consideration two main groups of stakeholders: the primary intended users of your evaluation findings, and interested and affected parties. With these groups identified, you can define the intended applications of the evaluation.

• Primary intended users are those who have the primary stake in the evaluation and who will commit to taking ownership of the findings by acting on them. They are interested and knowledgeable about the programme and are open and available for interactions.

• The primary intended applications of the evaluation are identified in a discussion between the evaluator and the primary intended users. It is important to recognize that many evaluations can have multiple uses, such as providing an opportunity for learning, generating input for improvements or assessing impact. TOOL 8 includes facilitation questions to help you to reflect on how the evaluation will influence decision-making (16).

CONSIDER INCLUDING A WIDER RANGE OF STAKEHOLDERS

For the evaluation to be credible, it is important to incorporate a wide range of stakeholders, as well as their perspectives and experiences, preferably from the beginning of the process.
1.5. **Plan for dissemination of findings**

Once you have identified the primary intended users and applications of the findings, you are ready to start drafting your dissemination plan. This should include an overview of stakeholders, channels and frequency of dissemination, as well as allotting responsibility for the dissemination plan.

- **Stakeholders.** With whom will you share the findings of your evaluation?
- **Channels.** What formats and channels will you use to share findings?
- **Frequency.** How frequently do you plan to share findings?
- **Responsibility.** Who is responsible for disseminating the findings?

You can adapt the communication plan in TOOL 5 by adding the columns needed for your specific task.

1.6. **Select information sources and analysis methods**

Next, the evaluators and the primary intended users should collaborate to determine which data and findings are needed to support decision-making, according to the intended uses. Determine the most appropriate method to collect data on answers to each main evaluation question and subquestions, bearing in mind feasibility, timeliness and other factors that might impact on the chosen methodology. Box 1 provides an overview of methods to assess and evaluate the impact of interventions (some of these methods are described in more detail in TOOL 9). Other types of evaluations, such as process evaluations, often involve qualitative methods and the collection of multiple qualitative and quantitative datasets. When the evaluation aims to understand how an intervention works, it is more appropriate to use a mixed methods approach: for example, qualitative data from interviews and observations can be combined with quantitative data on attendance rate or questionnaire data.

Using the evaluation matrix guides the analysis and ensures that all collected data are analysed and used to answer the evaluation questions. Remember to determine whether ethical review is needed for each evaluation method. If so, initiate this process as soon as possible.
STAGE 1
Design an evaluation plan

**BOX 1. INFORMATION SOURCES FOR IMPACT EVALUATIONS**

Typical information sources include:
- **surveillance** data;
- data from social and behavioural insights surveys;
- **health service utilization data**, including the use of quarantine facilities;
- population health data;
- **lifestyle data** and reports;
- **population health** and socioeconomic surveys, analysis and studies;
- data from the Global Health Observatory (17);
- multiple Indicator **Cluster Surveys**;
- surveys, strategies, action plans and **legislation** relevant to the target group;
- reports and evaluations of previous interventions from national and international organizations, as well as academia;
- **media** and social media coverage within the time frame of the intervention; and
- peer-reviewed **academic publications**.

Evaluation methods cover a wide range of qualitative and quantitative data collection techniques and data analysis approaches. This framework recommends using mixed methods approach that integrates quantitative data from statistical databases, epidemiological reports, surveys or other sources with qualitative insights obtained from focus groups, interviews, observations or other sources. Use of a mixed method approach increases the likelihood that the evaluation will capture how aspects, processes and causal pathways interact in interventions in complex settings. It also improves robustness and comprehensiveness of the findings. TOOL 9 summarizes the different evaluation methods and recommends when to use each one.
### STAGE 1
Design an evaluation plan

**1.7. Start collecting evidence**

Begin by bringing together all of the evidence at hand. This will help you to reconstruct the theory of change and shed light on how closely the intervention delivery corresponded to the implementation plan. Helpful evidence sources include:

- **documentation** on the intervention
- **programme files** and policy briefing notes
- **implementer reports** or participant evaluation forms
- **registration lists** or forms.

Obtaining some of these materials might require extra time, money or special permission, including:

- requesting access to relevant health or socioeconomic databases;
- requesting materials from libraries;
- requesting unpublished data or reports from relevant researchers; and
- analysing epidemiological or other primary health data.

**1.8. Conduct socioeconomic assessments**

The best impact assessment of an intervention is obtained when the socioeconomic environment in the region or country is considered. For this reason, this framework recommends that you finalize your planning phase by completing a socioeconomic assessment and a gender equality checklist, as follows.

- **Health equity** indicators for rapid socioeconomic pre-assessment (TOOL 10).
- **Minimum requirements checklist** for integrating gender equality in A UN framework for the socioeconomic response to COVID-19 (14).

**OUTPUTS FOR STAGE 1**

Upon concluding STAGE 1, you will have created an evaluation plan that includes:

- **The team** – the roles, responsibilities, reporting structure and communication plan have been defined.
- **The approach** – the main evaluation questions, primary users and applications of findings, information sources and evaluation methods, budget and timeline have been prepared.
- **Checklists** – socioeconomic assessment and gender equality checklist have been completed.

**BE PRAGMATIC**

Rather than setting out to collect only hard facts, accept that the process is iterative and a learning experience. Be realistic and pragmatic: aim for the highest quality possible in the given context.
STAGE 2

Develop a theory of change

Objectives
If the intervention has already been implemented and a theory of change was used, the objective of STAGE 2 is simply to recall this. If no theory of change was used to design the intervention, then the primary aim is to reconstruct the theory of change that underlies the intervention logic. In the latter case, consider using the COM-B (18) or a similar health model as the starting point.

Steps
Describing a theory of change is a collaborative effort, requiring consultation among a range of stakeholders. Together, you will construct or reconstruct the logic behind the intervention: What was the challenge? What was the end goal? and What were the assumptions about each of the steps towards the goal? Once documented, this becomes a theory that you can revisit and refine during the evaluation process as new knowledge arises.

Tools
This section will refer to the following tools: TOOL 11. Theory of change – visualization
STAGE 2
Develop theory of change

2.1. Visualize the theory of change
Your first step in developing the theory of change is to create a visual model that captures the rationale behind the intervention (the challenge, the steps in the intervention and the interconnectedness among the steps) and how this is believed to lead to change. If there are competing hypotheses, draw a model for each of them. There are several models to choose from, including:
- a linear model – the simplest and oldest approach;
- a logical framework (or log frame) – similar to the linear model but in tabular form; and
- a causal model – a nonlinear approach with feedback loops that can be used for causal analysis at a later stage.

TOOL 11 shows a generic model for visualizing the theory of change. You can use this as guidance to create your own visualization.

2.2. Identify the influencing factors
Brainstorm to identify which influencing factors might have played a role in bringing about the expected results. As there are likely to be several, keep track of them in a table. Box 2 provides some facilitation questions as a helpful starting point; primary intended users should be able to draw upon their expertise to brainstorm potential influencing factors.

BOX 2. EXAMPLES OF INFLUENCING FACTORS

Other interventions
- Are there other ongoing interventions that could influence the results?

Media
- Were major scientific or policy announcements made during the intervention?

Socioeconomic forces
- Was there considerable societal or structural change during the time frame of your intervention?
- Are economic or labour indicators trending upwards or downwards?
- Were there any major political events (including scandals) during the intervention?
2.3. Identify the assumptions
Assumptions are the external conditions that need to be in place for the intervention logic to remain valid. As a group, discuss the assumptions made in your model and keep track of your results. As these assumptions will vary according to which theory of change you are using, be prepared to update this list as necessary (Box 3).

BOX 3. ASSUMPTIONS

Intervention delivery
- Did the intervention occur as intended?
- Were changes made along the way?

Influencing factors
- Did the influencing factors stay constant throughout the intervention?
- Were others introduced?

Causal links
- Does your intervention logic rest upon one or more causal links? For example, are you assuming that all members of your target group can communicate in the country’s official language?
- Causal links can also exist in the theory of change itself. For example, using the COM-B model(18) you might have identified that capability is the barrier to the behaviour that your intervention seeks to modify.

2.4. Identify the risks
Your assumptions might not remain valid under all conditions. Identifying risk is about identifying the conditions that will challenge your assumptions. Use the examples given in Box 4 to inspire your own identification of risks to your assumptions.

BOX 4. RISKS

Intervention delivery
- The intervention delivery did not go as planned.

Influencing factors
- Influencing factors did not remain constant throughout the intervention.
- Other interventions were introduced during the same period that targeted the same groups.

Causal links
- For example, in a communication intervention, migrant groups might not be fluent in the language of your communication strategy.
- For example, in the theory of change, the wrong barrier to or driver of behaviour might have been identified.
STAGE 2
Develop theory of change

OUTPUTS FOR STAGE 2

Upon concluding STAGE 2, you will have documented the theory of change including:

- **visualization** of the logic behind the intervention
- **influencing factors** in the context
- **assumptions** about the influencing factors
- **risks** to your assumptions.

THE THEORY OF CHANGE CAN BE REVISED

The theory of change is both an ongoing process of learning about how change happens and a model that illustrates the process of change. Develop, revise and be ready to adapt your theory of change so that it reflects the findings.
STAGE 3

Create an indicator framework

3.1. Revisit the evaluation matrix
3.2. Choose the indicators
3.3. Determine the baseline
3.4. Detail the methodology
3.5. Rate the evidence
3.6. Update the evaluation plan

Objectives
The objective of STAGE 3 is to select the indicators that will allow the evaluation team to answer the main evaluation questions and subquestions. Indicators are the types of data or information used to measure change. They provide information on whether the intervention has had the desired output (19).

Steps
There are two main types of indicators: process indicators and impact indicators. In this stage, you will create an indicator framework and select the relevant indicators for the main evaluation questions, as well as for the subquestions created in STAGE 1.

Tools
In this section you will encounter references to the following tools:
TOOL 12. Indicator Framework
TOOL 4. Terms of Reference for the Evaluation Team
3.1. Revisit the evaluation matrix
Think about how to best answer each question in the evaluation matrix you created in Step 1.3. At this point, the main evaluation questions and subquestions might need to be revised to reflect changing views or new information. Some typical examples of indicators that could be relevant for your intervention are given below.

Indicators to measure **protective behaviours** include:
- the area or level of skills achieved
- the area or level of knowledge achieved
- types and level of activities performed.

Indicators to measure **positive influencing factors** include the following types of **contribution**:
- contribution to well-being
- contribution to building trust
- contribution to social cohesion.

3.2. Choose the indicators
Consider the descriptions given above and follow best practice to understand which indicators will work best for you. In this process, you can review literature, including grey literature, and talk to subject matter experts. You can also draw upon both quantitative and qualitative sources of data to build credibility (20).

3.3. Determine the baseline
The baseline is the starting point for your measurements. If the baseline levels were **predetermined**, list sources of baseline data in TOOL 12. If baseline levels were **not predetermined**, explore alternative possible sources. While these data might not be appropriate for quantitative temporal comparisons, they will provide a useful assessment for the evidence review step. To explore alternative sources, you could:
- consult **public health databases**, as data on some indicators are collected frequently;
- talk to **colleagues in the field** who might have run surveys that indirectly covered your target group; and
- use insights from the **behavioural insights surveys**, if available.

---

**IT CAN BE DIFFICULT TO ESTABLISH A BASELINE DURING A HEALTH EMERGENCY**

In the case of a health emergency, it can be very difficult to establish a baseline for an intervention, particularly where good health monitoring was not previously in place.
3.4. **Detail the methodology**  
In Step 1.7 you selected your evaluation methods. Now it is time to start the more detailed planning by considering:  
- who will be the **target group** for each method?  
- do you need to add **additional groups** to measure potential unintended impacts?  
- **when** will the methods take place?  
- do you need **additional team members** to broaden the perspective (see TOOL 4)?  

This is also the time to make sure that facilitation guides are written, interviews scheduled and survey questions rigorously defined.

3.5. **Rate the evidence**  
Take some time to discuss how you would rate the evidence source used to answer the evaluation question for each indicator. This framework recommends using a three-point scoring system (weak, medium and strong) according to the weight placed on each evidence source. Insert your rating in **TOOL 12**.

3.6. **Update the evaluation plan**  
If new insights are raised during the indicator selection process, update the evaluation plan for the following areas:  
- information sources and evaluation methods  
- timeline and budget.

---

**OUTPUTS FOR STAGE 3**

Upon concluding STAGE 3, you will have completed your indicator framework including:  
- a breakdown of the focus of the main evaluation questions  
- specific indicators (20) for each question  
- the availability of baseline data  
- how these indicators will be assessed  
- a measure of the assessment or rating.
STAGE 4

Gather and analyse evidence

4.1. Test the theory of change
4.2. Review the influencing factors
4.3. Identify supporting evidence
4.4. Consider additional factors
4.5. Rank the influencing factors

Objectives
The objectives of STAGE 4 are to:
- **organize the evidence** to make it understandable and relevant to primary intended users
- **involve primary intended users** in interpreting the findings
- **examine the findings** and their implications with a focus on their intended uses.

Steps
This stage is where the team first critically examines data and information regarding the results and the theory of change. It is suggested that you start with a small amount of evidence and then expand outwards as needed when additional clarification and understanding is required. It is better to make **causal inferences** from a few good-quality sources than from many low-quality sources. Additional sources might include direct measurements, additional surveys, reviews of programme files or administrative data, and searches of public health databases.

Tools
In this section you will encounter references to the following tools:
- TOOL 13. Evidence table
4.1. Test the theory of change
Start out by analysing the results from the selected information sources, and then summarize the findings using an evidence table (TOOL 13) and state your conclusions based on the findings. Now compare the findings to your theory of change and to the targets set in the indicator framework. Use facilitation questions to determine whether these findings make sense and whether they show that the intervention has achieved its target levels for behaviours. If there is a high level of contestedness – or a lack of consistency between hypotheses and findings – seek additional information or revise the theory of change.

4.2. Review the influencing factors
Awareness of influencing factors help the team to assess if any possible change – success or failure – was due to the intervention itself or to influencing factors. These may include other activities, policies or interventions implemented in parallel; legislative changes; political or societal changes; structural changes; major scientific breakthroughs; instances of political corruption; and significant media announcements.

List all relevant factors and review their potential impact on the result. If possible, retroactively establish a baseline to compare with the situation after the intervention.

4.3. Identify supporting evidence
Conduct a literature search to determine whether the evidence suggests that any of the influencing factors could have brought about the results. Discuss as a group whether you believe that these factors could have brought about or influenced the results.

4.4. Consider additional factors
There might be further influencing factors that you have not yet identified. For that reason, it is suggested that you examine evaluation and research studies to see whether they suggest other influencing factors that might be relevant. Brainstorm with the primary intended users to see whether any additional influencing factors should be considered and, lastly, determine whether there is any evidence to support a role for each of the additional factors.

4.5. Rank the influencing factors
It is now time to rank the influencing factors, both the original and the newly considered ones. Use TOOL 14 to rank each influencing factor by how likely it is to have played a role.

Use a relative ranking system – it is not important to break ties.
STAGE 4
Gather and analyse evidence

OUTPUTS FOR STAGE 4

Upon concluding STAGE 4, you will have analysed your evidence and may have updated your theory of change. The analysis will be documented in evidence tables for:

- the results and the theory of change
- the influencing factors.

CHANGE IS A CONSTANT FACTOR IN A PUBLIC HEALTH CRISIS

In a public health crisis, there are no constants and things change day by day, including the context of the evaluation. This has implications for how to gather data and what data to gather. Even if the evaluation and data collection approach are carefully designed, be ready to adapt them to a rapidly changing context.
STAGE 5
Summarize and assess findings

5.1. Write the contribution claim
5.2. Rate the evidence
5.3. Rate the theory of change
5.4. Test the contribution claim
5.5. Write up the evaluation findings

Objectives
The primary objective for STAGE 5 is to make a contribution claim. In a contribution claim, the primary intended users state whether the intervention could have contributed to the observed results. Once the claim has been stated, test its credibility by seeing whether it stands up to challenges raised during a peer-to-peer review of the evidence.

Steps
This stage involves open and critical discussion of whether your intervention contributed to the outcome, bearing in mind the types of intervention failure that can occur. This stage does not involve assessing to what degree the intervention played a role, but merely to say whether it is possible that it contributed in any way. It is important to bear in mind the intended applications of the evaluation, as well as its overall purpose.

If required, the theory of change should be refined or an alternative model selected. At this point, a wider discussion outside the main evaluation team can take place. Involve individuals who hold opposing stakes in the evaluation and have diverse views because they may have good insights into whether influencing factors might have contributed to the results and whether the assumptions in the theory of change held true.

Tools
In this section you will encounter references to the following tools:
TOOL 3. Evaluation report template
TOOL 15. Evidence rating.
5.1. Write the contribution claim
The contribution claim is typically a one-page document in which you state whether you believe that the intervention made a contribution to the observed results. Explain your decision in clear terms, as this will form part of the final report.

5.2. Rate the evidence
Now, think critically about your evidence and use the following questions to categorize your evidence.
- **Type.** Is the evidence primary or secondary?
- **Claim.** Does the evidence refute or confirm your claim?
- **Contribution type.** Classify the finding by what type of contribution it was supposed to make to the expected results. Was it intended to be a contribution or was it just a condition to an intended contribution? Or was it neither — other contribution?
- **Evidence rating.** Does this piece of evidence have low, medium or high credibility?

Record it in TOOL 15.

5.3. Rate the theory of change
Rate each link in the theory of change based on your conclusions from the evidence rating. Record your responses as follows:
- which links in the theory of change are strong?
- which links in the theory of change are weak?

5.4. Test the contribution claim
Test your contribution claim using the following facilitation questions as a guide:
- Would a reasonable person find the claim credible?
- Do stakeholders agree with the contribution claim?
- Does the pattern of outcomes and the links between outcomes validate the contribution chain?
- Are any of the significant influencing factors likely to have had a noteworthy influence on the results observed?
- What are the main weaknesses in the claim?
- Where would additional data or information be useful?

Record your responses.

5.5. Write up the evaluation findings
Lastly, you are ready to write up the evaluation findings. The report can take many different forms, including a written document or a slide presentation. TOOL 3 includes a report template. At the very least, it should include:
- an executive summary;
- background information on the public health crisis and its response in the country;
- background information on the intervention;
- the aims of the evaluation;
- the evaluation plan;
- the theory of change;
- the methods;
- the findings;
- the limitations and challenges;
- evidence ratings tables;
- the contribution claim; and
- conclusions (summarizing the reasons for rejecting or accepting the contribution claim).
STAGE 5
Summarize and assess findings

OUTPUTS FOR STAGE 5

Upon concluding STAGE 5, you will have:

- completed the summary tables; and
- created a final report that highlights the main findings and how these will be translated into practice.

COMBINE QUANTITATIVE AND QUALITATIVE METHODS

Different methods can support each other by building upon each other and by triangulation of your findings. Triangulation consists of using different methods and sources of data to answer a research question. For example, the same issue or question can be explored from the perspectives of different participants or using different data collection tools: questionnaires, interviews or focus group discussions. Triangulation is important to increase the trustworthiness of your findings or to capture all the nuances of your findings.
STAGE 6
Share and follow up

6.1. Conduct the final team session

6.2. Disseminate the findings

6.3. Follow up

Objectives
The objective of STAGE 6 is to ensure that best practices, challenges and newly developed capacities are shared with the intended users.

Steps
In this stage you gather your team a final time to evaluate the entire process and ensure corrective actions. You plan for dissemination of the findings and schedule a follow-up session.

Tools
In this section you will encounter references to the following tools:

TOOL 16. Dissemination reach
6.1. Conduct the final team session
Conduct a final working group session with the evaluation team to create a sense of ownership and ensure that the corrective actions identified in previous steps are acted upon. This session should include an evaluation of the framework and methodology, including ways to improve its rapid implementation.

6.2. Disseminate the findings
Following this, dissemination of the evaluation findings should continue, using TOOL 16 as a guide. The recommendation is that you measure and record the reach of the dissemination such as number of downloads from a central repository, tracking report requests from colleagues, or attendance at presentations or workshops.

6.3. Follow up
Conclude the entire process by scheduling a follow-up session with the WHO Regional Office for Europe.

OUTPUTS FOR STAGE 6

Upon concluding STAGE 6, you will have:
- completed the evaluation of the process itself
- created a schedule for the dissemination of products
- set a date for a follow-up session
- delivered a report on the reach of the evaluation findings.
1. Indicators of well-being, trust and social cohesion
2. Main evaluation questions
3. Evaluation report template
4. Terms of reference for the evaluation team
5. Team communication plan
6. Summary timeline
7. Evaluation matrix
8. Influence on decision-making
9. Overview of evaluation methods
10. Health equity indicators for rapid socioeconomic pre-assessment
11. Theory of change – visualization
12. Indicator framework
13. Evidence table
14. Influencing factor rating
15. Evidence rating
16. Dissemination reach
17. Exercises
**TOOL 1. Indicators of well-being, trust and social cohesion**

Indicators for each of the factors – well-being, trust and social cohesion – are given below. The questions are drawn from survey questionnaires and can be used as starting points for focus group discussions or interviews.

For questions specifically related to Trust, please randomize the last three questions to mitigate order effects. The answers to all five questions may be averaged but, where present summary statistics and disaggregate where possible. If asked for clarification during the interview, clarify that “most people” in question 1 refers to anyone in the country and in question 2 refers to anyone the respondents know personally, i.e. not just family and friends, but anyone the respondent has a personal relationship with – not anyone they have met only once. Questions 3–5, refer to the institutions themselves – not to their current performance or leadership.

### Indicators of well-being

<table>
<thead>
<tr>
<th>Question</th>
<th>Scale</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happiness</td>
<td>10-point scale, where: 0 = Extremely happy 10 = Extremely unhappy</td>
<td>European Social Survey (21)</td>
</tr>
<tr>
<td>Taking everything together, how happy would you say you are?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life satisfaction</td>
<td>10-point scale, where: 0 = Extremely happy 10 = Extremely unhappy</td>
<td></td>
</tr>
<tr>
<td>All things considered, how satisfied are you with your life?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Indicators of trust

<table>
<thead>
<tr>
<th>Question</th>
<th>Scale</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>How much do you trust most people?</td>
<td>10-point scale, where: 0 = Not at all 10 = Completely</td>
<td>Rosenberg, 1957 (22) New Zealand General Social Survey (23) OECD guidelines on measuring trust (8) World Values Survey (24)</td>
</tr>
<tr>
<td>In general, how much do you trust most people you know personally?</td>
<td>10-point scale, where: 0 = Not at all 10 = Completely</td>
<td></td>
</tr>
</tbody>
</table>
**Indicators of trust contd**

<table>
<thead>
<tr>
<th>Question</th>
<th>Scale</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>The next questions are about whether you have trust in various institutions in [country]. Even if you have had very little or no contact with these institutions, please state your general impression of these institutions. How much you personally trust each of the following institutions: 1. [Country’s] parliament 2. the police 3. the civil service</td>
<td><strong>10-point scale, where:</strong> 0 = Not at all 10 = Completely</td>
<td></td>
</tr>
</tbody>
</table>

Source: OECD guidelines on measuring trust, 2017 (8).

**Indicators of social cohesion**

<table>
<thead>
<tr>
<th>Question</th>
<th>Scale</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regarding people who come from other countries to live in [your country], how would you evaluate the impact of these people on the development of [your country]?</td>
<td><strong>5-point scale, where:</strong> 1 = Very bad 3 = Neither good nor bad 5 = Very good</td>
<td>European Values Study (25)</td>
</tr>
<tr>
<td>To what extent do you feel concerned about the living conditions of: 1. people in your neighbourhood? 2. the people of the region you live in? 3. your fellow countrymen? 4. Europeans? 5. all people all over the world?</td>
<td><strong>5-point scale, where:</strong> 1 = Very much 3 = To a certain extent 5 = Not at all</td>
<td>European Values Study (25)</td>
</tr>
<tr>
<td>Have you done any voluntary work in the last 6 months?</td>
<td><strong>2-point scale, where:</strong> Yes or No</td>
<td>European Social Survey (21)</td>
</tr>
</tbody>
</table>
### Indicators of social cohesion contd

<table>
<thead>
<tr>
<th>Question</th>
<th>Scale</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many close friends do you have, i.e. people who are not your relatives but who you feel at ease with, can talk to about what is on your mind or can call on for help?</td>
<td>Number</td>
<td>Canadian General Social Survey 2020 (26)</td>
</tr>
<tr>
<td>Thinking of [your friend/all of your friends], in the past [] month[s] how often have you: 1. seen [your friend/any of your friends] in person? 2. talked with [your friend/any of your friends] by telephone? 3. communicated with [your friend/any of your friends] by text message? 4. communicated with [your friend/any of your friends] by email or via online social networks?</td>
<td>• Every day  • A few times a week  • Once a week  • 2 or 3 times a month  • Once a month  • Not in the past month</td>
<td>Canadian General Social Survey 2020 (26)</td>
</tr>
</tbody>
</table>

Repeat these questions for your relatives (including those who live in the same household).
TOOL 2. Main evaluation questions

1. What is the contribution or added value of this intervention to health outcomes or behaviours related to protective measures to contain the health crisis?
   - Can anything be said about the effectiveness of the intervention?
   - How sustainable are the effects likely to be?
   - Are there parts of the intervention that could be improved? What corrective actions could or should be taken?
   - Should this intervention be scaled up in its current country and/or be adopted by other Member States in the WHO European Region?
   - Are there any recommendations that could be made for interventions implemented during future health crises?

2. In the target group, what is the contribution or added value of this intervention to factors such as well-being, trust and social cohesion that are considered drivers to a successful health crisis response?
   - Can anything be said about the size and direction (positive or negative) of each of these effects?
   - Can anything be said about the potential sustainability of these effects?
   - Does the disaggregated data reveal any differences in the size and direction of these changes within the target population?
   - Are there corrective actions or recommendations that could be made to increase the positive impact of the intervention on these factors?
   - Are there corrective actions or recommendations that could be made to mitigate the negative impact of the intervention on these factors?

3. Were the effects of the intervention on populations outside the target group (i.e. influencing factors), in particular on groups experiencing disadvantage?
   - Did the intervention contribute to improving well-being, trust and social cohesion in these populations (positive influencing factors)?
     Did this exceed the effect on the target group (proportionate universalism)?
   - Did the intervention contribute to worsening well-being, trust and social cohesion in these populations (negative influencing factors)?
     Was this effect higher (i.e. the intervention did more harm than good) or lower than the effect on the target group?
TOOL 3. Evaluation report template

Title:
Evaluation report for [Name of the intervention]
[Country]
Date evaluation completed: [DD/MM/YY]

This template should be used by the designated report writer (if identified) to document and structure discussions arising from the evaluation and to identify analyses and recommendations. This report should be shared with all evaluation team members for comments before being disseminated to the primary intended users and other stakeholders.

Executive summary
Summarize the key findings from the evaluation report in one or two paragraphs, making sure to include a brief overview of the intervention and its theory of change; a summary of the findings of the evaluation, including gaps, challenges and recommendations; and a timeline for the intervention and evaluation process.

Important public health background information
Provide a brief overview of the public health context relevant to understanding the intervention. For instance, the intervention might have been designed in response to a public health crisis in the country. In this case, the background information section might include a timeline of the crisis; any countrywide interventions that were implemented; the number of affected individuals by key demographic category; known previous or concurrent interventions in the region, city or municipality; and any other relevant influencing factors, such as political factors, socioeconomic factors or media announcements.

Background information on the intervention
Briefly summarize the intervention and its underlying theory of change, including the target population, scope, reach and duration. Describe its theory of change, main causal links and risks and assumptions, as well as influencing factors.

Aims of the evaluation
Summarize the rationale for conducting the evaluation, including the intended applications of the evaluation findings. Describe how the scope of the evaluation fits into the timeline of the intervention, and whether the primary focus is on understanding its effectiveness or sustainability. Summarize how the evaluation findings will be disseminated to stakeholders and other interested parties.

Evaluation plan
Summarize the evaluation plan created in STAGE 1.

Theory of change
Include an illustration of the finalized theory of change that was created during the process of evaluating the intervention.

Methods
Briefly describe the methods, including:
• the format of the evaluation (debrief, working group, interviews with key informants, focus group discussions, surveys or mixed methods);
• participating organizations, cities, regional governments or city councils; and
• a description of background or reference materials, including a bibliography if a literature review was conducted.

Findings
Describe the main discussions conducted during the evaluation for each domain, making sure to highlight recommendations for the design of future interventions.

Limitations and challenges
Mention any gaps in the data collection and evidence review that limited the findings, including any steps that were particularly challenging for the evaluation team.

The evidence ratings tables
Think critically about your evidence and categorize each type of evidence using the following questions.
• Type of data. Is the evidence primary or secondary?
Claim. Does the evidence refute or confirm your claim?

Contribution type. Classify the findings by the type of contribution it was supposed to make to the expected results. Was it intended to be a contribution or just a condition to an intended contribution? Or was it neither (i.e., other contribution)?

Evidence rating. Does this piece of evidence have low, medium, or high credibility?

The contribution claim
The contribution claim is typically a one-page document describing how the intervention has contributed to the observed results. Consider the following questions.

- Would a reasonable person find the claim credible?
- Do stakeholders agree with the contribution claim?
- Does the pattern of outcomes and the links between them validate the contribution chain?
- Are any of the significant influencing factors likely to have had a notable influence on the observed results?
- What are the main weaknesses in the claim?
- Where would additional data or information be useful?

Conclusions
Summarize how the recommendations for improving the planning and design of the intervention could be implemented. Suggest improvements for conducting the evaluation and how they could best be incorporated into intervention planning. Suggest how recommendations arising from the evaluation could best be implemented and tracked to provide accountability.
**TOOL 4. Terms of reference for the evaluation team**

The evaluation team comprises the members necessary to fulfil the scope, aims and format of the planned evaluation. Note that, depending on the size and scope of the evaluation, not all of the roles described below will be required.

Table entries are given as examples only.

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
<th>Overall responsibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall evaluation lead</strong></td>
<td>A staff member from the Ministry of Health (or other relevant ministry) who initiates an evaluation through contact with WHO</td>
<td>• Develops the scope and aims of the evaluation&lt;br&gt;• Identifies and invites team members, including facilitators, interviewers (ideally third party for larger evaluations) and a report writer&lt;br&gt;• Organizes and leads the evaluation team&lt;br&gt;• Ensures that senior management provides coordination and support&lt;br&gt;• Collects background materials and disseminates these to the rest of the team&lt;br&gt;• Supervises the evaluation sessions, including the development of facilitation questions or interview schedules, as per the evaluation aims&lt;br&gt;• Compiles the final evaluation report and shares this with intended users and the WHO Regional Office for Europe</td>
</tr>
<tr>
<td><strong>Lead facilitator/interviewer</strong></td>
<td>Leads the overall facilitation of an evaluation in a focus group format or plans the schedule for one-to-one interviews</td>
<td>• Supports the overall evaluation lead to define the evaluation aims, scope, format and participants&lt;br&gt;• Develops facilitation questions or the interview schedule as per the evaluation aims and format&lt;br&gt;• Briefs and supports other facilitators and/or interviewers&lt;br&gt;• Conducts debriefings with facilitators and/or interviewers&lt;br&gt;• Coordinates the writing of the final report</td>
</tr>
<tr>
<td>Title</td>
<td>Description</td>
<td>Overall responsibilities</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Facilitators/ interviewers (may be more than one, depending on the scope of the evaluation)</td>
<td>Supports the lead facilitator/interviewer in guiding the discussion around the themes set out in facilitation guide or interview schedule, and helps to keep the discussion on track. Is impartial and not directly involved in the intervention, e.g., staff member from another department or ministry. May assist in managing interpersonal conflicts by maintaining a focus on the underlying key themes. Key competencies: experience in facilitating group discussions; sensitive to country and participant contexts; and speaks language of participants</td>
<td></td>
</tr>
</tbody>
</table>
|  |  | • Introduces the evaluation and maintains the discussion structure by clarifying issues and assumptions  
|  |  | • Encourages contributions from all participants while maintaining impartiality  
|  |  | • Summarizes the discussion points  
|  |  | • Contributes to writing the final report |
| Note takers | Key competencies: some familiarity with the intervention, as well as the country’s organizational structures |  
|  |  | • Ensures that comments and discussions from focus group(s) are recorded and documented  
|  |  | • If interviews are not recorded for transcription, note takers also ensure that the responses to questions are also documented |

Source: adapted from WHO, 2020 (27).
**TOOL 5. Team communication plan**

A communication plan helps to determine and maintain effective flows of information. The communication plan should cover stakeholder groups, communication goals, and the content and frequency of communication.

Table entries are given as examples only.

<table>
<thead>
<tr>
<th>Priority audience</th>
<th>Goals</th>
<th>What will be shared</th>
<th>Timetable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country office</td>
<td>To stay informed on the lessons learned from reflective cycles</td>
<td>Targeted notes</td>
<td>Weekly</td>
</tr>
<tr>
<td>Ministry of Health</td>
<td>To inform of progress and discuss challenges</td>
<td>Minutes from virtual meetings</td>
<td>Monthly</td>
</tr>
<tr>
<td>Unit colleagues</td>
<td>To inform colleagues not directly involved to provide opportunities for creating synergy</td>
<td></td>
<td>Monthly</td>
</tr>
</tbody>
</table>

*Source: adapted from Centers for Disease Control and Prevention, 2011 (28).*
**TOOL 6. Summary timeline**

It is essential that the findings be completed within a time frame that best aligns with the decision-making requirements. Complete the table with the summary timeline and key evaluation milestones.

Table entries are given as examples only.

<table>
<thead>
<tr>
<th>Main phases</th>
<th>Timeline</th>
<th>Tasks and deliverables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation</td>
<td>February–March 2021</td>
<td>Draft and final terms of reference&lt;br&gt;Evaluation team formed</td>
</tr>
<tr>
<td>Inception</td>
<td>February–March 2021</td>
<td>Desk review of the existing literature&lt;br&gt;Document review of intervention/ministry briefings&lt;br&gt;Draft and final inception note</td>
</tr>
<tr>
<td>Data collection and analysis</td>
<td>March 2021</td>
<td>Key interviews with intervention staff&lt;br&gt;Other evaluation methods (survey, focus groups, interviews)&lt;br&gt;Data analysis and review</td>
</tr>
<tr>
<td>Reporting</td>
<td>March 2021</td>
<td>Draft and final evaluation report</td>
</tr>
<tr>
<td>Management response and dissemination</td>
<td>March 2021</td>
<td>Management response to the evaluation recommendations</td>
</tr>
</tbody>
</table>
TOOL 7. Evaluation matrix

An evaluation matrix is an organizing tool to help to plan for the data collection, analysis and report writing phases. Use the evaluation matrix to guide the analysis and ensures that all collected data are analysed and used to answer the evaluation questions.

The below table includes some background information which can guide users in completing their own evaluation matrix.

<table>
<thead>
<tr>
<th>Evaluation questions and subquestions</th>
<th>Information sources</th>
<th>Evaluation methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>National population health surveys, analyses and studies</td>
<td>Focus groups</td>
</tr>
<tr>
<td></td>
<td>Relevant behavioural insights surveys that may have been conducted</td>
<td>Document review</td>
</tr>
<tr>
<td></td>
<td>Data from public databases</td>
<td>Surveys</td>
</tr>
<tr>
<td></td>
<td>Observational data/reports on lifestyle</td>
<td>Key informant interviews with implementation staff</td>
</tr>
</tbody>
</table>
TOOL 8. Influence on decision-making

The following facilitation questions can be used to challenge intended users to think about the intended influence of their evaluation on decision-making.16

- What decision, if any, are the evaluation findings expected to influence?
- How might the evaluation findings support future improvements to the intervention?
- When will decisions be made and by whom?
- When should the evaluation findings be presented to be timely and influential?
- What is at stake in the decisions and for whom?
- What controversies or issues surround the decisions?
- What is the history and context of the decision-making process?
- What other factors (e.g. values, politics and personalities) will affect the decision-making?
- How changeable is the decision-making environment?
- How much influence do you realistically expect the evaluation to have?
- What needs to be done to achieve a high level of influence?
- To what extent has the outcome of the decision already been determined?
- Who needs to be involved for the evaluation to have a high level of influence?
- How will you know afterwards if the evaluation was used as intended?
### TOOL 9. Overview of evaluation methods

The table provides a short list of methods for data collection, along with examples of when these methods are appropriate to use. In fact, a much wider range of tools can be used for process evaluation, including participant attendance sheets, questionnaires, observations and mapping. Similarly, there exists a wider range of tools for causal impact evaluation, such as field trials and quasi-experimental designs, which are not included here.

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>When to use</th>
</tr>
</thead>
</table>
| Focus groups/ facilitated workshops | • Moderated and facilitated conversations with a group of people from the same target group that are used to gain insight into their knowledge, perceptions, beliefs, attitudes and experiences about a certain topic  
• The facilitator follows a topic guide or interview schedule that has been created with the aims and outcomes of the evaluation design in mind  
• Can take different forms ranging from unstructured to structured, depending on the aims of the evaluation team and the design of the evaluation plan  
• Participants may be encouraged by the presence of others to share and exchange opinions and concerns, including myths, rumours or stories that may be circulating in a community  
• About 5–12 participants are usually involved in a focus group discussion | • These methods are useful for identifying social norms, and can reveal areas of both agreement and differences of opinion about a specific topic  
• Use when access is available to a skilled facilitator in that language  
• Suitable when the evaluator is interested in the interactions between people |
### Type

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>When to use</th>
</tr>
</thead>
</table>
| Individual in-depth interviews | - These are one-to-one conversations conducted in person, by phone or virtually in which the interviewer follows an interview schedule that was created to meet the aims and outcomes of the evaluation plan  
- Responses are recorded by the interviewer for later analysis and reporting                                                                 | - Useful when the target population is small (e.g. under 50) or when it is likely that a survey questionnaire will not be completed  
- Useful when the depth, rather than breadth, of information is required  
- Suitable when the evaluator is interested in subjective experiences  
- Useful when the views of key people are needed or when the participant has special knowledge or a unique point of view  
- Useful where the topic is sensitive and the participant may not feel comfortable speaking openly in a group  
- May also be used when it is difficult to bring a larger group together                                                                 |
| Surveys                     | - Data are collected through standardized questionnaires with predefined, probably closed-ended, questions  
- Survey participants can be invited to complete the questions on paper, electronically or by phone                                                                 | - Useful for gathering data on knowledge, perceptions, well-being and other barriers and drivers to behaviour  
- Best used when information is to be collected quickly and fairly easily from a large number of people (e.g. over 50)                                                                                                                                                                                                                                                  |
| Document review             | - The evaluation team reviews organizational documents to identify relevant information about the intervention and its underlying theory of change  
- Relevant documents might be in paper or electronic format  
- The research team extracts relevant information and summarizes the findings in an organized fashion for analysis and discussion by the larger evaluation team as a whole                                                                 | - Used when a historical perspective on the issue, policy or programme is needed, especially for those members of the evaluation team who are unfamiliar with the organization  
- Particularly useful when hard data is required on a certain aspect of the organization (i.e. the intervention)  
- Used only when relevant documents exist and are accessible to the evaluation team                                                                                                                                                                                                                                                     |
### TOOL 10. Health equity indicators for rapid socioeconomic pre-assessment

The table aligns the United Nations pillars for the socioeconomic impact of COVID-19 (14) and includes the impact of the main containment measures for a public health crisis and their related indicators. Where possible, sources for baseline data at national level are shown; disaggregated data at the regional and subregional levels may be available, depending on the country. This table is based on Annexes 3 and 5 of Health inequity and the effects of COVID-19; assessing, responding to and mitigating the socioeconomic impact on health to build a better future (29).

<table>
<thead>
<tr>
<th>United Nations pillar</th>
<th>Impact of containment measures for a public health crisis</th>
<th>Indicator</th>
<th>Baseline data source</th>
<th>Data obtained Yes or No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health first</td>
<td>–</td>
<td>Excess mortality</td>
<td>National statistics offices</td>
<td>☐ ☐</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>Self-reported health</td>
<td>European Social Survey, EU-SILC, HBSC, HED, WVS</td>
<td>☐ ☐</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>Rate of noncommunicable diseases</td>
<td>EHIS, GHO, HED, HFA, STEPS</td>
<td>☐ ☐</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>Avoidable hospitalizations</td>
<td>Administrative health data, Eurostat, HED</td>
<td>☐ ☐</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>Unmet need for health care</td>
<td>EU-SILC, HED, WVS</td>
<td>☐ ☐</td>
</tr>
<tr>
<td>Increased pressure on workers in the formal and informal sectors, including front-line health workers</td>
<td>Quality of health-care services</td>
<td>European Quality of Life Survey, HED</td>
<td>☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Reduced support for informal care</td>
<td>Informal caregiving</td>
<td>European Quality of Life Survey, HED</td>
<td>☐ ☐</td>
<td></td>
</tr>
<tr>
<td>United Nations pillar</td>
<td>Impact of containment measures for a public health crisis</td>
<td>Indicator</td>
<td>Baseline data source</td>
<td>Data obtained Yes or No</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Protecting people</td>
<td>Loss of income and increased poverty rates</td>
<td>Poverty, child poverty and in-work poverty rates</td>
<td>EU-SILC, HED</td>
<td>Yes</td>
</tr>
<tr>
<td>Reduction in life chances due to closure of early years facilities</td>
<td>Early years outcomes</td>
<td>National statistics offices</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Children locked out of learning</td>
<td>Educational performance</td>
<td>HED, PISA (OECD)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Young people not in education, employment or training</td>
<td>HED, ILO</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Housing deprivation</td>
<td>EU-SILC, HED</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Disruption/interruption of public transportation</td>
<td>Food insecurity</td>
<td>EU-SILC, European Quality of Life Survey, HED, WVS</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Fuel poverty</td>
<td>EU-SILC, HED</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Domestic violence and abuse</td>
<td>Feeling unsafe from crime or violence in the home</td>
<td>European Quality of Life Survey, HED, WVS</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Household debt</td>
<td>National survey data, OECD</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Acute insecurity of those already vulnerable</td>
<td>Unaffordable loans, household debt-to-income ratios</td>
<td>National statistics offices</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>United Nations pillar</td>
<td>Impact of containment measures for a public health crisis</td>
<td>Indicator</td>
<td>Baseline data source</td>
<td>Data obtained Yes or No</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------------------------------------------</td>
<td>-----------</td>
<td>----------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Protecting people</td>
<td>Disruption/interruption of water, sanitation and hygiene services</td>
<td>Adequate water and sanitation facilities</td>
<td>HED, Joint Monitoring Programme</td>
<td>□ □</td>
</tr>
<tr>
<td></td>
<td>Reduction in the substantiation of human rights</td>
<td>Monitor of human rights abuses</td>
<td>Fragile States Index – Human Rights Dimension</td>
<td>□ □</td>
</tr>
<tr>
<td></td>
<td>Closure of borders and/or reduction in consular services</td>
<td>Travel advisories and consular processing notifications</td>
<td>Embassies and consulates</td>
<td>□ □</td>
</tr>
<tr>
<td>Economic response and recovery</td>
<td>Loss of employment and work hours/opportunities</td>
<td>Unemployment rate</td>
<td>Eurostat, HED, ILO</td>
<td>□ □</td>
</tr>
<tr>
<td></td>
<td>Loss of employment and work hours/opportunities</td>
<td>Informal and part-time workers</td>
<td>Eurostat, HED, ILO</td>
<td>□ □</td>
</tr>
<tr>
<td></td>
<td>Risk of increased exposure to indoor air pollution</td>
<td>Indoor ambient air pollution</td>
<td>European Commission</td>
<td>□ □</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Incidence, coverage and adequacy of social assistance programmes</td>
<td>World Bank Atlas of Social Protection</td>
<td>□ □</td>
</tr>
<tr>
<td>Social cohesion and community resilience</td>
<td>-</td>
<td>Mental health (WHO 5-point scale)</td>
<td>European Quality of Life Survey</td>
<td>□ □</td>
</tr>
<tr>
<td></td>
<td>Social isolation</td>
<td>Suicides</td>
<td>WHO GBD, HFA</td>
<td>□ □</td>
</tr>
</tbody>
</table>
### United Nations pillar

Social cohesion and community resilience

<table>
<thead>
<tr>
<th>Impact of containment measures for a public health crisis</th>
<th>Indicator</th>
<th>Baseline data source</th>
<th>Data obtained Yes or No</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>Volunteering</td>
<td>EU-SILC, HED</td>
<td>□ □</td>
</tr>
<tr>
<td>Closure of public spaces and parks</td>
<td>Access to green space</td>
<td>European Quality of Life Survey, HED, OECD</td>
<td>□ □</td>
</tr>
<tr>
<td>-</td>
<td>Trust in others</td>
<td>European Quality of Life Survey, European Social Survey, HED, WVS</td>
<td>□ □</td>
</tr>
<tr>
<td>-</td>
<td>Not having someone to ask for help</td>
<td>EU-SILC, HED, WVS</td>
<td>□ □</td>
</tr>
<tr>
<td>Criminal exploitation, including organized crime</td>
<td>Rates of Internet crime</td>
<td>National statistics offices</td>
<td>□ □</td>
</tr>
<tr>
<td>Financial exploitations and unfair price inflation</td>
<td>Media stories and fake news monitors</td>
<td>University departments, media organizations</td>
<td>□ □</td>
</tr>
<tr>
<td>Fake or misleading news</td>
<td>Equal treatment under the law and absence of discrimination</td>
<td>HED, World Justice Project</td>
<td>□ □</td>
</tr>
<tr>
<td>Restrictions on a free and independent media</td>
<td>World Press Freedom Index</td>
<td>Reporters Without Borders</td>
<td>□ □</td>
</tr>
</tbody>
</table>

**TOOL 11. Theory of change visualization**

<table>
<thead>
<tr>
<th>Problem</th>
<th>Barriers</th>
<th>Divers (strategies)</th>
<th>Interventions</th>
<th>Output</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>The behaviour we wish to change</td>
<td>Various barriers identified – using various data collection methods and guided by a theoretical model</td>
<td>Opportunities to drive change – identified through the same methods as for the barriers</td>
<td>Interventions that address barriers and utilize drivers</td>
<td>The change in what they know, feel and experience and in their environments</td>
<td>The change in behaviours and overall effect on reaching health goals</td>
</tr>
<tr>
<td></td>
<td>• What they know</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• What they feel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Their experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The social and cultural environment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The structural environment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**EXAMPLE**

**Problem:** More women aged 50+ to attend breast cancer screenings.

**Barriers:** They have the knowledge, but they find it inconvenient, and they think it is scary and uncomfortable, and they prefer women health workers.

**Drivers:** They feel strongly about taking care of their families.

**Interventions:** Easy booking, drop-in screening; option to select the sex of the health worker. Invitation letter focuses on messages re doing this for someone else (not for themselves).

**Messages tested:** Fewer women perceive this as inconvenient; more women are aware of option to select sex of HW; more women perceive this as something they do for their families.

**Output:** More women get screened. More cases of breast cancer are detected at an early stage, leading to a lower number of deaths.
### TOOL 12. Indicator framework

Complete the table with indicators, data, sources, methods and ratings. Table entries are given as examples only.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Indicators</th>
<th>Baseline data (yes/no/source)</th>
<th>Information source</th>
<th>Evaluation method</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protective behaviours</td>
<td>Self-efficacy</td>
<td>No</td>
<td>Focus group discussions</td>
<td>Qualitative</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>Perceived risks</td>
<td>No</td>
<td>Online questionnaire</td>
<td>Qualitative and quantitative</td>
<td>Medium</td>
</tr>
<tr>
<td>Well-being</td>
<td>Life satisfaction</td>
<td>Yes</td>
<td>European Social Survey (21)</td>
<td>Quantitative</td>
<td>High</td>
</tr>
<tr>
<td>Intervention participant – well-being</td>
<td>No</td>
<td>Participant survey (see Tool 1)</td>
<td>Quantitative</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>Intervention participant – well-being</td>
<td>No</td>
<td>Focus group discussion</td>
<td>Qualitative</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>Trust</td>
<td>Interpersonal trust</td>
<td>Yes</td>
<td>World Values Survey</td>
<td>Quantitative</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Trust in government</td>
<td>Yes</td>
<td>OECD - Eurostat</td>
<td>Quantitative</td>
<td>High</td>
</tr>
<tr>
<td>Intervention participant – trust</td>
<td>No</td>
<td>Participant survey (see Tool 1)</td>
<td>Quantitative</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>Social cohesion</td>
<td>Intervention participant – social connectedness</td>
<td>No</td>
<td>Participant survey (see Tool 1)</td>
<td>Quantitative</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>Intervention participant – social connectedness</td>
<td>No</td>
<td>Focus group discussion</td>
<td>Quantitative</td>
<td>Medium</td>
</tr>
</tbody>
</table>
TOOL 13. Evidence table

Upon concluding STAGE 4, you will have analysed your evidence and may have updated your theory of change. The analysis will be documented in evidence tables for:
- the results and the theory of change
- the influencing factors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
<th>Evidence</th>
<th>Questions</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Findings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact indicators</td>
<td></td>
<td>Conclusions</td>
<td></td>
<td>Answers</td>
</tr>
<tr>
<td>Process indicators</td>
<td></td>
<td>Score</td>
<td></td>
<td>Contestedness</td>
</tr>
<tr>
<td>Assumptions</td>
<td></td>
<td></td>
<td>Are the assumptions valid? Can you find information about these assumptions in the literature?</td>
<td></td>
</tr>
<tr>
<td>Risks</td>
<td></td>
<td></td>
<td>Did the risks occur as predicted? If so, what result did they have?</td>
<td></td>
</tr>
</tbody>
</table>
## TOOL 14. Influencing factor rating

Rank the influencing factors, both the original and the newly considered ones. Rank each influencing factor by how likely it is to have played a role. Use a relative ranking system. Influencing factors can have the same relative rank – it is not important to break ties.

Table entries are given as examples only.

<table>
<thead>
<tr>
<th>Influencing factors</th>
<th>Hypothesized role</th>
<th>Evidence for or against this role</th>
<th>Relative rank of influence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nationwide intervention(s) conducted at the same time targeting similar behaviours</td>
<td>The effects of the intervention being evaluated could have resulted from other interventions that were being carried out simultaneously at a national level</td>
<td>Evaluation reports of national interventions describing an effect size (if available)</td>
<td>Medium</td>
</tr>
<tr>
<td>Background information from print, television, or social media providing conflicting information</td>
<td>Behaviours of participants could be influenced by information that cannot be isolated from the intervention itself</td>
<td>Focus group discussions can provide some self-reflexive evidence on this question</td>
<td>Medium</td>
</tr>
<tr>
<td>Importance of individuals associated with the intervention</td>
<td>The effect size of the intervention may be down to key personalities involved, which might make the intervention difficult to replicate</td>
<td>Focus group discussions can provide insights into the outsized impact of any individuals</td>
<td>High</td>
</tr>
<tr>
<td>Novelty of the intervention</td>
<td>Because the intervention is new or unusual, it is producing an exaggerated impact, which might not be replicable in settings where the intervention is not new or unusual</td>
<td>Literature on effect sizes of similar interventions in other settings</td>
<td>Low</td>
</tr>
</tbody>
</table>
**TOOL 15. Evidence rating**

Table entries are given as examples only. Please use this table as a template to rate your own evidence.

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Type</th>
<th>Claim</th>
<th>Contribution type</th>
<th>Evidence rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member State intervention with same target, magnitude and reported success</td>
<td>Secondary</td>
<td>Refuting</td>
<td>Other contribution</td>
<td>Medium</td>
</tr>
<tr>
<td>Interviewers cast doubt on the sustainability of the results</td>
<td>Primary</td>
<td>Refuting</td>
<td>Condition to intended contribution</td>
<td>Low</td>
</tr>
<tr>
<td>Focus group sessions with intended audience consistently confirm the impact of the intervention</td>
<td>Primary</td>
<td>Confirming</td>
<td>Intended contribution</td>
<td>High</td>
</tr>
</tbody>
</table>
**TOOL 16. Dissemination reach**

It is important to disseminate the evaluation findings to the priority audiences and key stakeholders identified in TOOL 5. The recommendation is that you measure and record the reach of the dissemination such as number of downloads from a central repository, tracking report requests from colleagues, or attendance at presentations or workshops.

Table entries are given as examples only.

<table>
<thead>
<tr>
<th>Product</th>
<th>Target date</th>
<th>Audience</th>
<th>Lead contributors</th>
<th>Status</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully designed evaluation report</td>
<td>15 April 2021</td>
<td>Ministry of Health; country office; unit colleagues</td>
<td>Tania Aase Dræbel</td>
<td>Complete</td>
<td></td>
</tr>
<tr>
<td>PowerPoint slide deck</td>
<td>23 April 2021</td>
<td>Ministry of Health; country office; unit colleagues</td>
<td>Tania Aase Dræbel</td>
<td>In progress</td>
<td></td>
</tr>
<tr>
<td>Press release or news item</td>
<td>10 May 2021</td>
<td>Interested public; academia; policy makers; public health professionals</td>
<td>Nils Fietje</td>
<td>Not yet started</td>
<td>Waiting for launch date</td>
</tr>
<tr>
<td>Webinar</td>
<td>20 May 2021</td>
<td>Policy makers; public health professionals and</td>
<td>Nils Fietje</td>
<td>In progress</td>
<td></td>
</tr>
</tbody>
</table>
**Exercise 1.**

**Conduct an evaluability assessment**

Evaluability refers to the extent to which an intervention can be evaluated reliably and credibly. By conducting an assessment of evaluability, it is possible to determine whether an evaluation should go ahead or should be delayed, or whether the intervention cannot be evaluated. This is accomplished by assessing whether the intervention design is adequate for what is seeks to achieve and whether the information to be used in the evaluation is available and of sufficient quality. In effect, it is a dry run of the evaluation.

Why spend time on this step, rather than launching into the full evaluation? Doing so will save time in the long run, regardless of whether the evaluation proceeds. If the assessment finds that the intervention design was inadequate or that there is insufficient, good-quality evidence, time will be saved by not proceeding to a full evaluation that would ultimately arrive at the same conclusion. On the other hand, if it is deemed useful to proceed to the full evaluation, then the assessment serves as a highly effective warm-up. It ensures that the team is well functioning, that there is consensus on the intended applications, and that the information required to conduct the evaluation can be obtained and analysed within the given time frame.

An assessment of evaluability can be determined through a simulation exercise using fictitious findings (28). These findings are quasi-real, meaning that they are close enough to the real expected outcomes that they (i) provide a credible learning experience for the primary intended users, (ii) provide sufficient information for an explicit decision to be made regarding the value of proceeding with the full evaluation and (iii) ensure that the primary intended users have ownership of the design and measures.

**Brainstorm findings**

**Make up results** that would be expected to be brought about by the intervention, using the information sources and evaluation methods set out in STAGE 1 of the evaluation plan. Make sure that these findings reflect the real-world setting by incorporating sufficient variability in the results. Incorporating variability should prompt discussion among the primary intended users.

**Interpret the fictitious findings**

**Interpret the findings.** Again, it is not necessary to reach consensus. Rather, it is important that the primary intended users feel comfortable voicing their opinions, while taking the simulation exercise seriously.

**Interpret the simulation experience**

**Reflect on what has been learned** with a view to the intended applications of the evaluation and its purpose.

1. **Intervention design**
   a. Does it make sense that the intervention could produce such results?
   b. Does it make sense that the intervention can be assessed as detailed in the evaluation plan?

2. **Information sources**
   a. Could it be useful to add additional sources of information? If so, which ones
   b. Were any information sources not as useful as the others?
   c. Has it become apparent that some evaluation methods are unfeasible, given time and financial constraints?

3. **Intended users and intended uses**
   a. Do the primary intended users have a better grasp of what a full evaluation will entail?
   b. Is it becoming more obvious that all primary intended users must play a role in the analysis of the evaluation findings?
**Make an explicit decision**

1. **Ask whether it is useful and cost-effective to proceed with the evaluation.**
   a. Are there modifications that could be made to the evaluation plan or intervention design?
   b. Would the findings provide sufficient information to meet the needs of the intended users of the evaluation?

2. **Come to consensus,** being sure not to assign blame if the decision is made not to proceed. Recognise the tendency to continue an endeavour once an investment in money, effort, or time has been made: it may be necessary to cut one’s losses in order not to “throw good money after bad”.
   a. Is it worth proceeding with the full evaluation?

---

**Record the results of the simulation exercise in this table.**

<table>
<thead>
<tr>
<th>Information source</th>
<th>Fabricated findings</th>
<th>Interpretation of these findings</th>
<th>Reflections from the exercise on</th>
<th>Decision to proceed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intervention design</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Information sources</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intended uses</td>
<td></td>
</tr>
</tbody>
</table>
Exercise 2.

Understanding causality
This exercise should be conducted as a group following STAGE 1 and the design of the evaluation plan. To save time, team members should complete this stage at home, and the results summarized in a table and discussed later as a group. Recognize that there is a trade-off between overly broad questions that are difficult to answer and overly precise questions that do not reflect the overall purpose of the evaluation. Keep track of the questions used and answers generated in a table (see the example below).

Think about attribution
Recall that attribution is central to how people understand the world: it explains why things occur and predicts the ability of the intervention to achieve the desired behavioural change. Attribution is necessary to determine whether the observed results were the direct results of the intervention.

Recognize that attribution is not the focus of the evaluation, but that asking attribution-type questions helps to build an understanding of how contribution is a very different concept from attribution. Sample questions are as follows.

1. What would convince you that the intervention has made a contribution that can be observed?
   a. Are there some types of evidence that you would find more convincing than others? Why?
      1. If these types of evidence were combined, would you find this more convincing?
   b. What kind of results do you hope that the intervention can bring about?
      1. Do you think this effect will be large enough to be observed?

2. If the intervention has already been completed, do you think that the intervention has made a difference?
   a. If yes, is this because the intervention:  
      1. influenced the observed result? or  
      2. made an important contribution to the result?

3. If the intervention has not yet been implemented, do you think that the intervention could make a contribution given:
   a. the type of intervention?
   b. the delivery of the intervention?
   c. the size and reach of the intervention?
   d. what the intervention seeks to achieve?

4. Can you think of some factors that might influence the results?
   a. Do you think these factors could have positively influenced the results?
   b. Do you think these factors could have negatively influenced the results?
Record the output of the take-home exercise in this table.

<table>
<thead>
<tr>
<th>Type of question</th>
<th>Evaluation questions used</th>
<th>Answers generated to these questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contribution</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Theoretical background

This document presents an evaluation framework that can be applied to behaviourally and culturally informed interventions in complex settings, either during the planning stage of the intervention or after implementation has been completed. It draws on evaluation models from the field of causal analysis that enable evaluators to map backwards map from results, if any, in order to explore causation.

The method of contribution analysis has been used to critically examine the cause–effect relationship between the intervention, its expected results and ongoing influencing factors that are outside experimental control.

A transparent methodology
An evaluation framework provides a rigorous and transparent methodology for evaluating an intervention of a similar type. It also provides:
- a shared language among stakeholders;
- a key set of main evaluation questions;
- logic models that help assess whether an intervention was successful in reaching its targets;
- documentation of the impact and cost–effectiveness of the intervention, which can lead to better programme development; and
- good causal reasoning, which helps to deliver research and policies that are rigorous and evidence based.

This framework was developed to provide a rapid but rigorous methodology to address the main evaluation questions while also addressing the objectives of accountability, learning and systematic social enquiry. This is done through a literature review and is guided by the WHO evaluation practice handbook (30), in accordance with the United Nations Evaluation Group’s Norms and standards for evaluation (20) and Ethical guidelines for evaluation (31).

Here, impartiality is achieved by drawing on a cross-section of information sources from various stakeholder groups and by using a mixed methods approach to ensure triangulation of information and decrease bias. Adherence to the pillars of a UN framework for the immediate socio-economic response to COVID-19, (14) which emphasise gender, equity and human rights, is addressed by first understanding the country situation and then assessing any positive or negative impacts that the intervention might have incurred on vulnerable or newly vulnerable groups, as well as on the target population.

Understanding evaluation
How does one determine whether any change occurred following an intervention and, if so, whether it resulted from the intervention and not because of influencing factors or changing contexts? Addressing such questions lies at the heart of the field of evaluation (30,32). Formal evaluations, such as the one detailed in this framework, aim to track results across Member States and generate knowledge. They also seek to provide accountability to governments and the public, as well as to donor partners, researchers and colleagues within research and development communities. Formalizing an evaluation practice helps to promote a culture of evaluation in everyday practice, in which evidence is used to test assumptions and to sustain a focus on results and understanding what works. The WHO Regional Office for Europe
works closely with the Member States to evaluate projects and programmes and to develop and refine methodologies to meet the complex needs of middle-income countries (Box 5).

Different types of evaluations can serve different purposes.

- **Formative evaluations** focus on improvement. They are open-ended, with a focus on gathering data about the strengths and weaknesses of the intervention.

- **Developmental evaluations** focus on providing rapid input throughout the evaluation process. They seek to capture concurrent feedback that can be used to improve implementation of the intervention in emergent and complex environments.

- **Summative evaluations** are usually conducted after an intervention has concluded. They are intended to judge its overall effectiveness, sustainability or worth; therefore, they are used to determine whether a particular intervention should be scaled up and replicated in other Member States.

**Methodology**

*Causality is hard to determine in a complex setting*

Although randomized controlled trials are considered the gold standard for evaluation in the field of public health, determining causality in real-world settings may not always be possible. Interventions and policies implemented in response to a complex public health challenge, particularly during an acute crisis such as a pandemic, are usually best evaluated as complex initiatives. In such evaluations, understanding the possible relationship between the intervention and the observed results relies on investigating and describing the intervening steps that includes the use of *generative causality* to make causal inferences (32,34,35).

If an intervention includes one or more of the following factors, it is considered to be complex (4,36–38):

- It comprises a number of different intervention components or types of intervention activities and strategies.
- It involves a number of different actors, such as partners and intermediaries.
- It includes a number of different pathways to impact.
- It is influenced by a number of influencing factors.
- The desired results are influenced by ongoing social or economic trends.
**Contribution analysis works well in complex settings**
Determining causality is a complex process. Fifth-generation evaluation methods have shifted from focusing on quantitative efforts (37) to assessing attribution (i.e., did the intervention cause the expected results to occur) to largely qualitative efforts to assess contribution (i.e., did the intervention play a role in bringing about the observed results) (33). Due to the changing nature of a public health challenge, particularly during crisis (such as a pandemic), establishing credible counterfactuals (what might have happened had the intervention not taken place) is a complex process. Therefore, this evaluation framework asserts that contribution analysis, which instead iteratively explores why and how an intervention has influenced change, may be a more appropriate, field-tested technique.

**Work backwards from the observed consequences**
Contribution analysis has increasingly been applied to large intervention portfolios (35). The goal of contribution analysis is to determine what and who the intervention has worked for, as well as to understand why and under which circumstances it worked (or did not). By seeing evaluation as a learning tool, insights can be gained at each step of the contribution analysis, even if not all steps are completed. Complex and changing settings are best evaluated adaptively, with expectations revised as new information comes to light. By working backwards from the observed consequences and learning from these, it is possible to better understand under which conditions the intervention might effect positive change and whether value has been added through delivery of the intervention and its subsequent evaluation.

**The theory of change articulates the underlying assumptions**
A theory of change describes a set of assumptions that underlies both the steps leading to the goal of intervention and the connections between these assumptions and the results achieved at each step of the process (39). This methodology is based on the belief that the inherent difficulty in evaluating interventions lies not in their complexity but rather in poor articulation of the assumptions underlying their design and implementation.

**Three types of intervention failures**
A theory of change provides an overview of the logic behind the intervention that spells out the challenge, the intervention and how the intervention addresses the challenge. Hence, better articulation of the theory of change should ensure that the designers of interventions in complex settings specify the underlying logic and include important factors that might also play a role in the generation of the final outcomes (i.e., influencing factors). Using a theory of change for both the design and evaluation of an intervention allows comparison between the predicted and actual outcomes of the intervention, enables better attribution of the outcomes to the intervention, and consideration of the lessons learned. Overall, conducting an evaluation based on a theory of change (as is contribution analysis) permits the evaluation team to determine whether an intervention did not work by distinguishing the three primary types of intervention failure (36):

- **theory of change failure** – implementing activities that have no effect or the wrong effect on a problem;
- **implementation failure** – carrying out the intervention ineffectively or not at all; and
- **evaluation failure** – using a flawed set of evaluation procedures.

**A collaborative effort**
Describing a theory of change is a collaborative effort that requires consultation with a range of stakeholders. Theories of change can be narrative or quantitative, linear or nonlinear, and can provide an overview or be detailed with explicit causal link assumptions. For complex interventions that target different groups or geographical areas, theories of change are typically nested. This framework asserts that contribution analysis, which instead iteratively explores why and how an intervention has influenced change, may be a more appropriate, field-tested technique.

**Use COM-B or similar robust theory of change**
As developing a “good” theory of change is expected to differ substantially based on the type of intervention, different theory of change models have emerged in different disciplines. In the field
of behavioural change interventions, the COM-B system asserts that behaviour is changed through the interaction of capabilities, opportunities and motivations (18). The COM-B theory of change has been increasingly adopted due to its implicit intuitiveness and basis in empirical evidence. The policy document, Pandemic fatigue – reinvigorating the public to prevent COVID-19. Policy framework for supporting pandemic prevention and management, (19) recommends selecting COM-B (18) or a similarly robust theory of change when designing and implementing an intervention.

A robust theory of change is structurally sound, is plausible and provides a reasonable expectation that the intervention that it underpins will produce the expected results, as long as the intervention is delivered as planned (see Box 6) (40).

**BOX 6. CRITERIA FOR A ROBUST THEORY OF CHANGE**

**Structural soundness**
- A causal explanatory approach has been used, with assumptions and risks defined for each causal link.
- The indicators used to measure the expected results, as well as the assumptions, are well defined.
- The logic of the theory of change is coherent and the predicted sequence of events follows a clear sequence.
- There are enough results for a reasonable analysis of contribution.
- The assumptions are necessary or probably necessary\(^a\) for the intervention to produce the expected results.

**Plausibility**
- A reasonable person would agree that the intervention would produce the desired results.
- The most important results and assumptions can be measured.
- The strength of the evidence used to measure the results, assumptions and risks can be determined by the evaluation team.
- Given the scale of the intervention, it is reasonable to predict that it could produce the expected results.
- It is likely that the assumptions remain valid throughout the delivery of the intervention delivery.\(^b\)

---

\(^a\) In causation, “necessary” is the condition that if A leads to B, then B could not have occurred without A taking place. For the purposes of an intervention, a necessary assumption is one that needs to be in place for the intervention logic to hold true: that the observed results predicted by the theory of change took place because of the intervention and not because of another influencing factor. Outside a controlled laboratory setting, it is recognized that not all assumptions are necessary, but rather are probably necessary.

\(^b\) This concept is often termed “sustainability” (41).

Source: adapted from Mayne, 2017 (40).
Evaluating interventions during a public health crisis

For several reasons, evaluating interventions implemented during a public health crisis differs substantially from evaluating those conducted during more normal times. These considerations can be largely grouped into the following categories:

- **Time**
- **Complexity**
- **Unintended impacts.**

**Time**

Because of the usually long time frame of a public health crisis and its accompanying response, an intervention may be modified during the course of its implementation in order to take into account of knowledge gained during the intervention period. In turn, this new knowledge is likely to impact on the underlying behaviour of the target population. In Contextual factors that vary as a function of time can also influence the observed results and, therefore, must be taken into account. Indeed, some of the changing contextual factors constitute independent variables in their own right. Moreover, in the context of a public health crisis, time is itself a variable because individuals implicitly consider time in their assessment of risk and of opportunity cost.

Logistically, time is also a factor in conducting the evaluation itself: health workers, public health authorities and other stakeholders who may form part of the evaluation team or be involved in some stage of the process are under increased workloads due to the public health crisis itself. Therefore, a balance must be struck between obtaining actionable information relevant to efforts to address the public health crisis (see Table 1) and minimizing the workload and potential emotional impact on participants (38).

| Table 1. Timing of evaluation in the intervention cycle: pros and cons |
|-----------------|-----------------|-----------------|
| **When**        | **Pros**         | **Cons**        |
| Early in the intervention | Highlights weaknesses in the theory of change, as well as in the types and quality of data. If conducted during the design stage, it can be used to brainstorm assumptions about the theory of change, identify gaps in evidence and modify the evaluation framework | Both the intervention and its influencing factors can vary over time; this is particularly true during a long intervention or when contexts are in flux. As a result, a more summative evaluation might need to be conducted later in the intervention cycle to ensure that the most current information is available |
| At the end of an intervention | Provides the most current information on how difficult an evaluation might be and whether it is even worthwhile proceeding. Useful for highlighting significant challenges that could hinder or delay an evaluation and, therefore, can be used to tweak the evaluation framework to find a way forwards | Limited scope for corrective action. If modifying the evaluation framework becomes necessary, this might require further discussions with stakeholders to decide how the findings will be used |
| During implementation | Allows new knowledge to be used to modify the theory of change, how the intervention is being delivered and the types of data collected. Permits agile management of the intervention [42] | Whether the new information can be used to modify the theory of change or implementation delivery depends on the level of commitment of the intended users |

*Source: adapted from Peersman et al., 2015 [2].*
**Complexity**
First, a public health crisis represents a natural experiment that has already created enormous socioeconomic changes. Secondly, interventions implemented during the course of a public health crisis response have created additional changes (from individual level to country level) from pre-2020 baseline behaviours. Thirdly, populations are increasingly subjected to an infodemic (surge of misinformation), which also contributes to modifying behaviours. Without proper accounting for influencing or confounding factors, it is difficult to ascribe causation to the intervention itself, to anticipate or measure unexpected phenomena, or to determine synergies.

**Unintended impacts**
There is mounting evidence that, in their drive to effect positive change, interventions can fail to produce the expected results and, in addition, can (and do) produce unintentional effects, some of which might adversely affect desired outcomes (Table 2) (37). Although the intervention might have been successful in reaching its goal, such unintentional effects might mean that, overall, the intervention has done more harm than good to the target population. For example, people may be frightened into changing their behaviours in the short term, but the negative impact of scare tactics on their feelings of well-being, trust and social cohesion may negatively affect their long-term motivation to uphold these behaviours (the outcome is offset by later behaviour). Understanding the environment is central to mitigating these types of intervention failure because (i) the environment might not support the types of behavioural change desired, (ii) the environment could trigger counteracting forces in response to positive behavioural change resulting from the intervention, or (iii) the intervention might inadvertently modify behaviours in non-target populations.

See Table 2 on the next page.
<table>
<thead>
<tr>
<th>Type</th>
<th>Effect</th>
<th>How to mitigate or control for these effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>No effect</td>
<td>The intervention did not modify the target behaviour relative to the baseline</td>
<td>Conduct a literature review to understand under what conditions and for which target groups the intervention will produce the desired results. Use this to determine or revise the assumptions of the theory of change.</td>
</tr>
<tr>
<td>Backfiring</td>
<td>The intervention changed the targeted behaviour, but in the opposite direction</td>
<td>Conduct a literature review to better understand the possible outcomes.</td>
</tr>
<tr>
<td>Negative side-effects</td>
<td>The expected outcomes were produced but were attenuated or eliminated because they were offset by negative side-effects</td>
<td>Consider multiple outcome criteria (triangulation) to better understand whether later compensatory behaviours could undo the results of the intervention.</td>
</tr>
<tr>
<td>Outcome not achieved, but a positive side-effect</td>
<td>The target behaviours remain unchanged by the intervention, but unforeseen positive consequences occurred</td>
<td>Anticipating that changing specific behaviours will generalize to other behaviours can be built into the theory of change. To assess whether this may occur requires additional indicators to monitor other behaviours.</td>
</tr>
<tr>
<td>Proxy changes only</td>
<td>Changes in the proxy for the expected result did not translate into changes in key target behaviours a</td>
<td>It is important to assess the sustainability of interventions over time according to outcome targets.</td>
</tr>
<tr>
<td>Intervention is offset by later behaviour</td>
<td>The intervention successfully modified the target behaviour, but this change was offset by later behaviour</td>
<td>This indicates the importance of understanding the underlying psychological processes. This effect can be mitigated by holding discussions with stakeholders and conducting a literature review.</td>
</tr>
<tr>
<td>Environment does not support change</td>
<td>The environment itself was unable to support behavioural change</td>
<td>It is imperative to consider whether, given the environment, the target behaviour can indeed be modified. Hold stakeholder discussions or conduct a literature review to understand whether the environmental factors necessary to produce the expected results are present</td>
</tr>
<tr>
<td>Intervention triggers counteracting forces</td>
<td>Positive effects are triggered by the intervention, leading to counteracting forces from the broader environment</td>
<td>It is important to consider whether context-specific influencing factors that could counteract interventions are present in the environment.</td>
</tr>
</tbody>
</table>

a This situation refers to the scenario in which proxies for the behaviour, rather than indicators for the behaviour itself, are used. In some cases, the proxies may change as a result of the intervention, but these fail to translate into a change in behaviour.

Source: adapted from Osman et al., 2020 (37).
Why might unintended impacts be of especial concern during a public health crisis such as, for example, the COVID-19 pandemic? The answer is that risk of exposure to disease, degree of severity on health, and social and economic consequences are unequally distributed across different population groups. As well as their negative impacts on groups experiencing the greatest health inequity (Box 7) and gender inequality, containment measures can also directly or indirectly create new vulnerable groups. Several documents highlight these effects and include recommendations for minimizing them, with a focus on COVID-19 pandemic:

- A UN framework for the immediate socio-economic response to COVID-19 (14);
- Health inequity and the effects of COVID-19: assessing, responding to and mitigating the socioeconomic impact on health to build a better future (29); and
- IANWGE compendium on integrating gender considerations in the response to COVID-19: key messages and actions from un entities (43).

Such negative impacts can directly increase the exposure of these groups to the public health crisis or can have indirect effects by worsening health outcomes and, thereby, increasing susceptibility to the disease. Although interventions generally operate at a smaller scale than countrywide containment measures, negative unintended impacts may similarly occur as a direct or indirect result of their implementation. Evaluating interventions that have already been implemented provides a valuable opportunity to determine whether this is the case. For interventions that are still at the planning stage or are currently being implemented, evaluation is essential to minimize potential harm to vulnerable groups. Completing a socioeconomic assessment using disaggregated data (TOOL 9) and the United Nations gender equality checklist (43) can provide a fuller understanding of the situation in the country or region and how it may be exacerbated directly or indirectly by the intervention.

### BOX 7. EXAMPLES OF GROUPS WHO EXPERIENCE HEALTH INEQUITY

- Women
- Ethnic minorities
- Indigenous peoples
- Slum dwellers, people in informal settlements, homeless people
- People in extreme poverty or facing insecure and informal work and incomes
- Older people
- Migrants, refugees, stateless and internally displaced people, conflict-affected populations
- People living with HIV/AIDS
- Groups that are particularly vulnerable and marginalized because laws, policies and practices do not protect them from discrimination and exclusion (e.g. LGBTQIA+ people)
- Children, adolescents and young people, especially girls and young women
- Small farmers, fishers, pastoralists, rural informal and formal workers, other people living in remote rural areas
- Urban informal workers and self-employed people who are dependent on a single platform
- People with pre-existing medical conditions
- People with disabilities or with mental health conditions
- People in detention
- People affected by food insecurity, particularly those in countries affected by prolonged conflict and crisis
- People in institutionalized settings (e.g. in psychiatric care, drug rehabilitation centres, old age homes)
Considerations of time, complexity and unintended impact have implications for the approach to evaluation and how evaluations are conducted. Throughout the planning and implementation of an evaluation, the evaluation team should follow the best practices for conducting qualitative research, especially if this is taking place during a public health crisis (summarized in Box 8)\(^\text{38}\).

**BOX 8. BEST PRACTICES FOR CONDUCTING QUALITATIVE RESEARCH DURING A PUBLIC HEALTH CRISIS**

- Reduce the research burden through better design of evaluation methods, including the use of novel rapid qualitative methods and of technological advances (such as real-time videoconferencing transcription software).
- Minimize the time required from staff.
- Keep interview guides brief and, ideally, conduct interviews at times that are convenient for participants, such as lunch breaks.
- Reduce the frequency of data collection as the public health crisis intensifies.
- Anticipate which steps of the evaluation will take longer than normal during a public health crisis (e.g. the research ethics review process, which prioritizes certain studies for review).
- Sift through evidence from a range of sources, rather than performing a “deep dive” into all available evidence.
Glossary

**Causality** (also known as causation or cause and effect) is the influence of the intervention on the observed results.
- **Attribution** explains why things occur and how the intervention influenced the desired change.
- **Causal inferences** are claims about cause and effect.
- **Contribution** explains how the intervention helped to achieve the observed results.
- **Generative causality** involves a description of the causal mechanism, as obtained by investigating the steps that occurred between the intervention and the observed results.

**Evaluation** is the act of calculating the utility, plausibility, generalizability and feasibility of an intervention, programme or policy. It may also assess the cost–effectiveness and sustainability of the results, among other factors.
- A **contribution analysis** uses generative causality to explain how a sequence of causal steps led from the intervention to the intended outcome. It makes causal inferences by testing the theory of change for an intervention against the observed effects, while also examining the impact of influencing factors.
- A **contribution claim** is the conclusion of a contribution analysis.

**Health outcomes** are changes in health that result from an intervention.

**Indicators** are types of data or information that are used to measure change.
- **Impact indicators** provide information on whether the intervention has had the desired output. The selection of impact indicators must be driven by an understanding of what can be clearly described. It is also important to note that while the emphasis during a pandemic is on short- and intermediate-term impacts, long-term impacts can be used to improve the understanding of future pandemics. As an example, intermediate-term impacts may refer to a measured increase in knowledge or behaviours, whereas long-term impacts may refer to the emotional effects of the intervention.
- **Process indicators** provide information about whether the planned activities took place or whether any essential components are missing. Process indicators allow the quality of intervention to be discussed and the best practices to be identified and incorporated into future interventions. Examples include holding meetings, conducting training courses, distributing medicines, and developing and testing health education materials.

**Inequity** is an instance of injustice or unfairness. Health inequities are systematic differences in the health status of different population groups. Such inequities have significant social and economic costs to both individuals and societies.

**Influencing factors** are those contexts or circumstances that are external to the intervention and cannot be controlled for, such as socioeconomic or political changes, but which are likely to influence the ability of the intervention to bring about the desired outcome.

**Mixed methods** is the intentional integration or combination of quantitative and qualitative methods in order to draw on the strengths of each in answering real-life research questions.

**Negative assurance** is the belief that a particular set of facts is accurate since no evidence has been found to the contrary.

**Rigorous thinking** is an approach that enables research methods to be scrupulously and meticulously carried out so that important influences can be recognized during the research process. It involves the use of a set of standards to evaluate the quality, trustworthiness and value of research.

**Stakeholders** are those who have a stake in the evaluation or will be affected by the issue and the findings.

**Targets** are the specific increases/decreases aimed for in the data measured. They are determined during the intervention design and are most relevant to interventions that have not yet been implemented.
A **theory of change** is the set of assumptions that underlies the steps leading to the goal of each intervention. It comprises a causal diagram or logic model, the underlying assumptions and risks, and key influencing factors. It is continually refined throughout the course of the intervention, as well as during the evaluation process.

- **Assumptions** are positive statements about the conditions that all need to be met if the activity is to stay on track, as well as the external conditions that need to be in place for the intervention logic to remain valid.
- **Risks** are negative statements about what might go wrong.

**Triangulation** increases the credibility and validity of research and evaluation through combining theoretical frameworks, evidence types and observers. It reduces bias that might arise due to reliance on a single model, method, evidence source or observer.
References*  


* All URLs were accessed 23 June 2022.


The WHO Regional Office for Europe
The World Health Organization (WHO) is a specialized agency of the United Nations created in 1948 with the primary responsibility for international health matters and public health. The WHO Regional Office for Europe is one of six regional offices throughout the world, each with its own programme geared to the particular health conditions of the countries it serves.

World Health Organization
Regional Office for Europe
UN City, Marmorvej 51,
DK-2100 Copenhagen Ø, Denmark
Tel: +45 45 33 70 00
Fax: +45 45 33 70 01
Email: eurocontact@who.int
www.who.int/europe

Document number: WHO/EURO:2022-6045-45810-65956

Member States
Albania Czechia Italy Norway Sweden
Andorra Denmark Kazakhstan Poland Switzerland
Armenia Estonia Kyrgyzzstan Portugal Tajikistan
Austria Finland Latvia Republic of Moldova Türkiye
Azerbaijan France Lithuania Romania Turkmenistan
Belarus Georgia Luxembourg Russian Federation
Belgium Germany Malta Serbia Ukraine
Bosnia and Herzegovina Greece Montenegro United Kingdom
Bulgaria Hungary North Slovakia Uzbekistan
Croatia Iceland
Cyprus

Israel

North Macedonia

Spain