EN

Evaluating drug policy

A seven-step guide to support the commissioning and managing of evaluations

Contents

Acknowledgements

| About this guide

This short document aims to provide a summary of the main issues that people engaged in developing drug policy, or in commissioning evaluations of drug policy, strategies and interventions, need to consider. There are many sources of more detailed information and guidance on conducting evaluations. This guide seeks not to duplicate or replace them but instead to act as an introduction, providing links to the wider literature and presenting the key issues for those managing rather than undertaking drug policy evaluations. Further reading and sources of more detailed information are provided at the end of this guide. There is also no single correct way to undertake an evaluation, and the choice of approach depends on many factors, including timing, objectives and the availability of resources. This publication is designed to assist people in choosing the best approach to suit their circumstances and to maximise the value of any evaluation.

There is a lot written about evaluation and sometimes the same terms are used to mean different things, so before starting it is important to clarify the definitions and concepts that the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) uses in this guide. The box below provides the definitions that we use in this guide for certain core concepts.

Definitions used in this guide

Drug policy is the overall direction and approach taken by a government to address (illicit) drug issues. It encompasses the whole system of laws, regulatory measures, courses of action and funding priorities concerning (illicit) drugs put into effect by a government or its representatives.

A **drug strategy** is generally a document, usually time bound, typically containing objectives and priorities alongside broad actions, and may identify at a high level the parties responsible for implementing them.

An **action plan** may accompany a drug strategy but is sometimes integrated into that strategy. It is typically focused on a relatively short period and identifies more detailed actions to implement the strategy, together with timings and responsible parties.

Policy evaluation can be simply defined as an evidence-based judgement about to what extent and how well a policy, strategy or intervention has been implemented, and/or whether the objectives have been achieved, together with any other effects it has had. There are a variety of criteria against which such assessments may be made. For example, an evaluation may consider whether or not the policy, strategy or intervention has:

- been *effective* (has it achieved its objectives?) and/or *efficient* (how do the resources used relate to the outcomes — positive and/or negative)?

- been *relevant* to both the identified needs and the policy objectives;

- been *coherent* both internally and in relation to other policy interventions; and

- achieved *added value* — has the existence of the policy/intervention improved outcomes over and above what might have been achieved anyway?

In addition, an evaluation might consider other criteria such as the quality and extent of *implementation* of the policy (process evaluation), whether or not it is *sustainable* and how fairly its effects are distributed across different stakeholders (*equity*). Evaluation can draw on scientific methods and the collection of empirical and measurable evidence to identify causal relationships between actions and outcomes.

However, in practice, these terms are often used interchangeably, other definitions may be used elsewhere for different purposes and the distinctions between them can be blurred, particularly with respect to drug policy and drug strategy. These definitions draw on the European Commission's guidelines and toolkit for evaluation and fitness checks (European Commission, 2015a) as well as published and unpublished work undertaken for the EMCDDA. Definitions for a wider range of terms can be found in the Glossary (page 22).

| Why evaluate drug policy?

The value of evaluation has been recognised in all EU drug strategies and in the strategies of many Member States. The 2013–20 EU Drugs Strategy (Council of the European Union, 2013) invites EU institutions, bodies and countries to 'recognise the role of scientific evaluation of policies and interventions (with a focus on outcomes achieved) as a key element in strengthening of the EU approach to drugs, and [to] promote its use both at national, EU and international level'.

Evaluation is essential for effective policymaking, helping ensure that policies and programmes have the desired effect, provide value for money and do not have negative unintended consequences. In particular evaluation can contribute to:

- Better planning of policy and services and the provision of timely and relevant advice to support decision-making and input to political priority setting. The assessment of the outcomes that have been achieved, and for whom, helps ensure that programmes meet the identified needs.

- Efficient resource allocation — identifying the most effective and efficient elements of a policy, or highlighting gaps in provision, can help ensure that scarce resources are used to maximum effect.

- Organisational learning — evaluations can not only identify areas for improvement but also encourage the sharing of lessons drawn from the assessments of both successes and failures. They also provide the opportunity to look for 'unintended' or 'unexpected' effects of actions.

- Transparency and accountability — all stakeholders and the general public have a right to know what has been spent on government drug policy and what it has achieved. This is important for developing and maintaining trust in government and public services.

Key messages

* There is no one 'correct' way to perform an evaluation of drug policy. What is best will depend on what you want to know, what data you have available or can obtain, and the resources and time available to you.

* Evaluation should not be seen as a one-off event but will be most useful if viewed as an ongoing process intertwined with policy or strategy development and implementation.

* Evaluation needs to be accompanied by a commitment to taking action on the findings, and the opportunity to do so. The timing and choice of evaluation design need to be realistic and take this into account; producing a detailed evaluation of a previous strategy only after a new strategy has been developed and implementation begun will limit its usefulness.

* Developing the expertise and data sources for drug policy evaluation over time will increase the ability to conduct evaluations in support of drug policy development and enhance action to address drug problems.

1 Preparing the ground

There are number of key factors that facilitate successful evaluation and which it is important to establish at the outset:

- **Leadership commitment to the process**

  This is essential for a successful evaluation. Ownership of the process by senior government officials and politicians is key to obtaining sufficient resources; without it, it will be difficult to ensure that the findings of the evaluation are acted upon. Where a culture of evaluation exists, leadership support is likely to be easier to achieve.

- **Adequate resources — finances and skills**

  A comprehensive and thorough evaluation of drug policy requires considerable skills and time, and hence is costly. When designing the evaluation it is important to be realistic about how much can be achieved within the resources available. A narrower, focused evaluation, which will be properly carried out, can be more useful than a more ambitious but under-resourced, and hence poorly performed, exercise.

- **Stakeholder engagement**

  The involvement of stakeholders in an evaluation contributes to its development, quality and transparency. Stakeholders' involvement can facilitate access to relevant data and individuals, and ensure that the approach taken is realistic, covers the most important aspects of the policy or strategy, and represents all relevant geographical areas and affected groups.

- **Establishing a steering group**

  It is considered good practice to establish a steering group to guide the evaluation process. To achieve the benefits described above, it is important to include representation from the following stakeholders:

  » policymakers (e.g. national drug policy coordinators, representatives of relevant ministries such as those of health and justice);

  » public officials (e.g. local drug coordinators, national focal points, administration, government agencies);

  » target groups and beneficiaries (e.g. health professionals, police forces, social workers, service user groups);

  » experts in drug policy and evaluation methodology (e.g. consultants, academics); civil society groups.

2 Deciding on the type and scope of the evaluation

Evaluation is often seen as something that takes place at the end of a strategy or action plan period. However, ideally evaluation is an ongoing process, intertwined with the policy cycle, and with different types of evaluation providing different types of information at different times. There is also no single 'correct' type or model of evaluation that will answer all questions or suit all circumstances. There are therefore a number of questions to be considered when deciding on the type and scope of evaluation that suits the circumstances in any particularly case:

- **To evaluate the overall policy/strategy, or individual programmes/elements?**

  Although occasionally the overall drug policy may be evaluated, the two main options are generally an overarching evaluation of the national drug strategy (general evaluation) and the evaluation of a number of key interventions (targeted evaluation). Factors that may influence the choice of what to evaluate include the purpose of the evaluation, the stage of the policy cycle at which it is being conducted and the resources available. For example:

  - » A general evaluation will be the preferred option when, as is often the case, the aim of the evaluation is to improve the quality, efficacy and efficiency of the drug strategy as a whole. This broad type of evaluation can also provide the opportunity to compare the effectiveness of a strategy's various elements, the interactions between them and the relative priority given to them during implementation.

  - » A targeted evaluation should be carried out when more in-depth assessment of one or a limited number of key interventions is needed. Narrowing the scope of the evaluation allows more detailed investigation of a particular intervention, and is often used when a new intervention approach is being developed or rolled out.

  However, it is important to note that these types of evaluation are not mutually exclusive. Different approaches may be used at different stages in the evaluation and the policy cycle, and a 'mixed approach' can also be adopted, for example assessing the implementation of all elements of a strategy while focusing on the effectiveness and efficiency of some key components.

- **At what stage(s) of the policy cycle will the evaluation be conducted?**

  The policy process can be viewed as cyclical: starting with the formulation of a policy, it continues through planning and resource allocation, programme design, implementation and the delivery of outputs and results. Evaluation follows a similar cyclical process and can be addressed through *ex ante* (before), interim or *ex nunc* (during) and *ex post* (after) evaluations. The policy and evaluation processes, therefore,

should be seen as parallel cycles influencing one another (Figure 1).

The differing types of evaluation provide a variety of information appropriate to the different stages of the policy cycle (see box on page 9), although the division between them is not absolute and, in practice, various combinations of these approaches can be used at different points in a strategy's life cycle. For example, an *ex post* evaluation at the end of one strategy time period may also act as, or form part of, an *ex ante* evaluation of a new strategy.

FIGURE 1
Parallel cycles
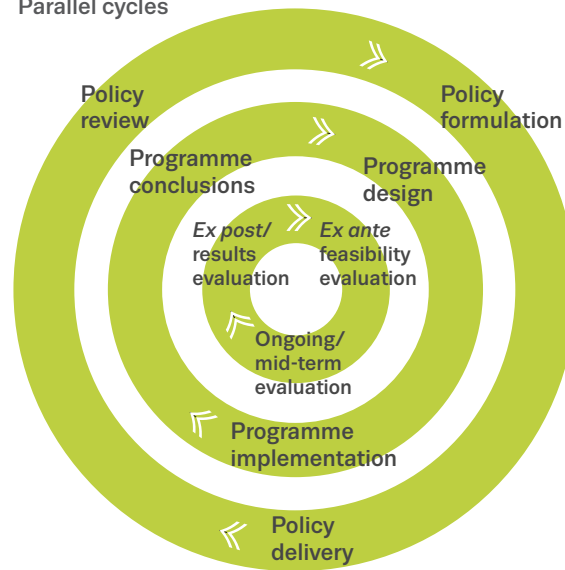
Source: European Commission, 2013.

■ **Feasibility considerations — resources, time, data**

The ideal evaluation is established right from the outset of a policy or programme, with data collection built in from the start and with sufficient resources and time set aside for detailed analysis by skilled professionals. However, a common criticism of evaluations of drug policy is that, as a result of data limitations, timeliness issues or under-resourcing, they can struggle to deliver useful results. It is important to be realistic in designing your evaluation and, if necessary, to design a more limited evaluation focused on key issues for the next policy phase rather than aim for a more extensive evaluation that is unachievable (evaluation design is discussed in more detail in steps 4 and 5).

Main types of evaluation and their different uses

There are a number of different, and to some extent overlapping, ways of categorising policy and strategy evaluations, as illustrated by the examples in the table below. One can identify them firstly by the stage of the policy process in which they take place, secondly by the type of evaluation and thirdly by the criterion against which the judgement is to be made. A further dimension that could be considered is the tool or method to be used in the assessment. The list given here is not exhaustive but illustrates the range of alternatives, including both data sources and analytical approaches.

When	Type of evaluation	Criteria assessed	Methods/tools
Ex ante evaluations	Formative evaluations	Relevance	Documentary analysis
Interim or *ex nunc* evaluations	Summative evaluations	Coherence	Focus groups/interviews
	Policy appraisals	Effectiveness	Surveys
	Impact assessments	Efficiency	Administrative data
Final or *ex post* evaluations	Process evaluations	Added value	Public expenditure studies
	Outcome/effectiveness evaluations	Equity	Social cost studies
	Impact evaluations	Sustainability	Cost-effectiveness analysis
	Economic evaluations		Cost–benefit analysis
			Statistical modelling

How these different aspects may be linked together within the overall evaluation approach is illustrated below:

(i)   *Ex ante* evaluations
Policy appraisals and **impact assessments** are types of *ex ante* evaluation and are carried out as part of the policy development process. They focus on policy content, assessing criteria such as **relevance** and **coherence**, but may also consider what the expected costs and impact will be, based on previous research or evaluation evidence, and make comparisons with alternative policy options. The purpose is to make sure that the policy is realistic, affordable and likely to achieve the desired outcomes before it is fully implemented, in order to maximise the likelihood that it will succeed.

(ii)   Interim (or *ex nunc*) evaluations
These take place part way through the policy or programme cycle with the aim of making improvements, and are often used for mid-term drug strategy evaluations, for instance after the end of a first action plan. These are usually **formative** or **process evaluations** that look mainly at the extent to which the policy or activity has been implemented and if this was done as intended. However, they can also include some consideration of outputs, outcomes and the extent to which different target groups have been reached, i.e. **effectiveness** and **equity** issues. These types of evaluation are particularly important for new programmes that may not work or may need adjustment when rolled out, as well as for longer-term strategies that may need to adjust to changes in patterns of drug use or the wider social or economic environment.

(iii)   Final or (*ex post*) evaluations
Outcome and **impact evaluations** are common components of ex post evaluations (i.e. evaluations conducted at the end of a programme or strategy). These are types of **summative evaluation**, which look at the extent to which a policy or programme has met its goals and had any other consequences, and draw lessons based on these assessments. The main focus at this stage is often on **effectiveness**, **added value** and **efficiency**, that is whether or not the objectives of the policy or programme were achieved and represented a good use of resources. Nevertheless, in addition they usually seek to identify other lessons for future policy development, such as issues that have acted as barriers and key factors for success, so they often also include process evaluations or implementation reviews.

Economic evaluations may form a part of evaluations conducted at any stage of the policy process. They are often used to assess the **efficiency** of the programme and/or compare alternative courses of action, relating the resources used to what is achieved in terms of outputs and the programme objectives. There are a number of different tools and analytical approaches that can be used in economic evaluations, including cost-effectiveness and cost–benefit analyses. Public expenditure and social-cost studies are often important components of economic evaluations of drug policy and strategies

3 Choosing an evaluation team

An evaluation may be conducted by an internal or external team, or a combination of the two. Whether an evaluation should be considered internal or external is also not necessarily clear-cut. In some cases the government may have a central department or unit with the remit to conduct policy evaluations or reviews, and while internal to the government it will be external to the departments responsible for drug policy. Similarly, external evaluation teams may be from organisations external to the government but from within the country concerned, e.g. academic institutions or evaluation professionals, or come from outside the country.

In choosing what is most appropriate for a particular situation there are a number of factors to consider:

- **The importance of independent evaluation**

  If an evaluation is to provide findings that are considered trustworthy, it is important that it is objective and independent. People who have been closely involved in implementing a strategy may find it difficult to be unbiased when judging the success of the different elements of the work, or may come under pressure to downplay problems. External evaluators may be less susceptible to these influences, but on the other hand may be more vulnerable to 'capture' by more vocal stakeholder groups and lack important contextual knowledge. Such a lack may have a negative impact on resources and timing, because of the need to bring the evaluators up to speed, and on the usefulness of their recommendations. Even when internal evaluators are objective, they may not be perceived as such, so in circumstances where the policy is contentious it may be considered preferable to have an external team.

- **The need for expertise and knowledge in both evaluation and drug policy**

  Expertise in evaluation is clearly important for both the design and conduct of the evaluation, but knowledge of drug policy issues is also necessary to ensure that the research questions and indicators used to assess outcomes are appropriate, and to guide the analysis and ensure useful conclusions and recommendations. Evaluators from another country or from international agencies will often have only limited knowledge of national drug policy issues so they are likely to need support and guidance about them.

- **The available resources and the time allowed for the evaluation**

  Both internal and external evaluations require adequate resourcing. If the time allocated for the evaluation is very limited then it may be difficult to contract an external team within the time frame.

These factors will need to be weighed against each other and the relative importance of each will vary depending on circumstances. For example, as discussed above, internal evaluation teams may be seen as biased and have limited evaluation expertise, but they know the drug policy context intimately, understand the data sources available and may be aware of occurrences that have not been documented. Similarly, external teams may be more trusted by the stakeholders, who in turn may be more honest with them, and they may find it easier to be objective and to deliver 'uncomfortable' lessons, but on the other hand they may need a lot of support in understanding local circumstances and data sources. Language issues may also be an important consideration; clearly, teams from within the country will speak the national language and thus be able to read the necessary documents and conduct interviews with all stakeholders, which may be a problem for some external evaluators.

In practice, some sort of joint evaluation team or approach may be the best solution. In these circumstances, the roles and responsibilities of internal and external team members will need to be clearly defined from the outset. However, in all cases it will be important to have clear terms of reference for the evaluation, as well as more detailed plans that are reviewed regularly.

4 Evaluation design: research questions and methods

In considering which methodology to adopt for the evaluation, a number of different factors need to be addressed:

- **First establish the research questions, i.e. what it is that you want to learn from the evaluation.**

  For example, do you want to know if all the actions envisaged have been undertaken; whether or not the expected outcomes have been achieved; or if one component of the strategy was more effective than another? The evaluation questions need to be appropriate to the stage in the policy cycle and the goal of the evaluation (e.g. is it an interim evaluation aimed at improving the implementation of a strategy or a particular intervention; is it being conducted to identify the impact that has been achieved?). In outcome or impact evaluations, these questions will obviously need to be linked to the objectives of the policy or programme being evaluated. It is common to have several different research questions for an evaluation, relating to different evaluation criteria, but too many or too complex questions may overburden the evaluation and cause it to lose focus. The questions also need to be specific, measurable, achievable, relevant and time bound (SMART):

  » specific — they must clearly stipulate what is to be evaluated (e.g. a strategy, part of a strategy, an intervention);

  » measurable — the questions must be translatable into measurable criteria;

  » achievable — it is important to be realistic about what can be achieved with the resources, time and data available;

  » relevant — questions must meet the concrete needs of policymakers/stakeholders and provide useful and useable information;

  » time bound — the time period to be covered needs to be made clear.

Examples of types of research questions addressing different criteria and relating to different types of evaluation are given in the box on page 13.

Examples of some types of research question

Considering *relevance*: 'To what extent were the goals and interventions identified in the 200X–201X drug strategy relevant to the drug problems being experienced at the beginning and end of the time period it covered?'

For a *process evaluation* looking at *implementation*: 'Was the drug intervention designed and managed in a way that ensured that it was implemented efficiently and according to the timetable set out at the start?'; or 'To what extent were the different actions planned under the 200X–201X drug strategy implemented, and what factors acted as barriers or facilitators to implementation?'

For an *outcome* evaluation: 'To what extent does the drug intervention, as it is now configured, satisfy the needs of its beneficiaries with regard to health and treatment?'

Relating to *impact*: 'To what extent did the 200X–201X drug strategy lead to a decrease in drug supply?'

- ■ **Then choose the appropriate evaluation approach to address the research questions.**

  Choice of evaluation method is likely to be an iterative process involving the evaluation team. If external consultants are involved, they may be requested to propose a methodology for answering the research questions in their proposal. There are a range of methods and tools that can be used to answer different types of question, and some provide more robust evidence than others. In outcome and impact evaluations, linking interventions causally with outcomes is a common goal, and some evaluation designs are better than others at demonstrating causal linkages (see the box on page 14). However, while experimental designs such as randomised controlled trials may be appropriate for evaluating individual interventions, these are not usually feasible when broader drug strategies are being considered, and drawing firm conclusions on cause and effect is extremely difficult. In such cases, alternative evaluative approaches that focus more on how and why interventions are working, and contextual factors that may affect outcomes, such as 'realist evaluation' approaches, may be more useful (Pawson and Tilley, 2004).

  Although the robustness of the findings from a particular evaluation method is a very important consideration, the choice will also depend on other factors such as data availability and quality, and time and resources available. A less ambitious but more focused and clearly targeted analysis that is achievable in the available time frame is likely to be more valuable than a more complex but under-resourced evaluation that fails to deliver or provides findings too late to be useful.

  There are many and varied tools and techniques for data collection and analysis that may be used in an evaluation. These include document reviews (e.g. of policy documents, budget reports, meeting records or research publications), case studies, participatory observation, interviews, focus groups, questionnaire surveys, panels of experts, spatial analysis, multi-criteria analysis and modelling. The selection of tools will depend on the type of data required, data availability, the research question to be answered, the availability of the necessary skills and resources, and other practical considerations.

Rating the strength of an outcome evaluation — the Maryland Scientific Methods Scale

For evaluations that seek to assess outcomes and impact, and to attribute these to a particular policy or intervention (causal inference), a common way of rating the strength of the evaluation design is the Maryland Scientific Methods Scale (Farrington et al., 2002). This was originally developed for reviewing evidence in the crime prevention field, and it grades evaluation designs from level 1, the lowest, to level 5, which provide the strongest evidence that the outcome seen was a result of the intervention.

The most common type of outcome evaluation is one that occurs after implementation, and involves comparing measures relating to the implementation of the intervention with measures of its supposed effects (e.g. on drug supply or use), and assessing the extent to which these are correlated. This is a level 1 type of evaluation. Like level 2 evaluations, which use data collected both before and after the intervention but with no control groups, it shows only association and not causation.

Levels 3 to 5 (randomised controlled trials) involve experimental designs that compare groups or areas that have received the intervention with similar groups or areas that have not. They provide evidence of causation but need to be integral to the intervention process and are really suited only to evaluation of specific interventions, such as a new treatment programme.

5 Evaluation design: logic models or cause-and-effect chains and data requirements
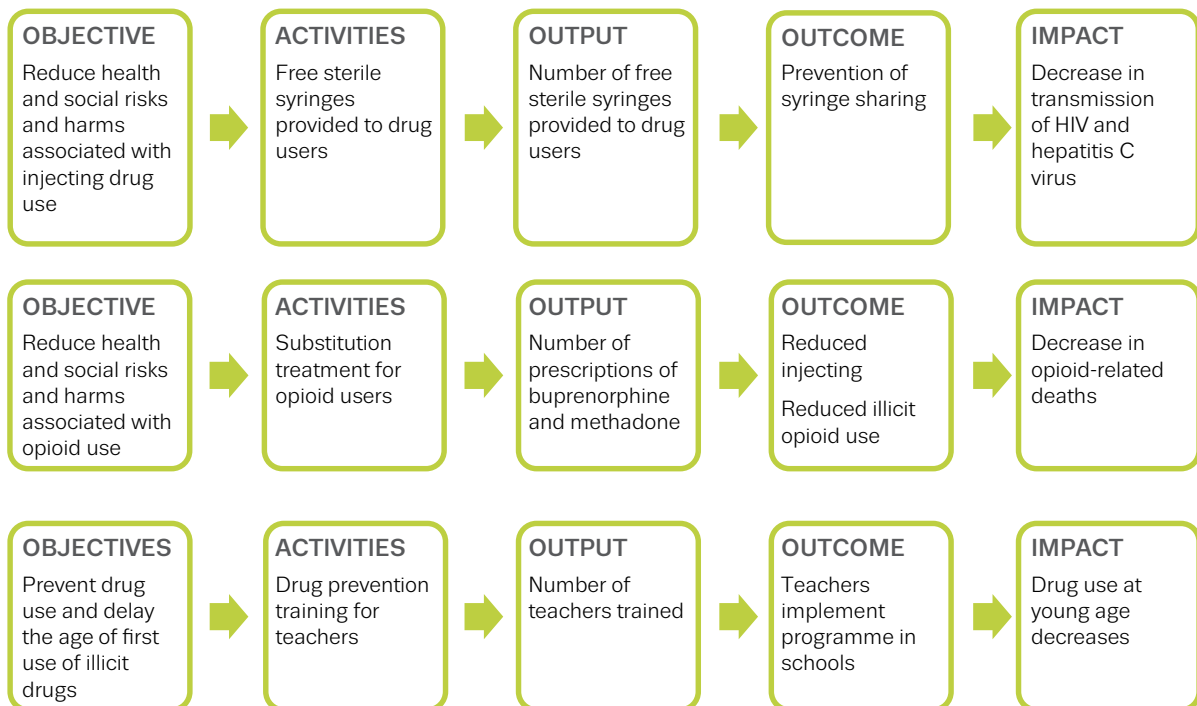
To decide what to measure to answer the research questions chosen, it is important to understand the components of the policy/strategy/intervention under consideration, and how it was expected to achieve its aims. This is also important for refining the research questions and making them SMART.

The following elements of the strategy or intervention to be evaluated need to be identified (W. K. Kellogg Foundation, 2004; OECD, 2010):

- the needs and problems that the intervention *aims* to solve;
- the *activities* which are a part of the intervention — actions to achieve the aims of the intervention;
- the financial and material *inputs*, and the organisational and regulatory *inputs* (including human resources, budgets and equipment), that make it possible to implement the intervention and deliver the intended results;
- the intended products or *outputs* of the intervention, such as the number of activities undertaken related to the aim of the intervention (e.g. numbers treated, drug seizures and training packages developed);
- the intended changes in behaviour and knowledge, or *outcomes* — the short- and medium-term results and consequences of the intervention which are related to its aim;
- the wider changes or *impacts*, such as health improvements or reduced crime, which are the long-term consequences of the intervention, direct and indirect as well as intended and unintended.

Taken together, these different components describe the **logic model** or **cause-and-effect chains** that underpin the strategy or intervention. In an ideal world, these would be clearly spelled out in the strategy or project documents. The more links in the objective-activities-output-outcome-impact chain that are clearly identified from the outset, where possible alongside proposals for monitoring them, the more amenable to evaluation the strategy will be. However, in the real world it is often necessary to construct them based on a review of a variety of documents, which in the case of drug strategies might include the EU's strategy and action plans; laws, decrees or other legal acts; documents supporting the planning, budgeting and management of the intervention; monitoring data; reports of audits; reviews; and previous evaluations. If the necessary information is lacking in official documents, additional key informants could be interviewed.

FIGURE 2
Examples of cause and effect chains

OBJECTIVE	ACTIVITIES	OUTPUT	OUTCOME	IMPACT
Reduce health and social risks and harms associated with injecting drug use	Free sterile syringes provided to drug users	Number of free sterile syringes provided to drug users	Prevention of syringe sharing	Decrease in transmission of HIV and hepatitis C virus

OBJECTIVE	ACTIVITIES	OUTPUT	OUTCOME	IMPACT
Reduce health and social risks and harms associated with opioid use	Substitution treatment for opioid users	Number of prescriptions of buprenorphine and methadone	Reduced injecting Reduced illicit opioid use	Decrease in opioid-related deaths

OBJECTIVES	ACTIVITIES	OUTPUT	OUTCOME	IMPACT
Prevent drug use and delay the age of first use of illicit drugs	Drug prevention training for teachers	Number of teachers trained	Teachers implement programme in schools	Drug use at young age decreases

Some examples of such models are shown in Figure 2, to clarify the terminology and underlying principles (Emmanuelli and Desenclos, 2005). In practice, there will often be a range of activities under each objective, several different outputs from an activity, and so on. In these cases a number of different models or a branching model may be appropriate.

These logic models and the research questions to which they relate help to identify the sort of information, or **indicators**, that will need to be collected to conduct the evaluation. These indicators include **measures of inputs, outputs, outcomes** and **impact**. Process evaluations will tend to focus mainly on describing the activities undertaken, the inputs and outputs, and contextual factors such as barriers and facilitators to implementation. Outcome and impact assessments will need information about the activities conducted, as well as measures of what has been achieved, in order to link one to the other as far as is possible.

The indicators used in evaluation could be existing datasets which are already monitored regularly. Such indicators can, for example, be found in routine government statistical publications or at monitoring centres. The epidemiological, supply-side and other indicators available in the many publications of the EMCDDA and its Reitox national focal points can be used in evaluation studies. These indicators cover both supply- and demand-side activities and include potential input, output, outcome and impact measures. Examples are:

- public expenditure (input);
- number of needle and syringe exchange programmes (input);
- number of drug-related research projects funded (output);
- number and quantity of illicit drug seizures (output);
- drug users entering treatment services (output);
- number of drug law offences (output);

- drug-related crime (output/impact);

- prevalence of drug use in specific populations (e.g. prisons) (outcome/impact);

- prevalence of drug use among the general population (impact);

- incidence of drug use (impact);

- age of initiation of drug use (impact);

- problem drug use (impact);

- drug-related infectious diseases (impact);

- drug-related deaths (impact);

- purity of drugs (impact);

- prices of drugs (impact);

- market size estimates (impact).

It is also possible to use ad hoc indicators. Especially for the evaluation of specific interventions, these indicators are likely to be necessary. However, unless introduced at the beginning of the programme, pre-existing baseline data may not be available for comparison, limiting the strength of the evaluation. Special data collections may also take time and delay the evaluation, so planning these in advance is needed to avoid delays. Examples of such indicators are:

- cost of harm-reduction programmes (input);

- number of sterile syringes sold and distributed (output);

- drug-related arrests (output);

- rate of syringe sharing among injecting drug users (outcome).

Specific data collection is also likely to be essential to obtain information on the activities undertaken and contextual factors that have a bearing on the implementation and outcomes achieved. In this context, interviews or focus groups with representatives of key stakeholder groups, documentary analysis and questionnaire surveys are all likely to play a role.

The research questions, the type of evaluation, the indicators and the tools used to obtain them all need to be linked, as shown in the following example:

An *outcome* evaluation: 'To what extent does the drug intervention satisfy the needs of its beneficiaries with regard to health and treatment?'

> CRITERION: Effectiveness
>
> TOOL: Questionnaire survey
>
> INDICATORS: Rate of satisfaction with quality and effectiveness of treatment services, from '*very dissatisfied*' (1) to '*very satisfied*' (5), with regard to waiting time, availability of staff, usefulness of information, outcomes of treatment, accessibility of infrastructure, equipment, etc.

6 During the evaluation

Even if the evaluation is being undertaken by external evaluators, those with responsibility for the evaluation will need to be involved beyond the design phase and throughout the evaluation process, in a range of ways including:

- **Facilitating access to stakeholders and data sources**

  While some of the routine statistical information may be readily available on the internet, much of the information about activities such as expenditure and throughput will need to be obtained for the evaluators. For external evaluators, assistance in translating data sources may sometimes be necessary. In addition, identifying potential stakeholders for interviews or surveys and obtaining their consent or, if appropriate, agreeing mechanisms for reaching them that maintain confidentiality is also likely to be necessary.

- **Monitoring progress**

  This is important to ensure that the evaluation remains on track and to deal with any practical difficulties — including access to key individuals and problems obtaining data — that may be encountered. Senior government officials and politicians with responsibility for the drug policy area, and other stakeholders, need to be kept informed of progress to ensure they remain engaged in the process. Establishing reporting mechanisms within the project plan for the evaluation, e.g. requirements for inception and interim reports, will facilitate this. If a steering committee has been set up to oversee the evaluation, this will be an important task for it.

- **Providing input or a reality check to recommendations, if appropriate**

  Although an evaluation needs to be objective and benefits from independence, it can be useful to discuss any recommendations being considered, to ensure that they are appropriate to the legal and administrative frameworks of the country involved. This may take place through comments on draft recommendations, or the evaluators may wish to engage the steering groups or stakeholder focus groups in developing recommendations based on the conclusions of the evaluation. This may facilitate the adoption of the recommendations and ensure they can be acted on. To avoid perceptions of undue influence on the findings and associated loss of independence, it might be useful to separate the initial evaluation, containing the judgement of the successes and failures of the strategy or intervention, from the process of developing recommendations for action. The evaluation findings could be presented as a separate report, or as one section of an overall report, to be followed by a separate section on 'taking the actions forward', setting the findings in the operating context of any upcoming strategy. This allows a mechanism for both the evaluators and those who must implement the results to contribute.

7 Using the evaluation results

Taking action and making changes

It is unlikely that any evaluation will conclude that the intervention or strategy is perfect, not least because the context in which the programme is operating is likely to have changed over time. The benefits from evaluation are obtained only when changes are made that improve the operation and outcomes of the intervention or strategy in question, or such changes are incorporated into subsequent interventions and strategies. Therefore, it is important to consider the timing and mechanisms for disseminating and implementing the recommendations of the evaluation right from the start. This might include taking account of the likely timeline for development of the next strategy when designing the evaluation of the current one, to ensure sufficient time for the evaluation findings to be used within the new strategy development process. Another example is that it might be important for the evaluation's budget to include costings for the dissemination of the evaluation findings to local governments and agencies.

Drawing lessons for the next evaluation

In the guidance above we have stressed the need to be pragmatic and to recognise that it may not be possible to undertake an 'ideal' evaluation of the highest quality, i.e. built in from the start, including interim evaluations and a full impact assessment that clearly demonstrates causation and assesses cost–benefit or cost-effectiveness. Evaluating a complex policy area such as the illicit drugs field is a very difficult undertaking and the hidden nature of drug use and supply makes it difficult to obtain accurate measures of many of the outcomes targeted. Nevertheless, performing the best evaluation possible under the circumstances can provide valuable information and is also likely to highlight the key data gaps and areas where improvements can be made. This can encourage the incorporation of evaluation from the beginning of the policy cycle and the allocation of resources to improving data sources.

Developing an evaluative culture

Over time this should lead to the development of an evaluative culture (OECD, 2016) and an improvement in the information and resources available for drug policy evaluation. Particular areas where quality improvements will support better evaluations include data collection systems, clearer logic models underpinning strategies and enhanced coordination. In turn, improved evaluations should contribute to more effective and efficient interventions.

Sources and further reading

Below can be found details of the documents referenced in the seven-step guide, as well as some other materials that may be useful to people involved in evaluations in the field of drug policy.

- Council of the European Union (2013), *European Union Drugs Strategy 2013–2020*, Publications Office of the European Union, Luxembourg (available at: http://dx.publications. europa.eu/10.2860/69835).

Evaluation guides and texts

- European Commission (2013), *EVALSED: The resource for the evaluation of socio-economic development*, European Commission, Brussels.

- European Commission (2015a), *Better regulation toolbox* (http://ec.europa.eu/smart-regulation/guidelines/docs/br_toolbox_en.pdf).

- European Commission (2015b), Better regulation guidelines, SWD(2015) 111 final (http://ec.europa.eu/smart-regulation/guidelines/docs/swd_br_guidelines_en.pdf).

- Farrington, D. P., Gottfredson, D. C., Sherman, L. W. and Welsh, B. C. (2002), 'Maryland Scientific Methods Scale', in Sherman, L. W., Farrington, D. P. et al (editors) *Evidence-based crime prevention*, pp. 13–21, Routledge, London.

- Fretheim, A., Oxman, A. D., Lavis, J. N. and Lewin, S. (2009), 'Support tools for evidence-informed policymaking in health 18: Planning monitoring and evaluation of policies', *Health Research Policy and Systems* 7, pp. 1–8.

- OECD (Organisation for Economic Co-operation and Development) (2010), *Quality standards for development evaluation*, OECD, Paris (available at: http://www.oecd-ilibrary.org/development/dac-quality-standards-for-development-evaluation_9789264083905-en).

- OECD (2016), *Evaluation systems in development co-operation: 2016 review*, OECD, Paris (http://www.keepeek.com/Digital-Asset-Management/oecd/development/evaluation-systems-in-development-co-operation_9789264262065-en).

- Pawson, R. and Tilley, N. (2004), *Realist evaluation* (http://www.communitymatters.com.au/RE_chapter.pdf).

- W. K. Kellogg Foundation (2004), *Logic model development guide*, W. K. Kellogg Foundation, Battle Creek.

Published examples of drug strategy evaluations

Emmanuelli, J. and Desenclos, J.-C. (2005), 'Harm reduction interventions, behaviours and associated health outcomes in France, 1996', Addiction 100(11), pp. 1690–1700.

Galla, M., Van Gageldonk, A., Trautmann, F. and Verbraeck, H. (2006), Evaluation of the implementation of the national strategy to combat drugs: A Hungarian-Dutch cooperation. Report of the external mid-term evaluation, Trimbos Institute, Utrecht.

Griffiths, P., Strang, J. and Singleton, N. (2016), Report of the rapid expert review of the national drugs strategy 2009–2016 (http://health.gov.ie/wp-content/uploads/2016/09/Rapid-Expert-Review-of-the-National-Drugs-Strategy-2009-2016.pdf).

Portuguese National Institute of Public Administration (PNIPA) (2004), External and independent evaluation of the national strategy for the fight against drugs and the national action for the fight against drugs and drug addiction, Horizon 2004, PNIPA, Lisbon.

Data sources for evaluation

EMCDDA Statistical Bulletin (available at: http://www.emcdda.europa.eu/data/stats2016).

Glossary of key terms used in evaluations

Activities — processes, tools, events, technology and actions that are part of the programme implementation. These interventions are used to bring about the intended programme changes or results, i.e. the actions taken or work performed to achieve the aims of the intervention.

Added value — the extent to which something happens as a result of an intervention or programme that would not have occurred in the absence of that intervention. Also known as 'additionality'.

Aim — the purpose of, for example, an intervention or a policy.

Causality — an association between two characteristics that can be demonstrated to be due to cause and effect, i.e. a change in one causes the change in the other.

Coherence — the extent to which intervention logic is non-contradictory or the extent to which the intervention does not contradict other interventions with similar objectives.

Control group — a group of participants in a study not receiving a particular intervention, used as a comparator to evaluate the effects of the intervention.

Criterion — character, property or consequence of a public intervention on the basis of which a judgement will be formulated.

Data — information; facts that can be collected and analysed in order to gain knowledge or make decisions.

Drug action plan — scheme or programme for detailed specific actions. It may accompany or be integrated into a drug strategy but typically focuses on a relatively short period and identifies more detailed actions to implement the strategy, along with timings and responsible parties.

Drug policy — overall philosophy on the matter; position of the government, values and principles; attitude, direction. It encompasses the whole system of laws, regulatory measures, courses of action and funding priorities concerning (illicit) drugs put into effect by governments.

Drug strategy — unifying theme; framework for determination, coherence and direction. It is generally a document, usually time bound, containing objectives and priorities alongside

broad actions, and may identify, at a top level, the parties responsible for implementing them.

Effectiveness — the fact that expected effects have been obtained and that objectives have been achieved.

Efficiency — the extent to which the desired effects are achieved at a reasonable cost.

Equity — the extent to which different effects (both positive and negative) are distributed fairly between different groups and/or geographical areas.

Evaluation — a periodic assessment of a programme or project's relevance, performance, efficiency and impact in relation to overall aims and stated objectives. It is a systematic tool which provides a rigorous evidence base to inform decision-making.

Evaluation criteria — aspects of the intervention which will be subject to evaluation. Criteria should fit the evaluation question. If all the criteria are put together, they should account for a good and complete measurement. Examples are relevance, efficiency and effectiveness.

Evaluation question — question asked by the steering group in the terms of reference and which the evaluation team will have to answer.

Evaluation team — the people who perform the evaluation. An evaluation team selects and interprets secondary data, collects primary data, carries out analyses and produces the evaluation report. An evaluation team may be internal or external.

Evidence-based — conscientiously using current best evidence in making decisions.

Evidence-informed policy — an approach to policy decisions that aims to ensure that decision-making is well informed by the best available research evidence.

Ex ante **evaluation** — an evaluation that is performed before implementation of an intervention. This form of evaluation helps to ensure that an intervention is as relevant and coherent as possible. Its conclusions are meant to be integrated when decisions are made. It provides the relevant authorities with a prior assessment of whether or not issues have been diagnosed correctly, whether or not the strategy and objectives proposed are relevant, whether or not there is incoherence between them or in relation to other related policies and guidelines, and whether or not the expected impacts are realistic.

Ex nunc **(or interim) evaluation** — an evaluation that is performed during implementation.

Ex post **(or final) evaluation** — evaluation of an intervention after it has been completed. It strives to understand the factors of success or failure.

External evaluation — evaluation of a public intervention by people not belonging to the administration responsible for its implementation.

Feasibility — the extent to which valid, reliable and consistent data are available for collection.

Impact — fundamental intended or unintended change and direct or indirect consequences occurring in organisations, communities or systems as a result of programme activities within 7 to 10 years, i.e. long-term consequences of the intervention.

Impact (or outcome) evaluation — evaluates whether the observed changes in outcomes (or impacts) can be attributed to a particular policy or intervention, i.e. determining whether or not a causal relationship exists between an intervention or policy and changes in the outcomes.

Indicator — quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to help assess the performance of a policy/intervention (to reflect the changes connected to an intervention, an output accomplished, an effect obtained or a context variable — economic, social or environmental).

Input — financial, human, material, organisational and regulatory means mobilised for the implementation of an intervention.

Internal evaluation — evaluation of a public intervention by an evaluation team belonging to the administration responsible for the programme.

Joint evaluation — evaluation of a public intervention by an evaluation team composed of both internal (people belonging to the administration responsible for the programme) and external evaluators.

Maryland Scientific Methods Scale — a system that provides an overview of evaluation designs.

Method — complete plan of an evaluation team's work. A method is an ad hoc procedure, specially constructed in a given context to answer one or more evaluative questions. Some evaluation methods are of low technical complexity, while others include the use of several tools.

Monitoring — a continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing intervention with indications of the extent of progress, achievement of objectives and progress in the use of allocated funds.

Need — problem or difficulty affecting concerned groups, which the public intervention aims to solve or overcome.

Norm — level that the intervention has to reach to be judged successful, in terms of a given criterion. For example, the cost per job created was satisfactory compared with a national norm based on a sample of comparable interventions.

Outcomes — the likely or achieved short- and medium-term effects of an intervention's outputs, relating to the aim of the intervention. Specific changes in programme participants' behaviour, knowledge, skills, status and level of functioning.

Outputs — direct products of programme activities which may include types, levels and targets of services to be delivered by the programme.

Process evaluation — one that focuses on programme implementation and operation. A process evaluation could address programme operation and performance.

Programme logic model — picture of how a policy/intervention works — the theory and assumptions underlying the programme. A programme logic model links outcomes (both short- and long-term) with programme activities/processes and the theoretical assumptions/principles of the programme.

Public managers — public (sometimes private) organisations responsible for implementing an intervention.

Random assignment — making a comparison group as similar as possible to the intervention group, to rule out external influences; randomly allocating individuals to either the intervention group or the control group.

Randomised controlled trial (RCT) — an experiment in which two or more interventions, possibly including a control intervention or no intervention, are compared by being randomly allocated to participants.

Relevance — the extent to which an intervention's objectives are pertinent to the needs, problems and issues to be addressed.

Scope — precise definition of the evaluation object, i.e. what is being evaluated.

Stakeholders — individuals, groups or organisations with an interest in the evaluated intervention or in the evaluation itself, particularly authorities that decided on and financed the intervention, managers, operators and spokespersons of the public concerned.

Steering group — the committee or group of stakeholders responsible for guiding the evaluation team.

Sustainability — the continuation of benefits from an intervention after major development assistance has been completed; the probability of continued long-term benefits.

Terms of reference — the terms of reference define the work and the schedule that must be carried out by the evaluation team. They recall the regulatory framework and specify the scope of an evaluation. They state the main motives for an evaluation and the questions asked. They sum up available knowledge and outline an evaluation method. They describe the distribution of the work and responsibilities among the people participating in an evaluation process. They fix the schedule and, if possible, the budget. They specify the qualifications required of candidate teams as well as the criteria to be used to select an evaluation team.

Tool — standardised procedure used to fulfil a function of evaluation (e.g. regression analysis or questionnaire survey). Evaluation tools serve to collect quantitative or qualitative data, synthesise judgement criteria, explain objectives, estimate impacts, and so on.

Validity — the extent to which the indicator accurately measures what is purports to measure.

Value for money — a value for money evaluation is a judgement as to whether the outcomes achieved are sufficient given the level of resources used to achieve them. It generally includes an assessment of the cost of running the programme, its efficiency (the outputs it achieves for its inputs) and its effectiveness (the extent to which it has achieved expected outcomes) and uses analytical approaches such as cost-effectiveness or cost–benefit analyses.

Sources

The principal sources of the definitions contained in the glossary are listed below.

Bureau of Justice Assistance (no date), *Guide to program evaluation* (https://www.bja. gov/evaluation/guide/bja-guide-program-evaluation.pdf).

Cochrane Collaboration (no date), *Glossary* (http://community-archive.cochrane.org/ glossary).

European Commission (1999), *Evaluating socio-economic programmes: glossary of 300 concepts and technical terms*, Office for Official Publications of the European Communities, Luxembourg.

European Commission (2000), *Focus on results: strengthening evaluation of Commission activities*, SEC (2000) 1051, European Commission, Brussels.

European Commission (2013), *EVALSED: the resource for the evaluation of socio-economic development*, European Commission, Brussels.

Farrington, D. P., Gottfredson, D. C., Sherman, L. W. and Welsh, B. C. (2002), 'Maryland Scientific Methods Scale', in Sherman, L. W., Farrington, D. P. et al. (editors), *Evidence-based crime prevention*, pp. 13–21, Routledge, London.

Fretheim, A., Oxman, A. D., Lavis, J. N. and Lewin, S. (2009), 'Support tools for evidence-informed policymaking in health 18: Planning monitoring and evaluation of policies', *Health Research Policy and Systems* 7, pp. 1–8.

OECD (2002), *Glossary of key terms in evaluation and results based management*, OECD Publications, Paris.

W. K. Kellogg Foundation (2004), *Logic model development guide*, W. K. Kellogg Foundation, Battle Creek.

Recommended citation:

About the EMCDDA

The European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) is the central source and confirmed authority on drug-related issues in Europe. For over 20 years, it has been collecting, analysing and disseminating scientifically sound information on drugs and drug addiction and their consequences, providing its audiences with an evidence-based picture of the drug phenomenon at European level.

The EMCDDA's publications are a prime source of information for a wide range of audiences including: policymakers and their advisors; professionals and researchers working in the drugs field; and, more broadly, the media and general public. Based in Lisbon, the EMCDDA is one of the decentralised agencies of the European Union.

Related publications and websites

EMCDDA
| Policy evaluation topic overview
http://www.emcdda.europa.eu/publications/topic-overviews/policy-evaluation

This publication is only available in electronic format.

Publications Office