
**Master of Public Health**

**Master international de Santé Publique**

# Evaluability of public health plans and programmes

May El Khoury

MPH2 (2011)

Haut conseil de la santé publique – HCSP

René Demeulemeester (HCSP)

Françoise Jabot (EHESP)

# Acknowledgements

This thesis would not have been possible without the guidance of my advisors, help from friends, and support from my family.

I am heartily thankful to my academic advisor, Françoise Jabot. Her insightful comments and constructive criticisms were thought-provoking and they helped me focus my ideas. Without her help and guidance it would have been next to impossible to write this thesis.

My deepest gratitude is to my professional advisor, René Demeulmeester. His encouragement, supervision and support from the preliminary to the concluding level enabled me to develop an understanding of the subject.

I am also grateful to the Haut conseil de la santé publique, especially Renée Pomarède, who gave me this opportunity to work on such an interesting project.

Finally, and most importantly, I would like to thank my family, and friends for their support and encouragement.

# Table of contents

# Acronyms

ARS : Agence Régionale de Santé

DREES : Direction de la Recherche des Etudes de l'Evaluation et des Statistiques

EA : Evaluability Assessment

EHESP : Ecole des Hautes Etudes en Santé Publique

HAS : Haute Autorité de Santé

HCSP : Haut Conseil de la Santé Publique

IGAS : Inspection Générale des Affaires Sociales

JCSEE: Joint Committee on Standards for Educational Evaluation

OECD: Organisation for Economic Co-operation and Development

OPEPS: Office Parlementaire d'Evaluation des Politiques de Santé

SEVAL: Swiss Evaluation Society

# Introduction

Evaluation of public policies in France has its origins in the 1970s through the introduction of evaluation in the public decision-making process by rationalizing budgetary choices (Rationalisation des choix budgétaires) with the aim of developing budgetary planning (Perret, 2008). However, it has been officially abandoned in 1984 having lost its influence regarding budgetary decisions (Perret, 2008). Though, a new beginning has been given to evaluation with the report of Patrick Viveret in 1989 which paved the way to an institutionalization of evaluation of public policies (Perret, 2008). Thus, the decree of January 22 1990 led to the creation of the "Comité interministériel de l'évaluation", the "Fonds national de développement de l'évaluation", and the "Conseil scientifique de l'évaluation". The purpose was to make evaluation a key element in the governmental decision-making process and the democratic debate as well as a mean to improve evaluation's methods and ethics (Perret, 2008). From 1990 to 1998, about fifteen evaluations have been undertaken. But the lengthy duration of evaluations, due in part to cumbersomeness of the procedures, gave rise to the decree of November 18 1998. Thus, the "Conseil scientifique de l'évaluation" has been replaced by the "Conseil national de l'évaluation" which was supposed to simplify the evaluation's procedures and make their results more easily usable and exploitable (Perret, 2008). Nevertheless, due to the heavy inter-ministerial procedures, the "Conseil national de l'évaluation" was dissolved in 2007 because it failed to fulfil its mission (Perret, 2008). The creation of the "Société française de l'évaluation" followed in 1999 which developed a charter (2006) with seven guiding principles for evaluation: plurality, impartiality, expertise, respect for persons, transparency, advisability, responsibility (my translation). More recently, the constitutional law of July 23 2008 makes an important step regarding institutionalizing evaluation of public policies: evaluation of public policies will thus be an integrant part of the Parliament missions (Cases, 2009).

In the meantime, the development of evaluation at a national level was followed by a major one at a regional level. This was mainly due to the European commission's directives imposing evaluation of programmes that are financed by structural funds (Cases, 2009).

Within this context, the evaluation of public health policies developed concomitantly with that of public policies. It was first mentioned in 1970 in the evaluation of the perinatal programme which was done upon request of the Ministry of Health (Cases, 2009). Furthermore, in the 1990s an assessment framework has been clearly envisaged for the evaluation of the French law against smoking and alcoholism (Evin, 1991) (Cases, 2009). An evaluation of this same law has also been done, later on, by the National council of evaluation (Perret, 2008).

The development of evaluation and its integration in the French political landscape led to the creation of numerous organisms and institutions involved in evaluation. These organisms and institutions deals with various fields. As mentioned above, the Parliament and the Court of Auditors have in their missions the evaluation of public policies (Cases, 2009). Indeed, in 2002 an "Office parlementaire d'évaluation des politiques de santé" (Opeps) has been created to evaluate public health policies, but it will be dissolved with the constitutional law of July 23 2008. A "Comité d'évaluation et de contrôle des politiques publiques" will, henceforth, be in charge of evaluating public policies (Perret, 2009). Likewise, the Court of Auditors, since 2008, has part of its missions the evaluation of public policies (Cases, 2009). Furthermore, the "Inspection Générale des Affaires Sociales" (IGAS), is the inter-ministerial audit and evaluation office for social, health, employment and labour policies[1]. Additionally, the "Direction de la recherche, des études, de l'évaluation et des statistiques" (DREES) promotes evaluation of results and social and economical impacts of public policies. It also coordinates the design, the validation and the implementation of the evaluation methods[2]. More specific agencies focus on evaluations pertaining to the health sector such as the French National Authority for Health (HAS) which has the specificity of medico-economic evaluation as well as the evaluation of professional practice[3]. At a regional level, the "Agence régionale de santé" (ARS) with its "Mission inspection, évaluation, contrôle" is in charge of evaluating regional health plans and programmes[4].

Furthermore, the law of August 9 2004 set up the "Haut conseil de la santé publique" (HCSP). This newly created institution is the result of the merge of the "Haut conseil de la santé with the "Conseil supérieur de l'hygiène publique de France" (Fouchard, 2010). The "Conseil supérieur de l'hygiène publique de France" has been created by the law of January 29 1906, its missions was related to issues pertaining to water and sanitation, hygiene, food, vaccines etc. However, the "Haut comité de la santé publique" was created by the decree n°91-1216 of December 3 1991. It was an instance of expertise for public health issues till 2001, when the law of March 4 2002 replaced it by the "Haut conseil de la santé". Yet, this newly created council will end up never seeing the light of day (Fouchard, 2010). Hence, the missions of the HCSP resume that of the "Haut comité de la santé publique" and the "Conseil supérieur de l'hygiène publique de France" (Law of August 9 2004). Despite its role of expertise, the HCSP has, among its missions, a role of evaluation of public health objectives as assigned by the law of 2004 (Law of August 9 2004). Indeed, among the six specialised commissions constituting the HCSP, the "commission évaluation, stratégie et prospective" is

[1] http://www.igas.gouv.fr/spip.php?article164
[2] http://www.sante.gouv.fr/IMG/pdf/Plaquette_DREES_oct2010.pdf
[3] http://www.has-sante.fr/portail/jcms/c_5443/english?cid=c_5443
[4] http://www.ars.sante.fr/La-Mission-inspection-evaluat.93955.0.html

in charge of evaluating national health plans and programmes. Thus, since 2007 – the official date of the establishment of the HCSP – several evaluations have been performed, such as the evaluation of the Cancer plan 2003 - 2007, the Cancer plan 2009-2013, the Programme against tuberculosis or the Rare diseases plan. So far, the decision of evaluating public health plans and programmes was either done upon request of the Minister of Health, or based on a self-referral. In the latter case, the HCSP decides by itself to undertake the evaluation of a particular plan or programme. The evaluations are mostly done after the programme is complete (*ex-post* evaluation), but other evaluations could also be performed such as midterm evaluations[5] or *ex-ante* evaluations[6].

Within this context and with the relatively high number of national public health plans or programmes carried out, the HCSP is looking for a typology of those plans and programmes that shall help the decision-making regarding the evaluation or not of those plans and programmes. A first demand has been done by asking for an elaboration of such a typology, based on a literature review and the identification of the plans and programmes to be considered for an evaluation. Some classification criteria have also been proposed by the HCSP such as the plans' objectives, the population concerned, and the type of funding.
Hence, several questions have been raised in order to clarify the request that has been made initially: what is the level of interest, the utility and the relevance of evaluating a health plan or programme? To which extent is the evaluation feasible and what are its assessment criteria?
Thus it was worth pointing out the importance of defining evaluation and the concept of evaluability. According to Patton (1997) "program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgements about the program, improve program effectiveness, and/or inform decisions about future programming". Whereas in France, the official definition for evaluation is given in the decree of November 18 1998: "the aim of evaluating public policies is to assess the efficiency of that policy by comparing its results to the objectives assigned and the means that have been used" (my translation). Thus evaluations could be initiated for many reasons. They may be intended for guiding programme improvement; gain knowledge about the programme effects; provide input to decisions about programme's structure, funding or administration; or respond to political pressure (Rossi, 2004).
On the other hand, evaluability is defined as the extent to which a programme can be evaluated in a reliable and credible fashion (OECD, 2002). The International Development

[5] Midterm evaluation is an evaluation that is performed towards the middle of the period of implementation of the programme.
[6] *Ex-ante* evaluation is an evaluation that is performed before implementation of a programme.

Research Centre (2006) defines evaluability assessment as “an assessment of whether an evaluation would be feasible and useful”.

The first documentary research performed shows that a tool that will help assess rapidly evaluability of public health plans and programmes does not exist.

Therefore, after negotiating with the HCSP, a new proposal has been done with the aim of building upon a multi-institutional team of experts with relevant experience in public health evaluation, so as to elaborate a practical tool that will help assess rapidly evaluability of public health plans and programmes. The aim of the project will be to construct this tool.

Therefore, through building this tool, the objective of this thesis will be to explore the applicability of the concept of evaluability in the health sector in France.

# Material and Method

To assess the evaluability of public health plans and programmes and to explore the applicability of this concept, a multi step method is needed.

### 1. Documentary research

Firstly a thorough documentary research has been done with use of the following keywords: “evaluability”, “evaluability assessment” and “evaluation standards”. A careful analysis of the literature has then been done so as to clarify the concept of evaluability and evaluability assessment, and to identify the criteria and methodology used to assess evaluability.

### 2. Delphi study

The two main items of evaluability assessment were identified (Utility and Feasibility) through the literature review. Therefore, in order to validate the concept of evaluability, we proposed to a group of fourteen experts to participate in a two rounds Delphi study. In the first round, experts were asked to give three feasibility criteria and three utility criteria which will assess, respectively, the feasibility and the utility of the evaluation of a public health plan or programme. In the second round, the experts had to choose, among all the answers given on the first round, two main criteria of feasibility and two of utility, and had to identify at least one indicator for each chosen criterion.

The chosen experts have relevant experience in evaluation of health programmes and come from different health institutions such as the “Direction générale de la santé”, the “Agence régionale de santé”, the “Direction générale de l’offre de soins”, the “Direction de la recherche, des études, de l’évaluation, et des statistiques”, the “Inspection générale des affaires sociales”, the “Institut national de prévention et d’éducation de la santé”, the “Institut national de la santé et de la recherché médicale”, the HCSP, and the Ecole des Hautes études en santé publique (EHESP).

### 3. Tool construction

The construction of the evaluability assessment tool was draw on:

- The findings of the literature review
- The criteria and indicators defined in the Delphi study

## 4. Pre-test of the tool

The tool has been pre-tested with four selected public health plans and programmes. These plans and programmes had already been evaluated by the HCSP. Thus, the evaluability assessment (EA) results obtained with the tool will be compared to the evaluations' findings in order to confirm the tool's relevance. Each of the chosen plans and programmes had been evaluated on different stages of the implementation of the programme: *ex-ante* evaluation, midterm evaluation[7] and *ex-post* evaluation[8]. This selection has been made in order to test the applicability of the constructed tool on different types of evaluation on one hand and to confirm the obtained results by the tool on the other.

The four selected plans and programmes are:

- Programme against tuberculosis 2007-2009 (*ex-post* evaluation)
- National stroke plan 2010-2014 (*ex-ante* evaluation)
- Cancer plan 2003 – 2007 (*ex-post* evaluation)
- Cancer plan 2009 – 2013 (midterm evaluation)

The technique to fill in the tool for each plan and programme consisted in two main steps:

- The analysis of the document describing the plan or programme.
- The referral including the evaluation's objectives.
- An interview with the programme manager or coordinator. Indeed, the person to be interviewed should have a very good knowledge of the plan or programme to answer our questions.

As a first step, the programme coordinator is identified in order to fix an appointment for the interview. Meantime, the evaluator fills in the tool and draws the programme logical model through reading and analysing the programme document and referral. The remaining answers are completed after doing the interview.

## 5. Validation of the tool with the experts

Throughout the tool construction and pre-test process, monthly meetings were organised with the working team of experts to present them the findings. The working team expertise helped fine-tune the tool and thus, validate it.

[7] Midterm evaluation is an evaluation that is performed towards the middle of the period of implementation of the programme.

[8] *Ex-post* evaluation is the evaluation of a programme after it has been completed.

# Results

## 1. Description of the tool

### a. Literature review

The literature review performed, revealed that EA developed in the United States in the 1970s for the aim of improving programmes' structure and making them, if necessary, more evaluable (Mathison, 2005). Wholey (2004) describes EA as a "process that helps evaluators to identify evaluations that might be useful, explores what evaluations would be feasible and design useful evaluations". He outlines six steps for conducting evaluability assessment: involving intended users of evaluation information; clarifying the intended programme from the perspective of policymakers, managers, those involved in service delivery, and other stakeholders; exploring programme reality, including the plausibility and measurability of programme goals; reaching agreement on any needed changes in programme activities or goals; exploring alternative evaluation designs; agreeing on evaluation priorities and intended uses of information on programme performance (Wholey, 2004). Hence, it is an analysis of the quality of the programme, a process that aims to clarify its design, explore its reality, and if necessary, help redraft it to meet the criteria to be evaluated.

Outcomes from EA often noted in the literature include the development of a programme theory that contains programme logic and performance measures, the explanation of programme goals and objectives, and the documentation of stakeholder perception, understanding, and interest in the programme (Rutman, 1980).

To explore how the evaluation would be feasible, Wholey (2004) outlined, in the different steps, the elements that one should take into account to assess its feasibility. Likewise, he points out the importance of involving stakeholders from the very beginning of the evaluation process in order to reach agreement among programme managers and policymakers on the specific uses that would be made of the evaluation findings (Wholey, 2004). Thus, the information produced by the evaluation is more likely to be used and therefore, increases the utility of the evaluation performed.

Many of the described steps above for assessing evaluability of programmes are present in many evaluation designs. As illustrated by Wholey (2004) this process might take months to be completed.

Moreover, the American Joint Committee on Standards for Educational Evaluation (JCSEE) (Yarbrough et *al*. 2011) in their guide presenting the programme evaluation standards, the utility and feasibility standards are on the top of the list. Those standards have been approved by the American National Standards and have been adopted – in their earlier

version (JCSEE, 1994) – as quality evaluation standards by many evaluation associations throughout the world, such as the Canadian Evaluation Association (2006), the African Evaluation Association (2002), and the Swiss evaluation association (SEVAL, 2000).

The feasibility standards suggested are "intended to increase evaluation effectiveness and efficiency" (Yarbrough et *al*. 2011). Four feasibility standards are identified. The first one focuses on a comprehensive evaluation plan that should include attention to all attributes and standards of evaluation quality. In addition, evaluation procedures should be practical and responsive to the way the programme operates. Evaluations are planned and implemented taking into account the cultural, political interests and needs of individuals and groups in order to obtain their cooperation and avoid any attempts to restrict the activities of the evaluation and a misuse of the findings (Contextual viability of the evaluation). Finally, evaluation should use resources effectively and efficiently (Yarbrough et *al*. 2011).

The utility standards, however, "are intended to increase the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs" (Yarbrough et *al*. 2011). The first standard addresses evaluator characteristics. The evaluation should be conducted by qualified people who establish and maintain credibility in the evaluation context. The other utility standards depicted, emphasize on the importance that all stakeholders that could be affected by the evaluation should be identified and involved in the evaluation process, and all the evaluation purposes should also be identified and negotiated based on the needs of stakeholders. Thus, the evaluation information should serve the identified and emergent needs of stakeholders. The activities, descriptions and judgements developed through evaluation should encourage participants to rediscover, reinterpret, or revise their understandings and behaviours. Besides, the evaluation information should be adapted to the needs of their multiple audiences and hence the evaluation report should be clear and concise to a better appropriation of the findings. To finish, evaluation results should be communicated to the intended users in a timely and appropriate manner (Yarbrough et *al*. 2011).

On the other hand, research on utilization of evaluation identified three main factors associated with the likelihood of greater evaluation utilization namely characteristics of potential users, the context in which evaluation takes place, and the evaluation itself (Mathison, 2005). Researchers found that utilization of evaluation depends on users' expectations for the evaluation, their prior experience and current disposition toward evaluation, and their perception of the risks they faced in having the evaluation performed. The context in which an evaluation takes places, such as characteristics of the programme, intraorganizational features, and factors external to the programme can have impact on use of evaluation information. Finally, the way in which the evaluation is conducted may also

exert influence. The different aspects of evaluation identified include the procedures used (methods and type of evaluation model), information dialogue (the amount and quality of interaction between evaluators and potential users), the nature of the evaluation information (its relevance and specificity), and evaluation reporting (the frequency and timing of the information and the style and format of the report) (Mathison, 2005).

Moreover, Patton (1997) describes what he calls “Utilization-Focused Evaluation” as an evaluation which goal is “intended use (of the evaluation results) by the intended users”. An evaluation-focused approach requires the identification and the involvement of intended users in the choice of the type of evaluation and the methods to be used to conduct the evaluation. Their involvement in interpreting findings, making judgments based on the data, and generating recommendations. Indeed, researches on use shows that intended users, when actively involved, understand and feel ownership of the evaluation process and findings, and thus, are more likely to use evaluation results. And by actively involving intended users, the evaluator is preparing the groundwork for use, and reinforcing the intended utility of the evaluation (Patton, 1997).

### b. Delphi results

Twelve answers have been collected in the first round and eleven answers in the second round of the Delphi study.
The first round pinpointed eight feasibility criteria and seven utility criteria.

<u>The feasibility criteria are:</u>

- Favourable conditions for evaluation
- Actors’ willingness to participate to the evaluation process
- Identification of programme managers and stakeholders
- Accessibility of data
- Quality of the indicators
- Resources available for the evaluation
- Extent to which the evaluation questions are realistic
- Design of the plan or programme

The utility criteria are:

- Reiteration of the plan or programme
- Relevance of evaluation for stakeholders
- Usefulness of evaluation for decision making
- Importance of the health issue underlying the plan or programme
- Importance of the plan or programme
- Importance of the allocated resources for the plan or programme
- Knowledge production

Three utility criteria and three feasibility criteria have been identified with their indicators at the end of the second round of the Delphi study. These criteria have been classified according to the number of respondents who have selected them. The indicators have been categorized by degree of similarity.

Feasibility criteria with their indicators:

- Access to data
  - Quality of the information system
  - Quality of the follow-up process
- Structure of the health plan or programme
  - Initial inventory
  - Clarity of the objectives
  - Coherence of the objectives
  - Quality of the plan or programme indicators
- Context of the evaluation
  - Stakeholders' willingness to participate in the evaluation process
  - Evaluation authority

Utility criteria with their indicators:

- Utility of evaluation for decision-making
  - Importance of utilisation of evaluation's results for decision-making
  - Decision-makers expectations from the evaluation
- Importance of the health plan or programme
  - Importance of allocated resources
    - Importance of the budget allocated
    - Number of actors mobilized
  - the problem underlying the plan or programme
  - Reiteration of the plan or programme
- Relevance of the evaluation for stakeholders
  - Level of commitment of stakeholders to the evaluation process
  - Plan or programme controversies

### c. The tool

The tool developed includes three main parts drawn on the criteria identified in the literature and the Delphi study. The three parts are:

- The feasibility assessment of the evaluation of a public health plan or programme
- The utility assessment of the evaluation of a public health plan or programme
- The stakeholders' commitment to the evaluation of a public health plan or programme

Each part contains the criteria and their indicators as identified by the Delphi study. Several questions have been developed for each of the selected indicators in order to construct the tool.

The third part of the tool: "Stakeholders' commitment to evaluation" is the result of the merger of two criteria identified by the Delphi study: "The context of the evaluation" included in the feasibility assessment part, and "The relevance of the evaluation for stakeholders" contained in the utility assessment part. Thus being a transversal criterion of evaluability assessment the decision was made to put it under a separate part.

Below is the translation from French of the framework of the tool. The questions have not been translated. The complete version of the tool in French is in the appendix.

*Question proposed for the evaluability assessment of public health plans and programmes*

The feasibility assessment of the evaluation of a public health plan or programme

- Access to data
  - *Quality of the information system*
  - *Quality of the follow-up process*
- Structuring of the plan or programme
  - *Reconstruct the logic model of the plan or programme*
  - *Initial inventory*
  - *Coherence of the objectives*
  - *Clarity of the objectives*
  - *Quality of the indicators*
- Availability of resources for the evaluation

The utility assessment of the evaluation of a public health plan or programme

- Utility of evaluation for decision-making
  - *Decisions conditioned by the evaluation*
  - *Explicit expectation of decision makers*
- Importance of the plan or program
  - *Importance of the intended and allocated resources*
  - *Importance of the problem underlying the plan or programme*
  - *Other important criteria*

The stakeholders' commitment to the evaluation of a public health plan or programme

- *Possibility of involving stakeholders*
- *Evaluation authority*
- *Willingness of actors to follow up the evaluation*

## 2. Pre-test of the tool

### a. Pre-test results

The pre-test was done on four plans and programmes. One of the tested programmes (Cancer plan 2009-2013) is currently being evaluated which will not allow the comparison of evaluability findings *versus* evaluation findings.

- **Program against tuberculosis**

Evaluability assessment results

The programme was not conceived in a way to be evaluated. The programme logic model built, shows gaps between the activities described and the expected outcomes. Few indicators were identified for the objectives. Furthermore, the quality of the information system, as assessed, is questionable. On the other hand, although stakeholders agreed on the evaluation decision, the intended use of the evaluation findings is limited to the deletion policy of BCG vaccination. Besides, there were no specific resources allocated to the programme, neither to the evaluation.

Thus an evaluation of the programme's outcomes is not likely to be feasible. However, an evaluation of the programme activities and its implementation would be feasible if the identified indicators have been installed. The utilization of evaluation findings was limited to the deletion policy of BCG vaccination.

- **National stroke plan**

Evaluability assessment results

The programme logical model is not explicit. The objectives detailed in the programme are not clearly defined. Very few indicators have been identified. However, there is a good quality of the information system, and a complete initial inventory is done that could be used for the evaluation. The stakeholders identified are willing to participate to the evaluation. Yet, the intended uses of findings concerns only the development of indicators.

- **Cancer plan 2003-2007**

Evaluability assessment results

Although an evaluation of the plan was expected, it is noticed that the plan has not been conceived in a way to facilitate its evaluation. There is no documentation on the follow up process of the plan combined with limitations of the information system. The programme logical model is not explicit with very few indicators identified. Further, the programme coordinator interviewed perceives evaluation as a judgemental approach and thus considers that his participation to the evaluation should only be limited so as to provide the necessary information needed. However, although a utilization of the findings was expected in order to provide recommendations and guidance for the next plan, the time allocated to give preliminary results was not sufficient (less than two months).

Hence, the evaluation of the plan is more or less feasible in regard to the available data and the programme's logic and the use of findings may be limited due to time constraint.

- **Cancer plan 2009-2013**

Evaluability assessment results

There is a good quality of the information system and follow up process thus providing required data for the evaluation, in addition to a full array of indicators. However, the programme's logic is not explicit and the objectives are not clearly defined. A utilization of findings is expected so as to reorient the objectives and develop new activities. The time allocated for the evaluation is more or less sufficient (six months).

Evaluation of this plan, compared to Cancer plan 2003-2007, seems to be more feasible with an intended uses of the evaluation findings.

**b. Evaluability assessment versus Evaluation**

The findings of evaluation were found to be coherent with those of evaluability assessment.

**Program against tuberculosis**

| Evaluability assessment | Evaluation |
|---|---|
| • Programme's logic model not explicit<br>• Quality of information system questionable<br>• Not enough indicators for the evaluation<br>• Limited resources (for programme and evaluation)<br>• Utilization of evaluations findings limited | • Lack of information on the programme implementation<br>• Lack of available and usable data for the evaluation<br>• Limited resources (for programme and evaluation)<br>• Programme steering was limited<br>• Utilization of evaluations findings limited |

**National stroke plan**

| Evaluability assessment | Evaluation |
|---|---|
| • Programme's logic model not explicit<br>• Few indicators identified<br>• Stakeholders willing to participate in the evaluation<br>• Utilization of evaluations findings limited<br>• Good quality of the information system | • Programme's logic model not explicit<br>• Few indicators identified<br>• Stakeholders participated to the evaluation<br>• Utilization of evaluations findings limited |

**Cancer plan 2003-2007**

| Evaluability assessment | Evaluation |
|---|---|
| • Limitations of the information system<br>• Few indicators identified<br>• Programme's logic model not explicit<br>• Utilization of evaluations findings limited | • Limitations of the information system<br>• Few indicators identified<br>• Programme steering was limited<br>• Utilization of evaluations findings limited |

# Discussion

Our study helped us construct a tool to assess evaluability of public health plans and programmes. Three main points will be discussed: pros and cons of the tool, pre-test process and other applications of the tool.

### 1. Pros and cons of the tool

Through pre-testing the tool has been fine-tuned with the experts working team. Thus, it has become more precise and relevant.

Yet, in the second part of the assessment framework, when assessing the utility of the evaluation, the “Importance of the problem underlying the plan or programme” was not found discriminating so as to decide if the evaluation was to be performed. Whether it was in the programme against tuberculosis, where the health issue is only of relative importance in the case of France, or in the Cancer plan where the health issue is among the leading causes of death, it has not affected the evaluation decision.

Furthermore, these pre-tests have proven the facility of use of the tool. The answers were easily found in the plan or programme document. The interview with the programme coordinator was important to corroborate the information previously found in the documents on one hand, and to complete the answers on the other.
The most time-consuming step when filling the assessment framework is the drawing of the programme logic model. Thus, the scope and size of the plan or programme will affect the time spent to draw its logic model. However, the amount of time spent to fill in the tool did not exceed two weeks. Hence, this tool meets our objective to assess rapidly the evaluability of the plan or programme.

During the pre-test on *ex-ante* evaluation, some questions were found to be irrelevant. Since in this case it is not appropriate to assess the evaluability of the plan or programme, the information gathered from the tool, allowed mainly through building the programme logic model, the identification of gaps between programme’s activities and expected outcomes. Hence the assessment framework highlights the elements of the programmes (Objectives, indicators...) that should be revised to make the programme more evaluable.

Selecting plans and programmes that have been evaluated permitted the comparison between tool’s results and previous evaluation findings. The conformity of the conclusions

made when using the assessment framework with the evaluation findings permitted to validate the results obtained with the tool and to assert its predictive value.

As a result, the assessment framework developed has many qualities and seems to meet the initial objectives.

## 2. The unexpected outcomes of the process

Testing the tool on the four selected programmes, and more particularly when drawing the programme logic model, demonstrated the actual gaps between the described activities and the expected outcomes of the programme. This problem is of significant importance when planning for future evaluation. In fact, building the logic model will be the first step of conducting an evaluation that allows a better understanding of the intervention's strengths and weaknesses on one hand, and analyzing whether the intervention is designed in a way that can logically produce the desired results (Brousselle, 2010).

Further, a participative approach of the evaluation by involving stakeholders to the process has been debated through meetings with the expert team. Indeed, during the interviews, we came to the conclusion that some stakeholders do not appreciate the benefits of participating to the evaluation process. Their vision of evaluation is that of a judgemental procedure, in that they consider that they do not have to participate to the evaluation process. In contrast, many studies highlight the importance of creating participatory spaces throughout the evaluation process as a mean of appropriating the results and thus contributing to a better utilization of findings (Bilodeau, 2009).

The referral, for its part, turned out to be a starting point to clarify evaluation purposes and context. In fact, it is of great importance to assess the aim of the evaluation and the context within which it takes place, as well as the intended use of evaluation findings, in order to choose the appropriate evaluation model and provide timely results.

Hence, through highlighting the importance of building the programme logic model, and sharing of a participatory culture of evaluation, the tool comes to be used for other purposes than the one it was originally assigned.

## 3. Thoughts on possible fallout of the tool

This tool was initially developed to assess evaluability of public health plans and programmes; however it might have other functionalities.

Firstly, its utilization is useful for the evaluator because it gives him an overview of the plans or programmes to be evaluated, helps formulate evaluation questions that the evaluation might address and thus choose appropriate evaluation procedures.

Furthermore, this tool originally designed for the specific case of national public health plans and programmes is amenable to evolve and therefore might adapt to smaller-scale programmes in other areas than that for which it was intended. Its spread to other fields of evaluation could be done, in a first step, by the members of the working group.

However, the context is an important element to take into account since it will have an impact on the applicability of the tool. Indeed, feasibility and utility of the evaluation are not only dependent on programme's quality, they are influenced by the context of the evaluation and the way the evaluation is undertaken. Thus, the tool should be tested on a larger number of programmes and plans, to better assess its applicability.
Furthermore, the applicability of the tool is also dependent on its user. In order to assess the extent to which the findings may vary between different users, it would be necessary to conduct inter-rater reliability studies. This will be useful to refine the tool and further make it generalizable.

# Conclusion

The tool developed through this project permitted to assess the extent to which an evaluation is feasible and useful. Thus, the evaluability assessment concept showed to be applicable in the French health sector. This tool will definitely evolve and therefore might adapt to other areas than that for which it was intended. In addition, its utilization could be done for other purposes than those assigned when initially developed such as giving feedback on how programmes are built through drawing the programme logic model, or deciding on evaluation procedures depending on the intended use of results.

However, the applicability and the scope of application of the tool are context and user dependent. Therefore, inter-rater reliability studies will help assess and improve these issues.

# Bibliography


African Evaluation Association. The African Evaluation Guidelines. (accessed on June 2011) *in* http://www.afrea.org/content/index.cfm?navID=5&itemID=204 (accessed on June 2011)

Bilodeau A, Allard D, Gendron S, Potvin L. Les dispositifs de la participation aux étapes stratégiques de l'évaluation. The Canadian Journal of Program Evaluation 2006 ; 21(3) :257-82 .

Brousselle A, Champagne F. Program theory evaluation: Logic analysis. Evaluation and program planning. 2011 ; 34(1):69-78.

Cases C, Gremy I, Perret B. L'évaluation en santé publique en France: diversification et consolidation *in* L'évaluation en santé publique. Actualité et dossier en santé publique. La documentation Française ; 2009.

Décret n°90-82 du 22 janvier 1990 relatif à l'évaluation des politiques publiques.

Décret n°91-1216 du 3 décembre 1991portant création du Haut Comité de la santé publique.

Décret n°98-1048 du 18 novembre 1998 relatif à l'évaluation des politiques publiques.

Fouchard A. Le Haut Conseil de la Santé Publique : Historique, missions, fonctionnement et comparaisons internationales. [Mémoire d'internat en santé publique]. Rennes : Ecole des Hautes Etudes en Santé Publique; 2010.

Haut conseil de la santé publique. Evaluation du programme national de lutte contre la tuberculose 2007-2009. La documentation Française ; 2010.

Haut conseil de la santé publique. Evaluation du plan Cancer 2003-2007 *in* http://www.sante.gouv.fr/IMG//pdf/rapport_cancer_230209.pdf

International Development Research Centre. Glossary, 2006 *in* http://www.idrc.ca/cp/ev-121349-201-1-DO_TOPIC.htm

Joint Committee on Standards for Educational Evaluation. The program evaluation standards. 2nd ed. Sage ; 1994.

Loi constitutionnelle n°2008-724 du 23 juillet 2008 de modernisation des institutions de la Ve république.

Loi n°92-32 du 10 janvier 1991 relative à lutte contre le tabagisme et l'alcoolisme.

Loi n°2004-806 relative à la politique de santé publique.

Loi n°2002-303 du 4 mars 2002 relative aux droits des malades et à la qualité du système de santé.

Mathison S. Encyclopedia of evaluation. Thousand Oaks, London: Sage; 2005.

Ministère de la santé et des sports. Plan cancer 2003-2007 (accessed on June 2011) *in* http://www.afsset.fr/upload/bibliotheque/564530097084408659387032910775/plan_cancer.ppd

Ministère de la santé et des sports. Plan cancer 2009-2013 (accessed on June 2011) *in* http://www.sante.gouv.fr/IMG/pdf/Synthese_plan_cancer_2009_2013.pdf

Ministère de la santé et des sports. Préparation du plan d'action AVC 2010-2014 (accessed on June 2011) *in* http://www.sante.gouv.fr/IMG/pdf/plan_actions_AVC_-_17avr2010.pdf

Ministère de la santé et des sports. Programme national de lutte contre la tuberculose 2007-2009 (accessed on June 2011) *in* http://www.sante.gouv.fr/IMG/pdf/prog_tuberculose_2007_2009.pdf

Organisation for Economic Co-operation and Development. Glossary; 2002 (accessed on June 2011) *in* http://www.oecd.org/dataoecd/29/21/2754804.pdf

Patton MQ. Utilization-Focused Evaluation : The Next Century Text, London : Sage Publications; 1997.

Perret B. L'évaluation des politiques publiques. Paris, La Découverte, collection Repères ; 2008.

Perret B. Acteurs, objets et méthodes de l'évaluation. *in* L'évaluation en santé publique Actualité et dossier en santé publique. La documentation Française ; 2009.

Rossi PH, Lipsey MW, Freeman HE. Evaluation: a systematic approach. 7th ed. Sage 2004.

Rutman L. Planning useful evaluations: Evaluability assessment. London. Sage; 1980.

Société Canadienne d'évaluation. Normes d'évaluation des programmes (accessed on June 2011) *in* http://www.evaluationcanada.ca/site.cgi?section=5&ssection=4&_lang=fr

Société Française d'Evaluation, Charte de l'évaluation. 2006 (accessed on June 2011) *in* http://www.sfe-asso.fr/intranet/ckfinder/userfiles/files/sfe_charte-evaluation_06-08.pdf

Société Suisse d'Evaluation, Standards SEVAL 2000 *in* http://www.seval.ch/fr/standards/index.cfm

Viveret P. L'évaluation des politiques et des actions publiques. Paris : La Documentation Française, coll. Rapports officiels. 1989.

Wholey JS, Hatry HP, Newcomer KE. Handbook of Practical Program Evaluation. San Francisco: Jossey-Bass; 2004.

Yarbrough DB, Shulha L M, Hopson R K, *et al.* The program evaluation standards: A guide for evaluators and evaluation users. 3rd ed. Thousand Oaks, CA: Sage; 2011.

# APPENDIX

# *Questions proposées pour l'appréciation de l'évaluabilité des plans et programmes de santé*

## I. Appréciation de la Faisabilité de l'évaluation d'un plan ou programme de santé :

### a. Accessibilité des données

#### *Qualité du système d'information*

- Existe-t-il un inventaire des sources de données susceptibles d'être utilisées pour l'évaluation ?
- Un protocole de recueil de données portant sur les indicateurs retenus pour le suivi du plan est il prévu ?
- Les méthodes de mesure prévues dans ce protocole sont elles fiables ?
- Les données disponibles sont ou seront elles intégrées dans un tableau de suivi ?

**COMMENTAIRES :**

#### *Qualité du suivi*

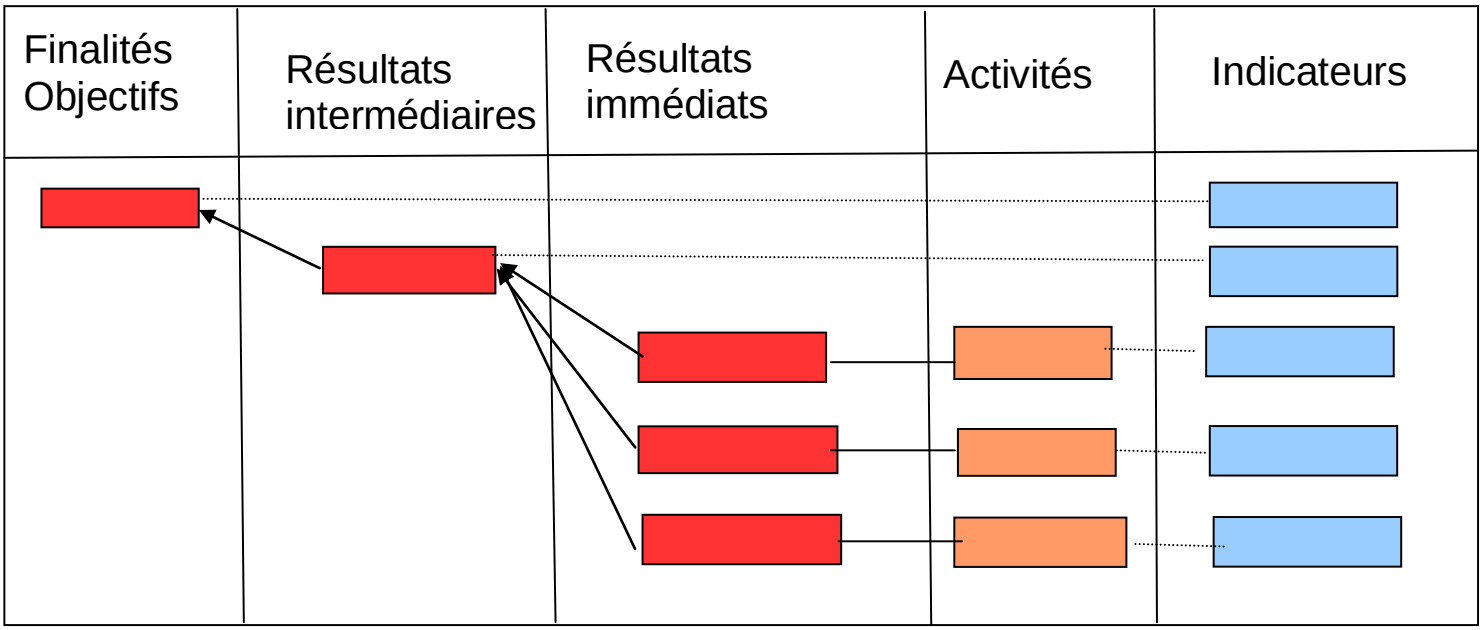
- Une instance de suivi du plan responsable du bon déroulement du programme et du processus de suivi est elle prévue ?
- Le travail de l'instance de suivi est il effectif ?
- Un suivi annuel des indicateurs est il prévu ?
- Ce suivi est il effectif ?
- Le suivi fournit-il ou fournira-t-il des informations sur :
  - Le niveau de mise en œuvre des actions ?
  - Les dépenses engagées ?
  - Les intervenants mobilisés ?
  - La population atteinte ?
  - L'opinion des intervenants et des populations concernées sur les activités ?

**COMMENTAIRES :**

## Structuration du plan ou programme

### Reconstituer le cadre logique du programme

Les réponses à cette partie du questionnaire nécessitent de reconstruire le cadre logique du programme, au minimum selon le modèle suivant :




### *Etat des lieux initial*

- Est ce qu'il existe un état des lieux :
  - de l'état de santé relatif à la question traitée par le plan ou le programme (incidence, prévalence, conséquences) ?
  - de la population concernée ?
  - des déterminants qui influencent cet état de santé ?
  - des leviers qui peuvent influencer ces déterminants ?
  - des intervenants mobilisables ?
  - des partenaires impliqués ?
  - des pratiques / organisations concernées par le plan ?
  - des ressources à engager ?

**COMMENTAIRES:**

### *Cohérence des objectifs*

- Les objectifs sont-ils cohérents avec l'état des lieux ?
- Existe-t-il différents niveaux d'objectifs ?
- Y a-t-il une cohérence entre les niveaux d'objectifs (cf. cadre logique) ?
- Les hypothèses qui sous-tendent les liens entre les niveaux d'objectifs sont elles explicites ?

**COMMENTAIRES:**

### *Clarté des objectifs*

- Les objectifs précisent-ils :
  - les facteurs à changer ?
  - la population concernée ?
  - l'importance des effets escomptés ?
  - le délai pour obtenir ces effets ?
- Quelle est la proportion d'activités ou de mesures pour lesquelles des objectifs opérationnels ont été formulés en termes de résultats à atteindre ?

**COMMENTAIRES :**

### *Qualité des indicateurs*

- Quelle est la proportion d'objectifs spécifiques :
  - assortis d'indicateurs ?
  - pour lesquels on dispose d'une mesure avant la mise en œuvre du plan ou programme?
  - mesurables par un système d'information pérenne ?
- Quelle est la proportion d'actions ou de mesures assorties d'indicateurs de suivi ?
- Quelle est la proportion d'indicateurs produits par des organismes reconnus dans la production de données et utilisés par ailleurs ?

**COMMENTAIRES :**

## b. Ressources nécessaires à l'évaluation

- Des ressources suffisantes sont-elles affectées à la réalisation de l'évaluation ?
- Un délai suffisant est-il disponible pour l'évaluation ?

**COMMENTAIRES**

# II. Appréciation de l'Utilité de l'évaluation d'un plan ou programme de santé

## a. L'utilité de l'évaluation pour la décision

### *Décisions conditionnées par l'évaluation :*

- Une utilisation des résultats de l'évaluation est-elle prévue ou annoncée (demande politique ou opinion publique) ?
- Les résultats de l'évaluation seront-ils utilisés dans un but de modification du plan ou programme :
  - Réorientation des objectifs ?
  - Mise en place de nouvelles mesures ?
  - Révision de la structure du programme ?
  - Révision du mode de pilotage ?
  - Modifications des ressources affectées au programme ?
- Les résultats de l'évaluation seront-ils susceptibles d'influencer des dispositifs ou actions de santé publique autres que ceux prévus par le plan ou programme ?
- Les résultats pourront-ils être communiqués aux utilisateurs prévus de telle sorte que ceux-ci puissent en faire un usage en temps utile ?

**COMMENTAIRES :**

### *Attente explicite des décideurs :*

- Y a-t-il une déclaration explicite du décideur (ministre, directeur,…) annonçant l'élaboration d'une nouvelle politique ou plan tenant compte de l'évaluation ?
- Des questions évaluatives ont-elles été formulées par le commanditaire ?
- L'évaluation demandée porte t elle sur :
  - La totalité du plan ?
  - Une partie du plan ?
  - Ou uniquement sur une mesure spécifique du plan ?

**COMMENTAIRES :**

## b. Importance du plan ou du programme :

### *Importance des moyens prévus et engagés :*

- Un budget est-il prévu pour le plan ou programme ?
- Les moyens financiers prévus pour le programme ont-ils été alloués ?
- Est-ce que l'ensemble des partenaires et acteurs de santé qui pourraient être impliqués dans le plan ou programme l'ont été ?

**COMMENTAIRES :**

### *Importance du problème à l'origine de plan ou du programme :*

- Le problème de santé à l'origine du plan ou programme est-il important au regard de critères tels que :
    - Mortalité ?
    - Morbidité ?
    - Années de vie perdues ?
    - DALYS ?

- Le coût engendré par le problème de santé à l'origine du plan ou programme est-il important :
    - Coût pour l'Assurance Maladie ?
    - Coût pour les patients ?
    - Coût pour leur famille ?

**COMMENTAIRES :**

### *Autres critères d'importance :*

- Le plan ou programme à évaluer constitue-t-il un plan stratégique de santé publique ?
- Une reconduction du plan ou programme est-elle prévue à l'avance ?
- Existe-t-il une thématique ou une stratégie jugée à l'origine, particulièrement innovante ou expérimentale dans le plan ou programme à évaluer ?

**COMMENTAIRES :**

## III. Adhésion des acteurs

### *Possibilité de participer des acteurs*

- Le pilote et les acteurs du plan ou programme sont ils clairement identifiés et joignables ?
- La décision d'évaluation fait elle l'objet d'un consensus de la part des acteurs ?
- Les intervenants ont ils contribué à l'élaboration des questions auxquelles l'évaluation devra répondre ?
- Sont-ils d'accord sur ces questions ?

### *Instance d'évaluation*

- Une instance d'évaluation est-elle prévue ?
- Sera-t- elle composée des parties prenantes (organismes de financement, les porteurs du projet, intervenants, population concernée, experts externes) ?
- Sera-t-elle présidée par une personnalité indépendante de l'institution commanditaire responsable du programme ?

### *Volonté des acteurs de donner suite à l'évaluation*

- Les acteurs (promoteur, gestionnaire du programme, chef de projet) :
    - Sont-ils prêts à participer à l'évaluation du plan ou programme ?
    - Sont-ils prêts à introduire des changements dans leurs pratiques (pilotage du programme, relation partenariale) ?

**COMMENTAIRES :**

# Abstract

**Introduction:** As part of its evaluation mission, the Haut conseil de la santé publique needs to assess evaluability of public health plans and programmes before undertaking an evaluation. The aim of the project is to elaborate a tool able to assess rapidly the evaluability of a public health plan or programme. The objective of the thesis will be to explore the applicability of the concept of evaluability in the health sector in France.

**Material and Method:** A literature review has verified that such a tool did not exist. It isolated the two major elements of the evaluability: feasibility and utility. A Delphi study was conducted with experts from institutions involved in evaluation. It identified the main criteria of utility and feasibility and related indicators. A tool for estimating the evaluability was developed. This tool is tested on various plans and health programmes in order to fine-tune it and verify its validity. The validation of the tool is done by the multidisciplinary team of experts.

**Results:** The tool consists of three parts drawing on the Delphi study and literature review: assessment of the feasibility, utility and stakeholders' commitment. A pre-test was conducted using four evaluated health programmes. The results obtained through pre-testing are consistent with those obtained by the evaluation. The tool is validated and refined through the pre-test by the working team of experts.

**Conclusion:** The tool fulfills its initial objectives. It is a practical tool that will definiteley evolve. The evaluability assessment concept showed to be applicable in the French health sector. However, applicability and the scope of application of the tool is context dependent.



**Key words:** public health plans and programmes – evaluation – evaluability – feasibility - utility

# Résumé


## L’évaluabilité des plans et programmes de santé

**Introduction :** Dans le cadre de sa mission d'évaluation le Haut conseil de la santé publique a souhaité se doter d'un outil permettant d'apprécier dans quelle mesure il convient de procéder à l’évaluation d'un plan ou programme de santé ou dans quelle mesure ce plan a été défini de façon à permettre son évaluation par la suite. Cet outil se doit d’être pragmatique et pédagogique. L’objectif du mémoire est d’étudier l’applicabilité du concept d’évaluabilité dans le champ de la santé publique en France.

**Matériel et méthode :** Une revue de littérature a vérifié qu'un tel outil n’existait pas. Elle a isolé les deux éléments majeurs de l'évaluabilité : la faisabilité et l’utilité. Une étude Delphi a été réalisée avec des experts d'institutions impliquées dans l'évaluation. Elle a identifié les principaux critères d’utilité et de faisabilité et les indicateurs afférents. Une grille d’appréciation de l’évaluabilité a été élaborée. Cette grille est testée sur différents plans et programmes de santé afin de l'affiner et d'en vérifier la validité. L'ensemble de ces travaux est supervisé et validé par un comité de pilotage multidisciplinaire.

**Résultats :** L’outil se compose de trois parties dérivées de l’étude Delphi et de la littérature : appréciation de la faisabilité, de l’utilité et l’adhérence des acteurs. Un pré-test a été effectué sur quatre programmes déjà évalués. Les résultats obtenus à travers l’application de la grille sont conformes à ceux obtenus par l’évaluation. L’outil est validé et révisé à travers le pré-test par le comité de pilotage.

**Conclusion :** Outil applicable avec finalités pratiques. Amené à évoluer, néanmoins des éléments de contexte vont influer sur son applicabilité et sur le cadre de son application.



**Mots clés :** plan et programme de santé – évaluation - évaluabilité – faisabilité – utilité