

*NOTE SUR LA RECHERCHE ET LES MÉTHODES /
RESEARCH AND PRACTICE NOTE*

UTILITÉ DE L'ÉVALUATION DE
L'ÉVALUABILITÉ DES POLITIQUES
GOUVERNEMENTALES DE LUTTE
CONTRE LE TABAGISME : L'EXPÉRIENCE
QUÉBÉCOISE DES CENTRES D'ABANDON
DU TABAGISME

Bernard-Simon Leclerc
Institut national de santé publique du Québec
Montréal, Québec

Résumé : Les politiques gouvernementales de lutte contre le tabagisme s'avèrent souvent multicomposites et panterritoriales. L'évaluation de leur implantation ou de leurs effets exige un investissement important en ressources, en temps, et en efforts. L'analyse de la faisabilité de l'évaluation, bien souvent négligée, est pourtant une étape cruciale de leur réussite. Elle vise à s'assurer qu'un programme remplit les conditions en permettant l'évaluation formelle. Dans le cas contraire, elle permet d'améliorer le programme lui-même. Cet article décrit les principes de l'évaluation de l'évaluabilité et présente un exemple d'application. Il conclut en exposant quelques leçons dégagées de l'expérience québécoise de l'évaluation des centres d'abandon du tabagisme.

Abstract: The government's tobacco control policies are often multifaceted and large-scale in nature. The process and outcome evaluation of such programs generally requires a significant investment in time, money, and effort. A critical but often neglected step prior to conducting a successful evaluation is the evaluability assessment. This tool is a method for examining a given program to determine whether it meets the criteria for a meaningful evaluation. Even when it fails to do so, it still contributes to the improvement of the program. This article describes the principles of the evaluability assessment and provides an example of its application. It concludes by presenting some lessons learned from the evaluation of Quebec's Quit Smoking Centres.

Correspondance à l'auteur : Bernard-Simon Leclerc, Centre de recherche du CSSS de Bordeaux-Cartierville–Saint-Laurent, centre affilié universitaire, 11 822, avenue du Bois-de-Boulogne, Montréal, QC, H3M 2X6; <bernardsimon.leclerc.bcstl@ssss.gouv.qc.ca>

Le secteur de la santé et des services sociaux n'échappe pas à la vague de fond néolibérale qui oriente les administrations publiques vers l'adoption d'un mode de gestion axée sur les résultats. La façon dont les planificateurs et les décideurs conçoivent l'évaluation y est d'ailleurs étroitement associée. La pression à laquelle sont soumis les évaluateurs entraîne parfois une évaluation prématurée des programmes, avant même qu'ils aient réalisé une analyse préliminaire de faisabilité de leur évaluation (Smith, 2005).

Joseph Wholey a introduit le concept d'évaluation de l'évaluabilité en 1979. Sa proposition était liée au constat que les programmes, souvent faiblement décrits et articulés, rencontraient rarement les critères permettant de mener un processus d'évaluation satisfaisant. L'évaluation de l'évaluabilité est devenue une pratique reconnue en évaluation, comme en témoigne la recension de Trevisan (2007) portant sur des évaluations publiées entre 1986 et 2006. Elle vise à assurer les chercheurs que le programme visé remplit les conditions permettant une évaluation d'implantation ou des effets. Dans le cas contraire, elle contribue au développement et à l'amélioration du programme lui-même.

Les politiques gouvernementales de lutte contre le tabagisme s'avèrent particulièrement indiquées pour ce type d'évaluation, étant donné leur caractère novateur, multicomposite, et panterritorial, de même que l'investissement important en ressources, en temps, et en efforts qu'exige leur évaluation formelle. Cet article décrit les principes de l'évaluation de l'évaluabilité et présente un exemple d'application concret. Il conclut en exposant quelques leçons importantes dégagées de l'expérience de l'*Évaluation préliminaire des centres d'abandon du tabagisme* (CAT) implantés au Québec (Leclerc, 2009).

ÉVALUATION DE L'ÉVALUABILITÉ

Évaluabilité d'un programme

Le néologisme « évaluabilité », dérivé de l'adjectif « évaluable » auquel le suffixe « -abilité » est attaché, respecte les règles morphologiques de la langue française. L'évaluabilité d'un programme représente pour ainsi dire la possibilité pour un programme d'être évalué. Cette définition sous-tend, comme la plupart des glossaires de termes utilisés en évaluation le soulignent, que l'exercice peut être fait « de manière fiable et crédible ».

L'évaluabilité, telle que définie précédemment, est essentiellement liée au fait que le programme est clairement articulé et explicite, que

les objectifs et les effets attendus sont précisés et réalistes, que le lien logique entre le programme, les objectifs, et les effets est exposé, et, enfin, que l'on puisse raisonnablement concevoir que les informations dont on a besoin peuvent être colligées (Potvin, 2009).

Évaluation de l'évaluabilité d'un programme

L'évaluation de l'évaluabilité suppose un examen de la façon dont les activités planifiées permettront d'atteindre les résultats prévus. Elle revient à examiner le programme afin de vérifier si ses objectifs sont définis de façon adéquate et si les résultats escomptés sont vérifiables. La modélisation de la logique théorique du programme est ainsi au cœur de la démarche et sert de canevas à l'analyse. L'adéquation entre les objectifs de l'intervention et les moyens mis en place pour les atteindre, le degré de conformité entre les principes préconisés par les concepteurs et les croyances des acteurs du milieu, ainsi que la disponibilité des données requises aux fins d'évaluation sont exprimés sous la forme de facteurs pouvant empêcher ou favoriser une évaluation de programme adéquate.

Une évaluation de l'évaluabilité comporte cinq composantes cruciales (Kaufman-Levy & Poulin, 2003) :

- Étude du contexte, de la conception, et du fonctionnement du programme;
- Observation du programme en action;
- Détermination de la possibilité de recueillir et d'analyser les données nécessaires à l'évaluation;
- Évaluation de la probabilité que le programme atteigne ses buts et objectifs;
- Démonstration des causes pour lesquelles une évaluation pourra ou non aider le programme et répondre aux attentes du requérant et des autres parties prenantes.

La procédure consiste globalement à déterminer si le programme réunit les critères nécessaires pour en permettre l'évaluation à proprement parler.

DESCRIPTION DU PROGRAMME

Centres d'abandon du tabagisme

Des CAT ont été implantés dans toutes les régions sociosanitaires du Québec. Ces services, généralement situés dans un établissement exploitant un centre local de services communautaires (CLSC), offrent

à tout fumeur qui en fait la demande de l'information et du soutien pour l'aider à se libérer du tabac. Leur mise en place répondait aux impératifs organisationnels de services de soutien à l'arrêt tabagique édictés par le ministère de la Santé et des Services sociaux dans son Plan québécois d'abandon du tabagisme (PQAT). La volonté d'intensifier la disponibilité des services d'abandon du tabagisme, d'offrir une gamme étendue de services communs dans chacun des territoires du Québec, et d'harmoniser les pratiques des divers intervenants constituaient le cœur des préoccupations du plan. Le Ministère a confié à l'Institut national de santé publique du Québec (INSPQ) le mandat d'en réaliser l'évaluation.

Les évaluateurs se sont appliqués à dresser un bilan sommaire du développement et du fonctionnement des CAT en 2005–2006, soit deux ans après le début de leur implantation, et à étudier la faisabilité et la légitimité de leur évaluation formelle. Cette démarche d'évaluation formative se situe dans un processus d'accompagnement et de soutien au développement des services d'abandon du tabagisme. L'objectif était d'éclairer le Ministère à l'égard de la conception, de la structure organisationnelle, et de la mise en œuvre de ces CAT, de contribuer à leur amélioration, de concourir à en orienter l'évolution future, et de suggérer les suites à donner en matière d'évaluation.

Réalisation de l'évaluation de l'évaluabilité du programme

L'examen des CAT a fait appel à des notions d'analyse de la faisabilité de l'évaluation d'un programme ainsi qu'à des concepts d'évaluation des processus.

L'évaluation des CAT était basée sur des travaux préliminaires réalisés dans le cadre d'une réflexion plus large sur l'évaluation du PQAT. De fait, un devis d'évaluation de l'implantation et des effets du PQAT a été élaboré en 2005 par l'INSPQ et révisé à la suite d'une consultation auprès des planificateurs tabac dans les directions régionales de santé publique. Le devis proposé touche l'ensemble des composantes du PQAT enchâssées dans un modèle logique développé expressément pour les fins de l'évaluation. L'auteur du devis y a défini les principales questions auxquelles l'évaluation visait à répondre, puis y a précisé les indicateurs et la méthodologie de collecte d'information envisagée pour y parvenir. Une hiérarchisation des priorités d'évaluation en fonction de l'état d'avancement de la mise en œuvre du PQAT ainsi que de la disponibilité des sources de données complète l'ouvrage. C'est dans ce contexte que l'évaluation des CAT a été privilégiée.

L'évaluation préliminaire a en effet donné l'occasion d'évaluer la disponibilité des données existantes et la possibilité de recueillir de nouvelles données nécessaires au processus d'évaluation formelle des CAT. Elle a permis de conclure à un certain nombre d'insuffisances dans l'information disponible au sujet de la clientèle et des services des CAT. Les données des systèmes d'information tout autant que celles colligées localement et manuellement par les intervenants des CAT se sont avérées peu utiles en raison de leur manque d'exhaustivité et d'uniformité. L'absence de données clés a donc dû être comblée par la collecte d'information au moyen d'approches conçues et administrées par les évaluateurs de l'équipe d'évaluation de l'INSPQ.

Plusieurs sources de données, tant quantitatives que qualitatives, et plusieurs outils ont été utilisés dans le cadre de cette étude. Ils comprennent une analyse de la documentation et des données existantes sur les CAT, la collecte du point de vue de différents acteurs clés (intervenants en soutien à l'arrêt tabagique et planificateurs dans les directions régionales de santé publique), ainsi qu'un sondage téléphonique auprès de fumeurs et de personnes ayant arrêté de fumer au cours des deux dernières années. En dernier lieu, une étude pilote sur l'utilisation d'une feuille de collecte de données harmonisée dans les CAT a été effectuée.

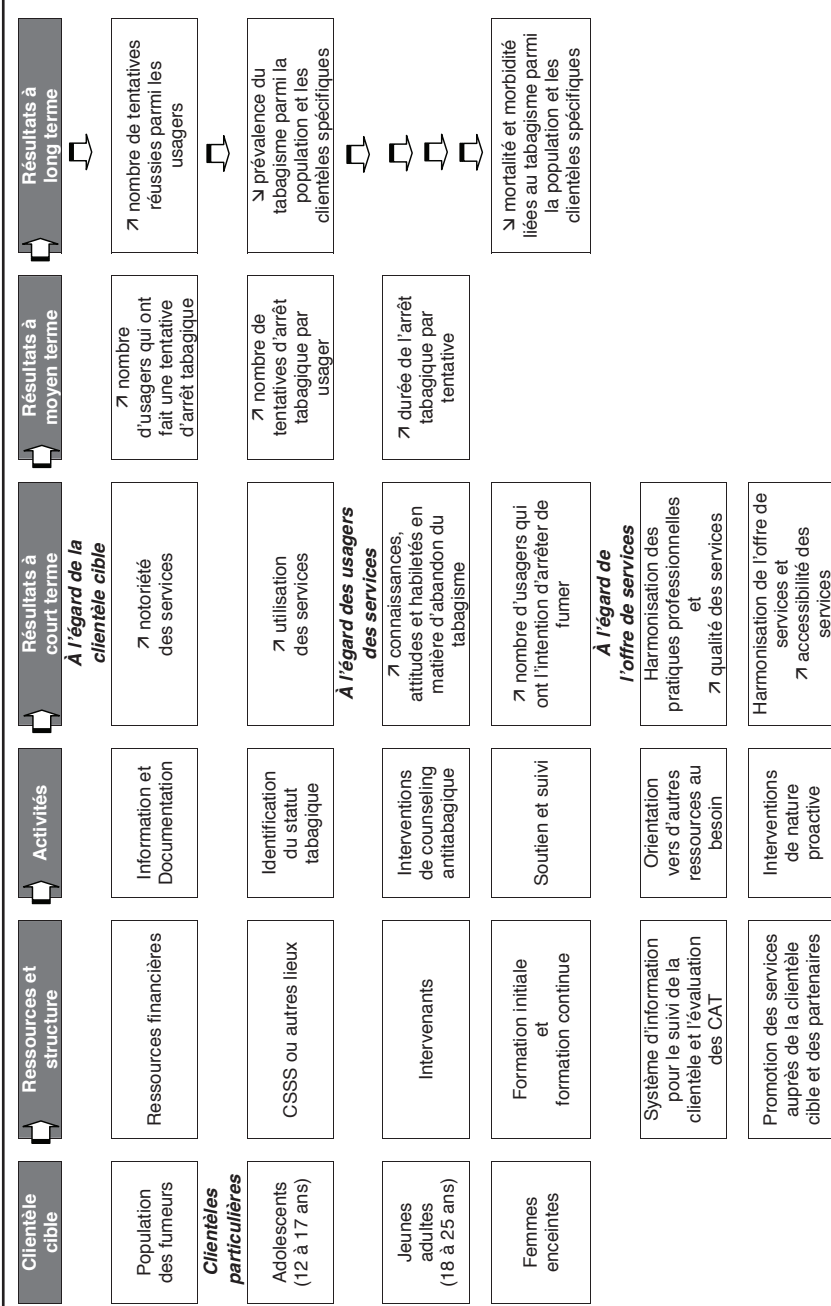
La Figure 1 présente le modèle de la logique théorique des CAT et clarifie les intentions initiales poursuivies par les concepteurs du programme. Les lecteurs intéressés qui souhaitent obtenir davantage de détails sur la méthodologie et les résultats sont priés de consulter le rapport complet disponible à <http://www.inspq.qc.ca/pdf/publications/899_RapportCAT.pdf>.

LEÇONS ET RECOMMANDATIONS

Leçons tirées de l'expérience

L'évaluation de l'évaluabilité des CAT a mis en lumière plusieurs retombées de la mise en œuvre du PQAT, notamment une intensification de la disponibilité d'une gamme étendue de services communs d'abandon du tabagisme. Néanmoins, les résultats de l'évaluation ont fait ressortir le besoin d'une réflexion et d'une mise au point actualisée des CAT et du PQAT dans son ensemble pour faire progresser l'initiative au-delà de cette première phase d'implantation. Finalement, l'évaluation a démontré que le développement des CAT était insuffisant sur le plan programmatique pour justifier la réalisation

Figure 1
Modèle logique théorique des centres d'abandon du tabagisme



d'une évaluation des effets. Similairement, elle a dévoilé que les CAT ne rencontraient pas les conditions minimales nécessaires pour permettre l'évaluation formelle de leur implantation.

L'une des principales retombées de l'élaboration du modèle logique aura été de rappeler au requérant que le programme ne visait pas seulement à offrir des services à des clients dans le besoin. Comme le résume la Figure 1, les plans initiaux aspiraient également à répondre à des objectifs de nature organisationnelle à l'égard de l'offre de services, et l'intervention devait se conformer à une intensité de services et à un mode d'intervention prédéfini.

À la lumière de ses travaux d'évaluation et de ses observations de la situation qui prévalait en 2005–2006, l'INSPQ a recommandé au Ministère :

- de procéder à la mise à jour du cadre de référence du PQAT, de définir des objectifs opérationnels et des cibles nationales à atteindre;
- d'entreprendre une réflexion d'ensemble sur la place des CAT au sein des services d'abandon du tabagisme du PQAT;
- de réaliser une étude approfondie pour déterminer les besoins réels des fumeurs québécois en matière d'abandon du tabac et les paramètres des services de soutien qui y répondraient;
- d'analyser la pertinence et la qualité des données saisies dans le système d'information clientèle des CLSC à des fins de suivi et d'évaluation des CAT ainsi que de favoriser l'implantation de la feuille de collecte harmonisée proposée par l'INSPQ dans l'ensemble des CAT du Québec;
- d'établir et de favoriser l'utilisation de lignes directrices à l'attention des intervenants des CAT, de bonifier et de normaliser la formation offerte aux participants, à l'embauche et en cours de pratique;
- de faciliter la réalisation d'une évaluation de la qualité de l'exercice professionnel des intervenants en CAT en collaboration avec les ordres professionnels concernés.

Par ailleurs, les apprentissages réalisés lors de cette expérience enrichiront sans aucun doute la pratique future des évaluateurs et du requérant, en premier lieu en rappelant à chacun que l'évaluation de l'évaluabilité de par la nature même de l'activité de générer un jugement sur la valeur d'un programme est potentiellement génératrice de tensions entre décideurs et évaluateurs. La procédure appelle en

conséquence à un processus continu de négociations entre les parties de même qu'à une éthique professionnelle qui exige de maintenir une relation centrée sur le désir d'amélioration du programme.

Nous avons produit une description détaillée, articulée, et explicite des hypothèses qui constituent le rationnel de la façon dont le programme devait opérer afin d'atteindre les résultats attendus. La définition du projet par les instances ministérielles elles-mêmes a servi de norme à son évaluation. Le principe de la théorie du programme se base sur le fait que si les buts et les objectifs du programme ne sont pas raisonnablement reliés à ce que le programme veut changer ou améliorer, ou si les hypothèses ou présupposés du programme ne représentent pas une approche crédible, il y a peu de chance que le programme soit efficace, qu'il produise les effets attendus, et qu'il soit évaluable. Évaluateurs, planificateurs, et décideurs doivent garder à l'esprit qu'il est cependant rare que les programmes offrent une théorie du programme suffisamment explicite et détaillée permettant d'emblée d'entreprendre une évaluation. Dans plusieurs cas, le programme s'appuie plutôt sur une théorie implicite, qui demande à être découverte ou dévoilée. L'évaluateur doit alors travailler à extraire et à décrire, par différents moyens, cette théorie implicite (Potvin, 2009).

Par ailleurs, dans le cas où l'évaluation conclut à l'insatisfaction des conditions préalables à l'évaluation formelle du programme, l'évaluateur doit obligatoirement collaborer avec les planificateurs et les gestionnaires pour clarifier les aspects qui exigent des ajustements. Ce qui paraîtra moins clair est la distance par rapport à son objet d'étude qu'un évaluateur—de surcroît mandataire de l'État—devra maintenir à l'étape de la planification initiale de futurs programmes de son client. L'évaluateur doit-il collaborer à l'élaboration d'une politique qu'il sera éventuellement appelé à évaluer? La question demeure entière et non résolue. Guba et Lincoln, dont les propos sont rapportés par Alain (2009), estiment qu'un évaluateur ne peut espérer demeurer neutre s'il entretient une relation privilégiée avec les gestionnaires et promoteurs du programme, pas plus d'ailleurs que s'il le fait avec les intervenants et les usagers de ce même programme.

La diffusion ou le transfert de connaissances est au cœur de la mission d'un centre d'expertise et de référence en matière de santé publique comme l'INSPQ. Il est conséquemment de la responsabilité des professionnels qui y travaillent de faire en sorte que les praticiens de la santé publique et les décideurs concernés par le contrôle du tabagisme puissent se servir des résultats de recherche pour l'élaboration et la mise en œuvre de politiques pratiques concrètes.

Les évaluateurs doivent également prendre conscience de leur responsabilité dans le transfert de connaissance de nature disciplinaire et méthodologique, c'est-à-dire dans l'amélioration de la *littératie de recherche* chez les planificateurs et décideurs—définie comme la capacité de comprendre et d'utiliser la recherche (Vanderheyden, Verhoef, Scott, & Pain, 2006).

Recommandations à l'intention des décideurs

Nous avons présenté l'évaluation de l'évaluabilité comme un outil susceptible de favoriser le développement de programmes optimaux et leur évaluation efficiente. Ce type d'évaluation ne doit toutefois pas servir à pallier à un manque de planification et de programmation initial ou à relayer à d'autres la responsabilité de ces tâches. Pour cette raison, nous croyons que les planificateurs et les décideurs doivent consacrer une grande attention aux préoccupations énoncées dans le présent article. À cet égard, nous leur recommandons fortement de recourir à une liste de vérification qui permette l'auto-évaluation de l'évaluabilité des programmes (Annexe).

CONCLUSION

Nous avons tenté de décrire les principes et l'utilité de l'évaluation de l'évaluabilité. L'évaluation quelle qu'elle soit constitue un outil de production de connaissances pour éclairer et soutenir la prise de décision. Son utilité ne pourra ultimement être appréciée que par la prise en compte, par les décideurs, de ses résultats et de ses recommandations. Nous espérons que notre « note sur la recherche et les méthodes » encouragera les planificateurs et les décideurs des politiques gouvernementales de lutte contre le tabagisme, ainsi que d'autres évaluateurs, à envisager l'évaluation de l'évaluabilité comme outil de compréhension et d'amélioration des programmes.

REMERCIEMENTS

L'auteur remercie son collègue Michael Cantinotti, qui s'est aimablement offert à réviser et à commenter le présent article, ainsi que Daniel Desrochers pour son minutieux travail de révision linguistique. Des remerciements s'adressent ensuite aux nombreuses personnes qui ont contribué de près ou de loin à la réalisation de l'évaluation de l'évaluabilité sous-jacente. Notre reconnaissance s'adresse enfin aux membres du Service de lutte contre le tabagisme du Ministère

de la Santé et des Services sociaux pour leurs commentaires sur les résultats de l'évaluation des CAT.

RÉFÉRENCES

- Alain, M. (2009). Des systèmes de classification des modèles d'évaluation de programmes d'intervention psychosociales à une proposition de modèle intégrateur. Les trois phases du cube : avant-pendant-après. Dans M. Alain & D. Dessureault (Éds.), *Élaborer et évaluer les programmes d'intervention psychosociale* (pp. 9–28). Québec, QC : Presse de l'Université du Québec.
- Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review*, 70(11), 35–36.
- Kaufman-Levy, D., & Poulin, M. (2003). *Evaluability assessment: Examining the readiness of a program for evaluation*. Washington, DC : Office of Juvenile Justice and Delinquency Prevention.
- Leclerc, B. S. (2009). *Évaluation préliminaire des centres d'abandon du tabagisme : période 2005–2006*. Montréal, QC : Développement des individus et des communautés, Institut national de santé publique du Québec.
- Potvin, P. (2009). Éléments et critères d'évaluabilité d'un programme d'intervention psychosociale. Dans M. Alain & D. Dessureault (Éds.), *Élaborer et évaluer les programmes d'intervention psychosociale* (pp. 101–114). Québec, QC : Presse de l'Université du Québec.
- Smith, M. F. (2005). Evaluability assessment. Dans S. Mathison (Éd.), *Encyclopedia of evaluation* (pp. 136–139). Thousand Oaks, CA : Sage.
- Trevisan, M. S. (2007). Evaluability assessment from 1986 to 2006. *American Journal of Evaluation*, 28, 290–303.
- Vanderheyden, L., Verhoef, M., Scott, C., & Pain, K. (2006). Evaluability assessment as a tool for research network development: Experiences of the complementary and alternative medicine education and research network of Alberta, Canada. *Canadian Journal of Program Evaluation*, 21(1), 63–82.

ANNEXE

Liste de vérification pour l'auto-évaluation de l'évaluabilité des programmes

Cette liste d'items tirée de Potvin (2009, pp. 109–111) permet de vérifier si un programme est évaluable. Elle permet également d'identifier les aspects auxquels un soin particulier doit être accordé avant d'entreprendre une évaluation.

A. Analyse globale du programme

- L'accès à des documents qui décrivent clairement le programme est possible.
- La description du programme est bien articulée.
- Le programme constitue une solution appropriée pour résoudre le problème visé.
- Les composantes du programme, les activités et le fonctionnement du programme sont suffisamment définis.
- Les modalités d'intervention sont bien opérationnalisées (consignes, procédures, activités, déroulement, etc.)
- Le programme offre un cheminement à travers une série d'activités prédéterminées en vue de changer l'état de la clientèle cible.
- Le programme a fait l'objet d'évaluations lors de la mise à l'essai et d'ajustements quand cela était nécessaire.
- Le programme a été appliqué à quelques reprises et présente un niveau de maturité (une certaine stabilité) permettant son évaluation.

B. Clientèle cible

- La clientèle cible qui peut bénéficier du programme est bien identifiée et ses caractéristiques, bien définies.
- Les modalités de sélection ainsi que les critères de sélection et d'exclusion sont précisés.

C. Problématique de la clientèle cible et analyse de besoins

- La problématique associée à la clientèle est bien documentée et soutenue par une théorie sur le problème.
- Cette théorie explicative du problème est valide.
- Le programme est le résultat d'une analyse de besoins de la clientèle, et l'information sur les besoins prioritaires de la clientèle cible est bien définie.
- Le niveau de correspondance entre les caractéristiques de

la clientèle cible (problème, capacités) et les activités du programme est bien dosé et adéquat.

D. Objectifs et atteinte de résultats

- Le programme répond à des objectifs généraux liés à la mission du milieu.
- Le but et les objectifs sont bien définis, clairement énoncés et mesurables.
- Les objectifs disposent d'une définition opérationnelle des résultats attendus comprenant des indicateurs et des critères de succès.
- Une précision est apportée sur les changements attendus de comportements ou d'attitude chez la clientèle cible.
- Les facteurs qui peuvent faciliter ou entraver la réalisation des objectifs sont présentés.
- Les différents degrés de réussite sont précisés.
- La façon de mesurer les progrès est présentée.
- Le lien articulé entre le programme et les objectifs ou les résultats attendus est suffisamment solide pour justifier une évaluation.

E. Théorie du programme

- La « logique » du programme, ou « théorie du programme », est clairement présentée par le biais de la structure, du fonctionnement, et des procédures du programme.
- Cette théorie du programme explique pourquoi le programme fait ce qu'il fait et présente un rationnel qui explique pourquoi, en faisant les choses de cette manière, on obtiendra les résultats désirés.
- Il existe un lien logique entre les activités du programme et ses résultats, et les postulats de causalité sont plausibles en liant le programme aux résultats.
- En résumé, le programme repose sur une théorie justifiant le choix des interventions, sur une autre portant sur le problème et sur l'action ainsi qu'un rationnel sur lequel s'appuient les interventions. Cette théorie du programme est assez solide pour réussir une évaluation.

F. Ressources humaines et matérielles

- Les caractéristiques et les compétences que doit posséder le personnel qui applique le programme sont présentées clairement.

- Les ressources humaines, matérielles, et financières sont bien décrites, adéquates, et suffisantes pour la réalisation du programme.

G. Résultat de l'évaluation

- L'utilisation de l'information issue des résultats de l'évaluation des effets du programme est adéquatement précisée.

Liste de vérification de la formulation des objectifs

Cette liste d'items de l'acronyme mnémotechnique SMART (Doran, 1981) permet de vérifier si un objectif opérationnel est formulé adéquatement.

- Spécifique*. L'objectif doit être précis, compréhensible, et sans ambiguïté. Il doit décrire précisément le changement souhaité chez la clientèle envisagée.
- Mesurable*. L'objectif doit être mesurable et accompagné d'un indicateur. Il doit permettre de suivre son état d'avancement et de juger objectivement de son atteinte.
- Approprié*. L'objectif doit être approprié et acceptable pour les parties en cause.
- Réaliste*. L'objectif doit être réaliste et réalisable. Il doit être atteignable avec les moyens disponibles.
- Temporellement défini*. L'objectif doit être associé à une échéance. La durée doit être déterminée et un délai, fixé.

Bernard-Simon Leclerc est chercheur, épidémiologiste, et spécialiste en évaluation des interventions au Centre de recherche du CSSS de Bordeaux-Cartierville–Saint-Laurent, conseiller scientifique à la Direction du développement des individus et des communautés de l'Institut national de santé publique du Québec, et chercheur associé au Centre de recherche de l'Institut universitaire de gériatrie de Montréal. Il est également professeur adjoint de clinique au Département de médecine sociale et préventive de l'Université de Montréal, et il offre de la formation continue en ligne aux praticiens de la santé publique pour le compte de l'Agence de la santé publique du Canada.