

# EVALUATION DE L'IMPACT D'UN PROGRAMME D'AMELIORATION DES PRATIQUES DE SEVRAGE : SCHEMAS ET CONTRAINTES

Yves MARTIN-PREVEL, Serge TRECHE, Francis DELPEUCH

Laboratoire de Nutrition Tropicale (UR44), Centre ORSTOM,  
Montpellier (France)

## 1. INTRODUCTION

Comme pour toute action de Santé Publique, l'évaluation d'un programme dans le domaine de la nutrition est un élément fondamental pour orienter les interventions futures. Nous allons traiter ici de l'évaluation de l'**impact** d'un programme nutritionnel, c'est-à-dire de l'évaluation cherchant à juger des effets de ce programme au niveau de son objectif final, qui doit être l'amélioration de l'état nutritionnel de la population cible. Ceci doit être distingué de l'évaluation dite "de processus", ou "de fonctionnement" du programme, qui consiste notamment à vérifier si les actions prévues ont été ou non effectuées, à mesurer la pénétration de la stratégie au sein de la population, à identifier et analyser les principales sources de blocage etc. Bien entendu l'évaluation d'impact n'exclut pas et ne doit pas occulter celle de fonctionnement.

Nous envisagerons tout d'abord, dans une première partie, les bases méthodologiques pour l'évaluation d'impact dans le domaine de la nutrition. Nous avons considéré ici le cas d'interventions tournées vers la malnutrition protéino-énergétique, mais les principes méthodologiques développés restent bien entendu valables pour les études d'impact sur d'autres formes de malnutrition. En revanche nous avons délibérément restreint notre propos à des évaluations de type transversal, les plus adaptées à l'évaluation de l'impact au niveau de la santé publique, les méthodes longitudinales, plus précises mais plus complexes à mettre en oeuvre, étant davantage tournées vers des évaluations de type recherche.

Nous présenterons ensuite dans une seconde partie, à titre d'illustration, les schémas mis en place au Congo pour l'évaluation de deux stratégies d'amélioration des pratiques de sevrage.

## 2. BASES METHODOLOGIQUES

### 2.1. Principes Généraux

L'évaluation d'impact d'un programme nutritionnel ne peut pas se limiter à la seule description de l'état nutritionnel de la population ayant bénéficié de l'intervention. Une comparaison est absolument nécessaire avec l'état nutritionnel d'un ou plusieurs autre(s) groupe(s), servant de témoins, et qui peuvent être issus d'une autre population, ou bien de la même population avant intervention, ou encore des deux à la fois.

Cette comparaison doit alors permettre fondamentalement de répondre à la question suivante : « *Quelle différence, ou quelle part de la différence, observée entre les groupes, est attribuable à l'intervention ?* ». La réponse à cette question n'est jamais facile à obtenir. Le degré de plausibilité, pour imputer à l'intervention une différence observée, est d'autant plus important que l'étude d'évaluation permet de supprimer les facteurs de confusion potentiels, ou pour le moins de mesurer leur influence pour en tenir compte dans l'analyse. Ceci a pour premier et essentiel corollaire que l'évaluation d'impact doit être prévue et mise en place en même temps que l'intervention elle-même.

Une réflexion préalable est donc nécessaire, non seulement pour définir le schéma d'étude et les variables à recueillir afin de contrôler les éventuels facteurs de confusion, mais encore pour choisir les indicateurs de résultat pertinents, déterminer la taille et la nature de l'échantillon, établir le budget nécessaire etc. Nous allons passer en revue ces différents points de façon théorique, mais souvent, en pratique, une étude de faisabilité devra être effectuée pour déterminer à quelles conditions une évaluation d'impact peut être réalisée avec des chances raisonnables d'aboutir.

### 2.2. Les facteurs de confusion

On appelle facteur de confusion tout élément qui influence l'état nutritionnel de la population cible de l'intervention, et qui, bien que ne faisant pas partie du programme, est lié à sa réalisation. Ils constituent l'écueil majeur à éviter, ou au minimum à contrôler, dans les évaluations d'impact.

Un exemple type : si l'intervention consiste en une supplémentation alimentaire des enfants, mais que la participation au programme entraîne par ailleurs un meilleur accès aux soins, il sera impossible de distinguer, au niveau de la croissance des enfants, l'effet de la supplémentation de celui des soins. La modification de l'accessibilité aux soins est alors facteur de confusion pour l'évaluation de l'impact de la supplémentation.

Cet exemple est particulièrement démonstratif, mais en pratique les facteurs de confusion peuvent être beaucoup plus difficiles à identifier. Il est donc nécessaire de recueillir systématiquement un certain nombre de variables réputées influentes sur l'état nutritionnel, et dont on examinera par la suite si leur niveau est identique selon qu'il y a ou non intervention. On peut considérer 4 sources principales de facteurs de confusion :

- non comparabilité des groupes ;
- biais d'information ;
- effets du temps ;
- régression vers la moyenne.

### 2.2.1. Non comparabilité des groupes

L'idéal, pour une évaluation d'impact, est que les groupes comparés soient le plus possible semblables pour tous les facteurs pouvant influencer l'état nutritionnel, et ne diffèrent que par leur participation au programme considéré. Alors une différence observée entre les groupes peut être attribuée au programme. Il faut donc étudier soigneusement la comparabilité des groupes, à partir d'informations recueillies sur les facteurs de confusion potentiels. On peut distinguer trois niveaux :

- communautaire : équipements sanitaires, accès à l'eau potable et à l'électricité, environnement en général ;
- familial : taille, revenus, niveau d'études, habitudes alimentaires et toutes variables socio-économiques ;
- individuel : âge, sexe, statut vaccinal et autres informations sur l'état de santé.

En pratique les groupes ne sont évidemment jamais totalement comparables, ne serait-ce que par le choix le plus souvent limité des groupes d'intervention et témoin, et par les raisons éthiques ou politiques qui motivent ce choix. L'important est donc de connaître et de mesurer les différences pré-existant entre les groupes, puis de suivre leur évolution au cours de l'intervention. Il est alors possible de prendre en compte ces différences dans l'analyse des résultats par des techniques statistiques d'ajustement. Mais ces techniques d'ajustement entraînent une réduction de puissance statistique dans les comparaisons, soulignant la nécessité de choisir des groupes le plus possible comparables au départ de l'étude.

Plus important encore, il y a lieu de rechercher et d'éviter toute cause systématique et liée au programme de différence entre les groupes, notamment un éventuel phénomène d'auto-sélection, une répartition non aléatoire des perdus de vue ou un effet de l'âge dans le cas de comparaisons avant/après. Il s'agit alors de véritables facteurs de confusion, à l'origine de différences entre les groupes qu'il sera impossible de prendre en compte dans l'analyse.

### 2.2.2. Biais d'information

On désigne ainsi le résultat d'une différence qui existe non pas au niveau des groupes de l'étude, mais dans la façon dont l'information a été recueillie pour chacun d'eux. Le simple fait d'employer pour le recueil des données une méthodologie non strictement identique dans les différents groupes, ou bien des équipes d'enquêteurs différentes, ou encore des moyens de mesures non équivalents, peut suffire à introduire au niveau des résultats une différence artificielle entre les groupes. Il y aura alors confusion avec les effets de l'intervention.

### 2.2.3. Effets du temps

L'état nutritionnel des populations est, comme la quasi-totalité des phénomènes de santé, soumis aux effets du temps (à distinguer des effets de l'âge). Lorsque des mesures sont faites à des moments différents, il peut devenir difficile de savoir si les variations observées sont liées aux effets du temps ou à ceux d'une intervention.

On distingue principalement deux types d'effets du temps :

- La tendance séculaire, qui traduit une évolution générale de l'état nutritionnel dans une population et qui est le plus souvent modérée, sauf événements extérieurs importants (situation de guerre par exemple) ;
- La tendance saisonnière, qui traduit cette fois une évolution cyclique de l'état nutritionnel, et dont l'effet peut être important sur les groupes les plus vulnérables.

### 2.2.4. Régression vers la moyenne

Il s'agit du phénomène selon lequel lorsque des sujets sont sélectionnés aux extrêmes d'une distribution, la valeur de la variable mesurée chez ces sujets a tendance, lors d'une seconde mesure, à se rapprocher spontanément de la valeur centrale. La cause de ce phénomène est qu'une partie des individus sélectionnés l'auront été par hasard, à la suite de fluctuations de la variable ou des erreurs de mesure. Ceci a naturellement peu de chances de se reproduire pour les mêmes individus lors d'une deuxième mesure.

Notons que ce phénomène ne sera gênant pour l'évaluation d'impact que si le programme s'adresse à des individus, et non à des populations, et concerne des individus sélectionnés pour des valeurs extrêmes d'indices nutritionnels.

### **2.2.5. Remarque**

En développant les divers facteurs de confusion ci-dessus nous avons évoqué essentiellement le cas où ces facteurs conduisent à attribuer à tort à l'intervention une différence observée. Mais l'inverse est tout aussi possible, et le facteur de confusion empêche alors d'observer une différence pourtant réelle. Le risque de conclure à tort, lui, est de toutes façons aussi grave dans un sens ou dans l'autre.

## **2.3. Les différents schémas d'étude**

Le schéma extrême, maximaliste, est celui de l'expérience randomisée en double aveugle. C'est celui qui, bien conduit, offrirait les résultats les plus probants quant à l'effet d'une intervention sur l'état nutritionnel de la population cible en permettant le contrôle du maximum de facteurs de confusion, mais en pratique il n'est évidemment pas réalisable à l'échelle d'un programme de santé publique. D'ailleurs la randomisation ne sera quasiment jamais possible en pratique pour de nombreuses raisons, notamment éthiques et politiques, si bien que l'on ne peut concevoir que des schémas dits « quasi-expérimentaux ».

Nous envisagerons ici les principaux types de schémas réalisables, en décrivant brièvement leurs intérêts et leurs limites. Pour chacun d'eux des variantes existent, ou peuvent être inventées, mais c'est le principe général que nous retiendrons.

Ainsi que nous l'avons déjà suggéré, l'évaluation d'impact fait appel à la comparaison d'un groupe d'intervention avec un groupe témoin. La comparaison peut s'effectuer dans l'espace (« ici/ailleurs »), dans le temps (« avant/après »), ou à la fois dans le temps et dans l'espace (schémas mixtes, les plus performants mais les plus lourds à réaliser). La figure 1 illustre les différents types de schémas que nous allons présenter.

### **2.3.1. Schéma « Ici/Ailleurs »**

L'intervention est réalisée au niveau d'une population « A » ; Au bout d'un certain temps, dépendant bien sûr du type d'intervention, l'état nutritionnel du groupe cible de cette population « A » est comparé à celui d'une population « B », non soumise à l'intervention.

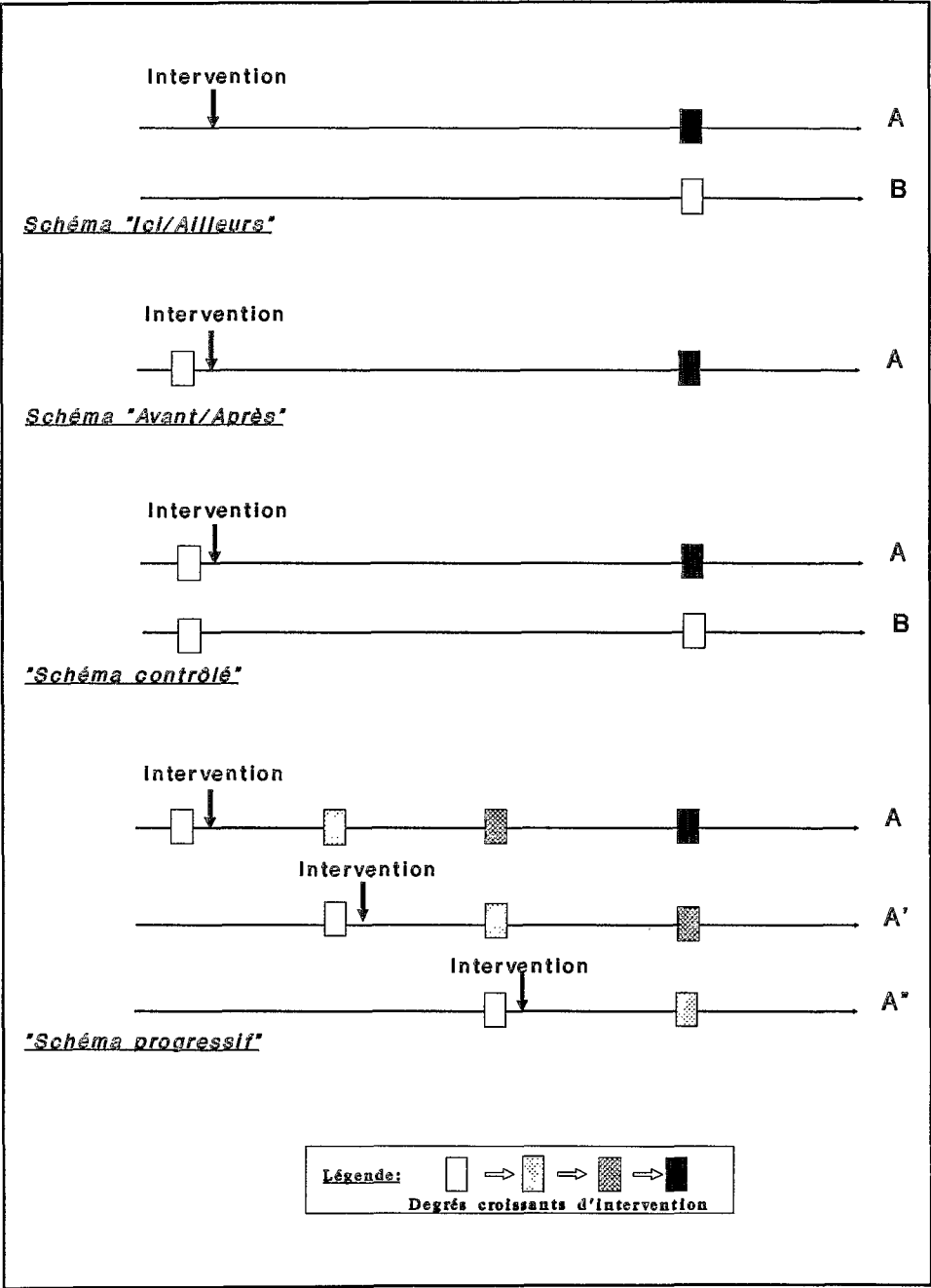


Figure 1  
Principaux schémas d'évaluation.

Comme on l'a détaillé plus haut, il est en premier lieu fondamental que les groupes A et B soient le plus possible comparables. Toutefois même si cette comparabilité est respectée, pour tous les facteurs influençant l'état nutritionnel, une différence observée entre les groupes ne pourra pas être imputée avec certitude à l'intervention dans la mesure où rien ne permet d'affirmer que cette différence n'existait pas avant l'étude.

Un autre problème de ce type d'étude réside dans le choix des populations devant être soumis, ou non, à l'intervention. Il intervient souvent à ce niveau des contraintes éthiques ou politiques qui vont à l'encontre de la comparabilité. D'autre part, il est souvent souhaitable, pour la bonne comparabilité, que les populations soient géographiquement proches. Mais la proximité des groupes d'intervention et témoin est source d'un phénomène dit « de contamination », traduisant le fait que les sujets de la zone témoin peuvent en fait être atteints par l'intervention. Ceci entraîne des difficultés pour identifier clairement participants et non-participants au programme. Enfin ce type d'étude est le plus exposé au biais d'auto-sélection.

### 2.3.2. Schéma « Avant/Après »

L'état nutritionnel du groupe cible de la population « A » est mesuré au temps zéro, c'est-à-dire juste avant le début de l'intervention, puis il est mesuré à nouveau au bout d'un certain délai. Ce type de schéma cherche à s'affranchir des problèmes de comparabilité en utilisant comme groupe de comparaison la population soumise à intervention elle-même. En revanche, ces schémas s'exposent aux effets du temps, pouvant biaiser l'étude de plusieurs façons :

- lorsque l'évolution de l'état nutritionnel est jugée sur les même enfants avant et après intervention, intervient l'effet de l'âge : pour tous les indicateurs anthropométriques les faibles valeurs par rapport aux références standard sont plus ou moins fréquentes selon l'âge, indépendamment de tout programme ;
- si le programme porte sur une partie seulement de la population, sélectionnée généralement par les faibles valeurs d'indices anthropométriques, on se heurte au phénomène de régression vers la moyenne ;
- enfin dans tous les cas, même lorsque l'évaluation porte sur l'ensemble d'une tranche d'âge dans la population avant et après intervention, subsiste au minimum l'effet de la tendance séculaire. Ce dernier peut être modéré si le délai d'étude est relativement court. On peut par ailleurs essayer de l'estimer indirectement (données sur la mortalité ou la morbidité, ou encore la surveillance de la croissance, les taux de vaccination etc.), mais cette estimation sera toujours partielle en l'absence de groupe témoin.

### 2.3.3. Schémas mixtes

Leur principe est de réaliser une comparaison à la fois dans le temps et dans l'espace, afin d'éviter au mieux les inconvénients propres à chaque type de comparaison.

Le « schéma contrôlé » consiste à mesurer l'état nutritionnel de la population cible avant et après intervention, et ceci à la fois dans le groupe d'intervention et dans un groupe témoin. La comparaison avant/après au niveau de ce dernier permet d'estimer la tendance séculaire, et donc de corriger éventuellement la comparaison avant/après au niveau du groupe d'intervention. Dans un tel schéma si la comparabilité des groupes est vérifiée avant le début du programme, puis contrôlée tout au long de son déroulement, alors on peut avoir une présomption de causalité vis à vis de l'intervention pour lui imputer une modification de l'état nutritionnel. Par ailleurs un certain nombre de facteurs de confusion peuvent alors être plus facilement identifiés et leurs effets mesurés. Ceci peut parfois être pris en compte dans l'analyse, mais il faut savoir que cela diminue le nombre de degrés de liberté et donc ampute la puissance des comparaisons effectuées. Ce « schéma contrôlé » est méthodologiquement le plus performant. Il est en contrepartie assez difficile à mettre en oeuvre, notamment en raison du choix des populations devant servir de témoin. Outre les problèmes évoqués pour le schéma simple « ici/ailleurs » à propos du choix géographique, il faut faire admettre, éthiquement et politiquement, qu'une population reste en dehors du programme pendant une durée suffisante pour permettre de juger des modifications de l'état nutritionnel dans le groupe soumis à intervention.

Le « schéma progressif » propose un compromis aux problèmes d'acceptabilité évoqués ci-dessus. Au lieu de laisser une population sans intervention pendant une longue durée, la mise en place du programme se fait de manière progressive, communauté par communauté, avec un délai plus acceptable. Ceci permet d'une part une comparaison « ici/ailleurs », d'autre part une estimation de la tendance séculaire par les mesures successives sur les nouveaux entrants (qui sont les témoins), et même l'estimation d'un « effet-dose ». Toutefois l'analyse est plus complexe et moins puissante que dans le schéma contrôlé. Par ailleurs la comparabilité est plus difficile à assurer et à contrôler sur les communautés successives. Enfin, et surtout, l'ordre d'inclusion des communautés dans l'étude est souvent lié à des problèmes d'accessibilité (géographique, culturelle etc.), dont on peut craindre qu'ils ne sont pas indépendants du risque nutritionnel. Ceci peut donc introduire un facteur de confusion.



#### **2.3.4. Remarque**

Nous avons essentiellement considéré jusqu'ici que les schémas d'étude s'adressent à des communautés. Les mêmes schémas peuvent se concevoir au niveau des individus, mais il s'ajoute alors des écueils supplémentaires. En particulier le biais d'auto-sélection devient beaucoup plus important, et même souvent impossible à préciser. Le problème de la répartition non aléatoire des perdus de vue devient également très difficile à contrôler. Enfin il apparaît une source supplémentaire de biais, non négligeable, liée aux erreurs de classification des individus dans les groupes d'étude.

#### **2.4. Indicateurs de résultats**

Dans la grande majorité des cas il s'agit d'indices anthropométriques, et nous nous limiterons à donner quelques principes pour leur utilisation dans le cadre d'une évaluation d'impact. Parmi les indices disponibles (poids/âge, poids/taille, taille/âge, périmètre brachial etc.), le choix d'un indicateur de résultat va dépendre :

- des objectifs du programme, ainsi que de ceux de l'évaluation ;
- de la situation de base (quel indice est à un faible niveau au départ ?) ;
- de la durée de l'étude (important selon la sensibilité des indices) ;
- des moyens techniques et financiers pour le recueil des mesures.

Au niveau de l'expression des indices, il est bon de rappeler :

- que les vitesses de croissance sont généralement plus sensibles aux interventions que les valeurs brutes des mesures ;
- que l'expression en centiles ou en Z-scores est standardisée sur une population de référence mais pas sur l'âge ;
- que les moyennes des indices sont moins sensibles aux erreurs (de mesure, d'estimation des âges etc.), et plus puissantes pour les comparaisons, que le pourcentage de sujets situés en dessous d'une valeur seuil ;
- que ce pourcentage en revanche est plus intéressant pour la prise de décisions de politique sanitaire.

#### **2.5. Echantillonnage**

L'élément le plus important est le choix de la tranche d'âge sur laquelle portera l'évaluation d'impact. Bien entendu ce choix doit se porter sur les âges qui ont le plus de chances d'être sensibles à l'intervention, et donc pas nécessairement sur toute la tranche soumise à cette intervention. Elargir le groupe d'étude au-delà des âges sensibles entraîne un effet de dilution, pouvant atténuer ou masquer l'effet réel ; à l'inverse, une tranche d'âge rétrécie revient à une perte de puissance.

La dimension de l'échantillon est une autre question importante. Il existe des méthodes de calcul du nombre de sujets nécessaire, que nous ne détaillerons pas ici. Nous soulignerons uniquement que ce calcul nécessite de connaître ou d'estimer un certain nombre de paramètres et de faire quelques hypothèses sur les résultats attendus. Ceci constitue fréquemment une raison supplémentaire pour réaliser une étude de faisabilité préalable.

Rappelons par ailleurs le rôle de la sélection de l'échantillon dans la plupart des facteurs de confusion évoqués plus haut (comparabilité des groupes, effets de l'âge, régression vers la moyenne, erreurs de classification).

## **2.6. Conclusion : choix, validité, interprétation**

L'évaluation d'impact d'un programme nutritionnel nécessite donc un schéma d'étude soigneusement élaboré. Un certain nombre de choix sont à faire, de décisions à prendre, qui doivent être réfléchis selon les objectifs du programme, les moyens disponibles, les considérations de terrain, les données de base etc. Tout ceci peut conduire à renoncer à cette évaluation d'impact lorsqu'il n'est pas possible de mettre en place un plan ayant des chances raisonnables de succès.

Dans le cas contraire, de toutes façons, la mise en place d'un schéma a priori performant ne constitue pas une garantie quant aux conclusions qui pourront en être tirées. En effet, il sera toujours nécessaire de vérifier que, malgré les précautions prises, des facteurs de confusion ne se sont pas glissés dans l'étude. Ceci est essentiel pour l'interprétation des résultats, ainsi que pour la validité interne de l'évaluation, c'est-à-dire pour établir que ce qui a été mesuré représente bien la réalité dans les groupes concernés.

Mais il faut aborder ensuite le problème de la validité externe, à savoir dans quelle mesure les conclusions de l'évaluation d'impact peuvent s'appliquer à d'autres populations. Il s'agit tout d'abord d'un problème de représentativité, mais il faut également savoir si l'intervention évaluée peut être reproduite facilement au niveau d'autres populations. A ce sujet il faut souligner que le management exceptionnel d'un programme au niveau de zones pilotes constitue une limitation importante à la validité externe d'une étude.

## **3. EXEMPLE : EVALUATION DE L'IMPACT NUTRITIONNEL DE DEUX STRATEGIES D'AMELIORATION DES PRATIQUES DE SEVRAGE AU CONGO**

Au Congo, deux stratégies pour l'amélioration de l'alimentation complémentaire du jeune enfant pendant la période de sevrage ont été mises en place dans des zones tests :

l'une vise le milieu urbain et l'autre est tournée vers le monde rural. Ce travail a été initié par le Laboratoire d'Etudes sur la Nutrition et l'Alimentation du Centre DGRST-ORSTOM de Brazzaville, en étroite collaboration avec la Direction de la Santé et de la Famille, dans le cadre d'un Projet d'Appui aux Activités en Nutrition financé par la Coopération Française et géré par l'UNICEF.

Nous ne fournissons pas ici de données chiffrées sur les résultats de l'évaluation d'impact (celle-ci n'étant de toute façon pas terminée), mais nous chercherons à illustrer les divers problèmes de méthodologie exposés dans la première partie.

### **3.1. Brève description des deux stratégies**

#### **3.1.1. La stratégie urbaine**

La stratégie urbaine s'adresse à des familles au sein desquelles l'aliment complémentaire donné aux nourrissons est quasiment toujours une bouillie achetée dans le commerce. L'intervention pilote repose sur la promotion et la commercialisation de la farine de sevrage « Vitafort » (Tchibindat et Trèche, 1995) dans un des quartiers anciennement urbanisé de Brazzaville, Poto-Poto. Son prix de vente est, à qualité nutritionnelle comparable, de 3 à 5 fois inférieur à celui des farines importées, et reste, rapporté au nombre de calories, équivalent à celui de la pâte de maïs fermentée utilisée localement pour la préparation des bouillies traditionnelles. La délimitation de la zone d'intervention a été calquée sur celle des circonscriptions socio-sanitaires correspondant à deux Centres de Santé Intégrés dans le nouveau Programme National de Développement Sanitaire. La stratégie d'intervention comporte une action de sensibilisation aux problèmes du sevrage de l'enfant pour les agents des centres de santé, et particulièrement ceux chargés de la surveillance de la croissance. Cette sensibilisation est bien entendu répercutée au niveau des mères et inclut la promotion de l'allaitement maternel, le respect du calendrier de sevrage et l'information sur la farine Vitafort.

#### **3.1.2. La stratégie rurale**

En milieu rural la zone d'intervention est constituée par le Plateau Kukuya, situé à environ 400 km au nord de Brazzaville. Les détails de l'intervention sont exposés par ailleurs (Moukolo *et al.*, 1995). Elle repose donc sur la formation d'« animatrices en éducation nutritionnelle et technologie alimentaire », recrutées localement, et qui ont une fonction essentielle de relais entre les structures de santé et la communauté. Leur rôle comporte d'une part une action d'éducation nutritionnelle, tournée principalement vers la promotion de l'allaitement maternel et le calendrier de sevrage, et d'autre part

la vulgarisation de technologies alimentaires permettant, à partir d'aliments locaux, la préparation de bouillies de haute densité énergétique enrichies en protéines.

### 3.2. Modalités d'évaluation : schéma général

Un schéma identique a été retenu pour l'évaluation de ces deux stratégies d'intervention, urbaine et rurale, comprenant :

- d'une part, une évaluation en terme de « processus », constituée par l'observation et l'analyse de la pénétration des stratégies au niveau des populations cibles. Les indicateurs retenus concernent le degré d'information sur les aliments de sevrage améliorés proposés, leur notoriété, leur taux d'utilisation, les modifications des pratiques de sevrage et d'alimentation du jeune enfant etc. La méthode employée est ici la réalisation d'enquêtes transversales successives sur un échantillon représentatif de la population soumise à l'intervention ;
- d'autre part, un évaluation en terme d'impact sur l'état nutritionnel des enfants soumis à l'intervention. Il s'agit d'un schéma du type « contrôlé » tel que nous l'avons décrit plus haut. Pour chaque zone d'intervention (urbaine et rurale) on définit une zone témoin qui lui soit le plus possible comparable pour toutes les caractéristiques pouvant influencer l'état nutritionnel. Des enquêtes dites « point zéro » ont été réalisées, dans les zones d'intervention et dans les zones témoins, permettant de préciser la situation nutritionnelle de départ et d'étudier la comparabilité entre les groupes. La répétition de ces enquêtes avec une méthodologie identique dans toutes les zones, deux ans après la mise en place des interventions, permettra de juger si l'évolution de la situation nutritionnelle dans la zone d'intervention s'avère différente de celle observée en zone témoin. Par ailleurs, les enquêtes transversales effectuées pour l'évaluation de processus permettront de juger du maintien de la comparabilité entre les groupes.

### 3.3. Aspects pratiques

Le schéma général de l'étude ayant été établi, il restait à résoudre tous les problèmes de choix des indicateurs et des variables potentiellement confondantes, de définition des populations cibles et des échantillons, et de délimitation des zones d'intervention et témoins.

Avant de passer en revue ces divers éléments, signalons que de nombreuses études nutritionnelles ont été réalisées depuis 5 ans au Congo, tant au niveau de Brazzaville qu'en milieu rural. Ces études nous ont fourni toutes les données de base nécessaires à la mise en place de notre programme, si bien que nous avons pu faire l'économie d'enquêtes de terrain lors de notre étude de faisabilité.

### 3.3.1. Choix des indicateurs

L'aspect essentiel de la malnutrition au Congo est une prévalence assez importante des retards de taille. L'hypothèse principale qui sous-tend notre étude est que ces retards de taille sont au moins en bonne partie liés à des mauvaises pratiques de sevrage, ainsi qu'au problème de faible densité énergétique des bouillies de sevrage. Il était donc logique de retenir comme principal indicateur de résultat du programme l'indice anthropométrique taille/âge. Bien entendu des mesures de poids sont également effectuées, de façon à pouvoir juger aussi de l'influence éventuelle de l'intervention sur les autres indices, et notamment sur les maigreurs. Mais c'est l'indice taille/âge qui a été retenu pour les calculs de nombres de sujets nécessaires.

Comme nous l'avons indiqué plus haut, cet indice sera exprimé préférentiellement en Z-scores. On considèrera à la fois la valeur moyenne de l'indice dans les différents groupes et le pourcentage de sujets situés en dessous de 2 déviations standard. Le premier indicateur sera plus performant dans les comparaisons, le second permettra de vérifier, avec l'étude de sa distribution, que l'intervention a bénéficié aux sujets les plus à risque.

### 3.3.2. Choix des co-variables

Les études préalables dont nous disposons nous ont permis d'identifier les principales variables, essentiellement d'ordre socio-économique ou concernant les pratiques de sevrage, dont il a été démontré l'influence sur le statut nutritionnel des jeunes enfants au Congo. Toutes ces variables ont donc été recueillies dans les enquêtes « point-zéro », afin de vérifier la comparabilité des groupes. D'autres variables réputées comme facteurs de confusion potentiels ont également été incluses par sécurité dans l'étude, même si les travaux précédents n'avaient pas démontré leur rôle au Congo. Le même questionnaire sera reproduit pour les enquêtes en fin d'intervention, et les principales variables sont également étudiées lors des enquêtes pour l'évaluation de processus, de façon à juger du maintien de la comparabilité des groupes.

### 3.3.3. Définition des populations cibles

La population cible de l'intervention est constituée par les enfants pendant la période allant de l'introduction d'un aliment de complément jusqu'à l'âge du passage au plat familial, soit grosso modo les enfants de 4 à 9 mois. Toutefois, vu le type d'intervention, il ne s'agit là que de ce qui est recommandé ; l'analyse de l'évolution des pratiques de sevrage au cours de l'intervention permettra de dire plus précisément quelle tranche d'âge a été concernée en réalité.

La population cible de l'évaluation d'impact a été choisie nettement plus large, puisque l'on constate au Congo une prévalence croissante des retards de taille jusque vers l'âge de 2 ans (même si l'on ne peut savoir actuellement si l'influence de mauvaises pratiques de sevrage se répercute jusqu'à cet âge). Etant donné que d'autre part la durée prévue du programme pilote est de deux ans, on a retenu comme cible pour l'étude d'impact la tranche de 4 à 27 mois. Ainsi toute la génération enquêtée en fin d'étude aura été concernée par l'intervention. Mais l'on est conscient que ce choix comporte un risque de dilution de l'effet éventuel de l'intervention si celui-ci était plus limité dans le temps.

### 3.3.4. Délimitation des zones d'étude

#### 3.3.4.1. *Stratégie urbaine*

Pour la stratégie urbaine la zone d'intervention a été définie à partir des Centre de Santé Intégrés. Au Congo le Plan National de Développement Sanitaire (P.N.D.S.) prévoit la réhabilitation des dispensaires en C.S.I., ou parfois la création de ces C.S.I. dans des zones où la couverture sanitaire était jugée insuffisante. Un des principes du P.N.D.S. est que chaque C.S.I. prend en charge une Circonscription Socio-Sanitaire géographiquement délimitée avec précision. La population de cette circonscription est recensée et bénéficie d'un certain nombre d'actions de Santé Publique. Il était donc tout à fait logique de calquer la délimitation des zones d'étude sur celle de ces circonscriptions, et c'est ce qui a été fait, à quelques exclusions près concernant des petits îlots d'habitat très différents du reste de la zone. Cette décision a été prise étant donné le très faible pourcentage de population dans ces îlots, et dans le souci de sauvegarder l'homogénéité socio-économique.

Quant au choix des C.S.I. retenus pour l'étude, il a été guidé par les éléments suivants :

- zone d'intervention suffisamment éloignée de la zone témoin pour limiter les phénomènes de contamination ;
- quartiers de niveaux socio-économiques équivalents pour avoir la meilleure comparabilité initiale possible ;
- même état d'avancement de la mise en place du P.N.D.S., avec comme corollaire que le soutien de cette mise en place relève de la même agence de développement ;
- éviter les interférences avec d'autres programmes de recherche dans le domaine de la nutrition qui se déroulent également à Brazzaville.

Compte tenu de tous ces éléments le choix était très restreint. Ce sont finalement deux quartiers anciennement urbanisés de Brazzaville qui ont été retenus, Bacongo et Poto-

Poto, dans lesquels la mise en place du P.N.D.S. est soutenue par l'UNICEF et prévoyait la mise en service de respectivement 3 et 2 CSI courant 93.

#### *3.3.4.2. Stratégie rurale*

Cette fois le problème était très différent puisque la zone d'intervention a été définie très tôt. Il s'agit du Plateau Kukuya, situé environ à 400 km au nord de Brazzaville, pour lequel des données de base concernant l'état nutritionnel et les habitudes alimentaires étaient disponibles. L'intervention, d'ailleurs, est spécifiquement conçue pour cette zone, notamment au niveau des technologies alimentaires.

La difficulté résidait donc d'abord dans le choix d'un schéma d'étude, puis dans celui d'une zone témoin. La première solution envisagée a été de diviser le Plateau en deux zones (une d'intervention et une témoin) ; mais cela s'est avéré impossible pour des raisons d'acceptabilité par la population, et également à cause de la faiblesse de l'effectif disponible. Il a donc été décidé de prendre pour zone témoin le Plateau voisin de Djambala, dont la population est très proche du point de vue ethnique. La proximité peut être source de contamination, mais celle-ci devrait être négligeable étant donné le type d'intervention. La mise en place du P.N.D.S., qui concerne tout le Congo et pourrait être un facteur de confusion important, est heureusement programmée dans une même phase pour l'ensemble de la région. Toutefois ce choix a posé des problèmes d'échantillonnage que nous évoquons ci-dessous.

#### 3.3.5. Echantillonnage

##### *3.3.5.1. Nombre de sujets nécessaire*

C'est le premier aspect à considérer. Disposant de données sur les retards de taille dans les populations concernées par l'étude, nous avons dressé le tableau 1 pour la construction duquel nous avons, en quelque sorte, effectué les calculs « à l'envers », de façon à présenter, pour divers niveaux d'effectifs, les hypothèses qui doivent être faites sur les résultats de l'intervention pour envisager un résultat statistiquement significatif. Cela permet de donner un ordre de grandeur et de juger rapidement des chances de succès de l'étude. Par la suite un calcul « à l'endroit » peut être réalisé pour préciser la taille de l'échantillon, en fonction des hypothèses de résultats qui seront finalement retenues. Dans notre étude ceci n'a pas été fait, d'une part, car les hypothèses de résultats ne sont pas faciles à faire, d'autre part car nous avons rapidement décidé de recourir, nous y reviendrons plus loin, à une étude exhaustive dans les populations cibles.

Tableau 1

Hypothèses de différences nécessaires entre deux populations selon l'effectif de l'échantillon, pour une mise en évidence significative au seuil 5 % et avec une puissance de 90 %.

n/ groupe	Z-score moyen	Prévalence Initiale							
		12,5%	15 %	17,5 %	20 %	22,5%	25%	27,5%	30%
250	- 0,38	- 7,3	- 8,1	- 8,8	- 9,3	- 9,9	- 10,4	- 10,8	- 11,2
500	- 0,27	- 5,5	- 6,0	- 6,5	- 6,9	- 7,2	- 7,6	- 7,9	- 8,0
750	- 0,22	- 4,5	- 5,0	- 5,4	- 5,7	- 6,0	- 6,3	- 6,5	- 6,7
1000	- 0,19	- 4,0	- 4,3	- 4,7	- 5,0	- 5,2	- 5,5	- 5,7	- 5,8
1500	- 0,15	- 3,3	- 3,6	- 3,9	- 4,1	- 4,3	- 4,5	- 4,6	- 4,7

Dans le tableau 1 nous lisons par exemple, à l'intersection de la colonne 30 % de prévalence initiale et de la ligne 500 sujets par groupe, la valeur « - 8 % ». Ceci signifie que si la prévalence réelle du retard de taille dans la population soumise à intervention est de 30 % au temps zéro, et que cette prévalence a chuté de 8 % après intervention, un effectif de 500 sujets par groupe sera nécessaire pour mettre en évidence cette différence réelle entre les deux populations comparées, au seuil 5 % et avec une puissance de 90 %.

Il faut bien comprendre que cela ne signifie pas qu'une différence d'au moins 8 % doit être observée sur les échantillons pour que l'étude soit concluante. Cela signifie bien que si la différence réelle (mais qui restera inconnue) entre les populations (et non les échantillons) est de 8 %, les fluctuations d'échantillonnage font que, avec une distribution normale de la variable, la réalisation de 100 échantillons de 2 x 500 personnes permettrait statistiquement 90 fois (puissance = 90 %) la mise en évidence d'une différence significative au seuil 5 %.

Dans notre étude nous avons considéré qu'en milieu urbain, où la prévalence initiale des retards de taille est estimée à 15 %, un effectif de 1500 enfants par groupe était souhaitable. En effet cela permettrait la mise en évidence, aux conditions fixées, d'une différence réelle de 3,6 % dans la population. Compte tenu du fait que l'adhésion de la population à l'intervention ne sera certainement pas massive, il s'agirait là d'un résultat déjà satisfaisant. En milieu rural en revanche la prévalence initiale est estimée



aux alentours de 30 %, et l'on peut espérer une participation nettement plus forte à l'intervention. Dans ces conditions une diminution de 8 à 10 % de la prévalence peut être envisagée, et un effectif de 500 sujets par groupe nous a semblé suffisant.

### *3.3.5.2. Critères d'inclusion/exclusion*

Le seul critère qui nous a paru important concerne la durée de résidence dans la zone retenue. En effet, il est préférable de faire porter l'évaluation uniquement sur des enfants ayant été soumis, ou du moins ayant pu être soumis, à l'intervention. Pour cela il est nécessaire que ces enfants aient résidé dans la zone depuis leur naissance. Nous avons toutefois laissé une tolérance d'absence temporaire fixée à 1 mois avant l'âge de 9 mois et à 3 mois de 9 à 27 mois. Bien entendu les mêmes critères s'appliquent à la zone témoin, de façon à préserver la comparabilité. En milieu rural en revanche, où la population est beaucoup moins mobile, il n'a pas été jugé nécessaire d'imposer ce type de critère.

### *3.3.5.3. Constitution des échantillons*

La limite a été ici la taille des populations dans les zones d'étude. En effet, étant données les difficultés rencontrées pour délimiter les zones d'étude, celles-ci se sont avérées être de taille modeste.

En milieu urbain, essentiellement pour des raisons d'homogénéité, il n'était possible de retenir que les circonscriptions socio-sanitaires correspondant à deux C.S.I. pour le quartier Poto-Poto et trois C.S.I. pour Bacongo. La population résidant dans ces zones était estimée aux alentours de 30 000 habitants pour chaque quartier, soit un potentiel d'environ 2000 enfants de 4-27 mois. Compte tenu d'une proportion importante d'exclusions à prévoir, l'enquête exhaustive restait le seul moyen d'atteindre l'objectif fixé de 1500 enfants par groupe. L'échantillon était donc constitué de tous les enfants âgés de 4 à 27 mois et répondant aux critères de résidence dans la zone considérée. Pour approcher le plus possible l'exhaustivité il a été nécessaire de dresser des plans des quartiers, parcelle par parcelle, et de faire un gros travail de recensement. Finalement moins de 1 % des sujets recensés n'ont pu être enquêtés, et moins de 2 % ont refusé de participer.

Dans la zone d'intervention en milieu rural, c'est-à-dire sur le Plateau Kukuya, une importante enquête nutritionnelle avait déjà été effectuée en Avril 1992. La méthode retenue avait été celle du sondage en grappes, avec une fraction de sondage de 0,4. La population totale étant de l'ordre de 16000 habitants, c'est finalement un peu plus de 400 enfants de 4 à 27 mois qui se sont retrouvés dans l'échantillon, dont 25 % provenant de la petite ville de Lékana. Pour des raisons évidentes d'acceptabilité, il était impossible de pratiquer une nouvelle enquête juste avant l'intervention, début

1993. Nous avons donc dû nous contenter de cet échantillon, un peu inférieur aux 500 sujets souhaités. Dans la zone témoin, peuplée environ de 16 000 habitants également, le problème était que près de 60 % des sujets résident dans la ville de Djambala. La reproduction de la même méthode d'enquête qu'en zone d'intervention (sondage en grappes) aurait conduit à un échantillon conservant grosso modo 60 % de sujets « urbains », en disproportion donc avec les 25 % dans l'enquête faite sur le Plateau Kukuya. Nous avons donc été conduits à stratifier notre échantillon sur la résidence pour respecter cette répartition 25-75 % entre « centre » et « périphérie ». Le sondage en grappes a été reproduit pour la ville de Djambala, mais pour la périphérie le calcul donnait une estimation de 350 à 400 enfants de 4-27 mois au total. De nouveau tout sondage était inutile et l'enquête a été exhaustive pour les enfants des villages hors de Djambala.

### 3.4. Evolution et limites

#### 3.4.1. Evaluation d'un « Package »

Notre étude vise l'évaluation d'impact d'un programme dont les stratégies ont été définies en fonction de l'analyse préalable de la situation nutritionnelle. Cependant la mise en place simultanée du Programme National de Développement Sanitaire nous a imposé de moduler l'intervention prévue. L'impact évalué sera donc celui d'un « package », l'intervention au niveau des pratiques de sevrage couplée à un programme sanitaire beaucoup plus vaste, en comparaison avec l'impact du seul programme sanitaire. Ceci pose différents problèmes :

- risque de biais lié à un avancement du P.N.D.S. différant d'une zone à l'autre ; c'est malheureusement ce que l'on a constaté au niveau rural, le Plateau de Lékana ayant fait l'objet de davantage d'attention dans la mise en place du P.N.D.S. que le plateau voisin. Toutefois, il faudra attendre l'évaluation finale pour savoir si le niveau des co-variables retenues a effectivement été modifié de façon sensible dans une zone par rapport à l'autre ;
- risque de dilution de l'effet de l'intervention nutritionnelle : en cas d'impact de l'intervention nutritionnelle mineur par rapport à celui du programme sanitaire, la différence d'évolution entre zones d'intervention et témoin peut ne plus apparaître significativement ;
- problème d'interprétation : on ne dispose pas d'éléments pour supposer l'indépendance des impacts, au niveau anthropométrique, de chacun des programmes (l'inverse est même plus vraisemblable). Toute conclusion quant à l'effet de la seule intervention nutritionnelle devient donc impossible.

### 3.4.2. Comparabilité des Groupes

La comparabilité initiale des groupes est satisfaisante en milieu urbain même si, comme attendu, des différences pré-existent entre les deux quartiers de Brazzaville. En revanche la différence de méthodologie pour la constitution des échantillons en milieu rural, ajoutée au fait que les deux zones n'ont pas été enquêtées la même année, pose un problème. Même si l'analyse établit la comparabilité initiale des diverses variables retenues, il restera le point d'interrogation de la tendance séculaire pour l'année séparant les deux enquêtes. Si elle a été négligeable, il n'y a pas de problème. Si elle a été dans le sens d'une aggravation de l'état nutritionnel, cela risque de masquer un impact réel du programme. Si elle a été dans le sens d'une amélioration de l'état nutritionnel, cela risque à l'inverse d'exagérer l'effet de l'intervention. On devra donc essayer d'estimer par ailleurs cette tendance séculaire.

Le maintien de la comparabilité des groupes au cours de l'étude doit être vérifié. Au niveau rural, ainsi que nous l'avons signalé au paragraphe précédent, on redoute ici un facteur de confusion produit par l'implication plus précoce des agences de développement pour la mise en place du P.N.D.S. au niveau du Plateau de Lékana, par rapport à celui de Djambala. Mais au niveau urbain la situation est bien pire, en raison des importants événements socio-politiques qui ont secoué la capitale congolaise depuis la mise en place de l'étude. Ceux-ci ont tout d'abord considérablement freiné le déroulement de l'intervention, la production et la diffusion de la farine Vitafort ayant été interrompues plusieurs mois. Mais surtout ces événements ont provoqué d'importants mouvements de population, concernant à la fois la zone d'intervention et la zone témoin, y compris avec passages d'un quartier à l'autre. Il a donc fallu renoncer à une évaluation d'impact en milieu urbain, qui, à l'évidence, ne peut plus donner de résultats interprétables.

### 3.4.3. Validité externe

Au niveau rural l'implication des différents acteurs, pour la réalisation pratique de l'intervention nutritionnelle dont nous cherchons à évaluer l'impact, a été importante. Si cette évaluation s'avère concluante, malgré les problèmes possibles d'interprétation des résultats évoqués plus haut, il faudra se poser la question de la reproductibilité de l'intervention à une échelle supérieure. En effet, bien que la stratégie évaluée ait été conçue pour s'intégrer dans la politique nutritionnelle nationale, les moyens humains et financiers mis en oeuvre à une échelle pilote, ainsi que la motivation des acteurs, sont souvent des facteurs importants de succès d'une intervention.

#### 4. CONCLUSION

L'évaluation d'impact d'un programme nutritionnel s'avère donc relativement complexe du point de vue méthodologique et, nous l'avons illustré, les contraintes de terrain rendent sa réalisation encore plus difficile. Finalement, la mise en oeuvre d'une telle étude est essentiellement une affaire de compromis entre ce qui est méthodologiquement acceptable et ce qui est possible en pratique.

Le problème de la reproductibilité doit être relativisé dans la mesure où les délais d'étude nécessaires font que, de toutes façons, l'évaluation d'impact sera rarement utile au programme lui-même. Il s'agit donc de pouvoir dire avant tout si le type d'intervention qui est évalué donne ou non des résultats probants en terme d'impact sur l'état nutritionnel de la population cible. Par ailleurs, l'évaluation d'impact s'accompagne d'une évaluation dite « de fonctionnement », qui fournira des informations précieuses pour adapter le programme à une autre situation, et pour l'intégrer dans une politique nationale.

Se donner les moyens d'évaluer l'impact d'un programme paraît logiquement fondamental pour faire évoluer les politiques d'interventions nutritionnelles. Il est pourtant rare qu'à l'issue d'une intervention dans le domaine de la nutrition on puisse savoir dans quelle mesure les objectifs visés ont été atteints. Ceci est lié essentiellement aux difficultés méthodologiques de l'évaluation d'impact, à son coût, et à la durée de l'étude qui la rend vulnérable vis à vis de modifications importantes et non maîtrisables du contexte général. On insistera donc pour conclure sur une double nécessité :

- celle de concevoir l'évaluation d'impact d'un programme en même temps que le programme lui-même, de façon à définir un schéma d'étude adapté au mieux aux particularités du programme comme aux contraintes budgétaires ;
- celle d'explorer et de développer toutes les méthodes d'évaluation dites « qualitatives » qui, si elles ne peuvent remplacer les résultats quantitatifs d'une étude d'impact, sont plus légères et plus facilement mises en oeuvre.

#### REFERENCES

BOUYER J., HEMON D., CORDIER S., DERRIENIC F., STÜCKER I., STENGEL B., CLAVEL J., 1993 - *Epidémiologie : Principes et méthodes quantitatives*, Paris, Editions INSERM, 498 p.

*Schémas et contraintes pour l'évaluation d'un programme d'amélioration*

- HAAGA J., CLARK L., EDMONSTON B., HABICHT J.-P., KENRICK C., KURZ K., MASON J., TEST K., 1985 - *Evaluating Effects of Child Feeding Programs (Report of Workshop)*, Ithaca, New York, Cornell University, 29 p.
- JENICEK J., CLEROUX R., 1983 - *Epidémiologie*, Edisem, Québec, 454 p.
- KLEIN R.E., READ M.S., RIECKEN H.W., BROWN JR J.A., PRADILLA A., DAZA C.H., 1979 - *Evaluating the Impact of Nutrition and Health Programs*, New York, Plenum Press, 462 p.
- MASON J.B., HABICHT J.-P., TABATABAI H., VALVERDE V., 1987 - *La Surveillance Nutritionnelle*, Genève, OMS, 212 p.
- MOUKOLO A., TCHIBINDAT F., TRECHE S., MARTIN-PREVEL Y., PEZENNEC S., GAMI N., LOUYAT DE DIBANTSA Y., 1995 - « L'opération d'éducation nutritionnelle et de transfert de technologie sur le plateau Kukuya (Congo) ». In Trèche S., de Benoist B., Benbouzid D., Delpuech F., éd: *L'alimentation de complément du jeune enfant*, Paris, Editions Orstom: sous presse.
- RUMEAU-ROUQUETTE C., BREART G., PADIEU R., 1985 - *Méthodes en Epidémiologie (3è éd.)*, Paris, Flammarion Médecine-Sciences, Paris, 398p.
- TCHIBINDAT F., TRECHE S., 1995 - « Vitafort: une farine infantile de haute densité énergétique au Congo ». In Trèche S., de Benoist B., Benbouzid D., Delpuech F., éd: *L'alimentation de complément du jeune enfant*, Paris, Editions Orstom: sous presse.
- WHO, 1983 - *Measuring changes in nutritional status*, Geneva, World Health Organization, 104 p.