

Aspects statistiques du ciblage des politiques et programmes nutritionnels dans les pays en développement

Pierre Traissac ^a Bernard Maire ^a Francis Delpuech ^a

Des subventions généralisées aux programmes ciblés

Pour maintenir ou accroître les revenus des populations tout en facilitant un meilleur approvisionnement alimentaire, nombre de pays classés parmi les pays en développement se sont orientés vers des programmes de large soutien à la consommation en subventionnant systématiquement les prix des denrées de base. Le bilan de ces politiques est cependant mitigé : malgré un certain taux de réussite, ils ont des coûts élevés et entraînent des fuites importantes vers les non nécessaires.

Les programmes de stabilisation économique et d'ajustement structurel, qui sont plus que jamais à l'ordre du jour, conduisent de surcroît à diminuer de façon parfois rapide les subventions alimentaires lorsqu'elles représentent une part significative des dépenses des Etats. Néanmoins, l'ambition de la plupart des programmes d'ajustement économique n'est pas tant d'opérer une simple réduction des dépenses publiques que d'effectuer une véritable réallocation de ressources vers les plus démunis dans les domaines de la nutrition, de la santé ou de l'éducation au travers d'un ciblage efficace (Pinstrup-Andersen, 1991).

Des arguments plus directement d'ordre nutritionnel militent également pour la mise en œuvre de politiques et programmes nutritionnels ciblés. En effet, dans un nombre croissant de pays du Sud (pays à revenus intermédiaires et de manière générale dans les grands centres urbains des pays en développement), on voit s'amorcer une "transition nutritionnelle", avec émergence d'une population présentant des problèmes de surcharge

^aORSTOM ; Laboratoire de Nutrition Tropicale. Montpellier — Centre Collaborateur de l'OMS pour la Nutrition — BP 5045 34032 MONTPELLIER Cedex 1
e-mail : traissac@mpl.orstom.fr

pondérale et des troubles qui y sont liés, alors même que subsistent faim, malnutrition et carences spécifiques, pour une autre partie de la population (Maire *et al.* 1992; Popkin 1994).

Le contexte du ciblage

De manière générale, qu'il s'agisse de lutter contre la pauvreté ou d'améliorer directement la situation nutritionnelle, le ciblage est un processus qui s'inscrit dans un ensemble, par exemple une politique ou un programme ; le choix de la population cible et de la stratégie de ciblage ne peut se faire indépendamment de cet ensemble.

Si la nature du ciblage dépend du programme mis en œuvre, la réussite d'une politique ou d'un programme est avant tout fonction de la pertinence des interventions projetées, par rapport aux problèmes identifiés. Ainsi, en ce qui concerne l'amélioration de la situation nutritionnelle, le concept évident d'un lien entre malnutrition et pauvreté à un niveau global, a conduit trop fréquemment à une vision réductrice des programmes d'intervention nutritionnelle en les transformant en programmes d'assistance aux plus pauvres. Or, il a été montré, à de nombreuses occasions, que les causes de malnutrition sont complexes, différent selon l'environnement physique, socioculturel ou économique des communautés, des ménages ou des individus et varient dans le temps. Outre les disponibilités alimentaires insuffisantes (quantité, qualité) au niveau des ménages, peuvent être mises en cause les maladies infectieuses, l'accès aux soins, la prise en charge insuffisante des membres vulnérables du ménage etc. Seule une analyse préalable des déterminants des malnutritions et la définition précise du "risque nutritionnel"¹ dans le contexte étudié pourront conduire à la caractérisation des groupes cibles du programme, d'indicateurs pour les identifier, et d'une stratégie pour les atteindre.

Stratégies classiques de ciblage

Si le concept même de ciblage s'impose facilement, sa réalisation n'en est pas pour autant une tâche aisée. Son succès dépend en effet des réponses apportées aux questions évoquées ci-dessus : qui sont les personnes à risque et comment les atteindre ? Il n'existe pas de méthodologie simple qui permette de définir immédiatement ces personnes, quel est leur niveau

¹Risque nutritionnel est à prendre ici au sens large, *i.e.*, risque de décès, de maladie, de déficit fonctionnel (capacité physique, immunité, développement mental) ou plus simplement de faim, par suite d'une alimentation insuffisante ou inadaptée.

de risque, ni comment on pourra les faire bénéficier d'un programme spécifique, et à quel coût. Faute de réponses toutes faites, les responsables politiques peuvent être saisis d'une appréhension légitime lorsqu'ils envisagent des programmes ciblés, malgré les bénéfices espérés. Néanmoins, les expériences accumulées dans un passé récent fournissent aujourd'hui des éléments d'appréciation suffisamment variés et nombreux pour être à même de dresser un tableau assez complet des choix possibles et de leurs implications (Grosh, 1994).

Si l'on excepte les programmes de subvention des produits de base, qui correspondent pratiquement à une absence de ciblage, tous les autres programmes s'adressent en fait à l'une des deux grandes modalités de ciblage : l'auto-ciblage, par lequel les individus à risque se portent eux-mêmes sur un produit ou un service programmé à leur intention, et le ciblage direct de ces personnes par le programme lui-même. Dans ce dernier cas, le repérage peut se faire sur des groupes de personnes comme sur des individus selon la nature du programme : ciblage sur critères socio-économiques de quartiers ou de ménages, ciblage des groupes vulnérables tels que femmes enceintes ou jeunes enfants, jusqu'au ciblage individuel selon l'état de santé (Maire *et al.* 1995). En pratique, on utilise souvent les différentes stratégies de façon combinée : par exemple, ciblage initial de quartiers pauvres, puis, au sein de ces quartiers, ciblage spécifique des femmes enceintes et des enfants de moins de 5 ans.

Evaluation et mesure de la performance d'une stratégie de ciblage

Que ce soit en matière de subventions ou d'interventions alimentaires et nutritionnelles, l'objectif de toute stratégie de ciblage est d'améliorer l'efficacité par rapport aux coûts, en touchant spécifiquement des groupes et/ou des individus à risque. Les critères d'évaluation doivent donc mesurer la capacité du ciblage à inclure la population cible dans l'intervention mais aussi sa capacité à exclure la population non-cible. Les méthodes d'analyse proposées pour l'évaluation du ciblage (Habicht *et al.* 1984) sont proches des méthodes d'analyse de sensibilité/spécificité et de valeurs prédictives qui sont utilisées en épidémiologie pour mesurer la performance d'un test de dépistage², les capacités de prédiction d'une fonction de risque ou encore les

²Si l'on s'intéresse par exemple au dépistage d'une maladie M ($M+$: individu malade, $M-$: individu non malade) au moyen d'un test T ($T+$: test positif, $T-$: test négatif) on a les définitions suivantes : sensibilité du test $Se = P(T+ / M+)$ (capacité du test à détecter les malades), spécificité du test $Sp = P(T- / M-)$ (capacité du test à identifier les non-malades), valeur prédictive positive $VPP =$

qualités d'un système de surveillance (Dabis *et al* 1992). Par analogie, les ménages ou les individus d'une population concernée par une intervention ciblée vers une catégorie à risque peuvent être classés dans un tableau de contingence 2x2 (cf. tableau 1) selon qu'ils sont bénéficiaires ($B+$) ou non ($B-$) et à risque nutritionnel ($R+$) ou non ($R-$). Ainsi la proportion de personnes ou de ménages à risque qui bénéficient de l'intervention, ce qui correspond à la sensibilité ($Se = P(B + /R+)$), représente la couverture par le programme de la population à risque. On peut noter que ce critère de sensibilité ne sert cependant pas à évaluer directement la qualité du mécanisme de ciblage. Il évalue surtout la capacité du programme à toucher les personnes à risque et donc dépend avant tout de la nature et de la mise en œuvre de l'intervention : ainsi, dans le cas d'une intervention passant par l'utilisation d'un système de santé, la couverture du programme dépendra de la couverture de la population par ce système, mais aussi de la fréquentation de ce système par les personnes à risque, ainsi que de la façon dont ce système identifie ces personnes pour leur attribuer l'intervention.

Tableau 1 *Classification selon le critère de ciblage et la participation au programme.*

		Risque nutritionnel	
		Oui $R+$	Non $R-$
Bénéficiaires du programme	Oui	$B+$ a	b
	Non	$B-$ c	d

Le deuxième critère utile, qui correspond à la valeur prédictive positive ($VPP = P(R + /B+)$), mesure le taux de ciblage de l'intervention : c'est la proportion des bénéficiaires du programme qui sont des personnes à risque. Un problème limite néanmoins l'interprétation de ce critère ; en effet, plus la proportion de personnes à risque dans la population est élevée, plus grande est la probabilité que le taux de ciblage soit élevé. Aussi, il a été proposé d'ajuster le taux de ciblage par rapport à la proportion de personnes à risque dans la population en calculant un rapport relatif de ciblage qui est obtenu en divisant la prévalence de personnes à risque chez les bénéficiaires du programme par la prévalence de personnes à risque chez les non bénéficiaires (Yip *et al.* 1991) soit $k' = P(R + /B+)/P(R + /B-)$.

$P(M + /T+)$ (probabilité d'être effectivement malade lorsque déclaré tel par le test) et de manière symétrique valeur prédictive négative $VPN = P(M - /T-)$. On utilise également les rapports $k = Se/(1 - Sp)$ (un individu a k fois plus de chances d'être positif au test s'il est malade que s'il est non malade) et $k' = VPP/(1 - VPN)$ (risque relatif de maladie : un individu a k' fois plus de chances d'être malade s'il est positif au test que s'il est négatif).

Ce rapport relatif de ciblage peut s'interpréter comme un risque relatif en épidémiologie.

Très souvent, la règle de décision sur laquelle est basée le ciblage découle du choix d'une valeur seuil sur une grandeur continue : indice de niveau économique ou ligne de pauvreté dans le cas d'un ciblage au niveau des ménages, probabilité de risque nutritionnel pour un ciblage individuel par exemple. A chaque valeur seuil correspond une répartition différente des individus dans les 4 cases du tableau de contingence 2x2 et donc une valeur de sensibilité et de spécificité. On représente généralement les résultats associés aux différents seuils de coupure par une courbe dite courbe ROC (Receiver Operating Characteristic Curve) ROC (Hanley et McNeil 1982) comportant en abscisse le taux de faux positifs soit $1 - Sp = P(B + / R -)$ et en ordonnée le taux de vrais positifs ou sensibilité $Se = P(B + / R +)$ (cf. figure 1).

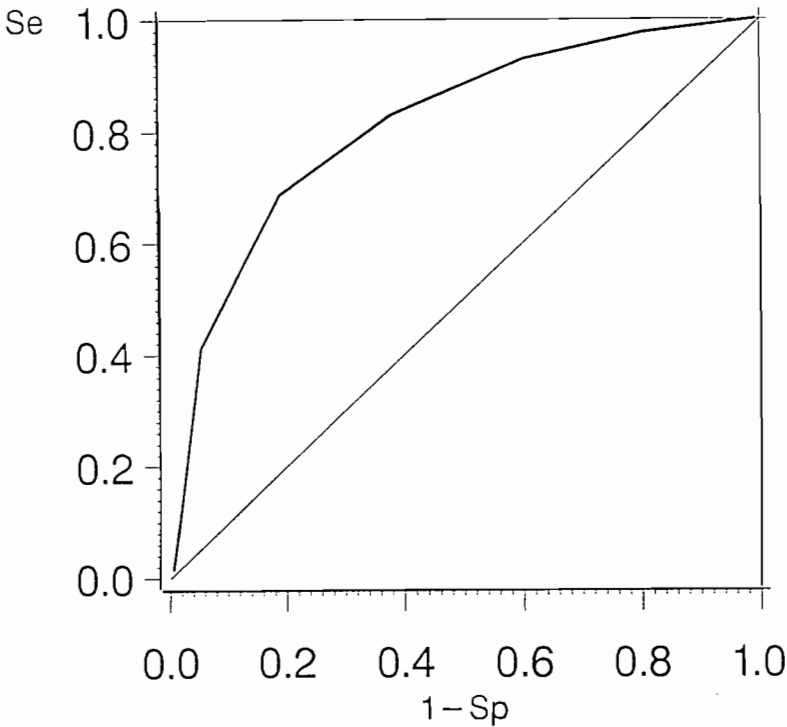


Figure 1

Exemple de courbe ROC.

Le choix du seuil dépend de l'aspect que l'on souhaite privilégier (sensibilité ou spécificité) et des coûts —au sens mathématique et général du

terme— associés aux mauvais classements. Si l'on sait évaluer les "coûts" associés aux "mauvais classements" soit $C(FP)$ pour un faux positif (individu atteint à tort par le programme) et $C(FN)$ pour un faux négatif (personne à risque mais n'ayant pas bénéficié du programme), on peut montrer que le "coût" global est

$C = P(R+)C(FN)(1 - Se) + (1 - P(R+))C(FP)(1 - Sp)$ si $P(R+)$ est le taux de personnes à risque dans la population. Le choix optimal (coût global minimum) est le point de la courbe ROC dont la pente est $[C(FP)/C(FN)] \times [(1 - P(R+))/P(R+)]$. Si les "coûts" sont égaux, le choix optimal est le point de la courbe de pente $[1 - P(R+)]/P(R+)$: si la prévalence d'individus à risque est faible (*ie* : pente élevée), il ne faut pas se placer dans le haut de la courbe car cela conduirait à un nombre très grand de faux positifs. Par contre, si on pense qu'il est très important de faire bénéficier du programme les personnes à risque et peu important de cibler à tort des personnes non à risque, on choisira des valeurs de coûts $C(FP)$ faibles et $C(FN)$ élevés. La formule nous incite alors à choisir un un seuil de coupure déplacé vers le haut de la courbe (pente plus faible) car on privilégie la sensibilité par rapport au cas précédent (coûts égaux).

Si le choix du point de la courbe ROC minimisant le coût global est mathématiquement facile, il est par contre difficile de pondérer en termes financiers et humains, les coûts associés aux deux types de mauvais classements sur lesquels est basé le calcul.

Collecte des données pour le ciblage des interventions nutritionnelles

Pour la mise au point d'un programme d'intervention ciblé et/ou son évaluation il est nécessaire de disposer d'une information récente et fiable. Cette source d'information doit porter sur la totalité de la population de la zone concernée par le programme ou sur un échantillon représentatif. Une solution serait de faire appel à des systèmes d'information existants (systèmes d'information sanitaire, surveillance nutritionnelle), lorsqu'ils fonctionnent sur une base représentative. Un grand nombre de pays ne disposent pas de tels systèmes et il convient alors d'envisager l'évaluation sur la base d'enquêtes qui peuvent être relativement légères mais qui doivent être réalisées sur un échantillon aléatoire représentatif de la population de la zone d'extension du programme.

La conception même de ces enquêtes par sondage doit tenir compte des contraintes propres aux pays en développement : on ne dispose pas toujours des éléments nécessaires ni de la précision requise pour la construction et

la réalisation du plan d'échantillonnage (recensements absents ou anciens, modifications rapides du tissu urbain par exemple). Des techniques particulières doivent être envisagées au cas par cas. Par exemple, échantillonnage sur photos aériennes comme cela a pu être fait dans des camps de réfugiés à la structure très régulière, ou, bien qu'ils aient été utilisés dans un autre contexte, les sondages aréolaires utilisant des images satellites comme base de sondage en milieu urbain (Dureau *et al.* 1989). D'autre part, la nécessité de prise en compte du plan de sondage, parfois complexe, pose des problèmes d'estimation des paramètres des modèles utilisés pour l'analyse des résultats de l'enquête³ (Pfefferman 1993).

De plus, certaines informations, en particulier de nature économique, comme les revenus, sont particulièrement difficiles à apprécier quand il s'agit par exemple de personnes travaillant dans le secteur informel. Se pose alors le problème d'estimer le niveau économique à partir d'indicateurs indirects du revenu (Traissac *et al.* 1997).

Enfin, l'étude de la perception des programmes et interventions nutritionnelles par les populations elles mêmes, fait souvent appel, en plus des enquêtes quantitatives classiques, à des techniques qualitatives issues des sciences sociales telles que groupes focalisés ou entretiens semi-dirigés (techniques souvent regroupées sous le nom de méthodes 'RAP' : Rapid Assessment Procedures ; Scrimshaw et Gleason 1992 ; Kumar 1994). La gestion et l'analyse des informations posent ici des problèmes particuliers puisque les données de base recueillies lors de ces enquêtes sont pour une large part de nature textuelle (Lebart et Salem 1994 ; Miles et Huberman 1994).

Les coûts supplémentaires entraînés par de telles enquêtes sont à considérer par comparaison au coût total des interventions. Ils peuvent et doivent être intégrés dans la conception même des programmes.

Pour ce qui concerne le point particulier de la validation des modèles, notamment statistiques, utilisés pour le ciblage, même si l'idéal est de collecter un deuxième échantillon complètement indépendant du premier pour valider la règle de classement, ces coûts et les difficultés pratiques de collecte de données font apparaître l'intérêt de techniques telles que celles de l'échantillon test ou de la validation croisée. En effet, pour la construction

³En effet, la plupart des modèles et des formules présentés dans les ouvrages classiques de statistique de même que les logiciels couramment utilisés pour leur mise en œuvre font implicitement l'hypothèse que les unités statistiques ont été échantillonnées selon un sondage aléatoire simple. Lors des enquêtes sur le terrain, ce mode d'échantillonnage étant, pour diverses raisons, impossible à mettre en œuvre, il est souvent utilisé des plans de sondage à un ou plusieurs degrés, en général associés à une stratification préalable de la zone d'étude.

d'une règle de ciblage, c'est-à-dire du processus permettant de décider à partir d'un certain nombre de ses caractéristiques si une personne (ou un ménage) va être ou non bénéficiaire du programme, on procède en général de la façon suivante : dans un premier temps à partir d'un échantillon (dit échantillon de base ou d'apprentissage) d'unités statistiques (personnes ou ménages) pour lesquelles on connaît à la fois les caractéristiques potentiellement prédictives du risque nutritionnel et la réalité de ce risque, on bâtit un modèle expliquant au mieux ce dernier en fonction des caractéristiques.

L'objectif final est bien évidemment d'utiliser de façon opérationnelle la règle ainsi construite pour décider, à partir d'un certain nombre de caractéristiques simples, si la personne ou le ménage va être bénéficiaire de l'intervention, sans observer directement son état nutritionnel. Néanmoins, dans un deuxième temps, on s'attache à valider le modèle en comparant ce qu'il prédit à ce qui est observé en réalité, pour des unités statistiques pour lesquelles on connaît à la fois les caractéristiques et le risque nutritionnel, mais n'ayant pas été utilisées pour la construction de la règle.

En effet, si l'on utilise les unités statistiques de l'échantillon de base pour valider le modèle et notamment pour estimer des quantités telles que sensibilité, spécificité ou encore le taux d'erreur de classement (ou coût total) il est clair que l'on va obtenir des estimations biaisées des valeurs réelles (*ie* celles concernant la population à laquelle on s'intéresse), puisqu'on utilise les observations mêmes qui ont servi à construire la règle. La méthode de l'échantillon test consiste à tirer au hasard un certain pourcentage (par exemple 30%) d'unités statistiques de l'échantillon de départ qui ne participeront pas à la construction de la règle (échantillon test) et seront utilisées pour la validation. Une généralisation de cette approche est la validation croisée⁴ : on divise l'échantillon de départ en un certain nombre de sous échantillons (par exemple 10) et l'on répète autant de fois la procédure décrite ci-dessus, chacun des sous-échantillons étant à tour de rôle échantillon test. On fait ensuite la moyenne des valeurs obtenues pour estimer les quantités d'intérêt (spécificité, sensibilité etc.).

Exemples d'utilisation de méthodes statistiques (discrimination, classement), pour construire une règle de ciblage

Les méthodes utilisées pour la construction d'une règle de ciblage doivent être évaluées sur un ensemble de critères incluant leurs qualités sta-

⁴Cette méthode est en particulier intéressante lorsque l'échantillon dont on dispose n'est pas suffisamment grand pour permettre de constituer un échantillon test sans trop réduire la taille de l'échantillon d'apprentissage.

tistiques mais prenant également en compte la nature opérationnelle du ciblage. Dans cette optique, on doit privilégier la simplicité d'emploi du modèle utilisé sans néanmoins que cela soit au détriment des qualités de la règle de classement. La recherche du meilleur compromis simplicité/performance doit prendre en compte les points suivants :

- nombre de prédicteurs aussi réduit que possible : nécessite une analyse conceptuelle approfondie préalablement à la collecte de l'information de base, même si dans une certaine mesure les procédures de sélection de variables type pas à pas peuvent aider à simplifier certains modèles statistiques,
- facilité d'obtention des valeurs des prédicteurs : la règle de ciblage ne sera véritablement opérationnelle que si le coût de l'information à collecter pour décider si une personne ou un ménage doit être bénéficiaire du programme est faible,
- forme du modèle : dans une optique opérationnelle on pourra, à performances égales, préférer une règle de ciblage plus intuitive ou plus simple pour les décideurs ou les personnels chargés de la mise en œuvre sur le terrain.

Pour ce qui concerne les méthodes statistiques elles mêmes, si l'on s'intéresse, par exemple, au ciblage de certains groupes ayant un risque nutritionnel spécifique (enfants avec un retard de croissance, enfants de faible poids de naissance, adultes présentant une insuffisance pondérale), ou de ménages comportant ces personnes à risque, on se situe dans le cadre général des problèmes de discrimination et de classement : on cherche à prédire, à partir d'un certain nombre de descripteurs socio-économiques du ménage et de caractéristiques de l'individu simples à collecter, la nécessité de faire bénéficier ou non l'unité statistique considérée (individu, ménage) du programme ciblé. Des méthodes telles que la régression logistique (Hosmer et Lemeshow 1989), l'analyse factorielle discriminante (Tomassone *et al.* 1988), et les arbres de régression et de discrimination CART (Breiman *et al.* 1984) sont classiquement utilisées pour ce type de problème et disponibles en standard dans la plupart des logiciels statistiques spécialisés comme SAS (utilisé pour mettre en œuvre les deux premières méthodes) ou S+ (dont nous avons utilisé les fonctions dédiées à la construction d'arbres de régression).

Dans le cadre de leur utilisation pour construire une règle de ciblage, on a comparé (Traissac 1995) ces trois méthodes sur la base de leurs caractéristiques statistiques intrinsèques ainsi que leurs performances sur un échantillon de 3852 ménages (enquête nationale budget consommation réalisée

par l'Institut National de la Statistique de Tunisie et la FAO⁵ en 1990). Se basant sur les données de cette enquête, qui comportait également un volet d'anthropométrie nutritionnelle, on a cherché à établir une règle pour le ciblage des adultes présentant une insuffisance pondérale, soit $IMC^6 < 18,5$ kg/m² (OMS 1995), à partir de caractéristiques socio-économiques simples du ménage (habitat, assainissement, environnement, taille du ménage, etc.) et de l'individu lui-même (âge, éducation, etc.), i.e. une cinquantaine de variables au total. Sont présentés dans la suite uniquement les résultats concernant les hommes adultes (n=3138 répartis en 2198 pour l'échantillon de base et 940 pour l'échantillon test, avec une prévalence d'insuffisance pondérale de 4,5 % sur l'échantillon de base⁷).

Pour chacune des trois méthodes, le modèle est ajusté sur l'échantillon de base tandis que les performances de la règle d'affectation, notamment la courbe ROC, le choix d'un seuil à partir de cette courbe, les valeurs de sensibilité, spécificité et valeurs prédictives qui en découlent, sont estimées à partir de l'échantillon test. Pour l'évaluation des performances de la règle d'affectation, on a posé des "coûts" égaux pour les deux types d'erreurs de classement. Par la suite on a donc choisi des seuils de coupure qui ne privilégient pas un des aspects par rapport à l'autre (valeurs de Se et Sp proches).

Régression logistique

Si l'on décrit l'insuffisance pondérale par une variable binaire ($Y = 1$ si $IMC < 18,5$, $Y = 0$ sinon), on peut se poser le problème de discrimination en terme de modèle de régression logistique de Y en fonction des descripteurs X_i du ménage et de l'individu. Dans le modèle de régression logistique, qui appartient à la famille des modèles linéaires généralisés, pour chaque observation, la valeur Y_i de la variable réponse est supposée être une réalisation d'une variable aléatoire binomiale $B(1, p_i)$ dont l'espérance $E(Y_i) = p_i = P(Y_i = 1 | X_{1i}, X_{2i}, \dots, X_{pi})$ dépend des variables explicatives

⁵Organisation des Nations Unies pour l'alimentation et l'agriculture

⁶IMC : indice de masse corporelle aussi appelé indice de Quételet.

IMC = poids / taille² (taille en m, poids en kg).

⁷La Tunisie présente une situation caractéristique des pays en transition nutritionnelle avancée avec un fort pourcentage de personnes en surpoids soit $IMC \geq 25$ (OMS 1995) et une faible proportion d'individus présentant une insuffisance pondérale. Ceci explique en grande partie que l'âge prédomine dans les critères de ciblage de la maigreur comme nous le verrons par la suite. Cela ne remet pas en cause l'intérêt des méthodes présentées. Il nous semblait important de les illustrer sur un exemple concret, basé sur des données fiables et représentatives à l'échelle nationale.

sous la forme d'une fonction logistique :

$$P(Y_i = 1 | X_{1i}, X_{2i}, \dots, X_{pi}) = \frac{e^{b_0 + \sum_{j=1}^p b_j X_{ji}}}{1 + e^{b_0 + \sum_{j=1}^p b_j X_{ji}}} \quad (1)$$

ou, ce qui est équivalent,

$$\text{Log}[p_i / (1 - p_i)] = b_0 + \sum_{j=1}^p b_j X_{ji} \quad (2)$$

Elle est adaptée à modéliser un certain nombre de situations où la variable réponse prend ses valeurs dans l'intervalle $[0, 1]$. Les variables X_i peuvent être indifféremment quantitatives ou les indicatrices de variables qualitatives. On peut également introduire dans le modèle des variables codant des interactions spécifiques. Un des intérêts de cette méthode pour les personnes de culture épidémiologique est le suivant : dans le cadre d'un modèle additif (*i.e.* en l'absence d'interactions) si b_i est le coefficient de la variable X_i dans le modèle, on peut montrer que $e^{b_i(d-c)}$ est l'odds-ratio⁸ mesurant la modification du risque (ici d'insuffisance pondérale) quand la valeur de X_i passe de c à d , ajusté pour la présence des autres variables dans le modèle. L'estimation des paramètres est faite par la méthode du maximum de vraisemblance. L'utilisation d'un algorithme de type pas à pas ascendant a conduit à retenir 6 variables. Aucun terme d'interaction n'a été inclus dans le modèle pour des raisons de simplicité. Dans une optique de prédiction, pour chaque individu, en utilisant les paramètres estimés, et à partir de ses valeurs pour les variables explicatives, on peut calculer la valeur de la fonction [1] qui lui est associée. On obtient une probabilité estimée d'insuffisance pondérale compte tenu de ses caractéristiques. On peut construire un score de ciblage simple d'emploi à partir des estimations des coefficients b_i ⁹ (cf. tableau 2).

La courbe ROC (cf. figure 3) correspondant aux divers choix de seuils possibles (probabilité à partir de laquelle on considère que l'individu doit

⁸Un indice d'association permettant de juger de l'effet d'un facteur de risque F dichotomique ($F+$: facteur présent, $F-$: facteur absent) sur une maladie M est le risque relatif $RR = P(M+/F+)/P(M+/F-)$. Dans la mesure où le RR n'est pas calculable dans tous les types d'études épidémiologiques, on utilise d'autres indicateurs tel que l'odds-ratio

$OR = [P(M+/F+)/P(M-/F+)]/[P(M+/F-)/P(M-/F-)]$. Bien que ce dernier indice ait dans tous les cas une interprétation spécifique, il est aussi souvent utilisé car il donne une bonne approximation du RR lorsque la maladie étudiée est rare dans la population (Bouyer *et al.* 1995). Ces notions se généralisent sans difficulté dans le cas d'une exposition à plusieurs niveaux.

⁹Les coefficients sont multipliés par 100 pour simplifier la présentation. Une valeur positive correspond à un risque de maigreur augmenté.

être ciblé) pour la règle d'affectation basée sur la régression logistique, a conduit à retenir une valeur associée aux estimations¹⁰ de Se , Sp , VPP , VPN présentées dans le tableau 3. Avec ce choix, tout individu présentant un total supérieur ou égal à 300 points est considéré comme devant bénéficier du programme ciblé.

Tableau 2 *Score de ciblage, régression logistique.*

Score de départ		+264
Individu		
Age < 20 ans	+99	
Activité nulle ou légère	+57	
Enfant ou petit enfant du chef de ménage	+52	
Ménage		
Nombre d'enfants dans le secondaire public ≥ 2	+52	
Radio	-58	
Machine à laver	-106	
	Total	

Analyse factorielle discriminante

Le vocable analyse discriminante regroupe un grand nombre de méthodes parfois assez différentes du point de vue de leurs fondements mathématiques. L'analyse factorielle discriminante utilisée ici est basée sur la recherche d'une combinaison linéaire des variables explicatives (appelé axe discriminant) maximisant un critère de séparation entre les deux groupes, par exemple le rapport variance inter / variance intra. L'affectation d'une unité statistique à l'un des deux groupes en fonction de ses valeurs pour les différentes variables se fait en calculant la coordonnée de sa projection sur l'axe discriminant puis en fonction du choix d'un seuil. En général l'affectation se fait au groupe dont elle est le plus proche du barycentre sur cet axe. Cette règle d'affectation géométrique (distance de Mahalanobis) a également une interprétation probabiliste (maximum de vraisemblance) si la distribution des variables explicatives est multinormale. Bien que cette méthode soit initialement prévue pour des variables quantitatives, dans une optique de ciblage on aura souvent un mélange de descripteurs quantitatifs et qualitatifs. Une possibilité simple est de réaliser un codage disjonctif

¹⁰Bien que nous ne les distinguons pas par des notations spéciales, les valeurs de ces paramètres présentées dans la suite sont obtenues à partir de l'échantillon considéré et non des valeurs en population telles que présentées au paragraphe 4, dont elles sont des estimations.

complet des variables qualitatives (c'est ce qui a été utilisé).

De même que pour la régression logistique, un algorithme de type pas à pas ascendant (i.e. avec introduction progressive des variables) a été utilisé pour rechercher la meilleure discrimination possible à partir d'un nombre réduit de variables : 5 variables ont été retenues. De même la construction d'un score de ciblage pour utilisation sur le terrain peut se faire simplement à partir des coefficients de la combinaison linéaire définissant l'axe discriminant. Nous ne présentons pas dans le détail les résultats obtenus avec cette méthode. Néanmoins, l'étude de la courbe ROC (cf. figure 3) correspondant aux divers choix de seuils possibles a conduit à retenir une valeur associée aux estimations de Se , Sp , VPP , VPN présentées dans le tableau 3.

Arbre de régression et/ou de classification (CART)

On cherche à prédire la variable dichotomique codant l'insuffisance pondérale à partir des variables explicatives, en utilisant un arbre binaire au lieu d'une équation de régression. Le principe général de toutes les méthodes de segmentation, dont CART fait partie, est de construire un tel arbre à partir des séries de nœuds que représentent l'ensemble, puis les sous-ensembles successifs, d'unités statistiques que délimite chaque segmentation. Les notions essentielles ici sont celles de coupure et d'impureté des nœuds :

- Coupure : étant donné une variable continue X , on appelle coupure un critère binaire du type $X < c$ (par exemple âge < 18) qui permet de répartir sans ambiguïté les individus en deux groupes : ceux pour lesquels $X < c$ et ceux vérifiant $X \geq c$. Si la variable X est qualitative une coupure sera définie par un sous-ensemble des modalités.
- Impureté d'un nœud : un nœud est dit d'autant plus pur, pour ce qui concerne la variable dichotomique étudiée (ici l'insuffisance pondérale), qu'il contient une classe plus largement majoritaire. Différents choix sont possibles pour l'expression mathématique de cette impureté. Il semble que ce choix n'ait en général pas grande influence sur les résultats.

A partir de ces éléments, l'algorithme fonctionne de la façon suivante : à l'étape 1 on dispose de l'ensemble des individus de l'échantillon de base qui constituent le premier nœud de l'arbre. On recherche alors la coupure qui permet de scinder en deux l'ensemble des individus de façon à obtenir

une réduction maximale de l'impureté¹¹. La recherche se fait de manière exhaustive par examen de toutes les variables explicatives et balayage de l'ensemble des valeurs prises par chacune de ces variables pour déterminer la coupure qui permette une réduction maximale de l'impureté. L'algorithme se déroule ensuite de manière récursive : pour chaque nouveau nœud ainsi créé, on recherche la question binaire qui favorise un découpage en deux nouveaux nœuds et conduise à une réduction maximale d'impureté etc. Il est à noter que dans chacun des deux nœuds créés lors d'une coupure, des variables différentes peuvent être utilisées pour les coupures suivantes : c'est-à-dire que les interactions entre variables sont gérées de manière implicite par la méthode, sans qu'il soit nécessaire de les spécifier comme c'est le cas par exemple dans un modèle de régression logistique lorsqu'on souhaite les prendre en compte¹².

Se pose le problème de l'arrêt du découpage : un arbre de bonne taille est un arbre fournissant le taux d'erreur de classement le plus petit possible avec le nombre de nœuds le plus réduit possible. Si l'on considère l'échantillon de base, le taux d'erreur le plus petit possible sera obtenu en poursuivant le découpage jusqu'à ce que chaque nœud terminal ne contienne qu'un seul individu. Ce choix n'est néanmoins pas acceptable. Il conduit à des arbres très grands, peu pratiques à utiliser. De plus, l'inconvénient majeur est que pour tout autre échantillon, l'utilisation de cet arbre risque de conduire à un taux d'erreur très important. Il est en effet peu probable de trouver deux échantillons ayant exactement la même structure. On procède de la façon suivante : on laisse se développer l'arbre assez loin, chacun des nœuds terminaux ayant un effectif très faible. Par utilisation d'un échantillon test ou par validation croisée, l'arbre est ensuite progressivement élagué par le bas en remontant. On obtient ainsi une suite d'arbres emboîtés que l'on compare pour choisir l'arbre de taille adéquate optimisant le critère choisi (taux d'erreur de classement par exemple).

¹¹Soit $i(t)$ l'impureté d'un nœud t et tg et td respectivement les nœuds gauche et droite créés par la coupure. La réduction d'impureté est $Di = i(t) - p(tg)i(tg) - p(td)i(td)$ où $p(tg)$ et $p(td)$ sont les proportions d'individus du nœud t qui tombent dans tg et td .

¹²Pour permettre à un modèle de régression (logistique ou autre) de prendre en compte de telles interactions, même si bien évidemment elles ne sont en final pas toutes retenues dans un modèle élaboré avec une procédure de type pas-à-pas, celles-ci doivent être spécifiées au départ pour tous les couples possibles de variables (soit $p(p-1)/2$) pour ce qui concerne seulement les interactions d'ordre 2. Lorsque l'on dispose d'un grand nombre de variables explicatives potentielles (une cinquantaine dans le cas de l'exemple présenté), cela conduit à des modèles qui ne sont pas faciles à spécifier.

Pour chacun des nœuds terminaux de l'arbre ainsi créé, on obtient une estimation de la probabilité d'insuffisance pondérale par la proportion d'individus de l'échantillon test présentant la caractéristique. Pour ce qui concerne la prédiction, partant du sommet, l'individu chemine le long de l'arbre en fonction des questions binaires et des valeurs qu'il prend pour chacune des variables retenues, jusqu'à ce qu'il arrive dans un nœud terminal. On obtient alors une estimation de la probabilité d'insuffisance pondérale de l'individu en fonction des variables explicatives. L'affectation à une des deux sous-populations est alors liée au choix du seuil fait après examen de la courbe ROC (cf. figure 3). Voir également le tableau 3 pour les caractéristiques de la règle d'affectation choisie.

Un graphique très simple d'emploi peut-être directement déduit de l'arbre de classification produit par la méthode (figure 2) pour savoir quelles sont les personnes qui devraient bénéficier d'un éventuel programme ciblé sur l'insuffisance pondérale.

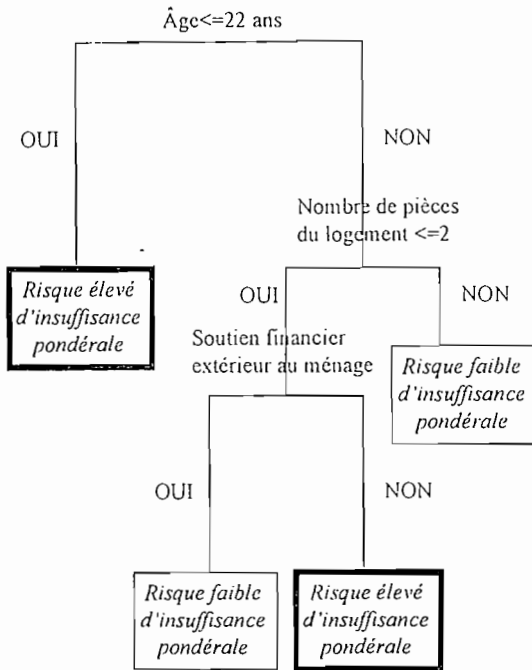


Figure 2

Arbre de décision pour le ciblage, CART.

Interprétation des résultats et comparaison des différentes méthodes :

Il n'est pas dans les objectifs de cette présentation de faire une interprétation détaillée des résultats obtenus : on peut néanmoins remarquer que pour les trois méthodes (y compris l'analyse factorielle discriminante dont les résultats ne sont pas présentés) sont mis en évidence une combinaison de critères individuels (âge, activité physique) et d'indicateurs indirects du niveau économique du ménage (taille du logement, possession de biens, etc.). De manière générale, si l'on rapproche les valeurs obtenues pour des indicateurs tels Se , Sp de celles habituellement obtenues pour des tests diagnostiques, celles-ci peuvent effectivement sembler faibles. Cela pourrait vouloir dire qu'on ne dispose pas nécessairement des informations adéquates pour le ciblage. Néanmoins, il faut nuancer cette conclusion car d'une part le contexte n'est pas le même, et d'autre part, il s'agissait d'utiliser au mieux dans une optique de ciblage, des données déjà existantes.

La comparaison fine des propriétés mathématiques des différentes méthodes ainsi que de leurs performances respectives sur de nombreux jeux de données (O'Gorman *et al.* 1991, Breiman *et al.* 1984) sort également du cadre de cette présentation. Néanmoins, dans le contexte qui nous occupe, pour ce qui concerne la possibilité de création de scores de ciblage simples d'emploi, la régression logistique et l'analyse discriminante semblent équivalentes. Par contre l'interprétation des paramètres de la régression logistique en terme d'odds-ratio, qui est une des raisons de la popularité de cette méthode chez les personnes de culture épidémiologique pourra la faire préférer dans certains contextes. Pour la présentation de résultats à un décideur et pour l'utilisation sur le terrain, la compréhension de l'arbre de régression/discrimination issu de CART est encore plus aisée pour des personnes qui n'ont pas une culture statistique et/ou épidémiologique.

Du point de vue du type de données, il est clair que l'on est dans les limites extrêmes de validité de la méthode lorsqu'on utilise l'analyse factorielle discriminante sur les indicatrices de variables qualitatives comme cela a été fait ici. La régression logistique et CART permettent par contre de gérer sans problèmes des prédicteurs quantitatifs et qualitatifs. L'avantage majeur de CART est de gérer de façon implicite la sélection des variables et les interactions entre les prédicteurs, avec tous les avantages que cela comporte en termes pratiques et de performances.

D'autre part, on pourrait s'attacher à comparer précisément les performances des trois méthodes, par exemple par des tests basés sur les surfaces délimitées par les courbes ROC (Hanley et McNeil 1982). Sans rentrer dans

tous les détails, en ce qui concerne l'exemple de la Tunisie, on peut commenter les performances respectives des méthodes telles que représentées par les courbes ROC (figure 3) et pour le choix des couples (Se, Sp) qui ont été faits (tableau 3) : si les résultats ne sont pas très différents d'une méthode à l'autre, l'analyse discriminante semble logiquement donner les résultats les moins bons (cf. remarque ci-dessus), suivi par la régression logistique et enfin CART. Il est moins aisé de comparer objectivement ces modèles sur la base de plusieurs critères simultanément tels que parcimonie, performances, lisibilité des résultats mais on peut remarquer que le modèle qui donne la meilleure prédiction (CART) est également celui qui utilise le moins de variables et dont la forme d'expression des résultats est la plus simple.

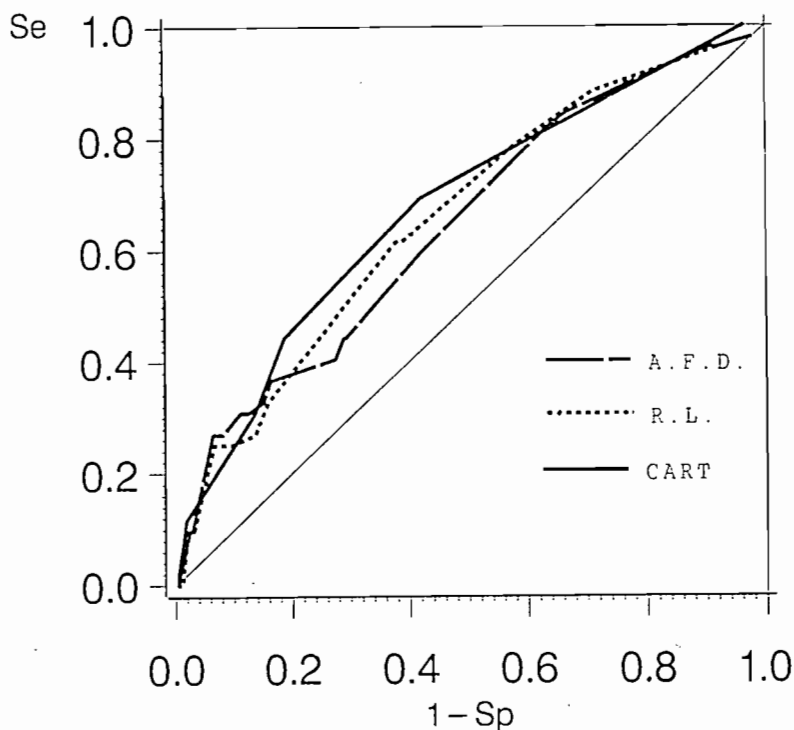


Figure 3

Comparaison des courbes ROC.

Tableau 3 : *Comparaison des performances des règles d'affectation (seuils choisis après examen des courbes ROC).*

	Se	Sp	k	VPP	VPN	k'
Analyse discriminante	0,60	0,58	1,41	0,08	0,95	1,41
Régression logistique	0,61	0,62	1,62	0,09	0,97	2,48
CART	0,69	0,58	1,64	0,09	0,97	2,83

Conclusion

Le ciblage des programmes et des interventions nutritionnelles peut contribuer à une meilleure équité dans un contexte général de ressources limitées et notamment de réduction des ressources affectées à ces programmes : crise économique, ajustement structurel. Il est souvent l'une des conditions pour l'accès de tous à la santé et à une alimentation adéquate en quantité comme en qualité.

Différents types de ciblage peuvent être envisagés ; le choix repose sur une caractérisation précise des groupes cibles, sur les indicateurs disponibles et les stratégies d'atteintes utilisables selon le contexte. Mais quel que soit le type de ciblage envisagé, celui-ci doit être basé sur des données fiables, récentes et représentatives, avec les problèmes de collecte spécifiques qui peuvent se poser dans les pays en développement. De même, quel que soit le type de ciblage, on doit se poser le problème de l'évaluation et de la mesure des performances de la stratégie de ciblage, en particulier en terme de coûts.

Les méthodes statistiques classiquement employées pour les problèmes de discrimination peuvent être utiles pour construire des outils de ciblage. Au-delà du type de données qu'il est possible de traiter et des performances, certaines caractéristiques comme par exemple la possibilité de calculer un odds-ratio, ou encore une meilleure lisibilité des résultats pour les décideurs, peuvent faire préférer une méthode à une autre suivant le contexte, comme cela a été illustré sur l'exemple tunisien présenté ci-dessus. D'autres méthodes peuvent encore être envisagées, telles que les méthodes neuronales. Elles devront être évaluées suivant les mêmes critères.

A chaque niveau de mise en oeuvre des méthodes statistiques (de l'observation jusqu'à l'analyse), au-delà des aspects purement techniques, le choix de l'une ou l'autre, dans les pays en développement, se fera souvent sur la base de critères et de contraintes spécifiques.

Références bibliographiques

- Bouyer J., Hémon D., Cordier S., Derrienic F., Stücker I., Stengel B., Clavel J. Epidémiologie. Principes et méthodes quantitatives. Les Editions INSERM, Paris, 1995.
- Breiman L., Friedman J.H., Olshen R.A., Stone J.C. Classification and Regression Trees. Wadsworth Inc., 1984.
- Dabis F., Drucker J., Moren A. (eds.), Epidémiologie d'Intervention. Arnette, Paris, 1992.
- Dureau F., Barabary O., Michel A., Lortic B. Sondages aréolaires sur image satellite pour des enquêtes socio-démographiques en milieu urbain. Manuel de formation. Orstom Ed., Paris, 1989.
- Grosh M.E. Administering targeted social programs in latin America. From platitudes to practice. The World Bank, Washington D.C., 1994.
- Habicht J-P., Mason J. B., Tabatabai H. Basic concepts for the design of evaluation during programme implementation. *In* : Methods for the Evaluation of the Impact of Food and Nutrition Programmes, Sahn D. E., Lockwood R. and Scrimshaw N., eds., The United Nations University, Tokyo; Food and Nutrition Bulletin, 1984; Suppl 8 :1-25.
- Hanley J. A., Mc Neil J. B. The meaning and use of the area under a Receiver Operating Characteristic (ROC) Curve. *Radiology*, 1982 ; 143 : 29-36.
- Hosmer D. W., Lemeshow S. Applied Logistic Regression. J. Wiley, New-York, 1989.
- Kumar K. Rapid appraisal methods. The World Bank, Washington D.C., USA; 1993.
- Lebart L., Salem A. Statistique Textuelle, Dunod, Paris, 1994.
- Maire B., Delpeuch F., Cornu A., Tchibindat F., Simondon F., Massamba J-P., Salem G., Chevassus-Agnès S., Urbanisation et transition nutritionnelle en Afrique sub-saharienne : les exemples du Congo et du Sénégal, *Rev. Epidémiol. Santé Publ.*, 1992 ; 40 : 252-58.
- Maire B., Delpeuch F., Padilla M., Le Bihan G. Le ciblage dans les politiques et les programmes nutritionnels. *In* : Les politiques alimentaires en Afrique du Nord ; d'une assistance généralisée aux interventions ciblées. 35-57, FAO/CIHEAM-IAMM/ORSTOM, Karthala, Paris, 1995.
- Miles M. B. and Huberman A. M. Qualitative data analysis. An expanded source-book. 2nd ed., SAGE Publ., Thousand Oaks, USA, 1994.
- O'Gorman T. W., Woolson R. F. Variable selection to discriminate between two groups : stepwise logistic regression or stepwise discriminant analysis ? *The American Statistician*, 1991 ; 45(3) : 187-93.

- OMS, Utilisation et interprétation de l'anthropométrie, Organisation Mondiale de la Santé, Genève, 1995.
- Pfeffermann D. The role of sampling weights when modeling survey data, *International Statistical Review*, 1993 ; 61 : 317-37.
- Pinstrup-Andersen P. Targeted nutrition interventions, *Food Nutr Bull*, 1991 ; 13(3) : 161-9.
- Popkin B. M. The nutrition transition in low-income countries : an emerging crisis. *Nutr Rev* 199 ; 52(9) : 285-98.
- Scrimshaw N. S. and Gleason G. R., (eds). RAP : rapid assessment procedures. Qualitative methodologies for planning and evaluation of health related programmes. International Food and Nutrition Foundation for Developing Countries (INFDC), Boston, Ma, USA; 1992.
- Tomassone R., Danzart M., Daudin J. J., Masson J-P. Discrimination et classement, Masson, Paris, 1988.
- Traissac P., Delpeuch F., Maire B., Martin-Prével Y., Cornu A., Trèche S. Construction d'un indice synthétique de niveau économique des ménages dans les enquêtes nutritionnelles. Exemples d'applications au Congo. *Rev. Epidemiol. Santé Publ.* 1997 ; 45(s1)114-5.
- Traissac P. Eléments de méthodologie statistique pour la discrimination et le classement. Application au ciblage des interventions. In : Les politiques alimentaires en Afrique du Nord; d'une assistance généralisée aux interventions ciblées. 393-431, FAO/CIHEAM-IAMM/ORSTOM, Karthala, Paris, 1995.
- Yip R., Fleshfood L., Spillman T.C., Binkin N.J., Wong F.L., Trowbridge F.L. Using linked program and birth records to evaluate coverage and targeting in Tennessee's WIC program. *Public Health Rep.*, 106(2) : 176-181, 1991.

Société Française de Biométrie

De l'observation à l'analyse,
implication de la biométrie
dans les pays
de développement

1998

Textes présentés lors des journées réalisées par
la Société Française de Biométrie
à Carcassonne, les 29 et 30 mai 1997

N° 15