

JSFS

Vie de la société

Journal de la société statistique de Paris, tome 136, n° 3 (1995), p. 3-15

http://www.numdam.org/item?id=JSFS_1995__136_3_3_0

© Société de statistique de Paris, 1995, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

I

VIE DE LA SOCIÉTÉ

PRIX DU DOCTEUR NORBERT MARX

Un prix **Norbert Marx** de 25 000 francs est décerné par la **Société de Statistique de Paris** récompensant les travaux d'un jeune chercheur (âgé de moins de 34 ans au 31 mai) "*contribuant à l'avancement des statistiques médico-sociales ou interprétant de telles statistiques*". Plus précisément, ces travaux doivent :

- au plan **thématique**, être de nature à faire progresser les connaissances soit sur les causes médico-sociales des maladies (aspects médicaux), soit sur les coûts et avantages d'une politique sociale de santé (aspects économiques) ;
- au plan **méthodologique**, reposer à la fois sur l'utilisation de données statistiques relevant du domaine de la santé et de techniques statistiques permettant de traiter efficacement ces données.

Le Jury comprend les neuf membres suivants de la SSP : son Président (Président du Jury), ses trois vice-Présidents, ses trois précédents Présidents, sa Secrétaire Générale et le Rédacteur en chef du *Journal* de la SSP. Le Jury demandera l'**avis de spécialistes reconnus dans le domaine**, membres ou non de la SSP, pour évaluer la valeur scientifique des travaux des candidats. (A l'avenir, les rapporteurs extérieurs feront partie du jury).

Le Jury se réunit à l'automne pour désigner le (ou les) lauréat(s).

Les deux articles qui suivent présentent et résument les travaux des deux lauréats de la première édition du Prix.

TESTS D'HOMOGENÉITÉ POUR LES MODÈLES LINÉAIRES GÉNÉRALISÉS.

Application à l'étude des variations géographiques du déficit cognitif chez les personnes âgées

Hélène JACQMIN-GADDA
INSERM U330 – Université de Bordeaux II

En épidémiologie, les données recueillies sont fréquemment groupées (par sujets, par familles, par zones géographiques, par hôpitaux...) et les observations appartenant à un même groupe ne sont généralement pas indépendantes. Ce problème a suscité le développement de modèles de régression pour données corrélées qui permettent d'étudier la relation entre une variable réponse et plusieurs variables explicatives en tenant compte de la corrélation intragroupe. Cependant peu de méthodes ont été proposées pour tester l'homogénéité de la distribution de la variable réponse entre les groupes.

L'objectif de ce travail est de développer, dans le cadre de la famille de distribution exponentielle, un test d'homogénéité entre groupes qui permette l'ajustement sur des facteurs de confusion potentiels, qu'il s'agisse de variables caractérisant les groupes ou de variables caractérisant les sous-unités. Cette méthode permet par exemple de tester l'existence d'une corrélation intrafamille ou d'une hétérogénéité entre zones géographiques pour une pathologie donnée. D'autre part, en testant l'indépendance des observations, ces procédures aident à choisir la meilleure méthode d'analyse entre un modèle pour données indépendantes et un modèle pour données corrélées toujours plus lourd à mettre en œuvre. Le premier test proposé est un test du score d'homogénéité développé dans le cadre des modèles linéaires généralisés à effets mixtes, le deuxième est un test analogue au test du score pour les modèles marginaux.

Les modèles à effets mixtes et les modèles marginaux sont les deux principaux types de modèles de régression pour données corrélées. Les modèles marginaux (Liang et Zeger, 1986) permettent l'étude de la distribution marginale de la variable réponse en considérant la corrélation intragroupe comme une nuisance. Les modèles à effets mixtes (Stiratelli, Laird et Ware, 1984) permettent l'étude de la distribution de la variable réponse conditionnellement à un effet aléatoire qui traduit l'hétérogénéité entre les groupes. Les modèles conditionnels constituent une troisième approche, mais ils ne sont utilisables que lorsque les groupes sont de petites tailles ou lorsque les observations sont ordonnées (étude longitudinale).

I Test du score d'homogénéité pour les modèles à effets mixtes

Soit Y_{ij} la variable réponse pour l'observation j , $j = 1, \dots, n_i$ du groupe i , $i = 1, \dots, k$. On suppose que Y_{ij} suit une distribution de la famille exponentielle (McCullagh et Nelder, 1989) :

$$f_{ij}(Y_{ij}; \beta, \phi, D) = \exp \left\{ \phi^{-1} [\theta_{i,j} Y_{ij} - g(\theta_{i,j})] + C(Y_{ij}, \phi) \right\} \quad (1)$$

La définition du modèle linéaire généralisé à effets mixtes est complétée par :

$$E(Y_{ij} | \alpha_i) = \mu_{i,j} = g'(\theta_{i,j}), \quad \theta_{i,j} = X_{i,j}^T \beta + Z_{i,j} \alpha_i \text{ et } \alpha_i = \alpha + D^{1/2} v_i.$$

Les v_i sont des variables i.i.d. de fonction de répartition G , de moyenne 0 et de variance 1. Le vecteur β est le vecteur des effets fixes, $X_{i,j}$ est un vecteur de variables explicatives et $Z_{i,j}$ est une variable explicative associée à l'effet aléatoire.

L'hypothèse nulle d'homogénéité entre les groupes, ou d'indépendance des observations, est définie par " $H_0 : D = 0$ " et l'hypothèse alternative est " $H_1 : D > 0$ ". La statistique du score pour le paramètre D peut être décomposée en deux termes :

$$S(\beta, \phi) = S_1(\beta, \phi) + S_2(\beta, \phi) \quad (2)$$

où

$$S_1(\beta, \phi) = \sum_{i=1}^k \sum_{j=1}^{n_i} \sum_{j' \neq j}^{n_i} (Y_{i,j} - \mu_{i,j})(Y_{i,j'} - \mu_{i,j'}) Z_{i,j} Z_{i,j'} \quad (3)$$

et

$$S_2(\beta, \phi) = \sum_{i=1}^k \sum_{j=1}^{n_i} [(Y_{i,j} - \mu_{i,j})^2 - \sigma_{i,j}^2] Z_{i,j}^2 \quad (4)$$

S_1 est une statistique de corrélation par paire et S_2 est une statistique de surdispersion qui compare deux estimateurs de la variance, la variance empirique de l'échantillon et la variance dérivée du modèle sous H_0 .

On démontre que les statistiques :

$H_{S_1} = S_1(\hat{\beta}) / \text{Var}(S_1)^{1/2}$, $H_{S_2} = S_2(\hat{\beta}) / \text{Var}(S_2)^{1/2}$ et $H_S = S(\hat{\beta}) / \text{Var}(S)^{1/2}$ suivent asymptotiquement une loi normale centrée réduite. Les variances asymptotiques de $S_1(\hat{\beta})$, $S_2(\hat{\beta})$ et $S(\hat{\beta})$ sont données dans Jacqmin-Gadda et Commenges (1995) et dans Jacqmin-Gadda (1994), pour le cas où le paramètre de dispersion ϕ est connu et le cas où il doit être estimé. Le test d'homogénéité proposé consiste à rejeter l'hypothèse nulle lorsque la statistique H_S est supérieure à la valeur seuil indiquée par la table de l'écart-réduit pour un test unilatéral.

Ce test d'homogénéité présente deux avantages majeurs par rapport au test du rapport des vraisemblances proposé par Self et Liang (1987). D'une part, il ne nécessite pas l'estimation du modèle à effets mixtes puisque tous les paramètres sont estimés sous H_0 ; d'autre part, il est indépendant de la distribution de l'effet aléatoire.

II Test d'homogénéité pour les modèles marginaux

Définissons un modèle linéaire généralisé marginal. La distribution de la variable réponse appartient à la famille exponentielle (1). L'espérance marginale des Y_{ij} est précisée par $E(Y_{ij}) = \mu_{ij} = g'(\theta_{ij})$ et $\theta_{ij} = X_{ij}^T \beta$, et la covariance des Y_{ij} est spécifiée par une matrice $n_i \times n_i$, $V_{i,11}(\delta)$ qui dépend d'un paramètre δ . On introduit un second ensemble de variables $s_{ijj'} = (Y_{ij} - \mu_{ij})(Y_{ij'} - \mu_{ij'})$ d'espérances $\Sigma_{ijj'}(\delta) = \text{cov}(Y_{ij}, Y_{ij'})$. Les paramètres de régression β et de corrélation δ peuvent être estimés en résolvant les équations d'estimation généralisées (Liang et Zeger, 1986 ; Prentice et Zhao, 1991), soit :

$$\sum_{i=1}^k D_{i,11}^T V_{i,11}^{-1} (Y_i - \mu_i) = 0 \quad (5)$$

$$\sum_{i=1}^k D_{i,22}^T V_{i,22}^{-1} (s_i - \Sigma_i) = 0 \quad (6)$$

où

$$s_i = (s_{i12}, \dots, s_{i1n_i}, s_{i23}, \dots, s_{in_i-1n_i}, s_{i11}, \dots, s_{in_i n_i})^T,$$

$$\Sigma_i = (\Sigma_{i12}, \dots, \Sigma_{i1n_i}, \Sigma_{i23}, \dots, \Sigma_{in_i-1n_i}, \Sigma_{i11}, \dots, \Sigma_{in_i n_i})^T,$$

$$D_{i,11} = \frac{\partial \mu_i}{\partial \beta} \text{ et } D_{i,22} = \frac{\partial \Sigma_i}{\partial \delta}.$$

Σ_i est la matrice $V_{i,11}$ sous forme vectorielle, $V_{i,22}$ est la matrice de covariance de travail des s_i . Sous H_0 , l'équation (6) se réduit à :

$$T(\beta, \phi) = T_1(\beta, \phi) + T_2(\beta, \phi) \quad (7)$$

avec

$$T_1(\beta, \phi) = \sum_{i=1}^k \sum_{j=1}^{n_i} \sum_{j' \neq j}^{n_i} D_{i,22,jj'}(\delta_0) \frac{(Y_{ij} - \mu_{ij})(Y_{ij'} - \mu_{ij'})}{\sigma_{ij}^2 \sigma_{ij'}^2} \quad (8)$$

$$T_2(\beta, \phi) = \sum_{i=1}^k \sum_{j=1}^{n_i} 2D_{i,22,jj}(\delta_0) \frac{(Y_{ij} - \mu_{ij})^2 - \sigma_{ij}^2}{\mu_{4ij} - \sigma_{ij}^4} \quad (9)$$

où μ_{4ij} désigne le moment d'ordre 4 de Y_{ij} .

Par analogie avec le test du score, on peut prendre $T(\hat{\beta}, \hat{\phi})$ comme statistique de test pour $\delta = 0$, comme l'ont suggéré Liang, Zeger et Qaqish (1992). Les statistiques $H_{T_1} = T_1(\hat{\beta})/\text{Var}(T_1)^{1/2}$, $H_{T_2} = T_2(\hat{\beta})/\text{Var}(T_2)^{1/2}$ et $H_T = T(\hat{\beta})/\text{Var}(T)^{1/2}$ suivent asymptotiquement une loi normale centrée réduite. Les variances asymptotiques sont données dans Jacqmin-Gadda et Commenges (1995) et dans Jacqmin-Gadda (1994).

Cette deuxième approche permet de proposer un test d'homogénéité en précisant l'hypothèse alternative par une structure de covariance au lieu d'un modèle à effet aléatoire. Lorsque l'alternative peut être représentée dans les deux approches, les deux statistiques de test sont identiques.

III Propriétés

Une étude par simulations a permis de mettre en évidence les avantages de ce test par rapport aux tests précédemment proposés (Self et Liang, 1987 ; Hamerle, 1990). Nous avons également montré que le risque de première espèce du test H_S peut augmenter considérablement lorsque les données présentent une surdispersion de nuisance (c'est-à-dire non due au regroupement des données étudié), alors que le test H_{S_1} est plus robuste. De plus, les puissances des deux tests sont peu différentes. On recommande donc l'utilisation du test H_{S_1} lorsque l'hypothèse alternative intéressante est l'hétérogénéité entre les groupes.

IV Application

Le test H_S a été utilisé pour étudier les variations géographiques du risque de déficit cognitif chez les personnes âgées, à partir des données de l'enquête Paquid réalisée à l'unité INSERM 330. L'échantillon est constitué de 3777 sujets de 65 ans et plus vivant à leur domicile dans 75 communes de Gironde et de Dordogne. Le déficit cognitif est défini comme un score inférieur à 24 sur 30 points au test Mini-Mental State Examination (Folstein *et al.*, 1975).

Sans ajustement, la statistique H_S prend la valeur 14.36, qui permet de rejeter l'hypothèse d'homogénéité du risque de déficit cognitif entre les communes ($p \ll 0.001$). Cependant, lorsque l'on ajuste sur l'âge, le sexe, le niveau d'études et la profession, la valeur de la statistique H_S chute à 1.948 ($p = 0.03$). On montre ainsi que l'essentiel des variations géographiques du risque de déficit cognitif observées dans Paquid est dû à une distribution hétérogène entre les communes de deux variables caractérisant les sujets : le niveau d'études et la profession.

A l'aide d'un modèle logistique à effets mixtes, nous avons ensuite montré que l'hétérogénéité résiduelle pouvait être expliquée par une relation entre le risque de déficit cognitif et la composition de l'eau de boisson de la commune : le risque de déficit cognitif semble augmenter lorsque la concentration en calcium de l'eau de boisson diminue, et la concentration en aluminium apparaît associée positivement avec le risque de déficit cognitif lorsque le pH de l'eau est faible, et négativement lorsque le pH est élevé (Jacqmin *et al.*, 1994).

V Perspectives

Des travaux supplémentaires sont nécessaires pour proposer une méthode générale qui permette de tester n'importe quelle structure de corrélation sur un échantillon. Par exemple, dans une étude de variation géographique, on pourrait envisager l'existence d'une corrélation entre des zones voisines qui dépende de la distance qui les sépare.

BIBLIOGRAPHIE

- FOLSTEIN M.F., FOLSTEIN S.E. et MCHUGH P.R. (1975) "Mini-Mental State. A Practical Method for Grading the Cognitive State of Patients for the Clinicians", *Journal of Psychiatric Research*, 12, pp 189-198.
- HAMERLE A. (1990) "On a Simple Test For Neglected Heterogeneity in Panel Studies", *Biometrics*, 46, pp 193-198.
- JACQMIN H., COMMENGES D., LETENNEUR L., BARBERGER-GATEAU P. et DARTIGUES J.F. (1994) "Components of Drinking Water and Risk of Cognitive Impairment in the Elderly", *American Journal of Epidemiology*, 139, pp 48-57.
- JACQMIN-GADDA H. (1994) *Tests d'Homogénéité pour les Modèles linéaires généralisés. Application à l'Etude des Variations géographiques du Déficit Cognitif chez les personnes âgées*, thèse de Biomathématiques, Université Paris 7.
- JACQMIN-GADDA H., COMMENGES D. (1995) "Tests of Homogeneity for Generalized Linear Models", *Journal of the American Statistical Association*, sous presse.
- LIANG K.Y. et ZEGER S.L. (1986) "Longitudinal Data Analysis Using Generalized Linear Models", *Biometrika*, 73, pp 13-22.
- LIANG K.Y., ZEGER S.L. et QAQISH B. (1992) "Multivariate Regression Analyses for Categorical Data", *Journal of the Royal Statistical Society B*, 54, pp 3-40.
- MCCULLAGH P. et NELDER J.A. (1989) *Generalized Linear Models*, New York, Chapman & Hall.
- PRENTICE D.E. et ZHAO L.P. (1991) "Estimating Equations for Parameters in Means and Covariance of Multivariate Discrete and Continuous Responses", *Biometrics*, 47, 825-839.
- SELF S. et LIANG K.Y. (1987) "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions", *Journal of the American Statistical Association*, 82, 605-610.
- STIRATELLI R., LAIRD N. et WARE J.H. (1984) "Random-Effects Models for Serial Observations with Binary Response", *Biometrics*, 40, 961-971.

**EVALUATION DES ACTIONS
DE DÉPISTAGE :
ASPECTS EPIDÉMIOLOGIQUES
ET ECONOMIQUES**

**Application à l'évaluation
du cancer côlo-rectal**

Benoît DERVAUX¹

Centre de Recherches Economiques, Sociologiques
et de Gestion de l'Université Catholique de Lille

Notre travail de thèse traite des problèmes épidémiologiques et économiques rencontrés lorsque l'on cherche à évaluer des actions de prévention en général, des programmes de dépistage en particulier. Impliqués depuis quelques années dans le travail d'évaluation de la campagne de dépistage du cancer côlo-rectal en région Nord - Pas-de-Calais - Picardie, nous retenons comme terrain d'application le dépistage de cette pathologie par recherche de sang occulte dans les selles.

La rapidité du progrès médical, notamment en médecine prédictive, et le resserrement des contraintes budgétaires rendent inéluctable et souhaitable, y compris d'un point de vue éthique, l'application du calcul économique au domaine sanitaire. L'évaluation économique des programmes de prévention s'avère particulièrement délicate car l'établissement de la preuve épidémiologique est difficile (il faut se prémunir de nombreux biais) et exige un temps considérable (il peut s'écouler 10 ou 20 ans entre le lancement d'une étude et la publication des résultats). Dans ces conditions, le recours à la modélisation permet non seulement de juger plus rapidement le bien-fondé d'un programme mais aussi de comparer différentes modalités de mise en oeuvre de ce dernier et de sélectionner la plus pertinente au regard des bénéfices de santé engendrés et des coûts supportés.

Dans notre travail de thèse, nous proposons une typologie des modèles d'évaluation des programmes de dépistage organisée autour des deux interrogations suivantes :

1. Thèse soutenue aux Facultés Universitaires Catholiques de Mons (B) en Avril 1992. Membres du jury : L. Eeckhoudt (Président), A. Bultez, G. Duru, J.-Cl. Sailly et G. Sturbois.

1. en premier lieu, comment le modèle décrit-il l'histoire naturelle de la maladie ? L'analyse de la littérature fait apparaître deux types de modèles : les premiers décrivent de manière continue le processus d'évolution de la maladie (par exemple, il s'agira en oncologie de décrire la croissance d'une cellule cancéreuse) [pour illustration, voir M. SHWARTZ, 1978] contrairement aux seconds qui, de manière discrète, considèrent uniquement les moments auxquels les différents tests parviennent à dépister la maladie [voir par exemple D. EDDY, 1980].

2. en second lieu, cherche-t-on à optimiser la distribution des tests sur le cycle de vie d'un individu donné (individu moyen ou possédant certaines caractéristiques) [R. KIRSH et M. KLEIN (1974), A. DWYER et alii (1983), D. KENT et alii (1987), S. OZEKICI et S. PLISKA (1991)] **ou se contente-t-on de calculer les bénéfices et les coûts induits en distinguant différents scenarii** [D. EDDY (1980), N. DAY et S. WALTER (1984)] ? La réponse à cette question amène à distinguer les modèles purement épidémiologiques des modèles économiques. Dans chaque groupe, nous classons les différents modèles en fonction de la pertinence des hypothèses sur lesquelles ils sont construits et, de façon concomitante, leur complexité.

Notre second apport renvoie à la conduite des analyses coût-efficacité. Nous montrons qu'il est possible de mener à bien ce type d'analyses lorsque les indicateurs de résultats sont multiples et difficilement agrégeables comme cela s'avère fréquent en matière de dépistage.

Nous établissons une correspondance étroite entre les concepts spécifiques utilisés en matière d'évaluation économique des programmes de santé (on parle d'analyses coût-efficacité, coût-utilité ou coût-bénéfice) et ceux de la théorie micro-économique du producteur (minimisation du coût, maximisation du revenu et/ou du profit). Si la transposition des concepts s'avère possible, il en va probablement de même pour les méthodologies d'estimation des fonctions de production, de coût et de profit. Ainsi, nous transposons au domaine de l'évaluation des programmes de santé une méthode d'estimation non paramétrique d'une frontière de production déterministe : la méthode DEA ("Data Envelopment Analysis"), méthode d'enveloppement des données fondée sur la programmation linéaire particulièrement bien adaptée au secteur non marchand caractérisé par l'absence de prix véritables, tant pour les facteurs que pour les produits.

Dans cet article, nous présentons l'analyse coût-efficacité du dépistage de masse du cancer côlo-rectal par Hémocult IITM que nous avons réalisée. Dans les deux premières sections, nous traitons des questions méthodologiques soulevées par ce travail, à savoir la détermination de l'ensemble des stratégies de dépistage pertinentes et la comparaison de l'efficacité de stratégies différentes lorsque cette efficacité se traduit dans

des indicateurs de résultat multiples difficilement agrégeables. Dans la troisième section, nous exposons les résultats auxquels cette analyse conduit.

Détermination des stratégies diagnostiques admissibles

Il s'agit du premier travail déterminant d'une analyse coût-efficacité. Cette opération est capitale car l'oubli d'une stratégie peut entraîner la ruine des conclusions avancées. Lorsque le nombre de tests à la disposition du milieu médical est important et qu'il est possible de les combiner, il faut concilier le souci de n'oublier aucune stratégie avec la préoccupation de ne pas avoir un nombre de stratégies trop important. Deux solutions s'offrent pour concilier ces deux objectifs : ne garder dans l'analyse que les stratégies se conformant aux pratiques médicales habituellement adoptées (mais on peut alors craindre l'existence d'un biais de sélection lié à l'existence d'"écoles", à l'imperfection de la tarification des actes...); définir un protocole de sélection des stratégies et ne retenir que celles répondant aux critères contenus dans ce protocole (critères qui, rendus explicites, peuvent être discutés). Dans ce travail, nous privilégions clairement la seconde méthode.

Pour diagnostiquer cancers et polypes, les gastro-entérologues disposent de divers examens : le lavement baryté en double contraste (LBDC), la sigmoïdoscopie et la coloscopie. Pour un dépistage de masse, compte tenu de l'importance de la population à surveiller (les 45-75 ans, en fonction des courbes d'incidence), ces examens s'avèrent trop invasifs, trop dangereux et/ou trop coûteux. Dès lors, le recours au test Hémocult IITM constitue, aux yeux des experts, une solution adaptée malgré la faible sensibilité de ce test dans le dépistage des petits polypes. Nous supposons donc que toute action préventive de masse utilise ce test en première intention. La question que nous traitons devient, dès lors, la suivante : *sur une population répondant positivement au test Hemocult IITM, comment utiliser efficacement les moyens diagnostics dont on dispose (coloscopie, sigmoïdoscopie, LBDC) ?*

Si on admet qu'un patient peut recevoir plusieurs fois un même test, le nombre de séquences d'examen pouvant être pratiquées combinant le LBDC, la sigmoïdoscopie ou la coloscopie devient très grand, voire infini. Aussi convient-il de restreindre au maximum ce nombre, d'autant que de nombreuses séquences sont, à l'évidence, sans intérêt médical ou impraticables. Pour réaliser cette sélection, nous avons adopté les critères de sélection ci-dessous :

- après une coloscopie (complète ou non), le LBDC et la sigmoïdoscopie perdent tout intérêt ;

– après un LBDC positif, un examen endoscopique s'impose car la coloscopie et la sigmoïdoscopie autorisent certains gestes thérapeutiques essentiels (polypectomies, biopsie sur la tumeur) ;

– un patient ne subit jamais deux fois un même test (compte tenu de l'inconfort lié à la réalisation des examens).

L'application de ces trois critères "de bon sens" à l'ensemble de toutes les stratégies possibles laisse subsister 18 stratégies diagnostiques.

Une fois cette sélection réalisée, il convient de rassembler pour chaque stratégie deux types d'éléments : le nombre de lésions (cancers et polypes) qu'elles permettent de découvrir (l'efficacité) et le nombre d'examens qu'elle entraînent (le coût).

Estimation de la fonction de coût minimal

S'il est possible de mesurer le coût et l'efficacité d'un ensemble de stratégies à l'aide de critères univoques, la fonction de coût minimal se définit alors comme la courbe réunissant les stratégies ou les combinaisons de stratégies telles que pour un coût donné, on obtienne le maximum d'efficacité ou, inversement, pour une efficacité donnée, on doive consentir le minimum de coût. Cette courbe constitue l'enveloppe des stratégies coût-efficaces.

Lorsque l'efficacité des stratégies se mesure à l'aide d'indicateurs non agrégeables, la construction de cette courbe enveloppe doit s'effectuer dans un espace à plusieurs dimensions. A chaque stratégie diagnostique correspondent désormais deux vecteurs : le premier décrivant les examens réalisés, le second dénombrant les divers résultats obtenus en termes de santé. Dans le cas du dépistage CCR, si la détermination du premier vecteur ne pose pas de difficulté (il s'agit du dénombrement des LBDC, des sigmoïdoscopies et des coloscopies effectués), l'estimation du résultat soulève des problèmes délicats quant à la nature des indicateurs retenus et quant à leur agrégation. Pour notre part, le résultat de santé que nous tentons de mesurer est un résultat intermédiaire, à savoir le nombre de lésions décelées (polypes de différentes tailles et au potentiel de dégénérescence variable, cancers), l'absence de données nous interdisant d'estimer le résultat final en termes d'années de vie gagnées.

Afin de construire la courbe enveloppe permettant de désigner les stratégies diagnostiques offrant le meilleur "rendement", nous avons adapté une méthodologie proposée par CHARNES, COOPER et RHODES (1978) à partir des travaux de FARELL (1957). Ces derniers ont mis au point une méthode autorisant l'évaluation de l'efficacité productive des entreprises non marchandes et la qualité de leur gestion en comparant

leurs dotations factorielles et leurs productions sans faire référence à un quelconque système de prix (ou de pondération). Cette technique (appelée DEA), bien que développée en vue d'applications fort éloignées de notre champ, paraît bien adaptée à notre problème.

Soit une entité j ; dénommons $X_{i,j}$ sa dotation en facteur i , et $Y_{r,j}$ sa production en bien r . Cette dernière sera déclarée efficace si et seulement si le ratio $X_{i,j}/Y_{r,j}$ s'avère minimal pour au moins un intrant (i) par rapport aux autres entités (j) réelles ou fictives², produisant au moins $Y_{r,j}$.

Le programme linéaire correspondant à cette définition s'écrit de la manière suivante :

Min μ_0

sous les contraintes :

$$1) \mu_0 X_{i0} - \sum \mu_j X_{i,j} - S_i^+ = 0 \text{ pour tout } i ;$$

$$2) Y_{r0} - \sum \mu_j Y_{r,j} + S_r^- = 0 \text{ pour tout } r ;$$

$$3) \sum \mu_j \leq 1$$

avec les solutions optimales : μ_0^* , μ_j^* , S_i^{+*} , S_r^{-*} .

Dans ce problème, l'indice 0 correspond à l'entité courante qui se trouve comparée à toutes les autres, dont elle-même (il faut donc résoudre n programmes linéaires de ce type), et ce de manière à garantir l'existence d'une solution et à borner supérieurement le coefficient d'efficacité μ_0 à l'unité. Les contraintes 1) interdisent à la solution optimale de consommer trop d'intrants ; les contraintes 2) l'obligent à une production au moins égale à celle de l'entité 0 et ce, pour chaque bien considéré isolément.

La solution optimale ci-dessous caractérise une entité techniquement efficace :

- $\mu_0^* = 1$: en d'autres termes, l'utilisation d'au moins un facteur est minimale ;

- $S_i^{+*} = 0$, $S_r^{-*} = 0$ (pour tout i et r) : les autres intrants ne sont pas "gaspillés" et le niveau de production est maximal. Il s'agit donc d'un optimum de Pareto.

A l'inverse, nous aurons pour une entité inefficace :

- $\mu_0^* < 1$: on peut réduire la dotation factorielle de l'entité en appliquant à tous les intrants le facteur de réduction $(1 - \mu_0^*)$ sans que cela affecte le niveau de production atteint ;

2. Fictives, en ce sens qu'elles constituent des combinaisons linéaires de plusieurs entités.

– $S_i^{+*} > 0$ et/ou $S_r^{*-} > 0$ (pour quelques i ou r) : certaines dotations factorielles peuvent être réduites au delà de cette contraction radiale.

En résumé pour toute entité efficace μ_0^* prend la valeur un et toutes les variables d'écart s'annulent.

Pour reconstituer les éléments de coût et d'efficacité de chaque stratégie, nous utilisons les données issues de la littérature ou des travaux d'évaluation déjà menés en matière d'évaluation du dépistage du CCR et élaborons plusieurs hypothèses.

Les données collectées concernent la prévalence des lésions recto-coliques dans une population présentant un test HémoCult IITM positif, la sensibilité et la spécificité des différents tests, la localisation des lésions le long du cadre colique, enfin la taille et le potentiel de dégénérescence des polypes en fonction de leur taille, de leur histologie ou de leur degré de dysplasie. Nous admettons au maximum la coexistence de deux cancers et de trois polypes chez un même patient, supposons que la sensibilité du LBDC et de l'HémoCult IITM varie seulement avec la taille des polypes et que la sensibilité des examens endoscopiques dépend uniquement de la répartition des lésions le long du cadre colique et de l'expérience de l'endoscopie.

Pour chacune des 18 stratégies retenues, en balayant les arbres de décision, nous pouvons calculer le nombre d'examens prescrits et le nombre de lésions dépistées, polypes et cancers. Une fois effectué ce travail, nous pouvons utiliser la technique d'enveloppement des données présentée ci-dessus pour déterminer les stratégies décelant le plus de lésions au moindre coût.

Résultats

Parmi les 18 stratégies diagnostiques recensées, 7 apparaissent techniquement inefficaces. Quel que soit le système de prix relatif des examens, ces 7 stratégies s'avèrent toujours dominées par une combinaison linéaire des autres séquences diagnostiques. En introduisant le prix des examens selon deux modes de valorisation, le coût de production réel obtenu à partir d'analyses monographiques en milieu hospitalier et les prix officiels issus de la nomenclature générale des actes professionnels, nous montrons que seules trois stratégies s'avèrent coût-efficaces : la sigmoïdoscopie d'emblée, la sigmoïdoscopie suivie d'une coloscopie si l'examen de première intention détecte une lésion recto-colique, la coloscopie d'emblée. Ces résultats confirment l'intérêt des examens endoscopiques, notamment de la coloscopie, et interrogent quant à la pertinence du LBDC dans le diagnostic des lésions côlo-rectales. A cet égard, la sous-tarifcation des examens endoscopiques en général, de la coloscopie en particulier, pose deux problèmes : elle diminue artificiellement

le coût marginal de la stratégie "coloscopie d'emblée" et maintient la possibilité de prescription d'examens radiologiques se caractérisant par une faible efficacité technique. D'un point de vue méthodologique, cette étude démontre la possibilité de mener à bien des analyses de type coût-efficacité lorsque les indicateurs d'efficacité sont multiples et difficilement agrégeables avec la méthode DEA. En outre, l'utilisation de cette technique d'enveloppement des données permet de réaliser correctement les analyses de sensibilité post-optimales pour juger de la robustesse des résultats obtenus eu égard à l'incertitude qui entâche les données épidémiologiques et économiques communément utilisées dans ce type d'évaluation. La programmation linéaire stochastique constitue une autre façon d'intégrer l'incertitude dans l'analyse, nous continuons à travailler dans ce sens.

BIBLIOGRAPHIE

- CHARNES A., COOPER W.W., RHODES E. (1978) "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, 2 : pp 429-444.
- DAY N., WALTER S. (1984) "Simplified Models for Screening for Chronic Disease : Estimation procedures from Mass Screening Programmes", *Biometrics*, 40 : pp 1-14.
- DWYER A., PREWITT J., ECKER J., PLUNKETT J. (1983) "Use of the Hazard Rate to Schedule Follow-up Exams Efficiently. An Optimization Approach to Patient Management", *Medical Decision Making*, 3(2) : pp 229-244.
- EDDY D. (1980) *Screening for Cancer : Theory, Analysis and Design*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 304 p.
- FARELL M.J. (1957) "The Measurement of Productive Efficiency", *Journal of Royal Statistical Society*, A 120 : pp 253-281.
- KENT D., SHACHTER R., SOX H. ET ALII (1987) "Efficient Scheduling of Cystoscopies in Monitoring for Recurrent Bladder Cancer", *Medical Decision Making*, 9(1) : pp 26-37.
- KIRSH R., KLEIN M. (1974) "Surveillance Schedules for Medical Examinations", *Management Science*, 20(10) : pp 1403-1409.
- OZEKICI S., PLISKA S. (1991) "Optimal Sheduling of Inspections : A delayed Markov Model with False Positives and Negatives", *Operations Research*, 39(2) : pp 261-273.
- SHWARTZ M. (1978) "A Mathematical Model used to Analyse Breasr Cancer Screening Strategies", *Operations Research*, 26(6) : 937-954.